⌘ *Author's Choice*

Quantitative Proteomics Analysis of the Nuclear Fraction of Human CD4$^+$ Cells in the Early Phases of IL-4-induced Th2 Differentiation*[S]

Robert Moulder‡§, Tapio Lönnberg‡§‖, Laura L. Elo‡‖, Jan-Jonas Filén‡¶, Eeva Rainio‡, Garry Corthals‡, Matej Oresic‡‡, Tuula A. Nyman§§, Tero Aittokallio**, and Riitta Lahesmaa‡¶¶**

We used stable isotope labeling with 4-plex iTRAQ (isobaric tags for relative and absolute quantification) reagents and LC-MS/MS to investigate proteomic changes in the nucleus of activated human CD4$^+$ cells during the early stages of Th2 cell differentiation. The effects of IL-4 stimulation upon activated naïve CD4$^+$ cells were measured in the nuclear fractions from 6 and 24 h in three biological replicates, each using pooled cord blood samples derived from seven or more individuals. In these analyses, in the order of 800 proteins were detected with two or more peptides and quantified in three biological replicates. In addition to consistent differences observed with the nuclear localization/expression of established human Th2 and Th1 markers, there were changes that suggested the involvement of several proteins either only recently reported or otherwise not known in this context. These included SATB1 and among the novel changes detected and validated an IL-4-induced increase in the level of YB1. This unique data set from human cord blood CD4$^+$ T cells details an extensive list of protein determinations that compares with and complements previous data determined from the Jurkat cell nucleus. ***Molecular & Cellular Proteomics 9:1937–1953, 2010.***

As a response to antigen encounter and their cytokine environment, naïve CD4$^+$ cells differentiate into functionally distinct T helper (Th)[1] cell subsets, the best characterized of which are Th1 and Th2 cells (1–3). Th1 cells produce proinflammatory cytokines and are generally acknowledged to be involved in autoimmune diseases such as multiple sclerosis and diabetes, whereas Th2 cells produce proallergic cytokines, the dysregulation of which can lead to asthma and other atopic diseases (4–6). Because these subsets share a common progenitor, the early events that determine lineage are important in the understanding of the pathways leading to polarization and associated disease. DNA microarrays (7–12) and proteomics platforms (13–20) have been used in the characterization and comparison of these Th cell types.

Interleukin-4 (IL-4) is the key cytokine in Th2 differentiation; it is expressed by Th2 cells and also drives Th2 differentiation (21, 22). The binding of IL-4 to its receptor at the Th cell surface leads to the phosphorylation of signal transducer and activator of transcription 6 (STAT6) followed by STAT6 homodimerization and nuclear localization where it regulates transcription of its target genes (23, 24). The importance of IL-4 and STAT6 for the induction of the Th2 response is well documented, and similarly the transcription factor GATA3 plays an important role in several stages of Th2 cell development where it is required for the regulation of several Th2-specific cytokines (25–27). Th1 differentiation is induced by IL-12 stimulation and characterized by up-regulation of T box transcription factor TBX21 (T-BET) and interferon-γ (28–30).

The differentiation process leads to the specific and heritable gene expression profiles of the Th cell subtypes without alteration of the base sequence of the DNA. These thus

From the ‡Turku Centre for Biotechnology, FIN-20520 Turku, Finland, ¶National Graduate School in Informational and Structural Biology, FIN-20520 Turku, Finland, ‖National Graduate School in Computational Biology, Bioinformatics, and Biometry, FIN-00014 Helsinki, Finland, **Department of Mathematics, University of Turku, FI-20014 Turku, Finland, ‡‡VTT Technical Research Centre of Finland, FIN-02044 Espoo, Finland, and §§Institute of Biotechnology, University of Helsinki, FIN-00014 Helsinki, Finland

⌘ *Author's Choice*—Final version full access.

Received, October 15, 2009, and in revised form, March 17, 2010
Published, MCP Papers in Press, May 13, 2010, DOI 10.1074/mcp.M900483-MCP200

[1] The abbreviations used are: Th, T helper; FasR, Fas receptor; FDR, false discovery rate; FPR, false positive rate; IKZF1, Ikaros; IL, interleukin; iTRAQ, isobaric tags for relative and absolute quantification; MMTS, methyl methanethiosulfonate; SATB1, special AT-rich sequence-binding protein 1; SCX, strong cation exchange; STAT, signal transducer and activator of transcription; TCF7, transcription factor 7; YB1, nuclease-sensitive element-binding protein 1 (Y box-binding protein 1); GO, gene ontology; WB, Western blotting; Erk, extracellular signal-regulated kinase; Ct, threshold cycle; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; MAPK, mitogen-activated protein; siRNA, short interfering RNA; c-MAF, Transcription factor Maf; GATA3, trans-acting T-cell-specific transcription factor GATA-3; TBX21, T-box transcription factor TBX21; ThP, T helper cell precursor.

termed epigenetic mechanisms have been shown to play a key role in the determination of the fate of Th cell specification (31, 32). Proteomic changes involving signaling and organization at the nucleus are therefore important in the early phases leading to differentiation. The aim of the present work was to apply a quantitative proteomics approach to investigate changes in the nuclear proteome of naïve $CD4^+$ human cells during the early stages of Th2 cell differentiation. In transcriptomic measurements (12, 96), we observed a large number of distinct changes during the first 24 h of differentiation and on the basis of these selected to investigate proteomic changes at time points of 6 and 24 h. We used 4-plex iTRAQ reagents (33) to compare the abundances of proteins in the nuclear fractions of activated $CD4^+$ cells and those activated and IL-4-stimulated (at 6 and 24 h). With these measurements, we aimed to identify protein abundance changes associated with Th2 cell differentiation with potential mechanistic relevance to the early phases of this process. Three biological replicates were made with triplicate analysis of the sample material. Proteins with consistent and statistically significant changes were considered for further validation. Further evaluations were made in terms of known protein interactions and functions and GO annotation in general.

In addition to the expected changes in relative protein abundance for Th2- and Th1/interferon-related proteins that were indicated in these data, there were a number changes that involved proteins either not previously reported or fully characterized in association with human Th2 differentiation. Furthermore, although large scale identification experiments have been made previously for the nucleus of Jurkat cells (34, 35), to our knowledge, this is the first such study for human cord blood $CD4^+$ cells.

EXPERIMENTAL PROCEDURES

Cell Culture and Cytokine Activation

$CD4^+$ cells were isolated from the cord blood samples (collected from healthy neonates at Turku University hospital) by Ficoll-Paque (Amersham Biosciences) density gradient centrifugation and anti-CD4 magnetic beads (Dynal). The purity of the isolated cell population was estimated using CD4 surface staining and flow cytometry as described previously (19). The mean measured purity for the gated populations was 98.5% (as illustrated in supplemental File S1.1).

Cells were cultured in Yssel's medium (provided by Dr. Hans Yssel) (36) supplemented with 1% AB serum (Red Cross Finland Blood Service). Cells were activated with plate-bound anti-CD3 (500 ng/ml) and soluble anti-CD28 (500 ng/ml) (Immunotech) and stimulated with IL-4 (10 ng/ml; R&D Systems). Activation was performed within 16 h of $CD4^+$ cell isolation. Cells were harvested at 6 and 24 h after activation for the iTRAQ and Western blotting (WB) experiments and at 0, 0.5, 2, 4, 6, 12, 24, 48, and 72 h for the RT-PCR measurements. For the 72-h sample, an IL-2 supplement (40 units/ml; R&D Systems) was added at 48 h.

Three separate biological cell cultures were made for the iTRAQ experiments with an additional 16 cultures made for the validations. The cultures were made using $CD4^+$ cells derived from the cord blood of $\geq$7 individuals for the iTRAQ labeling and were typically from five individuals in the Western blot validation experiments. The counts of $CD4^+$ cells used in the cultures varied from 32 to 270 million. Details of these cell counts and protein isolation yields are in supplemental File S1.2.

Isolation of Nuclear Fraction

The isolation of nuclear/DNA-binding proteins was performed according to the protocol of Andrews and Faller (37) with slight modifications. In brief, the cells (in aliquots of 10–50 million cells) were harvested, washed with ice-cold PBS, and lysed for 10 min on ice using a lysis buffer of 20 mM HEPES, 0.2% Nonidet P-40, 0.5 mM DTT, 1.5 mM $MgCl_2$, 10 mM KCl, 1 mM NaF, 1 mM Na_3VO_4, and Complete Mini protease inhibitors (Roche Applied Science). Lysis was controlled by trypan blue staining for which virtually all cells were positive after the 10-min incubation, indicating disruption of the plasma membrane.

The cell suspension was centrifuged at 10,000 rpm for 1 min, and the supernatant (cytoplasmic fraction) was collected. The pellet was washed using the lysis buffer and resuspended in the nuclear extraction buffer (20 mM HEPES, pH 7.9, 420 mM NaCl, 25% glycerol, 0.5 mM DTT, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 1 mM NaF, 1 mM Na_3VO_4, and Complete Mini protease inhibitors from Roche Applied Science) and incubated on ice for 20 min. Following incubation, the suspension was centrifuged (13,000 rpm for 15 min), and the supernatant (nuclear extract) was collected.

Enrichment of nuclear proteins in this nuclear fraction was verified by Western blotting with PARP1 antibody as indicated in Fig. 4*b*. To confirm that IL-4 signaling was functional, immunodetections for phospho-STAT6 and GATA3 were performed. Flow cytometry measurements of CD69 (38) and immunodetection of phospho-Erk are detailed as markers of activation in supplemental Files S1.3 and S1.4.

iTRAQ Labeling and Peptide Fractionation

Based on the results of the Bradford assays (39), equal protein amounts (50–100 μg) from the four compared states, *i.e.* 6- and 24-h activation *versus* and 6- and 24-h activation with IL-4 treatment, were labeled with the separate forms of the iTRAQ reagents. Protein precipitation was achieved during 4 h at -20 °C after mixing with 6 volumes of acetone. The pelleted proteins were then dissolved in 40 μl of triethylammonium bicarbonate buffer containing 0.1% SDS and labeled with iTRAQ reagents as described in the manufacturer's protocol. Briefly, this included reduction with tris(2-carboxyethyl)phosphine and derivatization of free cysteines with methyl methanethiosulfonate (MMTS) followed by overnight digestion with trypsin (sequencing grade modified, Promega). The resulting peptides were labeled with the iTRAQ reagents for 1 h at room temperature. In all three biological replicates, the 114 and 116 reagents were used to label the peptides from the cells activated for 6 and 24 h, respectively, and likewise the 115 and 117 reagents were used to label the activated and IL-4-treated cells for 6 and 24 h, respectively. The labeled peptides were combined and acidified (pH 2.9–3.1). The resulting peptide mixtures were fractionated with a 200 $\times$ 4.6-mm-inner diameter polysulfoethyl A column (Poly LC Inc., Columbia, MD) strong cation exchange (SCX) column using a BioCAD™ chromatograph (PerSeptive Biosystems, Freiburg, Germany). The peptides were eluted at 0.7 ml/min from the cation exchange column during a two-step gradient from 0 to 30% B in 14 min and then to 100% B in 10 min (held for 15 min). The A and B phases consisted of 5 mM KH_2PO_4 and 25% acetonitrile, pH 3 with the B phase containing 0.6 M KCl. The eluted peptides were subsequently collected in 20 sequential fractions. The SCX fractions were dried in a HetoVac vacuum centrifuge (Heto-Holten A/S, Allerød, Denmark) and desalted using Empore™ octadecyl bonded silica sorbent material (3M, St. Paul, MN).

LC-ESI-MS/MS Analysis

Aliquots (2–6 μg) of the desalted fractions were dissolved in 0.1% HCOOH and then characterized by LC-MS/MS. The LC-MS/MS system consisted of a nanoflow LC system (Famos, Switchos-II, and Ultimate, LC Packings, Amsterdam, Netherlands) coupled to a QSTAR® Pulsar ESI-hybrid quadrupole-time of flight instrument (Applied Biosystems/MDS Sciex). Sample loading was performed using a 0.3 × 5-mm PepMap C_{18} μ-precolumn (LC Packings) from which the samples were back-flushed onto a reverse phase 15-cm × 75-μm-inner diameter fused silica capillary column packed with 5-μm Magic C_{18} (Michrom Bioresources, Inc., Auburn, CA). Peptide separation was achieved using a gradient from 5 to 40% B in 120 min at a mobile phase flow rate of 200 nl/min. The phase compositions were as follows: phase A, 5% ACN and 0.1% HCOOH; phase B, 95% ACN and 0.1% HCOOH. The mass spectrometer was set to perform survey scans of 1 s followed by two 3-s MS/MS scans of the two most intense peaks from the survey scan with dynamic exclusion for 5 min. Data acquisition and instrument control were performed using Analyst QS 1.1 software (Applied Biosystems).

Triplicate LC-MS/MS analyses were carried out for each cation exchange fraction in separate batches of 20. Overall, the data set for each biological experiment consisted of 60 runs (*i.e.* 3 × 20) with 180 LC-MS/MS analyses in total for the three biological replicates.

Data Processing

A data processing work flow was implemented that aimed to facilitate a broad comparison of the data from the biological replicates and indicate consistent changes in protein abundance in the nucleus of $CD4^+$ cells induced to Th2 differentiation compared with cells that were activated only. The elements of the data analysis process are indicated in the experimental schematic depicted in Fig. 1.

ProQuant—The LC-MS/MS data were analyzed using the Applied Biosystems Analyst script ProQuant (version 1.1, Applied Biosystems). For this study using human umbilical cord blood T lymphocytes, a Swiss-Prot database (release date, December 5, 2005; 13,303 human sequences) composed of human proteins plus known contaminants (trypsin fragments and BSA; 11 entries) was concatenated with its reversed counterpart and formatted to create an interrogator that included MMTS modification of cysteine; iTRAQ labeling at the N terminus, lysine, and tyrosine; and one missed tryptic cleavage permitted. With the exception of tyrosine labeling, all the latter modifications were treated as fixed. The precursor and product ion mass tolerances were set to 0.3 and 0.2 Da, respectively, and the following variable zone modifications were specified (40): *O*-phosphorylation of serine, threonine, and tyrosine; deamidation of asparagine and glutamine; and methionine oxidation.

ProGroup—Protein inference (*i.e.* determining which proteins best describe the peptide data where degenerate peptides are present) from the ProQuant search results was made according to the ProGroup algorithm (version 1.0, Applied Biosystems). ProGroup uses the peptide identification results to determine the minimal set of proteins that can be reported for a given protein confidence threshold. Data inclusion of peptide data down to 70% confidence was used to increase the opportunities for overlap of identification across replicates, particularly in cases where the lower intensity precursor ion led to poorer scoring. For the peptide data, only peptides of 6 or more amino acids were accepted. Using these criteria, the false positive rates (FPRs) for protein identification were estimated from the number of hits for reversed peptide sequences according to the method described by Gygi and co-workers (41, 42). The estimated FPRs were typically on the order of 9% for proteins detected with two or more peptides in biological replicates 1, 2, and 3 (9.4, 9.3, and 7.3%, respectively) and 5.3 and 0.5% for proteins detected in all three biological replicates by one and two sequences, respectively.

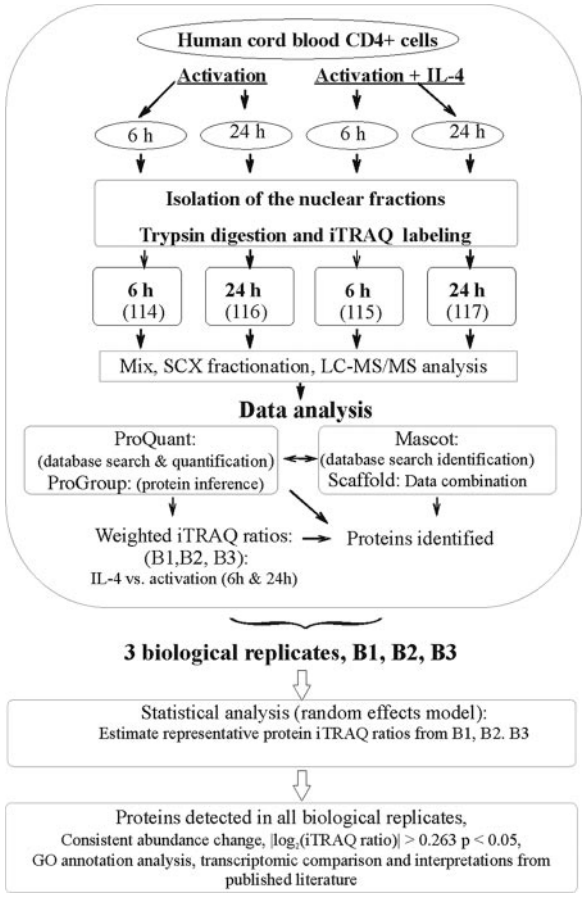

FIG. 1. **Schematic of experimental work flow.** Differences in protein abundance were determined for the nuclear fractions of IL-4-treated anti-CD3/CD28-activated *versus* anti-CD3/CD28-activated $CD4^+$ cells at 6 and 24 h by iTRAQ labeling with LC-MS/MS analysis. For identification, a combination of algorithms was used (ProQuant and Mascot). Relative quantification calculations for iTRAQ ratios (Th2/Th0) from each biological replicate were made with weighted protein abundance ratios and weighted standard deviations. A random effect model was used to assess the changes indicated in these biological replicates. A threshold of abundance change of 1.2-fold was used together with calculated of *Z* scores and associated *p* values to evaluate the significance of the changes indicated. The combination of consistent detection and statistical consistency of abundance changes, GO annotation, and comparisons with transcriptomics data and known literature using Ingenuity Pathway Analysis were used as criteria to select candidates of interest.

Mascot—To add an extra measure of confidence to the protein identifications and to represent the identification data in relation to a more frequently used identification algorithm, the data from the different biological data sets were also analyzed with Mascot (43). An in-house Mascot server was used (version 2.1, Matrix Science, London, UK), and the data were searched against the same concatenated database as used with ProQuant with the same mass tolerances and enzyme specificity. iTRAQ labeling of tyrosine, deamidation of asparagine and glutamine, and methionine oxidation were specified as variable modifications, and MMTS modification of cysteine and iTRAQ labeling at the N-terminus and lysine were specified as fixed modifications. Peak lists were created in *.mgf format using Protein-

Pilot 2.0 software (Applied Biosystems) as described previously (44). In brief, these were created without merging of any putatively similar spectra and with no restriction of mass range for precursors applied (beyond the constraints used during acquisition).

Based on the search results where reversed sequences were identified from the concatenated database, the inclusion criteria for protein identifications were chosen such that the overall FPR was on the order of 5%. Mascot protein scores greater than or equal to 40 were accepted, and peptides with an expectation score less than 0.05 and a peptide rank of 1 were counted. The false positive rates were 3.0, 5.0, and 5.3% for the biological replicates 1, 2, and 3, respectively. With these criteria, there were no protein hits based on reversed sequences detected by Mascot in all three replicates. Furthermore, for proteins identified with ProQuant in all three biological replicates and by Mascot in at least one replicate, the estimated FPR was 0.2%. Protein inference from the Mascot searches was made according to the Mascot *bold red* function; the bold red option assigns degenerate peptides to the highest scoring protein. These data were subsequently summarized using the Scaffold™ proteomic software (version Scaffold-02_05_01, Proteome Software Inc.) as described below.

Relative Quantification—Protein iTRAQ ratios and their associated standard deviations were calculated from each biological replicate. Although ProGroup provides this calculation, the ratios were purposely recalculated from the exported peptide data to establish the data used and the associated variance. These abundance ratios (IL-4-treated and activated *versus* activated for 6 and 24 h) were determined as weighted averages over the corresponding peptide-level iTRAQ reporter ion peak area ratios similarly to Gan *et al.* (45). More specifically, if x_i denotes the logarithmic ratio of the peak areas of the iTRAQ reporter ions for a peptide i in a particular biological replicate, then the protein-level ratio is defined as

$$d = \sum_{i=1}^{n} w_i x_i / \sum_{i=1}^{n} w_i \quad \text{(Eq. 1)}$$

where the weight w_i is the inverse of the percent error for the ratio of the peak areas for peptide i. The ratios and their percent errors were calculated by the ProQuant software. The corresponding weighted standard deviation was calculated as

$$s = \sqrt{v \sum_{i=1}^{n} w_i^2} / \sum_{i=1}^{n} w_i \quad \text{(Eq. 2)}$$

where v is the variance across the n peptide ratios x_i. In these calculations, only data from peptides with a sequence length greater than or equal to 6, with a ProGroup confidence of at least 70%, and without unlabeled residues (Lys or N terminus) were used. The exported peptide data were normalized by ProQuant to correct for possible systematic bias. In brief, the program excludes iTRAQ ratios that are blank, 0, or 9999 and those for which the sum of the peak areas is less than 40 counts. The program finds the median iTRAQ ratio and sets the applied bias so that the median ratio is 1. The normalization was confirmed for the selected data.

To identify consistent changes with these data, we used the so-called random effect meta-analysis model to estimate representative expression ratios for each protein from the three biological replicates. The approach is conceptually similar to the probe-level expression change averaging procedure that we have successfully applied to combine data across gene expression microarray experiments (46). For microarray applications, the method takes into account the probe-level information rather than a general measure of variability and enables determination of the variability of the estimated statistic directly from the data. A similar procedure has been applied to microarray data by other groups (47, 48). We selected this method to assess each value according to its own attributes and to facilitate the comparison of data of variable quality, particularly because protein quantifications are derived from unique peptides with their own physical characteristics and associated LC-MS/MS background.

The random effect model for the replicate measurements (m) can be written as

$$d_m = \mu + \gamma_m + \varepsilon_m \quad \text{(Eq. 3)}$$

where the Gaussian components

$$\gamma_m \sim N(0,\tau^2) \quad \text{(Eq. 4)}$$

and

$$\varepsilon_m \sim N(0,s_m^2) \quad \text{(Eq. 5)}$$

correspond to the between-replicate and within-replicate variability, respectively. Following the procedure described in detail by Choi *et al.* (49), the parameter μ and its variance were estimated as

$$\hat{\mu}_\tau = \frac{\sum_{m=1}^{3} (\tau^2 + s_m^2)^{-1} d_m}{\sum_{m=1}^{3} (\tau^2 + s_m^2)^{-1}} \quad \text{(Eq. 6)}$$

and

$$s_{\hat{\mu}_\tau} = \frac{1}{\sum_{m=1}^{3} (\tau^2 + s_m^2)^{-1}}, \quad \text{(Eq. 7)}$$

respectively. The parameter τ was estimated using the method of moments technique of DerSimonian and Laird (50) as is indicated in supplemental File S1.5. Finally, to test the null hypothesis H_0: $\mu = 0$ (*i.e.* that the iTRAQ ratio is unity), the statistic

$$Z = \hat{\mu}_\tau / s_{\hat{\mu}_\tau} \quad \text{(Eq. 8)}$$

that is distributed as $N(0,1)$ under the null hypothesis was used. A two-tailed significance threshold $p < 0.05$ was applied (*i.e.* $|Z| \geq 1.96$). These calculations were made for proteins that were detected and quantified in all three biological replicates. To consider potential changes for proteins with less spectral representation, iTRAQ ratios were calculated for proteins with at least five iTRAQ ratio calculations for the time point analyzed.

With iTRAQ applications in general, high confidence differences in expression on the order of 1.2-fold and greater have been reported (51–53). In an evaluation of the iTRAQ reproducibility, Gan *et al.* (45) have demonstrated technical variation of iTRAQ measurements to be on the order of 20% and observed greater biological variation on the order of 45%. In our iTRAQ measurements three biological replicates were made, although the material did represent a wider population with the cord blood from 34 individuals. Although individual specific biases could occur with such pooling (54), we have, however, selected a threshold for change of ~20%, *i.e.* a $\log_2$ ratio of ± 0.263, together with $p < 0.05$ to highlight potentially important results for further hypothesis and investigation. To estimate the false discovery rate (FDR), the Benjamini and Hochberg (55) correction was used. For this, each protein was assigned a rank according to their p values from smallest to largest, and then each was multiplied by the total number of proteins in the list and divided by its rank (56, 57). These calculations were performed using Microsoft Excel, the informatics program Kensington (InforSense Ltd., London, UK), and scripts written with R.

Data Annotation and Comparison—GO annotation of the identified proteins was done using PIGOK (58) and used for sorting and grouping. To gain an additional overview of these data in relation to published literature, the Ingenuity Pathway Analysis (version 6.5, Ingenuity Systems, Inc.) application was used.

In the context of other studies and T cell proteomics data sets, large scale measurements from the nucleus of human T cells have

mostly used the Jurkat cell lines (34, 35, 59). Han and co-workers (34) have studied nuclear fractionation methods and reported 1174 proteins associated with the nuclear fraction from Jurkat cells during apoptosis. These included 829 proteins from the Swiss-Prot database (current entries) of which there were 349 proteins with GO annotation for the nucleus. In a later study (35), extensive subcellular fractionation of human Jurkat A3 T leukemic cells was carried out, and 1,750 proteins were detected from the nuclear fractions. Using a hierarchical clustering-based method to indicate organelle-specific association, a subset of 768 proteins was defined, 520 of which had GO annotation for the nucleus. For the purpose of comparison, we considered the data from the latter two studies and selected data associated with Swiss-Prot identifiers. To accommodate for changes in nomenclature and deletion of entries, all the accession numbers entries have been updated using the software from the Universal Protein Resource (UniProt) with the database released March 3, 2009 (version 14.9). Altogether, the combined set from these Jurkat studies was determined to present 1439 proteins currently listed in the Swiss-Prot database of which 629 have GO annotation for the nucleus. For the comparison with our iTRAQ data, we compared the proteins detected in all three biological replicates with the former data sets. To improve the accessibility of these data, the Mascot search results (*.dat files) from the three biological replicates were summarized using Scaffold proteomic software (version Scaffold-02_05_01, Proteome Software Inc.) The Scaffold analysis includes an empirical probability-based validation of the protein identification data based on the PeptideProphet and ProteinProphet algorithms (60, 61). The protein assignment is based on Occam's razor approach that computes the minimal protein list that best describes the data (61). For these comparisons, protein summaries of the Mascot analyses of the iTRAQ data were created by Scaffold as lists of proteins detected in all three replicates with more than one peptide with a PeptideProphet probability of 95%. For broader comparisons, all proteins detected with more than one peptide and a ProteinProphet probability of 99.9% were also considered. The Scaffold result files for these analyses are available in supplemental File S1.6 together with results from searches of version 15.13 of the Swiss-Prot database (released date, January 19, 2010; 20,277 human sequence entries) such that the data can be freely explored with the Scaffold browser.

As an additional level of reference for these data, comparisons were made with previous transcriptomic measurements from studies of Th2 *versus* Th0 human cord blood cells (11, 12). These data were determined with three biological replicates using cultures of human naïve $CD4^+$ cells isolated from cord blood that were similarly activated with plate-bound anti-CD3 and soluble anti-CD-28 (Th0 state) with Th2 differentiation induced with the addition of IL-4. The cells were harvested at 2, 6, and 48 h as described previously (11, 12).

Real Time Quantitative RT-PCR

On the basis of the comparison of these proteomics data with previous transcriptomics results (11, 12), several targets were selected, and their expression was evaluated across a wider time series by RT-PCR. These RT-PCR analyses were carried out using a TaqMan ABI Prism 7900 HT instrument (Applied Biosystems) as described previously (62, 63). Primers and probes for *EF1α*, *GATA3*, *YB1*, *STAT1*, *STAT6*, *IKZF1*, *SATB1*, *TCF7*, and *TBX21* were designed using Primer Express (Applied Biosystems) or Universal ProbeLibrary ProbeFinder (Roche Applied Science) and are detailed in supplemental File S1.7. The quantitative measurements were determined as the threshold cycle (Ct) and normalized using the reference gene *EF1α* (64) as follows.

$$\Delta Ct_{gene} = Ct_{gene} - Ct_{EF1\alpha} \quad \text{(Eq. 9)}$$

Linear -fold differences between conditions were extrapolated from these normalized values using the following function.

$$\text{Fold difference } (a/b) = 2^{(\Delta Cta - \Delta Ctb)} \quad \text{(Eq. 10)}$$

Kinetic changes were measured by comparing each normalized ΔCt value with the respective base-line value from the T helper cell precursor (ThP) sample.

$$\Delta\Delta Ct = \Delta Ct_{sample} - \Delta Ct_{ThP} \quad \text{(Eq. 11)}$$

The statistical significance of the changes was assessed using a paired *t* test.

Western Blot Analysis

Western blot analyses were performed with the same sample material used in iTRAQ analysis, and 16 more cultures were made to support these data. An Odyssey system (LI-COR Biosciences, Lincoln, NE) with fluorescently labeled antibodies and direct infrared fluorescence detection was used for these determinations.

The isolated nuclear proteins (10–30 μg; quantified using the Bio-Rad detergent compatible protein assay) were boiled for 5 min with Laemmli sample buffer and resolved with 10% Bis-Tris gels (Criterion XT, Bio-Rad) using XT MOPS running buffer. Following electrophoresis, the proteins were either transferred onto nitrocellulose (Hybond ECL, Amersham Biosciences) or PVDF membranes (Immobilon FL, Millipore). For analyses with the Odyssey system, the IRDye 800- and Alexa Fluor 680-labeled secondary antibodies were used, and the signals were quantified using Odyssey software. Equal loading and transfer of proteins was confirmed by Coomassie staining (51), and quantification of the band intensities was performed with Microcomputer imaging device system software (InterFocus Imaging Ltd., Cambridge, UK) or, in case of whole cell lysate samples, by immunodetection of glyceraldehyde-3-phosphate dehydrogenase. Antibodies for the following proteins were used in these studies: IKZF1, TBX21, SATB1, and STAT1 from Santa Cruz Biotechnology; nuclease-sensitive element-binding protein 1 (Y box-binding protein 1 (YB1)) from Abcam; MAPK (Erk1/2), phospho-MAPK (Thr-202/Tyr-204), phospho-STAT6 (Tyr-641), and PARP1 from Cell Signaling Technologies; STAT6 and GATA3 from BD Biosciences; and TCF7 from Upstate.

RESULTS

Protein Identification and Relative Quantification—In this study, we investigated changes in the nuclear proteome of human naïve T helper cells in the early phases of Th2 cell differentiation using iTRAQ technology. Three biological replicates were made for iTRAQ analysis with triplicate analysis of the sample material from each experiment. Repeated analyses of this complex sample material were performed to improve the opportunities for identification and statistical analysis of the data (65). Furthermore, evaluation of the reproducibility of these protein ratios, from run to run, indicated good correlation between the values determined (detailed in supplemental Files S1.8 and S1.9). Fig. 2*a* illustrates the relationship between the results from the biological replicates and indicates a central core of proteins detected in all three experiments among a background of less frequently reported proteins and numerous single peptide hits. Of the 903 protein identities reported, 843 were characterized by two or more distinct peptides of which 815 were satisfactorily quantified in all three biological experiments using ProQuant. Typically,

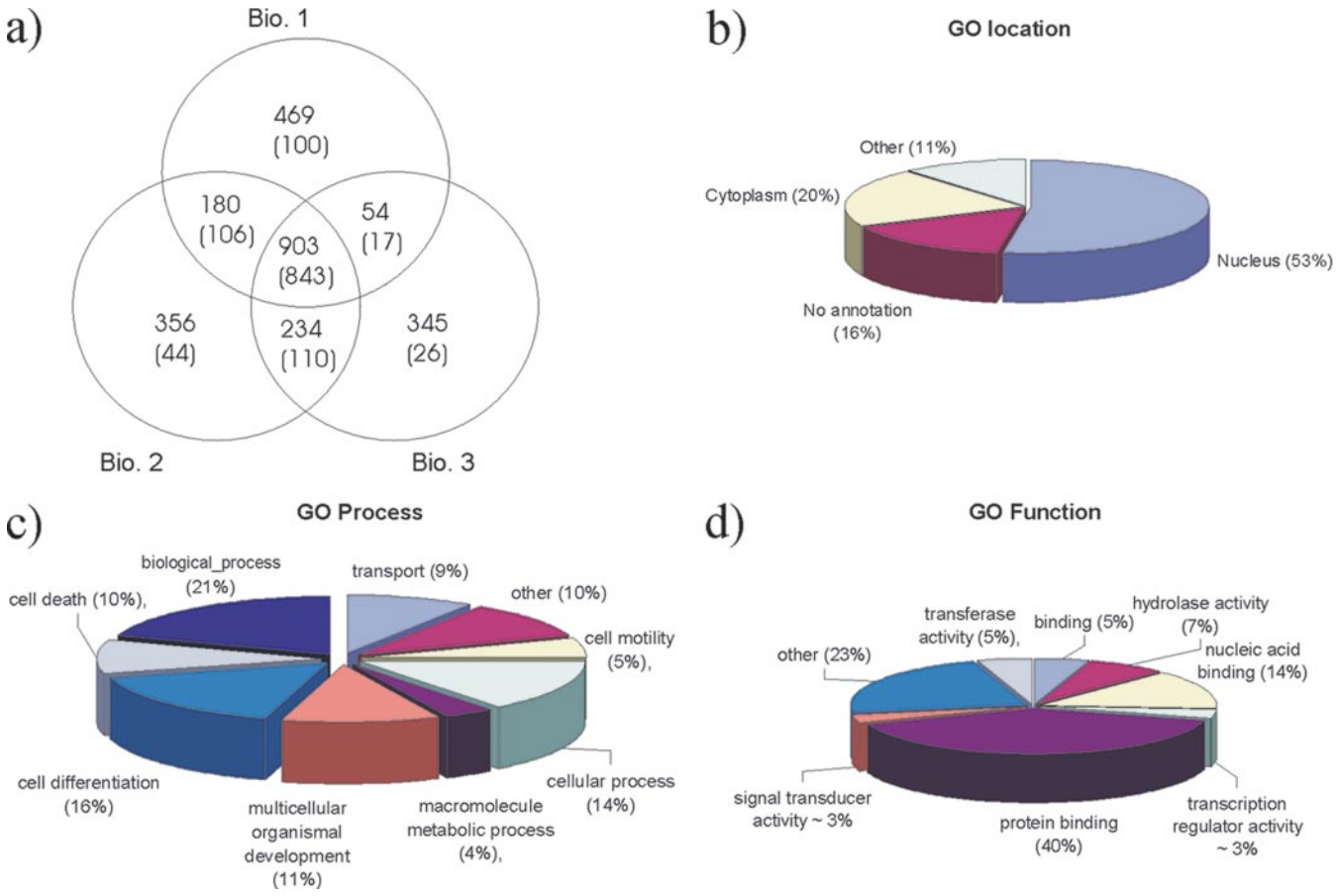

FIG. 2. **Comparison and classification of proteins detected in analysis of nuclear fraction of IL-4-treated anti-CD3/CD28-activated human CD4$^+$ cells.** *a*, the relationship between proteins from the biological replicates (*Bio. 1*, *Bio. 2*, and *Bio. 3*). The *numbers* in *parentheses* indicate the proportion of proteins identified with two or more peptide sequences by ProQuant; protein assignments were made in accordance with the ProGroup algorithm and Mascot bold red options. The numerical values indicated do not include those hits that were made from the reversed sequences in the concatenated database searches. *b*, the distribution of annotation for GO component for the consistently detected proteins. The classification is simplified to one annotation per protein. *c*, the distribution of GO process. The pie chart represents the distribution GO terms for process for proteins identified in all three biological replicates. The general group "other process" includes regulation of biological process (3%), metabolic process (2%), secretion (2%), electron transport (2%), behavior (0.7%), biosynthetic process (0.7%), and nucleobase (0.7%). *d*, the distribution of GO function. The pie chart represents the distribution GO terms for function for proteins identified in all three biological replicates. The general group "other" includes helicase activity, oxidoreductase activity, and structural molecule activity, ~3% each; molecular function, kinase activity, ligase activity, isomerase activity, and catalytic activity, ~2% each; motor activity, electron carrier activity, receptor activity, and transporter activity, ~1% each; and enzyme regulator activity, antioxidant activity, and protein transporter activity, ~0.5% each.

these protein determinations corresponded to a median of four peptides and seven calculations per biological experiment. The FPRs for identification were estimated from concatenated database searches to be 0.5% for the proteins that were detected and quantified by ProQuant in all three biological replicates with at least two unique peptide sequences. The attribute of specifying that these should also attain a Mascot score of ≥40 indicated an FPR of 0.2%. Moreover, the implementation of Mascot together with the Scaffold software supported several additional two-peptide identifications and provided an alternative measure of confidence with the calculation of ProteinProphet protein identification probabilities. In previous investigations, researchers have demonstrated that the use of multiple search algorithms can provide the advantage of consensus validation of the search results as well as additional identification data (20, 66–68). Similarly with these analyses, there was a high level of corroboration of the qualitative data as well as a small number of results that likely reflected differences in peptide scoring and protein inference. The list of repeatedly detected proteins was assessed in terms of their GO annotations for cellular location. The distribution of GO annotation for these proteins is indicated in Fig. 2*b*, and their function and process are indicated in Fig. 2, *b* and *c*. Although from this annotation it was apparent that a number of cytoplasmic proteins, among others, were co-enriched in this material, a large proportion of the proteins (~50%) are associated with the nucleus. The largest groups for GO processes and functions were described with

the broad terms *biological process* and *protein binding*, respectively, followed by *cell differentiation* and *nucleic acid binding*.

In the relation to other studies of the T cell nucleus, ~70% of the proteins detected in all three replicates were previously reported in the Jurkat nucleus data sets from Han and co-workers (34, 35). Also in similarity to the data from Han and co-workers (34, 35), proteins with GO annotation for the nucleus were well represented in these identifications (~450 proteins). Interestingly, approximately a quarter of the nuclear annotated proteins detected in all three replicates were not matched in the Jurkat data. Moreover, there were on the order of 230 nuclear annotated proteins from the Jurkat data sets not detected in our data. A comparison of the proteins that were specific to either set by associated GO terms for function and process indicated no major differences in their classifications. Notable differences from our data set were with the signaling proteins STAT1 and STAT6 (both classified with the GO function *signal transducer activity*) that are clearly related to the nature of our study. Some differences with proteins of the nuclear lamina (*i.e.* the lamin proteins) could be attributed to the isolation procedures used. Further representation of these data and their comparison as well as the full lists of peptide and protein identifications are included as supplement material to clarify the numbers of proteins reported and matched and their annotations (supplemental Files S2, S3, and S5).

To facilitate the investigation and summary of the quantitative data, we applied a statistical model that considered each protein and its individual variance in the three biological replicates. Adapting methodology we have successfully demonstrated in microarray comparisons (46), a random effect meta-analysis model was used to combine the individual estimates into an overall estimate of the -fold change and associated *Z* scores and *p* values for each protein. For these calculations the weighted protein iTRAQ ratios and their weighted standard deviations from each biological replicate were used. As a measure of the false discovery rate, the Benjamini and Hochberg (55) correction was applied. The distributions of estimated protein ratios are represented as volcano plots of $\log_2$(abundance change) *versus* $-\log_{10}$(*p* value) in Fig. 3. The iTRAQ ratios indicate changes in protein abundance in the nuclear fraction associated with IL-4 treatment of activated $CD4^+$ cells in comparison with cells that were activated only.

It was notable in the quantitative analysis of these data that, with the exception of a set of ribosomal proteins that were enriched in one of the biological replicates, the overall magnitudes of changes detected throughout these experiments were generally not greater than 2-fold. On the basis of previous observations in iTRAQ studies where expression differences greater than 1.2-fold have been confidently identified (51–53), we applied an abundance change threshold of >20% together with a *p* statistic <0.05 to define a group of proteins of interest and potential significance; these thresholds

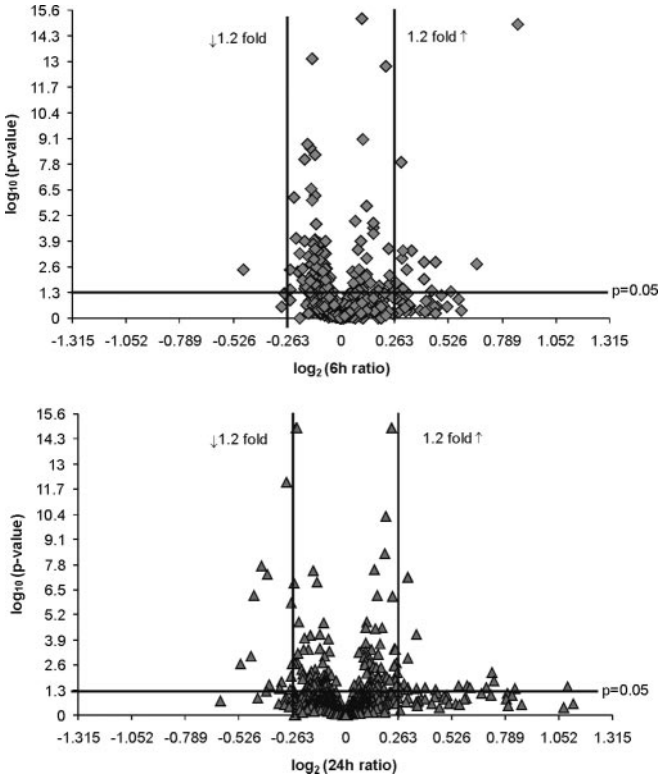

FIG. 3. **Protein abundance ratios determined from iTRAQ labeling of nuclear fraction of IL-4-treated anti-CD3/CD28-activated *versus* anti-CD3/CD28-activated $CD4^+$ cells.** The data from the random effect estimations of the $\log_2$(protein ratio) are plotted *versus* $-\log_{10}$(*p* value). *p* values were calculated from the normal distribution of the absolute value of the *Z* score. The *solid lines* are included to indicate the thresholds applied: *i.e.* a -fold change of 1.2 or more ($-0.263 \leq \log_2$ ratio ≥ 0.263) together with a $-\log_{10}$(*p* value) >1.3. ($p < 0.05$). The ratios represent the comparison Th2/Th0.

($-\log_{10}$(*p* value) > 1.3 and $|\log_2(\text{ratio})| > 0.263$) are indicated in Fig. 3. A list of the proteins defined by these criteria is included in Table I, although cytoplasmic ribosomal proteins and keratins are not included in this table but are available in supplemental File S2 as Table 1S. Using these thresholds, 5% of the data is represented (including the former). With the application of the Benjamini and Hochberg (55) correction with a 5% FDR threshold, 2% of the data are retained. Additional proteins that were noted during data processing or are otherwise of general interest in our research are listed in Table I. These include the Th1-specific transcription factor TBX21 (*i.e.* T-BET), which was detected in one replicate with a decreased abundance indicated in the IL-4-treated cells. Also included are the interferon-related proteins STAT1 and IFI16.

To provide a wider overview of the regulation of the proteins highlighted by their quantitative data, the transcriptomic profiles of eight selected targets were determined for the first 72 h of activation and treatment. These RT-PCR data are illustrated in Fig. 5*c*. The analyses indicated the accepted trends and differences for the general Th1 and Th2 transcription factors *GATA3*, *STAT1*, and *TBX21* and support expectations for

TABLE I

Comparison of protein abundance changes determined by iTRAQ labeling of nuclear fractions from IL-4-treated anti-CD3/CD28-activated CD4$^+$ cells versus anti-CD3/CD28-activated CD4$^+$ cells at 6 and 24 h with LC-MS/MS analysis

The iTRAQ ratios (Th2/Th0) are relative to the protein abundance of the activated cells (without IL-4) at the equivalent time points. The values are, unless otherwise stated, based on the weighted average from three biological replicates. These values were estimated using a random effect model. The inclusion criteria for these proteins are based on an absolutes *Z* score greater than 1.96 ($p < 0.05$), an abundance change greater than 20% (up or down), detection of the proteins in all three iTRAQ experiments, two or more associated peptides, and five or more iTRAQ ratios. Some exceptions are marked. The *p* values were calculated from the normal distribution of the absolute value of the *Z* score. PQ seq., number of unique peptides detected by ProQuant; Tot. calcs., total number of calculations used; Cov., coverage; NA, not applicable.

Accession	Name	GO location	iTRAQ 6-h ratio	+/−	6-h *p* value	6-h FDR	iTRAQ 24-h ratio	+/−	24-h *p* value	FDR	PQ seq.	Tot. calcs.	Cov.
													%
Th2 association													
STAT6_HUMAN (P42226)	Signal transducer and activator of transcription 6	Nucleus, cytoplasm	1.32	0.12	0.001	0.02	1.07	0.15	0.60	0.86	3	10	5
SATB1_HUMAN (Q01826)	DNA-binding protein SATB1	Nucleus	1.12	0.03	3·e−05	9·e−04	1.24	0.08	0.001	0.018	11	47	14
IKZF1_HUMAN[a] (Q13422)	DNA-binding protein Ikaros	Intracellular, nucleus	1.04	0.03	0.15	0.56	1.21	0.11	0.046	0.25	5	31	18
Interferon-related													
TRI22_HUMAN (Q8IYM9)	Tripartite motif-containing protein 22	Intracellular, nucleus	0.82	0.08	0.055	0.32	0.81	0.07	0.018	0.143	3	9	9
STAT1_HUMAN[b] (P42224)	Signal transducer and activator of transcription 1-α/β	Cytoplasm, nucleus	0.95	0.06	0.39	0.81	0.81	0.12	0.16	0.51	8	19	5
IF16_HUMAN[a,b] (Q16666)	g-Interferon-inducible protein IFI16	Nucleus, cellular component	0.99	0.02	0.75	0.95	0.85	0.04	0.001	0.012	22	142	39
TBX21_HUMAN[c] (Q9UL17)	T box transcription factor TBX21	Nucleus	0.98	0.09	0.83	NA	0.64	0.11	0.01	NA	2	5	4
TCF7_HUMAN[d] (P36402)	Transcription factor 7 (T cell-specific transcription factor 1)	Cytoplasm, membrane, nucleus	0.95	0.04	0.17	0.60	0.75	0.03	*2·e−08*	2·e−6	2	13	8
Other													
YBOX1_HUMAN (P67809)	Nuclease-sensitive element-binding protein 1 (Y box-binding protein 1)	Nucleus	1.23	0.10	0.11	0.50	1.65	0.30	0.006	0.069	4	18	30
G10_HUMAN[a] (P41223)	Protein BUD31 homolog (protein G10 homolog) (EDG-2)	Nucleus	0.97	0.05	0.68	0.92	1.24	0.05	7·e−08	5·e−06	3	10	26
AKAP8_HUMAN (O43823)	Protein kinase A anchor protein 8	Intracellular, membrane, nucleus, chromosome	1.10	0.06	0.26	0.70	1.24	0.13	0.045	0.25	3	9	6
RPB1_HUMAN[a] (P24928)	DNA-directed RNA polymerase II largest subunit	Nucleus, cellular component	1.45	0.17	0.046	0.29	0.86	0.33	0.70	0.90	2	5	3

TABLE I—*continued*

Accession	Name	GO location	iTRAQ 6-h ratio	+/−	6-h *p* value	6-h FDR	iTRAQ 24-h ratio	+/−	24-h *p* value	FDR	PQ seq.	Tot. calcs.	Cov.
TAF10_HUMAN (Q12962)	Transcription initiation factor TFIID subunit 10	Nucleus, cytoplasm, cellular component	1.59	0.14	0.002	0.03	0.90	0.06	0.095	0.42	2	14	2
RL5_HUMAN[a] (P46777)	60 S ribosomal protein L5	Intracellular, nucleus	1.36	0.13	0.042	0.27	1.66	0.35	0.016	0.13	6	23	27
RFC1_HUMAN[a] (P35251)	Replication factor C subunit 1	Nucleus, chromosome	1.01	0.03	0.75	0.94	1.24	0.13	0.038	0.23	7	17	10
MP2K3_HUMAN[e] (P46734)	Dual specificity mitogen-activated protein kinase kinase 3		1.82	0.08	1·*e*−15	0.00	0.92	0.04	0.086	0.38	1	11	9
CD3E_HUMAN[a] (P07766)	T cell surface glycoprotein CD3 ϵ chain precursor	Cell surface, membrane	1.25	0.10	0.004	0.05	1.16	0.07	0.016	0.132	2	10	14
CD5_HUMAN (P06127)	T cell surface glycoprotein CD5 precursor	Cell surface, membrane	0.90	0.07	0.142	0.55	0.73	0.07	0.001	0.015	3	9	15
CD45_HUMAN[a] (P08575)	Leukocyte common antigen precursor	Membrane	0.99	0.05	0.864	0.97	0.80	0.09	0.043	0.244	6	33	6
CD44_HUMAN (P16070)	CD44 antigen precursor	Cell surface, membrane	0.91	0.04	0.032	0.23	0.77	0.09	0.027	0.191	6	46	10
GLU2B_HUMAN[a] (P14314)	Glucosidase 2 subunit β precursor	Membrane	1.23	0.04	1·*e*−08	0.00	1.07	0.02	2·*e*−05	0.001	3	8	6
PSB8_HUMAN (P28062)	Proteasome subunit β type 8 precursor	Intracellular	1.03	0.06	0.670	0.92	0.83	0.03	2·*e*−06	7·*e*−05	1	8	
PSME3_HUMAN (P61289)	Proteasome activator complex subunit 3	Intracellular	0.93	0.07	0.341	0.79	0.73	0.05	6·*e*−07	3·*e*−05	4	10	19
GNAS_HUMAN (P63092)	Guanine nucleotide-binding protein G_s subunit α isoforms, short	Intracellular	0.72	0.08	0.004	0.05	1.14	0.39	0.688	0.900	2	10	14
ATP5J_HUMAN (P18859)	ATP synthase coupling factor 6, mitochondrial precursor	Cytoplasm	1.23	0.12	0.032	0.23	0.74	0.15	0.126	0.464	2	12	22
GRAP2_HUMAN (O75791)	GRB2-related adapter protein 2	Cytoplasm	1.00	0.01	0.816	0.96	1.27	0.08	6·*e*−05	0.002	6	18	12
APT_HUMAN[a] (P07741)	Adenine phosphoribosyl-transferase	Cytoplasm	0.99	0.02	0.723	0.93	0.83	0.08	0.045	0.243	4	24	25
HS105_HUMAN[a] (Q92598)	Heat-shock protein, 105 kDa	Cytoplasm	1.24	0.07	4·*e*−04	0.01	1.09	0.12	0.436	0.755	4	12	6

TABLE I—*continued*

Accession	Name	GO location	iTRAQ 6-h ratio	+/−	6-h *p* value	6-h FDR	iTRAQ 24-h ratio	+/−	24-h *p* value	FDR	PQ seq.	Tot. calcs.	Cov.
TXLNA_HUMAN[a] (P40222)	α-Taxilin	Cytoplasm, extracellular region	1.38	0.14	0.001	0.02	1.01	0.07	0.848	0.967	3	9	13
S10A9_HUMAN[a] (P06702)	Protein S100-A9 (S100 calcium-binding protein A9)	Extracellular space	1.27	0.09	4·e−04	0.01	1.82	1.09	0.289	0.653	4	7	31

[a] Detected in the published T cell nucleus data sets of Han and co-workers (34, 35).
[b] Included but did not pass inclusion criteria.
[c] In one biological replicate only.
[d] Shared peptides with LEF1 (Q9UJU2) and TF7L2 (Q9NQB0) support this change.
[e] Mascot provided two additional sequences.

SATB1 and *TCF7*. Induction by activation with IL-4 treatment, and activation alone was apparent for *YB1*, whilst consistent differences with IL-4 treatment were indicated for the *IKZF1* data. As the changes relating to STAT6 in Th2 differentiation are associated with its phosphorylation and translocation, the absence of transcriptional differences does not indicate an inconsistency.

iTRAQ Data Validation—On the basis of their iTRAQ ratios and associated statistics together with GO annotation, known mRNA expression (12), novelty, and antibody availability, a panel of proteins was selected for validation by WB. These were STAT6, SATB1, TCF7, IKZF1, and (least familiar in relation to T cell differentiation) YB1. Although the known Th1 transcription factors STAT1 and TBX21were not well represented in the MS data, these were also included as protein targets for WB analysis. During the primary handling of the material used for the mass spectrometry-based analysis, WB analyses were performed for phospho-STAT6 to confirm IL-4 signaling. Additional tests were performed for GATA3 to confirm Th2 differentiation status; these data are exemplified in Fig. 4*a*. Validation of nuclear fractionation was achieved using PARP1 (69) (Fig. 4*b*). The detection of GATA3 (for the nuclear fraction) also supported the efficacy of this fractionation (included as supplemental File S1.10). To improve the opportunities for detection and provide better discrimination between small changes, the Odyssey infrared fluorescence system was used for these WB measurements (70).

From the WB analyses, validations of quantification and identification were achieved for four of the proteins that were indicated from the iTRAQ data (Fig. 5). These results supported the changes in STAT6 and SATB1 in the original sample material and also in new cultures at these time points. The changes detected for YB1 in association with IL-4 treatment were also confirmed by WB as was the decreased nuclear abundance of TBX21. With the WB analysis of the DNA-binding protein Ikaros (IKZF1), multiple isoforms were detected, the strongest of which was a double band migrating at ~50 kDa. Although there appeared to be some consistency in the increase of these isoforms, the change was not statistically significant. With the mass spectrometry data for IKZF1, it was difficult to discern clear differences because of the absence of isoform-specific peptides. The decreased nuclear abundance indicated for TCF7 in the iTRAQ data was not statistically significant in the WB data. For STAT1, a familiar protein in the context of interferon-related pathways and Th1 cells, the changes in the iTRAQ and WB measurements were not significant for the nuclear fractions at 24 h.

DISCUSSION

Multiplex Relative Quantification of Early Phases of Th2 Differentiation—Using 4-plex iTRAQ methodology, protein identifications from the nuclear fraction of human $CD4^+$ cells under Th2-promoting conditions were attained with relative quantification compared with activation alone for 6- and 24-h

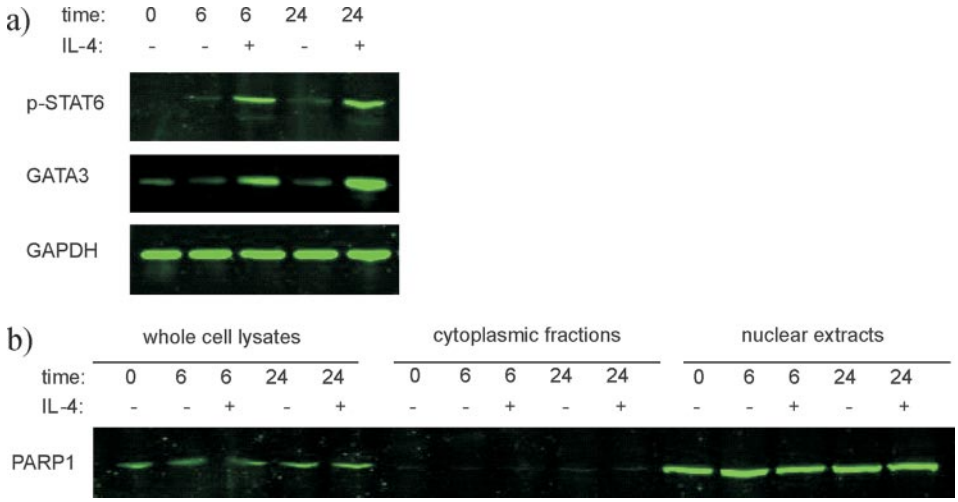

FIG. 4. *a*, Western blot analysis for IL-4 signaling and Th2 differentiation markers in anti-CD3/CD28-activated human $CD4^+$ cells. Immunodetection of phospho-STAT6 (*p-STAT6*) and GATA3 from whole cell lysates is shown. Clear IL-4-induced up-regulation is observed in both 6- and 24-h time points, confirming that the IL-4 signaling pathway is active in these cells. *b*, Western blot analysis for enrichment of the nuclear fraction using PARP1 as a nuclear marker. Immunodetection of PARP1 from whole cell lysates, cytoplasmic fractions, and nuclear extracts is shown. An equal amount of protein was loaded for each sample. As expected, PARP1 is strongly enriched in the nuclear extracts and virtually absent in the cytoplasmic fractions. *GAPDH*, glyceraldehyde-3-phosphate dehydrogenase.

time points. Whereas large scale proteomics studies of the T cell nucleus have been made with Jurkat cells, to our knowledge this data set presents the largest such set of identification data from the nucleus of human cord blood $CD4^+$ cells. Overall, there were on the order of 800 proteins detected with two or more peptides that were quantified in all three replicates, half of which had GO annotation for the nucleus. In spite of the differences in cell treatment and isolation of the nuclear fraction, in comparison with the extensive Jurkat data sets of Han and co-workers (34, 35) our measurements provide a complementary view of the T cell nuclear proteome. In consideration of all the proteins detected with more than one unique peptide and with a ProteinProphet identification probability of 99.9%, on the order of 1120 proteins were detected in these experiments of which 566 had GO annotation for the nucleus and 185 of which are not common in the comparisons with Jurkat data. Because of experimental dissimilarities, it is not appropriate to conclude that these differences indicate that the proteins may be less abundant or absent in Jurkat cells. Further comparisons with other Jurkat data, *e.g.* from targeted interactome studies in the Jurkat nucleus (71), did indicate additional matches (details are included in supplemental File S5).

With these measurements, we aimed to identify protein abundance changes associated with the differentiation process with potential mechanistic relevance to the early phases of the differentiation process. Notably, although some of the changes associated with this process may only be subtle or involve post-translational modifications or alternative isoforms not detected in our data, no large differences in protein abundance were observed in these analyses. Potential limitations to our observations were likely met by the challenges of the sample complexity *versus* the speed and dynamic range of the MS instrument. Moreover, with global peptide labeling, as achieved with the iTRAQ method, the presence of a background of highly abundant proteins can affect the normalization and thus reduce relative abundance differences calculated. Similarly, the transfer of multiple precursors from complex mixtures can affect the reporter ion signals, masking abundance changes (72). Targeted analysis, *e.g.* amino acid-specific labeling and fractionation, such as ICAT, or targeting post-translational modifications could have improved sensitivity. Similarly, changes in sample fractionation as well as the MS/MS data-dependent acquisition method may have resulted in additional data.

To represent the changes in the data as a whole, we used a random effect meta-analysis model and applied a significance threshold of $p < 0.05$ together with a fold change larger than 20% to indicate potentially interesting expression changes. In addition to the latter thresholds, we used the Benjamini and Hochberg (55) correction to estimate the FDR. On the basis of these criteria, a subset of proteins that indicated consistent change was defined (Table I). From the WB validation of a selection of these proteins, the results generally supported the direction of the iTRAQ-determined changes (*i.e.* increases and decreases). Notably, however, some of these determinations were close to the limit of discrimination, and accurate normalization was an important and limiting factor. The validations were based on numerous biological replicates and in general were of a similar magnitude. With STAT6, however, the iTRAQ-derived values were notably less than those suggested by the WB data. Although such differences can reflect technical reproducibility and biological variability, it was apparent that the iTRAQ reporter ions were

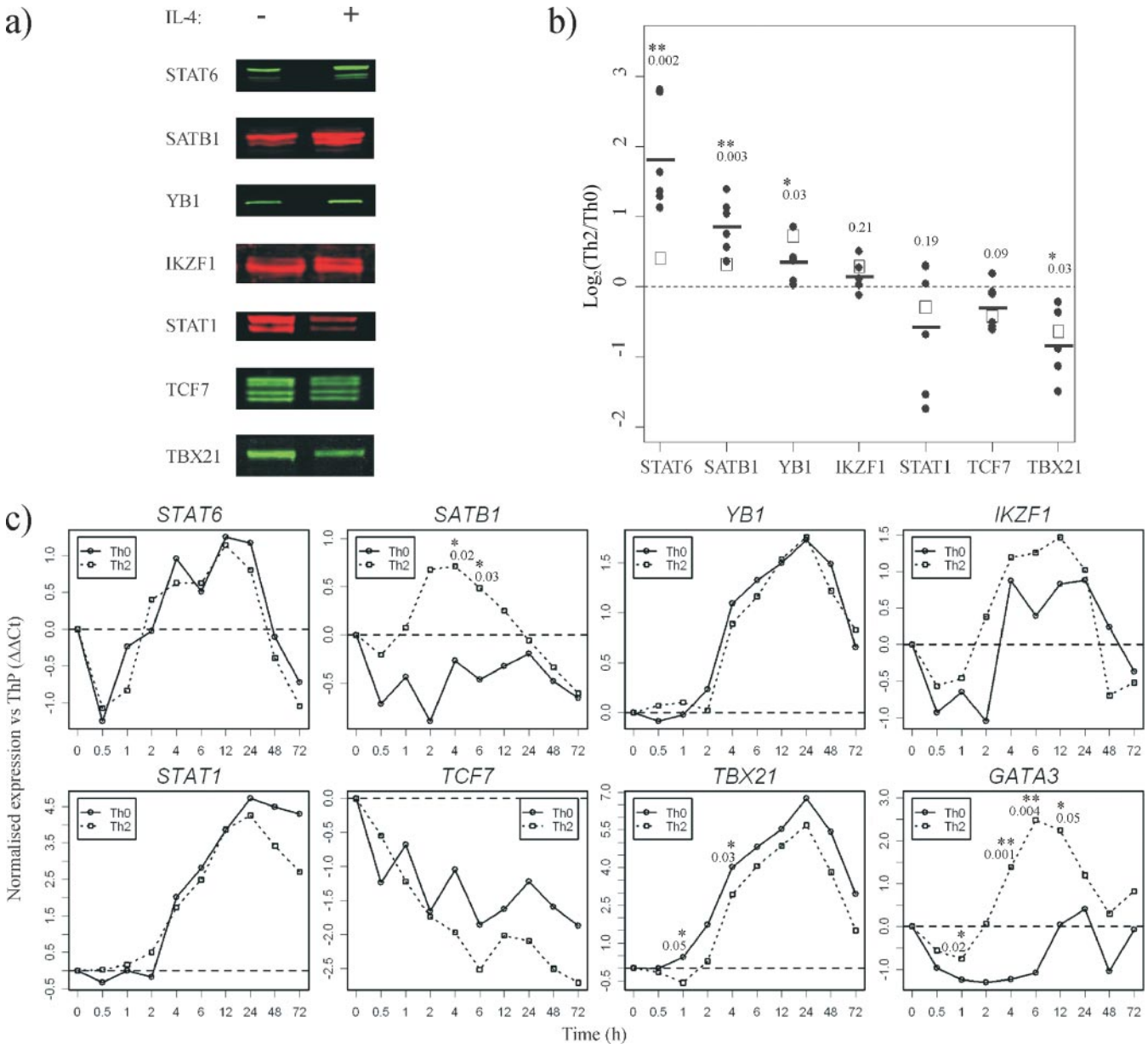

FIG. 5. **Proteomic and transcriptomic changes associated with IL-4 treatment of anti-CD3/CD28-activated human CD4$^+$ cells.** *a*, Western blot analysis for selected proteins in the nuclear fraction. The images show representative blots from cells cultured in the presence or absence of IL-4. With the exception of STAT6 (6 h), these blots are from the 24-h time point. Relative band intensities were quantified using the Odyssey infrared fluorescence system and normalized for equal loading by quantification of the total protein amounts from Coomassie-stained membranes. *b*, displays the data for the $\log_2$ ratio of Th2/Th0 for each WB protein determination (filled circles) with the median values represented with a horizontal line; the *p* values are as indicated. The weighted averages of the iTRAQ quantification values (n = 3) are inset as open squares. For the western blot data, the numbers of biological replicates were: STAT6 (n = 6), SATB1 (n = 6), YB1 (n = 6), IKZF1 (n = 5), STAT1 (n = 6), TCF7 (n = 6), TBET (n = 5). *c*, RT-PCR analysis for selected mRNA targets. On the basis of the comparison of proteomics and transcriptomics data, the temporal changes of eight transcription factors were determined during the first 72 h of activation and IL-4 treatment. The measurements were for *STAT6*, *SATB1*, *YB1*, *IKZF1*, *STAT1*, *TCF7*, *TBX21*, and *GATA3* at 0, 0.5, 1, 2, 4, 6, 12, 24, 48, and 72 h. The values are averaged from three biological replicates and determined as ΔΔCt values (see "Experimental Procedures"). A paired *t* test was used to calculate *p* values (*inset*).

close to the limit of quantification and only based on just a few peptide measurements (supplemental File S4). Similarly, it has been demonstrated previously that the precision of iTRAQ measurements is poorer with lower concentration analytes (73) in addition to the other limitations of iTRAQ quantification from complex mixtures as discussed above (72). Furthermore, to screen for changes of proteins with a lower spectral count, we permitted the inclusion of protein quantifications based on

five measurements. Although calculation of the FDR provides an extra level of control, in cases where the data is scarce, the values should be evaluated with scrutiny. Alternatively, on a wider scale, filters based on the minimum number of calculations for each replicate could be applied.

Included in the changes that were detected within the FDR criteria (<5%) were several T cell-related cytoplasmic and membrane proteins (*e.g.* CD3 and CD5); some of these were also reported in the data sets of Han and co-workers (34, 35) as indicated in Table I. For these observations, it is possible that with the concentration of lysis buffer used (0.2% Nonidet P-40) some of the hydrophobic membrane proteins were not sufficiently dissolved and thus co-isolated with the nuclear pellet. Among the proteins highlighted, there are also associations reported for GLU2B (74) and GRB2 (75) in T cell activation and differentiation. Other observations reported in Table I included an increased abundance of the dual specificity mitogen-activated protein kinase kinase 3 (MP2K3) indicated at 6 h that is known to catalyze phosphorylation of mitogen-activated protein kinase p38 and has been reported previously in the nucleus (76). An increase was also indicated for transcription initiation factor TFIID subunit 10 (TAF10) at 6 h. TAF10 is associated with several complexes involved with the modulation of transcription, and although a number of the other known components of these complexes were also observed in these data, *i.e.* TAF4, TAF6, TAF7, and TAF9, none of these showed significant changes in abundance. Among the data from the iTRAQ measurements, in general, there were no proteins that showed changes at both time points. Without further measurements, it is difficult to determine the significance of these seemingly transient changes.

In the assessment of these data, we also made comparisons with transcriptomic measurements from this system. Such proteomic and transcriptomic comparisons have been described previously (77, 78), including measurements from Jurkat T cells (35). Although the expression of proteins is often regulated at the level of transcription, differences in mRNA stability, translational regulation, and protein turnover limit the direct comparison of mRNA and proteomics data at a set point in time (79). Moreover, measurements may be difficult to account for or inappropriate when targeting subproteomes. However, from the comparisons made between our proteomics data and transcriptomic measurements across a wider time window, a number of the changes were reflected at the mRNA level (Fig. 5*c*). The differences are plotted as ΔΔCt as described above.

Proteomic Changes Characteristic of Th2 Differentiation—The changes observed with IL-4 treatment reflected the importance of STAT6 and the accumulation of nuclear phospho-STAT6 in Th2 differentiation (80, 81). However, although changes in the total level of nuclear STAT6 were indicated in our iTRAQ experiments, these lacked detail on the proportions of phospho-STAT6. Noticeably, in this overall data set, only a small number of phosphopeptides were detected, none of which were detected in all three biological replicates. It is likely that many of the phosphopeptides could have been lost in the SCX fractionation (82) and moreover may have been difficult to detect without any enrichment steps because of their relative stoichiometry. Nevertheless, changes for total STAT6 were confirmed by Western blotting, and similarly, the nuclear phospho-STAT6 levels were increased. In further support of the integrity of the cultures with respect to Th2 differentiation, the Western blots for GATA3 followed the expected line of change, *i.e.* an increased protein level with IL-4 stimulation at 6 and 24 h. The failure to detect GATA3 in the mass spectrometry data was attributed to its low level of expression in the stimulated naïve T cells at these time points and limitations of detection from such complex mixtures. Also, in consideration of Th1-related proteins, the decreasing level of TBX21 in the nuclear fraction in response to IL-4 treatment indicated in the mass spectrometric data was confirmed in the Western blot analysis. TBX21 controls the expression of the Th1 cytokine interferon-γ and directs Th1 lineage development. In keeping with this change, reduced levels of other interferon-related proteins would be expected (15–17) as indicated by a number of moderate changes in these data (Table I). With the corresponding transcriptomics data for these proteins, more distinct changes were observed along the extended time series (0.5–72 h), *e.g. GATA3*, *TCF7*, and *STAT1* (Fig. 5*c*).

Our earlier published results have also indicated that special AT-rich sequence-binding protein 1 (SATB1) is up-regulated at early stages of Th2 polarization of human $CD4^+$ T cells, both at the mRNA and total protein level (11). In keeping with this, an increased abundance of SATB1 in the nuclear fraction of the IL-4-treated human $CD4^+$cells was observed in the iTRAQ data and supported by the Western blot validations. Although the importance of SATB1 on IL-4-induced Th2 differentiation was first shown by its regulation of the expression of the Th2 cytokines IL-4, IL-5, and IL-13 in the mouse (83), recent studies with human $CD4^+$ cells suggest that SATB1 mediates the expression of GATA3 and Th2 cytokines in a wnt/β-catenin-dependent manner. In relation to this, it is notable that although lymphoid enhancer factor/TCF7 transcription factors associate with β-catenin in canonical wnt signaling leading to the transcription of its target genes, SATB1 has recently been shown to competitively influence the TCF7 binding to β-catenin (84). Taken together, it has been proposed that cell fate and differentiation could be influenced through the balance of expression, post-translational modifications, and interaction of TCF7, SATB1, and β-catenin (84) as indicated by the relative abundance of *TCF7* and *SATB1* in Th1 and Th2 cells (11). Additionally, in recent studies in mouse, it has been demonstrated that TCF7 initiates Th2 differentiation in activated $CD4^+$ cells through activation of the *GATA3* promoter (32), further indicating the role of wnt signaling in *GATA3* regulation and the balance of nuclear activity of TCF7 and SATB1 binding. Studies from our group on

human Th cell differentiation (96), including SATB1 siRNA measurements (97), indicate that of the proteins highlighted in our proteomic measurements *TRI22*, *IKZF1*, *IFI16*, and *PSB8* are both SATB1- and IL-4/STAT6-dependent, further supporting the relevance of the changes detected.

With these comparisons of activation and IL-4-induced Th2 polarization, an increased abundance of YB1 in the nuclear fraction was observed. YB1 is a multifunctional protein involved in processes such as cell proliferation, DNA repair, and stress responses as reviewed by Kohno *et al.* (85). In regard to these roles, it was observed from the transcriptomic measurements that *YB1* was induced similarly with IL-4 and activation alone (Fig. 5*c*). When translocated to the nucleus, YB1 represses genes associated with cell death, including the Fas cell death-associated receptor and the p53 tumor suppressor gene (85–88). In T cells, the triggering of the Fas receptor (FasR) by the Fas ligand, leading to activation of downstream caspase pathways (89–92), takes place during activation-induced cell death and is regarded as a method for maintaining T cell homeostasis and eliminating autoreactive cells. Notably, in relation to this, we have also observed that IL-4 treatment of activated human $CD4^+$ cells decreases caspase-3 activity and alters expression of several proteins involved in its upstream regulation, including FasR (19). In view of the observed changes in the nuclear levels of YB1 and its associations with FasR repression, we selected the use of YB1 siRNA (93) to investigate its effect upon FasR expression in these early phases of Th2 differentiation. However, using YB1 siRNA and measuring the expression of the FasR by fluorescence-activated cell sorting in both activated and IL-4-treated activated $CD4^+$ cells, we did not find any consistent change in cell surface FasR expression in response to YB1 knockdown at the 24-h time point (data not shown). Details of the siRNA constructs and methods are provided in supplemental File S1.11.

Recent studies in the mouse have shown that IKZF1 is a regulator of Th2 cell differentiation where it promotes chromatin accessibility in the nucleus and thus activates Th2 gene expression (*i.e.* IL-4, IL-5, and IL-13) and indirectly regulates the expression of Th2- and Th1-specific transcription factors (GATA3 and cMAF and TBX21 and STAT1, respectively) (94). Additional studies (in the mouse) have indicated that IKZF1 silences TBX21 expression and production during Th2 differentiation (95). In the transcriptomic comparison of IL-4 treatment of activated human $CD4^+$ cells, a greater expression of *IKZF1* was consistently observed (Fig. 5*c*). With the protein measurements from the nuclear fraction, the WB validations indicated isoform-specific changes with IL-4 treatment (Fig. 5*b*).

In summary, from these investigations of proteomic changes in the nuclear fraction of naïve human $CD4^+$ cells, a number of subtle differences were detected in the relative abundance of Th1- and Th2-related proteins in the early stages of IL-4-stimulated Th2 differentiation. In addition to differences in familiar proteins, some differences were indicated for proteins less well characterized or novel in this context, including changes in the nuclear abundance of SATB1, TCF7, and IKFZ1 that are in keeping with recent data from this system. Further proteomics studies of a wider time series could provide clearer detail of changes in the differentiation process. Targeted efforts, for example measuring changes in protein phosphorylation, would produce more specific information and reduce sample complexity. Overall, this data set presents a collection of 900 confident and reproducible protein identifications from human umbilical cord blood $CD4^+$ cells that complement and supplement previous proteomics data from the study of the T cell nucleus. The spectral data from the Mascot analyses of these data are available as supplemental material that can be previewed using the Scaffold browser, and the data are available in their raw format from the Tranche project (https://proteomecommons.org/dataset.jsp?i=NQhdNouKyMteqAmy5FwkeYRSt7XG1n0f3RGPvrwWy0vQEQjKAHl7KQYXqcMQ18ROm4ThCVQCc5fHaJHi0LdeHOp%2fiz8AAAAAAAAHwg%3d%3d; see also supplemental Files S1.6 and S1.12). Additional supplemental material is cited in the text.

Acknowledgments—We acknowledge Applied Biosystems for help and generosity (Rod Watson) and for advice with aspects of data processing (Sean Seymour and Alpesh Patel). Professor Sanjeev Galande of the National Centre for Cell Science, Pune, India, is thanked for sharing details of studies on SATB1. These studies were performed with the assistance of the Turku Centre for Biotechnology Proteomics core facility, Sarita Heinonen (RT-PCR), Marjo Hakkarainen, and the Finnish Microarray Centre.

* This work was supported by The National Technology Agency of Finland, the Academy of Finland (SysBio Program and Grant 8209083); the Sigrid Jusélius Foundation; the TurkuUniversity medical faculty research fund; The National Graduate School inComputational Biology, Bioinformatics, and Biometry; and EuropeanCommission Seventh Framework Grants EC-FP7-SYBILLA-201106, EC-FP7-NANOMMUNE-214281, and EC-FP7-DIABIMMUNE-202063.

S This article contains supplemental files S1 to S5. The supplements S1.1–S1.13 are included in S1, of which S1.13 details the contents of the other files. Table 1S is included in S2.

§ Both authors made equal contributions to this work.

¶¶ To whom correspondence should be addressed: Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, P. O. Box 123, FIN-20521 Turku, Finland. Tel.: 358-2-333-8004; Fax: 358-2-333-8000; E-mail: riitta.lahesmaa@btk.fi.

REFERENCES

1. Mosmann, T. R., Cherwinski, H., Bond, M. W., Giedlin, M. A., and Coffman, R. L. (1986) Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins. *J. Immunol.* **136,** 2348–2357
2. Del Prete, G. F., De Carli, M., Mastromauro, C., Biagiotti, R., Macchia, D., Falagiani, P., Ricci, M., and Romagnani, S. (1991) Purified protein derivative of mycobacterium tuberculosis and excretory-secretory antigen(s) of toxocara canis expand in vitro human T cells with stable and opposite (type 1 T helper or type 2 T helper) profile of cytokine production. *J. Clin. Investig.* **88,** 346–350
3. Lahesmaa, R., Yssel, H., Batsford, S., Luukkainen, R., Möttönen, T., Steinman, L., and Peltz, G. (1992) Yersinia enterocolitica activates a T helper

type 1-like T cell subset in reactive arthritis. *J. Immunol.* **148,** 3079–3085
4. Maggi, E., and Romagnani, S. (1994) Role of T cells and T-cell-derived cytokines in the pathogenesis of allergic diseases. *Ann. N.Y. Acad. Sci.* **725,** 2–12
5. Romagnani, S. (1994) Regulation of the development of type 2 T-helper cells in allergy. *Curr. Opin. Immunol.* **6,** 838–846
6. Romagnani, S. (1994) Lymphokine production by human T cells in disease states. *Annu. Rev. Immunol.* **12,** 227–257
7. Rogge, L., Bianchi, E., Biffi, M., Bono, E., Chang, S. Y., Alexander, H., Santini, C., Ferrari, G., Sinigaglia, L., Seiler, M., Neeb, M., Mous, J., Sinigaglia, F., and Certa, U. (2000) Transcript imaging of the development of human T helper cells using oligonucleotide arrays. *Nat. Genet.* **25,** 96–101
8. Lu, B., Zagouras, P., Fischer, J. E., Lu, J., Li, B., and Flavell, R. A. (2004) Kinetic analysis of genomewide gene expression reveals molecule circuitries that control T cell activation and Th1/2 differentiation. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 3023–3028
9. Bosque, A., Pardo, J., Martínez-Lorenzo, M. J., Iturralde, M., Marzo, I., Piñeiro, A., Alava, M. A., Naval, J., and Anel, A. (2005) Down-regulation of normal human T cell blast activation: roles of APO2L/TRAIL, FasL, and c-FLIP, bim, or bcl-x isoform expression. *J. Leukoc. Biol.* **77,** 568–578
10. Lund, R. J., Chen, Z., Scheinin, J., and Lahesmaa, R. (2004) Early target genes of IL-12 and STAT4 signaling in Th cells. *J. Immunol.* **172,** 6775–6782
11. Lund, R., Ahlfors, H., Kainonen, E., Lahesmaa, A. M., Dixon, C., and Lahesmaa, R. (2005) Identification of genes involved in the initiation of human Th1 or Th2 cell commitment. *Eur. J. Immunol.* **35,** 3307–3319
12. Lund, R. J., Löytömäki, M., Naumanen, T., Dixon, C., Chen, Z., Ahlfors, H., Tuomela, S., Tahvanainen, J., Scheinin, J., Henttinen, T., Rasool, O., and Lahesmaa, R. (2007) Genome-wide identification of novel genes involved in early Th1 and Th2 cell differentiation. *J. Immunol.* **178,** 3648–3660
13. Nyman, T. A., Rosengren, A., Syyrakki, S., Pellinen, T. P., Rautajoki, K., and Lahesmaa, R. (2001) A proteome database of human primary T helper cells. *Electrophoresis* **22,** 4375–4382
14. Rautajoki, K., Nyman, T. A., and Lahesmaa, R. (2004) Proteome characterization of human T helper 1 and 2 cells. *Proteomics* **4,** 84–92
15. Filén, J. J., Nyman, T. A., Korhonen, J., Goodlett, D. R., and Lahesmaa, R. (2005) Characterization of microsomal fraction proteome in human lymphoblasts reveals the down-regulation of galectin-1 by interleukin-12. *Proteomics* **5,** 4719–4732
16. Rosengren, A. T., Nyman, T. A., and Lahesmaa, R. (2005) Proteome profiling of interleukin-12 treated human T helper cells. *Proteomics* **5,** 3137–3141
17. Rosengren, A. T., Nyman, T. A., Syyrakki, S., Matikainen, S., and Lahesmaa, R. (2005) Proteomic and transcriptomic characterization of interferon-alpha-induced human primary T helper cells. *Proteomics* **5,** 371–379
18. Loyet, K. M., Ouyang, W., Eaton, D. L., and Stults, J. T. (2005) Proteomic profiling of surface proteins on Th1 and Th2 cells. *J. Proteome Res.* **4,** 400–409
19. Rautajoki, K. J., Marttila, E. M., Nyman, T. A., and Lahesmaa, R. (2007) Interleukin-4 inhibits caspase-3 by regulating several proteins in the fas pathway during initial stages of human T helper 2 cell differentiation. *Mol. Cell. Proteomics* **6,** 238–251
20. Filén, J. J., Filén, S., Moulder, R., Tuomela, S., Ahlfors, H., West, A., Kouvonen, P., Kantola, S., Björkman, M., Katajamaa, M., Rasool, O., Nyman, T. A., and Lahesmaa, R. (2009) Quantitative proteomics reveals GIMAP family proteins 1 and 4 to be differentially regulated during human T helper cell differentiation. *Mol. Cell. Proteomics* **8,** 32–44
21. Kopf, M., Le Gros, G., Bachmann, M., Lamers, M. C., Bluethmann, H., and Köhler, G. (1993) Disruption of the murine IL-4 gene blocks Th2 cytokine responses. *Nature* **362,** 245–248
22. Nelms, K., Keegan, A. D., Zamorano, J., Ryan, J. J., and Paul, W. E. (1999) The IL-4 receptor: signaling mechanisms and biologic functions. *Annu. Rev. Immunol.* **17,** 701–738
23. Kaplan, M. H., Schindler, U., Smiley, S. T., and Grusby, M. J. (1996) Stat6 is required for mediating responses to IL-4 and for development of Th2 cells. *Immunity* **4,** 313–319
24. Takeda, K., Tanaka, T., Shi, W., Matsumoto, M., Minami, M., Kashiwamura, S., Nakanishi, K., Yoshida, N., Kishimoto, T., and Akira, S. (1996) Essential role of Stat6 in IL-4 signalling. *Nature* **380,** 627–630
25. Ting, C. N., Olson, M. C., Barton, K. P., and Leiden, J. M. (1996) Transcription factor GATA-3 is required for development of the T-cell lineage. *Nature* **384,** 474–478
26. Zheng, W., and Flavell, R. A. (1997) The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. *Cell* **89,** 587–596
27. Pai, S. Y., Truitt, M. L., Ting, C. N., Leiden, J. M., Glimcher, L. H., and Ho, I. C. (2003) Critical roles for transcription factor GATA-3 in thymocyte development. *Immunity* **19,** 863–875
28. Szabo, S. J., Kim, S. T., Costa, G. L., Zhang, X., Fathman, C. G., and Glimcher, L. H. (2000) A novel transcription factor, T-bet, directs Th1 lineage commitment. *Cell* **100,** 655–669
29. Grogan, J. L., Mohrs, M., Harmon, B., Lacy, D. A., Sedat, J. W., and Locksley, R. M. (2001) Early transcription and silencing of cytokine genes underlie polarization of T helper cell subsets. *Immunity* **14,** 205–215
30. Fields, P. E., Kim, S. T., and Flavell, R. A. (2002) Cutting edge: changes in histone acetylation at the IL-4 and IFN-gamma loci accompany Th1/Th2 differentiation. *J. Immunol.* **169,** 647–650
31. Ansel, K. M., Lee, D. U., and Rao, A. (2003) An epigenetic view of helper T cell differentiation. *Nat. Immunol.* **4,** 616–623
32. Wilson, C. B., Rowell, E., and Sekimata, M. (2009) Epigenetic control of T-helper-cell differentiation. *Nat. Rev. Immunol.* **9,** 91–105
33. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., and Pappin, D. J. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **3,** 1154–1169
34. Hwang, S. I., Lundgren, D. H., Mayya, V., Rezaul, K., Cowan, A. E., Eng, J. K., and Han, D. K. (2006) Systematic characterization of nuclear proteome during apoptosis: a quantitative proteomic study by differential extraction and stable isotope labeling. *Mol. Cell. Proteomics* **5,** 1131–1145
35. Wu, L., Hwang, S. I., Rezaul, K., Lu, L. J., Mayya, V., Gerstein, M., Eng, J. K., Lundgren, D. H., and Han, D. K. (2007) Global survey of human T leukemic cells by integrating proteomics and transcriptomics profiling. *Mol. Cell. Proteomics* **6,** 1343–1353
36. Yssel, H., De Vries, J. E., Koken, M., Van Blitterswijk, W., and Spits, H. (1984) Serum-free medium for generation and propagation of functional human cytotoxic and helper T cell clones. *J. Immunol. Methods* **72,** 219–227
37. Andrews, N. C., and Faller, D. V. (1991) A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. *Nucleic Acids Res.* **19,** 2499
38. Lindsey, W. B., Lowdell, M. W., Marti, G. E., Abbasi, F., Zenger, V., King, K. M., and Lamb, L. S., Jr. (2007) CD69 expression as an index of T-cell function: assay standardization, validation and use in monitoring immune recovery. *Cytotherapy* **9,** 123–132
39. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72,** 248–254
40. Tang, W. H., Halpern, B. R., Shilov, I. V., Seymour, S. L., Keating, S. P., Loboda, A., Patel, A. A., Schaeffer, D. A., and Nuwaysir, L. M. (2005) Discovering known and unanticipated protein modifications using MS/MS database searching. *Anal. Chem.* **77,** 3931–3946
41. Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J., and Gygi, S. P. (2003) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J. Proteome Res.* **2,** 43–50
42. Elias, J. E., and Gygi, S. P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **4,** 207–214
43. Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20,** 3551–3567
44. Shilov, I. V., Seymour, S. L., Patel, A. A., Loboda, A., Tang, W. H., Keating, S. P., Hunter, C. L., Nuwaysir, L. M., and Schaeffer, D. A. (2007) The paragon algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol. Cell. Proteomics* **6,** 1638–1655
45. Gan, C. S., Chong, P. K., Pham, T. K., and Wright, P. C. (2007) Technical,

experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). *J. Proteome Res.* **6,** 821–827
46. Elo, L. L., Lahti, L., Skottman, H., Kyläniemi, M., Lahesmaa, R., and Aittokallio, T. (2005) Integrating probe-level expression changes across generations of Affymetrix arrays. *Nucleic Acids Res.* **33,** e193
47. Stevens, J. R., and Doerge, R. W. (2005) Combining Affymetrix microarray results. *BMC Bioinformatics* **6,** 57
48. Hu, P., Greenwood, C. M., and Beyene, J. (2007) Integrative analysis of gene expression data including an assessment of pathway enrichment for predicting prostate cancer. *Cancer Inform.* **2,** 289–300
49. Choi, J. K., Yu, U., Kim, S., and Yoo, O. J. (2003) Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics* **19,** Suppl. 1, i84–i90
50. DerSimonian, R., and Laird, N. (1986) Meta-analysis in clinical trials. *Control Clin. Trials* **7,** 177–188
51. Duthie, K. A., Osborne, L. C., Foster, L. J., and Abraham, N. (2007) Proteomics analysis of interleukin (IL)-7-induced signaling effectors shows selective changes in IL-7Ralpha449F knock-in T cell progenitors. *Mol. Cell. Proteomics* **6,** 1700–1710
52. Unwin, R. D., Smith, D. L., Blinco, D., Wilson, C. L., Miller, C. J., Evans, C. A., Jaworska, E., Baldwin, S. A., Barnes, K., Pierce, A., Spooncer, E., and Whetton, A. D. (2006) Quantitative proteomics reveals posttranslational control as a regulatory factor in primary hematopoietic stem cells. *Blood* **107,** 4687–4694
53. Seshi, B. (2006) An integrated approach to mapping the proteome of the human bone marrow stromal cell. *Proteomics* **6,** 5169–5182
54. Oberg, A. L., and Vitek, O. (2009) Statistical design of quantitative mass spectrometry-based proteomic experiments. *J. Proteome Res.* **8,** 2144–2156
55. Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* **57,** 289–300
56. Armenta, J. M., Hoeschele, I., and Lazar, I. M. (2009) Differential protein expression analysis using stable isotope labeling and PQD linear ion trap MS technology. *J. Am. Soc. Mass Spectrom.* **20,** 1287–1302
57. Hakimov, H. A., Walters, S., Wright, T. C., Meidinger, R. G., Verschoor, C. P., Gadish, M., Chiu, D. K., Strömvik, M. V., Forsberg, C. W., and Golovan, S. P. (2009) Application of iTRAQ to catalogue the skeletal muscle proteome in pigs and assessment of effects of gender and diet dephytinization. *Proteomics* **9,** 4000–4016
58. Jacob, R. J., and Cramer, R. (2006) PIGOK: linking protein identity to gene ontology and function. *J. Proteome Res.* **5,** 3429–3432
59. Blazek, E., and Meisterernst, M. (2006) A functional proteomics approach for the detection of nuclear proteins based on derepressed importin alpha. *Proteomics* **6,** 2070–2078
60. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74,** 5383–5392
61. Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75,** 4646–4658
62. Hamalainen, H., Meissner, S., and Lahesmaa, R. (2000) Signaling lymphocytic activation molecule (SLAM) is differentially expressed in human Th1 and Th2 cells. *J. Immunol. Methods* **242,** 9–19
63. Lund, R. J., Ylikoski, E. K., Aittokallio, T., Nevalainen, O., and Lahesmaa, R. (2003) Kinetics and STAT4- or STAT6-mediated regulation of genes involved in lymphocyte polarization to Th1 and Th2 cells. *Eur. J. Immunol.* **33,** 1105–1116
64. Hamalainen, H. K., Tubman, J. C., Vikman, S., Kyrölä, T., Ylikoski, E., Warrington, J. A., and Lahesmaa, R. (2001) Identification and validation of endogenous reference genes for expression profiling of T helper cell differentiation by quantitative real-time RT-PCR. *Anal. Biochem.* **299,** 63–70
65. Chong, P. K., Gan, C. S., Pham, T. K., and Wright, P. C. (2006) Isobaric tags for relative and absolute quantitation (iTRAQ) reproducibility: implication of multiple injections. *J. Proteome Res.* **5,** 1232–1240
66. Chamrad, D. C., Körting, G., Stühler, K., Meyer, H. E., Klose, J., and Blüggel, M. (2004) Evaluation of algorithms for protein identification from sequence databases using mass spectrometry data. *Proteomics* **4,** 619–628
67. Kapp, E. A., Schütz, F., Connolly, L. M., Chakel, J. A., Meza, J. E., Miller, C. A., Fenyo, D., Eng, J. K., Adkins, J. N., Omenn, G. S., and Simpson, R. J. (2005) An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: sensitivity and specificity analysis. *Proteomics* **5,** 3475–3490
68. Moulder, R., Filén, J. J., Salmi, J., Katajamaa, M., Nevalainen, O. S., Oresic, M., Aittokallio, T., Lahesmaa, R., and Nyman, T. A. (2005) A comparative evaluation of software for the analysis of liquid chromatography-tandem mass spectrometry data from isotope coded affinity tag experiments. *Proteomics* **5,** 2748–2760
69. Burns, D. M., Ying, W., Kauppinen, T. M., Zhu, K., and Swanson, R. A. (2009) Selective down-regulation of nuclear poly(ADP-ribose) glycohydrolase. *PLoS One* **4,** e4896
70. Picariello, L., Carbonell Sala, S., Martineti, V., Gozzini, A., Aragona, P., Tognarini, I., Paglierani, M., Nesi, G., Brandi, M. L., and Tonelli, F. (2006) A comparison of methods for the analysis of low abundance proteins in desmoid tumor cells. *Anal. Biochem.* **354,** 205–212
71. Gautier, V. W., Gu, L., O'Donoghue, N., Pennington, S., Sheehy, N., and Hall, W. W. (2009) In vitro nuclear interactome of the HIV-1 tat protein. *Retrovirology* **6,** 47
72. Ow, S. Y., Salim, M., Noirel, J., Evans, C., Rehman, I., and Wright, P. C. (2009) iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly". *J. Proteome Res.* **8,** 5347–5355
73. Hu, J., Qian, J., Borisov, O., Pan, S., Li, Y., Liu, T., Deng, L., Wannemacher, K., Kurnellas, M., Patterson, C., Elkabes, S., and Li, H. (2006) Optimized proteomic analysis of a mouse model of cerebellar dysfunction using amine-specific isobaric tags. *Proteomics* **6,** 4321–4334
74. Grigorian, A., Lee, S. U., Tian, W., Chen, I. J., Gao, G., Mendelsohn, R., Dennis, J. W., and Demetriou, M. (2007) Control of T cell-mediated autoimmunity by metabolite flux to N-glycan biosynthesis. *J. Biol. Chem.* **282,** 20027–20035
75. Koretzky, G. A., and Boerth, N. J. (1999) The role of adapter proteins in T cell activation. *Cell Mol. Life Sci.* **56,** 1048–1060
76. Furuno, T., Hirashima, N., Onizawa, S., Sagiya, N., and Nakanishi, M. (2001) Nuclear shuttling of mitogen-activated protein (MAP) kinase (extracellular signal-regulated kinase (ERK) 2) was dynamically controlled by MAP/ERK kinase after antigen stimulation in RBL-2H3 cells. *J. Immunol.* **166,** 4416–4421
77. Gygi, S. P., Rochon, Y., Franza, B. R., and Aebersold, R. (1999) Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **19,** 1720–1730
78. Li, L., Li, Q., Rohlin, L., Kim, U., Salmon, K., Rejtar, T., Gunsalus, R. P., Karger, B. L., and Ferry, J. G. (2007) Quantitative proteomic and microarray analysis of the archaeon Methanosarcina acetivorans grown with acetate versus methanol. *J. Proteome Res.* **6,** 759–771
79. Lian, Z., Kluger, Y., Greenbaum, D. S., Tuck, D., Gerstein, M., Berliner, N., Weissman, S. M., and Newburger, P. E. (2002) Genomic and proteomic analysis of the myeloid differentiation program: global analysis of gene expression during induced differentiation in the MPRO cell line. *Blood* **100,** 3209–3220
80. Tahvanainen, J., Pykäläinen, M., Kallonen, T., Lähteenmäki, H., Rasool, O., and Lahesmaa, R. (2006) Enrichment of nucleofected primary human CD4+ T cells: a novel and efficient method for studying gene function and role in human primary T helper cell differentiation. *J. Immunol. Methods* **310,** 30–39
81. Hebenstreit, D., Wirnsberger, G., Horejs-Hoeck, J., and Duschl, A. (2006) Signaling mechanisms, interaction partners, and target genes of STAT6. *Cytokine Growth Factor Rev.* **17,** 173–188
82. Ballif, B. A., Villén, J., Beausoleil, S. A., Schwartz, D., and Gygi, S. P. (2004) Phosphoproteomic analysis of the developing mouse brain. *Mol. Cell. Proteomics* **3,** 1093–1101
83. Cai, S., Lee, C. C., and Kohwi-Shigematsu, T. (2006) SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. *Nat. Genet.* **38,** 1278–1288
84. Notani, D., Gottimukkala, K. P., Jayani, R. S., Limaye, A. S., Damle, M. V., Mehta, S., Purbey, P. K., Joseph, J., and Galande, S. (2010) Global regulator SATB1 recruits beta-catenin and regulates T(H)2 differentiation in wnt-dependent manner. *PLoS Biol.* **8,** e1000296
85. Kohno, K., Izumi, H., Uchiumi, T., Ashizuka, M., and Kuwano, M. (2003) The pleiotropic functions of the Y-box-binding protein, YB-1. *BioEssays* **25,** 691–698
86. Lasham, A., Lindridge, E., Rudert, F., Onrust, R., and Watson, J. (2000)

Regulation of the human fas promoter by YB-1, puralpha and AP-1 transcription factors. *Gene* **252,** 1–13
87. Lasham, A., Moloney, S., Hale, T., Homer, C., Zhang, Y. F., Murison, J. G., Braithwaite, A. W., and Watson, J. (2003) The Y-box-binding protein, YB1, is a potential negative regulator of the p53 tumor suppressor. *J. Biol. Chem.* **278,** 35516–35523
88. Homer, C., Knight, D. A., Hananeia, L., Sheard, P., Risk, J., Lasham, A., Royds, J. A., and Braithwaite, A. W. (2005) Y-box factor YB1 controls p53 apoptotic function. *Oncogene* **24,** 8314–8325
89. Lenardo, M., Chan, K. M., Hornung, F., McFarland, H., Siegel, R., Wang, J., and Zheng, L. (1999) Mature T lymphocyte apoptosis—immune regulation in a dynamic and unpredictable antigenic environment. *Annu. Rev. Immunol.* **17,** 221–253
90. Siegel, R. M. (2006) Caspases at the crossroads of immune-cell life and death. *Nat. Rev. Immunol.* **6,** 308–317
91. Vincenz, C., and Dixit, V. M. (1997) Fas-associated death domain protein interleukin-1beta-converting enzyme 2 (FLICE2), an ICE/Ced-3 homologue, is proximally involved in CD95- and p55-mediated death signaling. *J. Biol. Chem.* **272,** 6578–6583
92. Wang, J., Chun, H. J., Wong, W., Spencer, D. M., and Lenardo, M. J. (2001) Caspase-10 is an initiator caspase in death receptor signaling. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 13884–13888
93. Basaki, Y., Hosoi, F., Oda, Y., Fotovati, A., Maruyama, Y., Oie, S., Ono, M., Izumi, H., Kohno, K., Sakai, K., Shimoyama, T., Nishio, K., and Kuwano, M. (2007) Akt-dependent nuclear localization of Y-box-binding protein 1 in acquisition of malignant characteristics by human ovarian cancer cells. *Oncogene* **26,** 2736–2746
94. Quirion, M. R., Gregory, G. D., Umetsu, S. E., Winandy, S., and Brown, M. A. (2009) Cutting edge: Ikaros is a regulator of Th2 cell differentiation. *J. Immunol.* **182,** 741–745
95. Thomas, R. M., Chen, C., Chunder, N., Ma, L., Taylor, J., Pearce, E. J., and Wells, A. D. (2010) Ikaros silences T-bet expression and interferon-gamma production during T helper 2 differentiation. *J. Biol. Chem.* **285,** 2545–2553
96. Elo, L. L, Järvenpää, H., Tuomela, S., Raghav, S., Ahlfors, H., Laurila, K., Gupta, B., Lund, R. J., Tahvanainen, J., Hawkins, R. D., Orešič, M., Lähdesmäki, H., Rasool, O., Rao, K. V., Aittokallio, T., and Lahesmaa, R. (2010) Genome wide Profiling of Interleukin-4 and STAT6 Transcription Factor Regulation of Human Th2 cell Programming. *Immunity.* **32**(6), 852–862.
97. Ahlfors, H., Limaye, A., Tuomela, S., Rihko, M., Rasool, O., Galande, S., and Lahesmaa, R. (2010) SATB1 dictates expression of multiple genes including IL5 involved in T helper cell differentiation. *Blood*, in press