

¹Institute of Cell Biology and Neurobiology (IBCN), CNR, via Ramarini 32, 00015 Monterotondo, Italy.

²Mouse Biology Unit, European Molecular Biology Laboratory, via Ramarini 32, 00015 Monterotondo, Italy

³Department of Medicine, University of Perugia, Perugia, Italy, 06132

⁴MRC Molecular Hematology Unit, Weatherall Institute of Molecular Medicine University of Oxford, John Radcliffe Hospital, Headington, Oxford OX3 9DS

¶ These authors contributed equally to this work

* corresponding authors: olga.ermakova@emma.cnr.it, claus.nerlov@imm.ox.ac.uk

Abstract

Background: C/EBPa and C/EBPb are transcription factors with tissue specific expression regulating several important cellular processes. They work by recruiting protein complexes to a common DNA recognition motif and both are able to compensate each other's absence in many cell types, thus showing functional redundancy. They also play distinct roles in specific cellular pathways and their abnormal functioning gives rise to different human pathologies.

Methods: To investigate the molecular basis of C/EBPa and C/EBPb specificity and redundancy we characterized their *in vivo* protein-protein interaction networks by Tandem Affinity Purification and Mass Spectrometry. To unravel the functional features of C/EBPa and C/EBPb proteomes we studied the statistical enrichment of binding partners related to GO terms and KEGG pathways.

Results: Our data confirmed that the C/EBPa and C/EBPb regulate biological processes like cell proliferation, apoptosis and transformation. We found that both C/EBPa and C/EBPb are involved in other cellular pathways such as RNA maturation, RNA splicing and DNA repair. Specific interactions of C/EBPa with MRE11, RUVBL1 and RUVBL2 components of DNA repair system were confirmed by co-immunoprecipitation assays.

Conclusions: Our comparative analysis of the C/EBPa and C/EBPb proteomes provides an insight for understanding both their redundant and specific roles in cells indicating their involvement in new pathways. Such novel predicted functions are relevant to normal cellular processes and disease phenotypes controlled by these transcription factors.

General Significance: Functional characterization of C/EBPa and C/EBPb proteomes suggests they can regulate novel pathways and indicate potential molecular targets for therapeutic intervention.

Keywords: CCAAT/enhancer binding proteins C/EBPa and C/EBPb;
large scale protein-protein interactions;
network functional analysis;
DNA repair protein complexes;
RNA metabolism;
adipogenesis

1. Introduction

The CCAAT/enhancer-binding proteins (C/EBP) is a family of six transcription factors (TFs) characterized by a highly conserved C-terminal bZIP domain formed by a dimerization region (leucine zipper, ZIP) next to a positively charged DNA binding domain (DBD). Among C/EBPs, C/EBPa and C/EBPb are the two most studied members.

Gene knockout studies and mutagenesis in mice showed separate and distinct functions for C/EBPa and C/EBPb in some tissues while in others, where they are co-expressed, one can

compensate for another assuming common function [1,2]. Genetic loss of function studies established the requirement of C/EBP α in granulopoiesis and lung epithelium physiology while both C/EBP β and C/EBP α are critical for normal adipogenesis. C/EBP α regulates cell quiescence of several cell types including hematopoietic stem cells and mutations in C/EBP α gene causes cancer of myeloid lineage [3]. C/EBP β is essential for the proper differentiation of mammary epithelial cells, granulosa cells, brown adipose tissue, chondrocytes and osteoblasts, reviewed in [2,4]. It also plays an important role in specific physiological processes of the central nervous system development, such as memory formation [5] as well as regulating the inflammatory responses in microglia [6].

The mechanisms by which C/EBP α and C/EBP β execute their instructive role while carrying out tissue specificity as well as those underlying their redundancy are poorly understood. However, a large amount of experimental data support the notion that their functions are sustained by a collaborating protein network or, in other words, by a grid of specific protein-protein interactions (PPI) [2].

The characterized binding partners of C/EBP α and C/EBP β could be divided into four main groups: (1) components of basal transcription machinery; (2) components of chromatin remodeling complexes; (3) lineage-specific transcription factors and (4) proteins that regulate cell cycle progression. For example, transcriptional activation from the promoters exhibiting the C/EBP recognition site needs C/EBP α and C/EBP β direct interactions with the general transcription factors such as TBP/TFIIB, CBP/p300, chromatin remodeling SWI/SNF (BRG1) and ISWI(SNF2H) complexes [7] and reviewed in [8]. The transcriptional activation of tissue-specific gene expression is obtained by combinatorial binding with other lineage-specific transcription factors: in myeloid cells both C/EBP α and C/EBP β interact with cMyb/PU.1 [9], in adipogenesis C/EBP α interacts with PPAR γ , in brown fat cells differentiation program is determined by C/EBP β -PRDM16 [10] and C/EBP β -Plac8 complexes [11] while both chondrocytes and osteoblasts differentiation are controlled by the C/EBP β -Runx2-ATF complex [12]. Moreover, C/EBP α and C/EBP β drive cell cycle arrest by direct protein-protein interactions with positive regulators of the cell cycle progression E2F thus inhibiting pro-proliferative E2F-dependent transcription [13,14]. Therefore, regulated protein-protein interactions act dictating the functional specificity and affinity of C/EBP α and/or C/EBP β to specific enhancer regions and drive their common and distinct gene expression programs.

Several studies demonstrated the substantial redundancy of C/EBP α and C/EBP β functions. C/EBP α or C/EBP β single knock out animals show complete embryonic development [2], while mice with combined C/EBP α and C/EBP β gene deletions die at early embryonic stage [15,16]. Also, C/EBP β expression from the C/EBP α locus can rescue perinatal embryonic lethality phenotype of C/EBP α deletion [15]. Conditional gene deletion studies provide additional evidences of functional redundancy of C/EBP α and C/EBP β in cortical neurons and epidermal keratinocytes differentiation [2,5]. Taken together, all the above data strongly support the notion that both C/EBP α and C/EBP β work within macromolecular protein complexes and that deregulation of their interactomes might lead to pathogenic phenotypes.

In this study we characterize C/EBP α and C/EBP β protein interactions by Tandem Affinity Purification (TAP) technology on an engineered cell line, Liquid Chromatography/Mass Spectrometry (LC-MS) and perform *in silico* network functional clustering to elucidate the distinctive and common functional pathways sustained by the proteomes associated with these two transcription factors.

2. Materials and methods

2.1 Cell cultures

Mouse embryonic fibroblast NIH-3T3 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Gibco), supplemented with 10% of FBS (Hyclone), 50units/mL penicillin and 50 μ g/ml streptomycin (Gibco) and grown at 37°C in a humidified atmosphere with 5% CO $_2$.

2.2 Establishing cell lines constitutively expressing C/EBP α -TAP and C/EBP β -TAP

Gene specific cassette containing TAP-tag was inserted in frame at the C-terminus of C/EBP α or C/EBP β cDNA and subcloned into BamHI-EcoRI restriction sites of pBabe-Puro

retroviral vector [17]. Retrovirus production and NIH-3T3 infection were performed as described [13].

2.3 Cell collection, lysis and nucleic acids removal for TAP purification

Stably expressing C/EBPa-TAP or C/EBPb-TAP NIH-3T3 cells were grown on 15mm dishes until confluence. Cells were trypsinized (0.5 % trypsin); trypsin was inactivated by adding serum containing DMSO medium; cells were pelleted by centrifugation at 1000g and washed twice with ice cold PBS. Cell pellet was fast-frozen in liquid nitrogen and stored at -70°C. For mass spectroscopy analysis, TAP purifications were performed from 1.12×10^9 C/EBPa-TAP expressing and 0.9×10^9 C/EBPb-TAP expressing NIH-3T3 cells, respectively. Cells were incubated in hypotonic cell lysis buffer: 10mM Tris-HCl pH 7.4, 1.5mM MgCl₂, 10mM KCl, 25mM NaF, 0.5mM Na₃VO₄, 1mM DTT and cocktail of protease inhibitors (Roche, cat:1 836 170), incubated on ice for 10 min, homogenized in dounce homogenizer and centrifuged for 10 min at 2,000g. The supernatant was saved on ice, cell pellet was re-suspended in nuclear extraction lysis buffer: 50mM Tris-HCl pH 7.4, 1.5mM MgCl₂, 20% (v/v) glycerol, 420mM NaCl, 25mM NaF, 0.5mM Na₃VO₄, 1mM DTT with protease inhibitors and incubated on ice for 30 min with frequent vigorous agitation. NaCl concentration was lowered to 146mM by adding 2 volumes of dilution buffer: 50mM Tris-HCl pH 7.4, 1.5mM MgCl₂, 0.26% NP40, 25mM NaF, 0.5mM Na₃VO₄, 1mM DTT with protease inhibitors. Extracts were incubated on ice for 10 min and cleared by centrifugation at 100,000g for 1 hour.

The nuclear extracts were collected, the NaCl concentration was lowered to 146mM NaCl with a buffer containing 50mM Tris-HCl pH 7.4, 1.5mM MgCl₂, 0.26% NP40, 25mM NaF, 0.5mM Na₃VO₄, 1mM DTT, protease inhibitors. To obtain nucleic acids removal, 10,000 units of Benzonase nuclease (Sigma Aldrich E1014) were added and incubation was performed for 30 min at 4°C with mixing.

2.4 Affinity purification using IgG beads

After NaCl concentration was adjusted to 146mM NaCl, nuclear extract and cytoplasmic fraction were combined, incubated for 2 hours at 40°C on rocking shaker with 400µl of unsettled rabbit IgG agarose (cat: A-2909, Sigma-Aldrich) and equilibrated with lysis buffer: 50mM Tris-HCl pH 7.4, 5% (v/v) glycerol, 0.2% (v/v) IGEPAL, 1.5mM MgCl₂, 100mM NaCl, 25mM NaF, 1mM Na₃VO₄, 1mM DTT, proteinase inhibitors. Beads were collected by centrifugation, transferred in Mobicol column (M1002; MoBiTec), washed by gravity flow first with 10ml of lysis buffer and then with 5ml of TEV buffer (10mM Tris-HCl pH 7.4, 100mM NaCl, 0.1% (v/v) IGEPAL, 0.5mM EDTA, 1mM DTT). Release of protein complexes was performed by incubation in 400µl of TEV buffer with 120 units of AcTevTM protease (cat:12575-015; Invitrogen) for 1 hour at 16°C with shaking.

For affinity purification, 400µl of Calmodulin affinity resin (cat: 214303-52; Stratagene) were washed 3 times with wash buffer: 10mM Tris-HCl pH 7.4, 100mM NaCl, 0.1% (v/v) IGEPAL, 1mM DTT, 2mM CaCl₂ and then transferred into Mobicol column. Tev eluates from the first affinity purification step were collected into eppendorf tubes where 1/3 volume of TEV buffer containing 6mM of CaCl₂ without EDTA was added in order to reach a final 2mM CaCl₂ concentration. The ongoing TEV reaction was left to go for 60 min at 4°C with shaking. The columns were washed with 10ml TEV buffer without EDTA and eluted with 600µl of 10mM Tris-HCl pH 8.0 and 5mM EGTA.

2.5 Protein identification by LC-MS

Protein identification was performed as described [18,19]. TAP purified samples were concentrated and analyzed on 4-12% NuPAGE gels (NPO 335, Invitrogen). After staining with colloidal Coomassie blue (cat: B-2025, Sigma), gels were sliced into 1.25mm bands along the entire separation range of each lane in order to include all potential interacting proteins without bias toward size and abundance. Cut slices were digested with trypsin using a previously described protocol [20]. The resulting tryptic peptide mixtures were analyzed by automated MALDI-TOF MS (Voyager DE-STR, Applied Biosystems). Proteins identification was accomplished by automated LC-MS analysis (Ultimate, LC Packings, QTOF2, Micromass) in conjunction with searches of the GenPept database (<ftp://ftp.ncbi.nlm.nih.gov/genbank/genpept.fsa.gz>) using the software tool Mascot (Matrix Science).

oxana bereshchenko 7/9/16 13:52

Comment [1]: You can use upscript and downscript that are on the right from bold, underlines etc buttons

All LC-MS experiments were performed on the mass spectrometry platform in Cellzome GmbH according to their quality control standards (Heidelberg, Germany).

2.6 Western blot analysis

Western blot analysis was performed as described [21]. For C/EBPa visualization, immunoblots were incubated with anti-C/EBPa antibodies (14AA, 1:1500, Santa Cruz), signals visualized with horseradish peroxidase-conjugated secondary antibodies (GE Healthcare) and enhanced by chemiluminescent (ECL) system (cat: RPN2108, GE Healthcare). For TAP-tagged proteins detection, immunoblots were incubated with PAP (Peroxidase-Anti-Peroxidase Soluble Complex) diluted 1:1000 (cat: P1291, Sigma) and directly imaged by ECL. Visualization of Flag-tagged proteins was performed with monoclonal anti-Flag antibodies (cat: F7425, 1:2500 dilution, Sigma), anti-HA antibodies (cat: 11867423001, Roche), polyclonal anti-RUVBL1 and RUVBL2 antibodies (a gift from Dr. Bruno Amati, Italian Institute of Technology), Mre11 antibodies (Novus, cat: NB 100-142) at 1:5000 dilution, RFC1 (Chemicon International, cat: 09-094) at 1:200 dilution. Nuclear extracts were prepared from NIH-3T3 stable expressing rat C/EBPa protein as described [22].

2.7 Co-immunoprecipitation

Nuclear extracts were prepared from NIH-3T3 stable expressing rat C/EBPa protein as described above (2x10⁶ confluent cells) but omitting the benzonase treatment. Co-immunoprecipitation (CoIP) was performed with C/EBPa antibodies (14AA, Sigma), and immunocomplexes were recovered on protein A Sepharose (cat: 22810, Pierce) equilibrated with CoIP lysis buffer (50mM Tris-HCl pH 7.4, 150mM NaCl, 0.18% NP-40, 10% glycerol, 1.5mM MgCl₂, 25mM NaF, 0.5mM Na₃VO₄, 1mM DTT) by incubation for 3 hours followed by five washes with 0.8ml of CoIP buffer and elution with Laemmli loading buffer. Western blots were incubated with anti-RFC-145 antibodies (cat: AB3770, Chemicon International) or anti-Mre11 antibodies (cat: NB100-142, Novus Biological) and developed as described above. When CoIPs were performed with HA or Flag-tagged proteins, NIH-3T3 cells were transfected by calcium-phosphate method with the RUVBL1-HA [23] and RFC5-flag-pCDNA expressing plasmids. CoIP was performed as described [24].

2.8 Interactomes statistical analysis

The list of interactors obtained by mass spectrometry in the form of IPI codes (145 for C/EBPa and 102 for C/EBPb) were filtered out by applying an *in silico* selection to get rid of those proteins (ribosomal, cytoskeletal and heat shock), which are widely considered to be false positives hits due to their abundance or inherent molecular stickiness [18,25]. To this purpose, the C/EBPa and C/EBPb raw hits were inquired through the CRAPome repository (<http://www.crapome.org>) [26]. After such selection, filtered sets of 64 and 51 interacting proteins for C/EBPa and C/EBPb respectively were obtained.

Filtered C/EBPa and C/EBPb interaction sets were analyzed by Cytoscape 3.2.1 and its plugins ClueGO and CluePedia [27,28] to scrutiny Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways repositories for the presence of functional enriched clusters. Statistical dependencies (correlation) for proteins in the input list were calculated by hypergeometric tests with Bonferroni P-value correction. The resulting networks were generated by Cytoscape with connection nodes above the 0.4 kappa threshold.

3. Results

3.1 Identification of C/EBPa and C/EBPb interactomes by TAP/LC-MS

To define the C/EBPa and C/EBPb interaction networks we used TAP coupled to LC/MS on cell extracts of NIH-3T3 cells [18,19]. To this aim, we tagged both C/EBPa and C/EBPb at C-terminus by adding a TAP-tag composed of two affinity moieties, Calmodulin Binding Peptide (CBP) and protein A separated by a TEV-protease cleavage site [18].

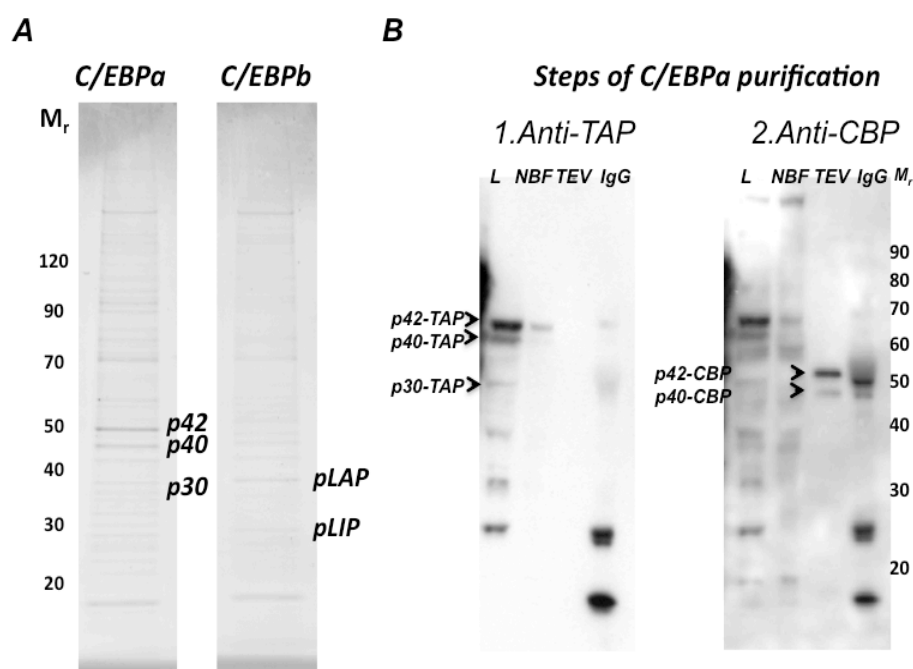


Figure 1. TAP of C/EBPa and C/EBPb partners. **A.** SDS-PAGE gel stained with colloidal Coomassie showing TAP complexes associated with C/EBPa and C/EBPb. **B.** Western blot analysis of TAP steps of C/EBPa. (1) Western blot analysis was performed using PAP antibodies, which detect protein A inside the affinity tag. (2) Western blot analysis was performed with anti-calmodulin peptide antibodies. L – total lysate; NBF – non bound fraction after incubation with IgG sepharose; IgG – fraction of IgG beads after TEV elution; TEV – fraction of TEV elution.

In human and mice, C/EBPa and C/EBPb are single exon containing genes. Yet, different translation initiation sites give rise to several isoforms: C/EBPa is translated as three isoforms named p42, p40 and p30, while C/EBPb is translated as LAP*-C/EBPb, LAP-C/EBPb and LIP-C/EBPb [29,30]. Normal cells mainly produce the long isoforms p42-C/EBPa and LAP-C/EBPb [29]. We engineered NIH-3T3 cell lines with constitutive C/EBPa-TAP or C/EBPb-TAP longest isoforms expression by using a retroviral gene integration system [13]. Placing the TAP-tag at C-terminus, we forced the expression and translation of all C/EBPa and C/EBPb isoforms with the TAP tag on. To accomplish TAP of C/EBPs associated protein complexes, the cells were grown till contact inhibition and collected before the adipogenic differentiation program would begin. The TAP purification scheme yields protein complexes with adequate efficiency and specificity under physiological conditions [31]. Purified proteins belonging to C/EBPa and C/EBPb associated complexes were separated by SDS-PAGE gel and visualized by Colloidal Coomassie staining. We detected a substantial enrichment of all C/EBPa and C/EBPb isoforms during purification (**Figure 1A**). Western blot analysis on purified fractions confirmed such observation (**Figure 1B**).

TAP/LC-MS allowed us to identify 137 and 99 interactors for C/EBPa and C/EBPb respectively (**Supplementary Tables 1 and 2**). These raw datasets were subjected to selection of spurious hits through CRAPome database filtering, a repository database that keeps track of the contaminant proteins found in over 300 negative control affinity purification and mass spectrometry experiments. Because non-specific interactions are inherently bait independent, this approach allows for an effective exclusion of possible false positive hits. CRAPome datasets inspection

allowed us to discriminate against known sources of noise, mainly cytoskeletal and ribosomal proteins, often found in affinity based experiments. After CRAPome background removal, the datasets listed 64 and 51 interacting proteins for C/EBPa and C/EBPb respectively

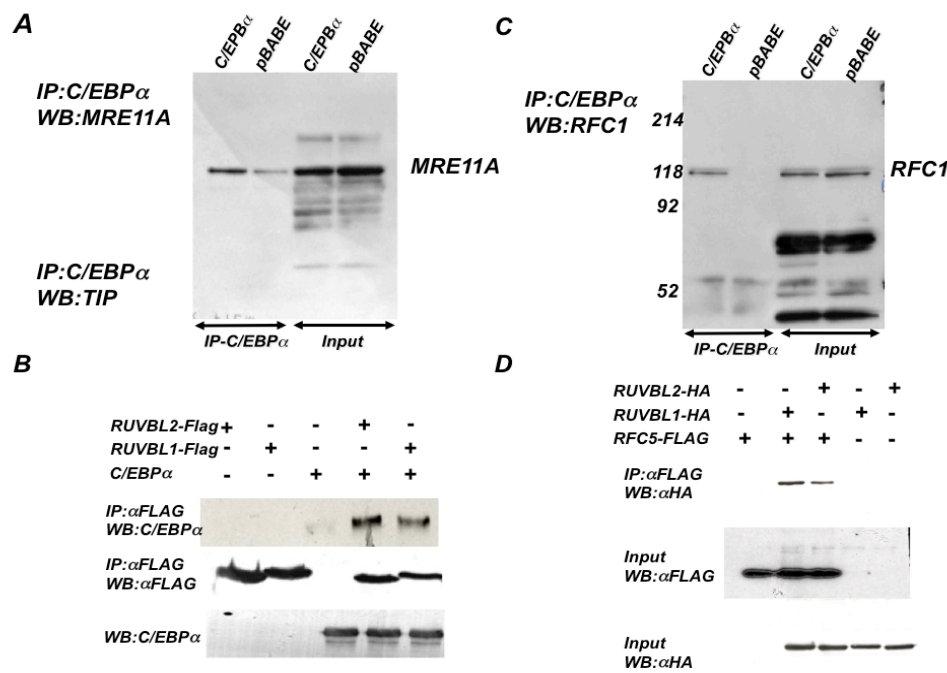
(Supplementary Table 3).

Affinity purification and LC/MS studies on C/EBPa protein network were previously accomplished in mouse myeloid progenitor cell line FDPC-1 using several single and double tags [32]. This paper authors remarked that both the nature of tags and the following purification procedures might significantly affect the composition of proteome associated with C/EBPa. Therefore, we compared the list of the C/EBPa interactors obtained from the six diverse affinity purifications in FDPC-1 cells with the C/EBPa-TAP proteome presented in this report. 92 proteins detected in our non filtered and 35 in our filtered data sets are also found in one or more of the affinity purifications listed in [32] **(Supplementary Tables 1 and 3).**

Analysis of C/EBPb interactome has been also previously performed in mouse and human cells by CoIP [33,34]. A comparison with these data shows that our non filtered set shares 32 proteins while the filtered one 18 proteins **(Supplementary Tables 2 and 3)**. Thus, the large overlaps with other independent results confirm the efficacy of our experimental approach and identifies novel protein-protein interactions.

3.2 C/EBPa is an in vivo binding partner of DNA repair proteins MRE11 and RUVBL1

In order to validate some new interactions identified by high-throughput TAP/LC-MS, we performed selected CoIP experiments on specific binders. First, we confirmed that C/EBPa interacts with the MRE11 protein, a component of the DNA repair complex. While large-scale TAP purification using mouse cell line FDPC-1 showed that MRE11 and C/EBPa co-purify [32] a validation of this finding by immunoprecipitation analysis was missing. Our CoIP experiment results demonstrate a substantial enrichment of MRE11 associated with C/EBPa confirming both our novel and previously published results **(Figure 2A)**.



Cirilli M., Bereshchenko O. et al., Figure 2

Figure 2. Coimmunoprecipitation analysis of the proteins identified by TAP-affinity purification.
A. C/EBP α interacts with MRE11A protein, components of DNA repair cellular machinery. CoIPs from stably transformed with C/EBP α -pBABE or empty pBABE vector NIH-3T3 cells were performed with anti-C/EBP α antibodies and western blots were incubated with MRE11A antibodies. **B.** C/EBP α co-precipitated with RUVBL1 and RUVBL2. Co-IPs were performed on NIH-3T3 cells transiently expressing RUVBL1-Flag or RUVBL2-Flag and C/EBP α . Antibodies for CoIP and western blotting analysis are indicated. **C.** C/EBP α interacts with RFC1 protein. Immunoprecipitation with anti-C/EBP α antibodies of nuclear extract from stable transformed C/EBP α -pBABE or empty pBABE vectors NIH-3T3 cells, followed by Western blot analysis with RFC1 antibodies. **D.** Replication protein RFC5 forms a complex with RUVBL1 and RUVBL2 components of chromatin remodeling complexes. NIH-3T3 cells were transfected with RUVBL1-HA or RUVBL2-HA and RFC5-Flag carrying plasmids. Antibodies for the Co-IPs and western blot analysis are indicated.

Likewise, we confirmed the RUVBL1-C/EBP α interaction obtained in our TAP purifications. Since RUVBL1 forms heterodimers with RUVBL2, we also tested whether both proteins are in complex with C/EBP α . Our immunoprecipitation results demonstrated that C/EBP α indeed forms a complex with both proteins, RUVBL1 and RUVBL2 (**Figure 2B**).

3.3 C/EBP α forms a complex with Replication Factor C subunits and RUVBLs proteins

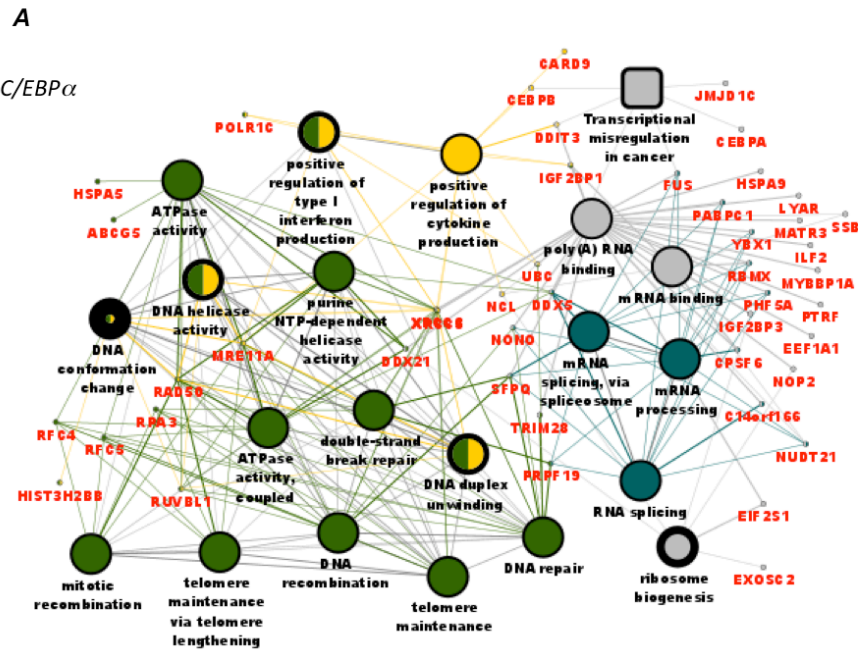
Our proteomic experiments detected two of the five subunits of Replication factor C (RFC) complex, namely RFC4 and RFC5. RFC is a heteropentameric AAA+ protein complex directly involved in DNA replication with a loading function for PCNA processivity factor on DNA. Within this complex the large subunit RFC1 interacts directly with both RFC4 and RFC5 subunits [35]. Therefore, we hypothesized that C/EBP α could interact with all the above three RFC subunits and have examined whether also the RFC1 subunit interacts with C/EBP α . CoIP experiments demonstrated that a C/EBP α -RFC1 complex is actually present in stably transfected NIH-3T3 cells with C/EBP α gene but not in the NIH-3T3 cells which do not express C/EBP α (**Figure 2C**). The

ColP results and TAP purification data strongly suggest an *in vivo* association between the RFC complex and C/EBPa. Since TAP methods detects both direct and indirect protein-protein interactions, we questioned whether the replication factors detected by us and others within C/EBPa proteome [32] could bind components of the chromatin remodeling complexes containing RUVBL1 and RUVBL2 as well. Immunoprecipitation analysis demonstrated that RFC5 interacts with both RUVBL proteins (**Figure 2D**).

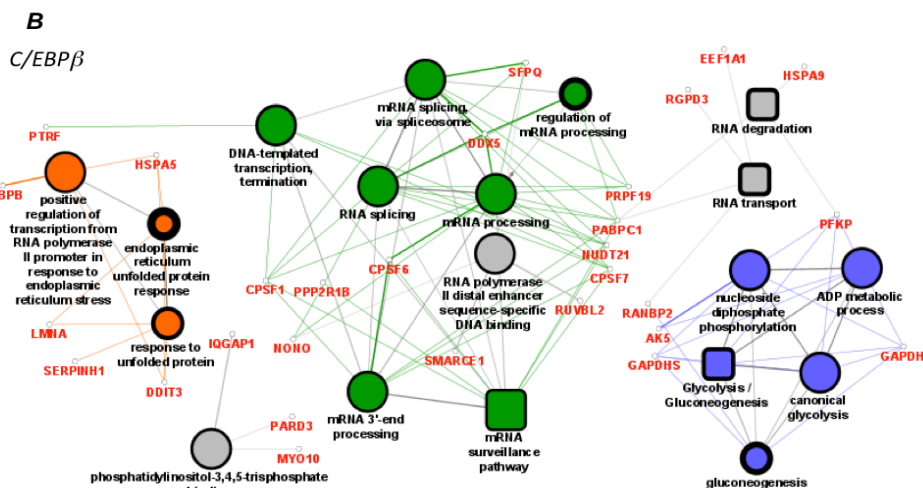
Taken together, our TAP/LC-MS and ColP results suggest that C/EBPa is probably a component of large multi-protein complexes containing C/EBPa, RUVBL1, RUVBL2 and subunits of RFC complex.

3.4 Analysis of cellular pathways associated with C/EBPa and C/EBPb interactomes

An annotation analysis on our proteomic datasets let us to highlight the functional peculiarities of C/EBPa and C/EBPb proteomes. We searched, within the C/EBPa and C/EBPb interacting proteins sets, for significant over-represented Biological Process and Molecular Function GO terms and KEGG cellular pathways associated with the identified proteins. This analysis gave the possibility to annotate and compare C/EBPa and C/EBPb functions based on their protein binding partners. Integration of GO dictionary with other sources of *a priori* knowledge, such as the KEGG database of signaling pathways, has been obtained by using the Cytoscape network representation capabilities together with some of its plugins features. ClueGo/CluePedia plugin allowed us to extract non-redundant biological information from our datasets by integrating several cellular signaling pathways *in silico* data and experimentally assessed protein-protein interactions available in curated databases. In this analysis we used the GO Biological Processes and Molecular Function terms and the genomic, chemical and systemic functional information provided by KEGG [36,37]. **Figures 3A** and **3B** show the maps of statistically significant GO terms and KEGG pathways functionally associated with C/EBPa and C/EBPb. The same results with additional statistics (P-values for terms and groups, number of associated genes and more) for both C/EBPa and C/EBPb are available in separated tables (**Supplementary Table 4**).



Cirilli M., Bereshchenko O. et al., Figure 3A



Cirilli M., Bereshchenko O. et al., Figure 3B

Figure 3. Analysis of the statistical significant pathways detected in C/EBP α and C/EBP β proteomes.

Statistically significant GO terms in each network are mapped as large circular nodes while KEGG pathways by squares. A colour code is used to show clustering by functional similarity. Ungrouped functional outliers are shown in grey. Large node thickness is inversely proportional to the associated corrected probability value. Small nodes represent the proteins, labelled with red geneID names, associated with the shown statistical significant GO and KEGG terms. The networks also illustrate the functional relationships (the edges) between the nodes whose thickness is directly proportional to the association significance score.

A. Gene Ontology analysis of C/EBP α proteome; **B.** Gene ontology analysis of the C/EBP β proteome.

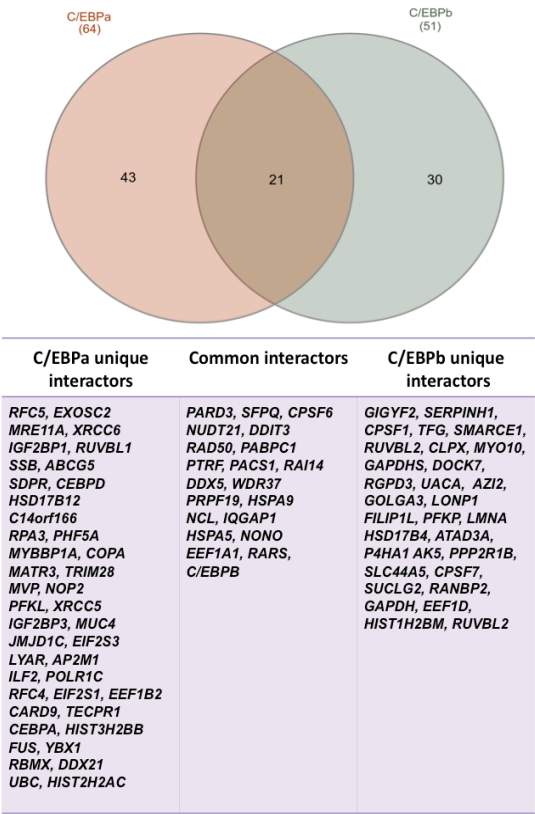
Inspection of the statistically significant GO/KEGG terms associations showed differences between C/EBP α and C/EBP β proteomes. Among the most significant GO/KEGG clusters associated with C/EBP α proteome, there are proteins involved in several functions related to DNA metabolism (DNA recombination, double-stranded break repair, telomere maintenance, mitotic recombination) (**Figure 3A**) and RNA metabolism, such as RNA splicing, RNA splicing via spliceosome and RNA processing. The known role of C/EBP α as transcription factor is delineated by GO terms such as DNA helicase activity, DNA duplex unwinding (shared with the DNA metabolism GO terms) and positive regulation of the type I interferon and cytokine production as well as transcriptional regulation in cancer. Separate non-clustered categories, such as polyA RNA binding, mRNA binding and ribosome biogenesis also occurred (**Figure 3A**).

Analysis of the C/EBP β proteome shows fewer significant enriched GO terms compared to C/EBP α proteome. Its dataset conveys a substantial enrichment of GO terms associated with RNA metabolism (mRNA splicing via spliceosome, RNA splicing, mRNA processing, mRNA3' processing, regulation of mRNA processing, DNA-template transcriptional termination, mRNA surveillance) (**Figure 3B**). Our analysis also identified a large cluster of GO terms related to unfolded protein response, suggesting a potential role of C/EBP β in regulating apoptosis. RNA polymerase II distal enhancer sequence-specific binding GO term is in accordance with the known

functional role of C/EBPb as a transcriptional regulator. In addition, inspection of C/EBPb network for statistically significant items showed unexpected GO clusters related to metabolic processes, such as glycolysis/gluconeogenesis, canonical glycolysis, nucleotide diphosphate phosphorylation and ADP metabolic process.

3.5 Differential proteomic and annotation analysis of the C/EBPa and C/EBPb interactomes

The size of our datasets allowed us to use same statistical methods of functional enrichment to address specificities and redundancies of C/EBPa and C/EBPb experimental networks. We first analyzed C/EBPa and C/EBPb proteomic data to identify unique and common hits in each dataset. Overall, C/EBPa and C/EBPb datasets share 21 proteins, accounting for 30% and 42% of their totals. We also detected that 70% of C/EBPa associated and 58% of C/EBPb associated binding partners are exclusive in each dataset (**Figure 4**).



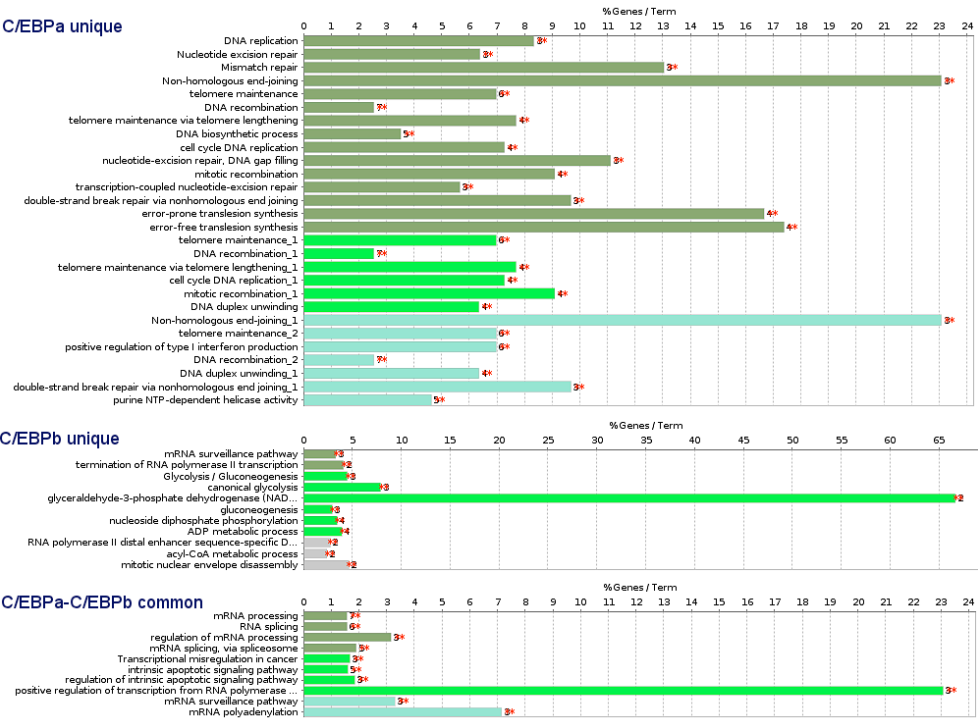
Cirilli M., Bereshchenko O. et al., Figure 4

Figure 4. Comparative analysis of C/EBPa and C/EBPb protein partners. Venn diagram showing the relationships between the 64 C/EBPa and 51 C/EBPb binding partners identified by TAP-LC/MS and CRAPOME filtering. The table lists unique C/EBPa associated (43 proteins), C/EBPb associated (30 proteins) and common (21 proteins) interactors.

In order to get an objective evaluation of the divergence and common traits in C/EBPa and C/EBPb proteomic networks, we performed statistical assessments of functional grouping, like

those described above, for interactors belonging to specific differential subsets: unique C/EBPa (i.e. all the proteins interacting with C/EBPa but not with C/EBPb), unique C/EBPb (i.e. all the proteins interacting with C/EBPb but not with C/EBPa) and C/EBPa-C/EBPb common (i.e. all the proteins interacting with both C/EBPa and C/EBPb) (**Figure 5 and Supplementary Table 5**). The GO and KEGG terms obtained analyzing the common C/EBPa-C/EBPb interactors display their activity in transcriptional driven biological processes such as positive regulation of transcription and regulation of the processes involved in cell proliferation, apoptosis and transcriptional dis-regulation in cancer. Moreover, this analysis showed that both C/EBPb and C/EBPa might be involved in pathways, such as surveillance, polyadenylation and splicing, important for proper mRNA production and maturation.

The GO terms unique to C/EBPa and C/EBPb suggest specific and distinctive functions for each group of C/EBPa and C/EBPb interactors (**Figure 5**).



Cirilli M., Bereshechenko O. et al. Fig. 5

Figure 5. Comparative enrichment analysis of TAP-purified C/EBPs proteomes. Results of the enrichment tests for differential subsets of proteins associated exclusively with the C/EBPa (Unique C/EBPa) or C/EBPb (Unique C/EBPb) and the proteins associated with both C/EBPa and C/EBPb proteomes (common) are shown. Only statistically significant GO genes/term associations are presented. Percentage of gene/GO terms are represented by bar lengths, the numbers at the top of the bar indicate the number of the genes associated with a specific GO term. Within each set analysis, each colour represents a distinct significant functional GO or KEGG cluster. Therefore, terms with identical colour belong to the same higher order ontology group (**Supplementary Table 5**). Same colours from different analysis are not related each other.

From close inspection of the results obtained from these sets, C/EBPa preferentially associates with nuclear processes like transcription, telomere maintenance and lengthening, DNA repair, whereas C/EBPb does not and exhibiting specific links to different pathways instead.

4. Discussion

Gene knockout and mutagenesis studies in mice showed distinct C/EBPα and C/EBPβ functions in some tissues, while in others, where they are co-expressed, one can compensate for another assuming common function [1,7,2]. Also, the mechanisms underlying C/EBPα and C/EBPβ tissue specificity as well as their redundancy are poorly understood. Many pieces of evidence lead to the hypothesis that the environment of the cell, represented by a specific collaborating protein network, sustain and determine their functional behaviour [2]. Thus, high-throughput approaches aimed to the characterization of C/EBPα and C/EBPβ associated PPIs are valuable tools to address how these interactomes define the specificity of their functional properties. We decided to use a TAP/LC-MS scheme in this study because it allows for high-throughput identification of purified complexes present at very low concentrations without lack of selectivity, purity and yield typically associated with large-scale experiments [39]. However, when using affinity capture methods to study PPI networks it is necessary to be aware of both the inherent limitations of such experiments and the outcome variability dependent on elements like the type of tags, the cell lines used, the natural isoform variants and the number of post-translational modifications characteristic of the particular bait under study, the dynamic and transient binding of some interactors. The above issues coupled to variable experimental setups and incomplete filtering of false-positive hits, lead to a marked reduction of both the number of identified interaction partners and the overlap among independent experiments. As a result, a partial and fragmented coverage of the real *in vivo* network is obtained. Although incomplete and sparse in nature, such experimental PPI networks are still a precious source of knowledge because they allow novel functional hypotheses to be proposed when analyzed by proper statistical tools such as enrichment tests.

Our data analysis shows a remarkable overlap between C/EBPα and C/EBPβ TAP/LC-MS datasets denoting a significant degree of operational redundancy. Indeed, the 30% of C/EBPα and 40% of C/EBPβ binding partners are shared between the two proteomes. The analysis of statistically significant enriched GO terms and KEGG pathways for the common partners highlighted their well known role as transcriptional regulators in cancer. These findings are supported by the presence in the C/EBPα-C/EBPβ common set of C/EBPα-DDIT3 and C/EBPβ-DDIT3 interactions in combination with the presence DDX5, NONO, SFPQ proteins. Even though those proteins are not classified as transcription factors, they are critical regulators of transcription from different promoters [40,41,42]. In addition, DDX5, NONO and SFPQ proteins were found in C/EBPα proteome of mouse myeloid progenitor cells (FDCP.1) [32], suggesting that these interactions are not restricted by tissue specificity.

Here, we validated the presence of an intracellular RUVBL1-RUVBL2-C/EBPα ternary subcomplex. We demonstrated that both RUVBL1 (by immunoprecipitation) and RUVBL2 (proteomic data set) are interactors of C/EBPα, while only RUVBL1 is a C/EBPβ binding partner (proteomic data set). RUVBL1 and RUVBL2 are known to heterodimerize forming a platform for the assembly of several multi-subunit chromatin remodeling complexes, such as NuA4, INO80, SWR/SRCAP [43] and complexes controlling transcription factor activity (e.g. cMyc, β-catenin) [43,44] and cell survival [45]. Of note, RUVBLs have been shown to interact physically with several transcription factors, such as ATF2, Myc, E2F1, HIF1α and β-catenin, thereby having an active role in the transcriptional control. Based on these functional patterns, it is straightforward to suggest the involvement of C/EBPα in both chromatin remodeling processes and transcriptional regulation. We hypothesise that C/EBPα could recruit RUVBL1/RUVBL2 complex to C/EBPα binding loci eventually altering the chromatin folding state, a mechanism already reported for E2F mediated transcriptional activity [46].

Production and inspection of C/EBPα and C/EBPβ proteomic data aimed to a system-wide assessment of the physiological roles played by these proteins. By statistical analysis of both datasets we identified significantly enriched shared pathways such as RNA metabolism, including RNA splicing, RNA surveillance, RNA processing and RNA polyadenylation indicating a novel possible role of C/EBPα and C/EBPβ in RNA metabolism. Although the direct role of the C/EBPα and C/EBPβ in downstream processes of RNA production and maturation has not been demonstrated yet, careful examination of the proteins belonging to the enriched groups suggest that C/EBPα and/or C/EBPβ could be constitutive or transient components of large multi-protein complexes (larger spliceosome complex; Nop56p-associated pre-rRNA complex; snRNP-free U1A

oxana bereshchenko 7/9/16 14:49

Comment [2]: Non ho capito bene

or SF-A complex, composed of DDX5, NONO, SFPQ and SNRPA proteins) controlling several steps of the RNA maturation pathway [44]. By analogy with several studies showing how MYC transcription factor regulates expression of the genes critical for the activity of spliceosome complex [47,48] and because our data are substantially enriched in proteins that form complexes involved in the mRNA splicing, we hypothesize that C/EBPs might be engaged in regulating RNA metabolism by physical association with transcript processing machinery.

In this study we also aimed to identify unique physiological roles of C/EBPa and C/EBPb. The striking enrichment of functional terms associated with DNA repair, replication and telomere maintenance processes imply that C/EBPa plays a role in these processes. Indeed, our proteomic and immunoprecipitation experiments demonstrated that C/EBPa interacts with three subunits of the replication complex, namely RFC1, RFC4 and RFC5. These results confirm previously published data demonstrating a synergism between C/EBPa and RFC1 in increasing the transcriptional activity of reporter gene promoter revealing a role of RFC1-C/EBPa complex in transcriptional activation [49]. However, we also found, by CoIP using RUVBL antibodies, that both C/EBPa and subunits of the replication complex were immunoprecipitated demonstrating the presence of large chromatin remodeling complexes containing C/EBPa, replication factors RFCs and RUVBL1/2 at the same time. These hypothetical complexes might be tethered by subunits to C/EBPa DNA recognition sites and facilitate DNA replication/repair activity, like already showed for ILF2 and ILF3 transcription factors [50]. This functional arrangement could be a paradigm for tissue specific organization of DNA replication in mammalian cells where specific DNA sequences acting as DNA replication origins are very few; instead large DNA regions could have origin firing activities which are regulated in tissue specific manner [51, 52].

C/EBPa proteome, but not the C/EBPb one, is statistically enriched in functional terms associated with the DNA repair machinery. In fact, we found two (RAD50 and MRE11) of the three subunits of MRN complex [53] (RAD50- MRE11-NBS1) being C/EBPa interactors. These findings were confirmed by immunoprecipitation analysis. Moreover, two components of the non-homologous end-joining (NHEJ) pathway, XRCC5 and XRCC6 [54], have been isolated in C/EBPa proteome as well. It remains to be determined whether C/EBPa is actively involved in the regulation of DNA repair processes or works as hub for the components of the complex to fulfil its role as transcriptional activator. Indeed, it has been recently reported a synergy between XRCC6/XRCC5 complex and Runx2 in order to activate the osteocalcin promoter [55]. Such results are in agreement with the observation that components of DNA repair complexes might have a role in initiation of transcription by changing chromatin accessibility for the transcriptional factors [56]. On the other hand, the tissue specific transcriptional factor LRF (LRF-leukemia/lymphoma related) has been recently shown to be directly involved in DNA repair processes by binding and stabilizing the XRCC5/XRCC6 complex on sites of the double stranded breaks [57]. Our results point to an analogous function for C/EBPa implying that it might be engaged in DNA repair processes.

The analysis of C/EBPb proteome provides fewer significant unique functional terms compared to C/EBPa. Such outcome could be a sign of higher functional diversity of C/EBPa in the cells compared to C/EBPb. In both C/EBPa and C/EBPb proteomes we did not detect a link with the other instructive transcription factors other than DDIT3, which belongs to C/EBP family and able to heterodimerize with bZIP proteins [58]. Strikingly, we noticed that, unlike the C/EBPa network, GO terms associated to the C/EBPb one are significantly enriched for proteins involved in the control of cellular metabolism, such as gluconeogenesis and glycolysis. While the role of the C/EBPb in the regulation of metabolic processes via transcriptional regulation of the genes that control cellular metabolism is well documented [59], the notion that C/EBPb proteome includes metabolic enzymes is rather unexpected and requires additional experimental evaluations.

5. Conclusions

In conclusion, we characterized C/EBPa and C/EBPb *in vivo* interactomes in NIH3T3 cells at the preadipocytes differentiation stage using a proteomic high-throughput approach with the aim to find novel roles of C/EBPa and C/EBPb in cellular pathways. We statistically compared the proteomes of C/EBPa and C/EBPb to evaluate their functional similarity and differences. Such analysis provided the background to propose new hypothesis aimed to understand the mechanisms underlying the instructive roles of C/EBPa and C/EBPb in normal physiological

conditions as well as their dis-regulation. Furthermore, we proved that C/EBPa binds directly with pivotal proteins of DNA repair machinery.

Authors contribution

C.N. and O.E designed the experiments. O.E. performed the TAP experiments. O.B. performed the immunoprecipitation experiments. M.C. performed statistical data analysis. M.C. and O.E. wrote the manuscript.

Funding

This work was supported by funds from EMBL. O.E. was supported by the European Commission FP5 EUMORPHIA. O.B. was a recipient of an HFSP post-doctoral fellowship. M.C. was supported by the Italian National Council of Research (CNR, Italy).

Acknowledgments

We are grateful to Dr. Andreas Bauer (Novartis Institute for BioMedical Research) for help to design and perform Tandem Affinity Purification experiments. Special thank to EMBL proteomics core facility for the LC/MS analysis of the protein gels.

Conflict of interest statement

The authors have declared no conflict of interest.

Supplementary Table 1. LC/MS derived data set of proteins isolated by C/EBPa –TAP affinity purification.

Supplementary Table 2. LC/MS derived data set of proteins isolated by C/EBPb –TAP affinity purification.

Supplementary Table 3. Data set of the C/EBPa and C/EBPb interacting proteins filtered using CRAPome database.

Supplementary Table 4. Functional annotation of C/EBPa and C/EBPb interacting proteins using Biological Processes and Molecular Functions Gene Ontology (GO) terms and KEGG cellular pathways. Only statistically significant values are showed in the table.

Supplementary Table 5. Comparative functional annotation of the interactomes of C/EBPa and C/EBPb via GO terms and KEGG pathways. Significantly enriched GO terms unique for C/EBPa, unique for C/EBPb and common for both transcription factors are listed.

References:

- [1] Nerlov, C. (2007). The C/EBP family of transcription factors: a paradigm for interaction between gene expression and proliferation control. *Trends Cell Biol* 17, 318-324.
- [2] Nerlov, C. (2010). Transcriptional and translational control of C/EBPs: the case for "deep" genetics to understand physiological function. *Bioessays* 32, 680-686.
- [3] Reckzeh, K. et al. (2012) Molecular and cellular effects of oncogene cooperation in a genetically accurate AML mouse model. *Leukemia* 26, 1527-1536.
- [4] Rosen, E.D. (2005). The transcriptional basis of adipocyte development. *Prostaglandins Leukot Essent Fatty Acids* 73, 31-34.
- [5] Calella, A.M., Nerlov, C., Lopez, R.G., Sciarretta, C., von Bohlen und Halbach, O., Bereshchenko, O., and Minichiello, L. (2007). Neurotrophin/Trk receptor signaling mediates

C/EBPalpha, -beta and NeuroD recruitment to immediate-early gene promoters in neuronal cells and requires C/EBPs to induce immediate-early gene transcription. *Neural Dev* 2, 4.

[6] Pulido-Salgado, M., Vidal-Taboada, J.M., and Saura, J. (2015). C/EBPbeta and C/EBPdelta transcription factors: Basic biology and roles in the CNS. *Prog Neurobiol* 132, 1-33.

[7] Nerlov, C., and Ziff, E.B. (1995). CCAAT/enhancer binding protein-alpha amino acid motifs with dual TBP and TFIIIB binding ability co-operate to activate transcription in both yeast and mammalian cells. *EMBO J* 14, 4318-4328.

[8] Nerlov, C. (2008). C/EBPs: recipients of extracellular signals through proteome modulation. *Curr Opin Cell Biol* 20, 180-185.

[9] Nagamura-Inoue T1, Tamura T, Ozato K. Transcription factors that regulate growth and differentiation of myeloid cells. *Int. Rev Immunol.* 2001 Feb;20(1):83-105.

[10] Oishi, Y., Manabe, I., Tobe, K., Tsushima, K., Shindo, T., Fujiu, K., Nishimura, G., Maemura, K., Yamauchi, T., Kubota, N., *et al.* (2005). Kruppel-like transcription factor KLF5 is a key regulator of adipocyte differentiation. *Cell Metab* 1, 27-39.

[11] Kajimura, S., Seale, P., Kubota, K., Lunsford, E., Frangioni, J.V., Gygi, S.P., and Spiegelman, B.M. (2009). Initiation of myoblast to brown fat switch by a PRDM16-C/EBP-beta transcriptional complex. *Nature* 460, 1154-1158.

[12] Hirata, M., Kugimiya, F., Fukai, A., Saito, T., Yano, F., Ikeda, T., Mabuchi, A., Sapkota, B.R., Akune, T., Nishida, N., *et al.* (2011). C/EBPbeta and RUNX2 cooperate to degrade cartilage with MMP-13 as the target and HIF-2alpha as the inducer in chondrocytes. *Hum Mol Genet* 21, 1111-1123.

[13] Porse, B.T., Pedersen, T.A., Xu, X., Lindberg, B., Wewer, U.M., Friis-Hansen, L., and Nerlov, C. (2001). E2F repression by C/EBPalpha is required for adipogenesis and granulopoiesis in vivo. *Cell* 107, 247-258.

[14] Slomiany, B.A., D'Arigo, K.L., Kelly, M.M., and Kurtz, D.T. (2000). C/EBPalpha inhibits cell growth via direct repression of E2F-DP-mediated transcription. *Mol Cell Biol* 20, 5986-5997.

[15] Chen, S.S., Chen, J.F., Johnson, P.F., Muppala, V., and Lee, Y.H. (2000). C/EBPbeta, when expressed from the C/ebpalpha gene locus, can functionally replace C/EBPalpha in liver but not in adipose tissue. *Mol Cell Biol* 20, 7292-7299.

[16] Begay, V., Smink, J., and Leutz, A. (2004). Essential requirement of CCAAT/enhancer binding proteins in embryogenesis. *Mol Cell Biol* 24, 9744-9751.

[17] Morgenstern, J.P., and Land, H. (1990). A series of mammalian expression vectors and characterisation of their expression of a reporter gene in stably and transiently transfected cells. *Nucleic Acids Res* 18, 1068.

[18] Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., *et al.* (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415, 141-147.

[19] Bouwmeester, T., Bauch, A., Ruffner, H., Angrand, P.O., Bergamini, G., Croughton, K., Cruciat, C., Eberhard, D., Gagneur, J., Ghidelli, S., *et al.* (2004). A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat Cell Biol* 6, 97-105.

- [20] Shevchenko, A., Jensen, O.N., Podtelejnikov, A.V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Boucherie, H., and Mann, M. (1996). Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc Natl Acad Sci U S A* 93, 14440-14445.
- [21] Towbin, H., Staehelin, T., and Gordon, J. (1992). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. 1979. *Biotechnology* 24, 145-149.
- [22] Timchenko, N.A., Wilde, M., and Darlington, G.J. (1999). C/EBPalpha regulates formation of S-phase-specific E2F-p107 complexes in livers of newborn mice. *Mol Cell Biol* 19, 2936-2945.
- [23] Bauer, A., Chauvet, S., Huber, O., Usseglio, F., Rothbacher, U., Aragnol, D., Kemler, R., and Pradel, J. (2000). Pontin52 and reptin52 function as antagonistic regulators of beta-catenin signalling activity. *EMBO J* 19, 6121-6130.
- [24] Pedersen, T.A., Bereshchenko, O., Garcia-Silva, S., Ermakova, O., Kurz, E., Mandrup, S., Porse, B.T., and Nerlov, C. (2007). Distinct C/EBPalpha motifs regulate lipogenic and gluconeogenic gene expression in vivo. *EMBO J* 26, 1081-1093.
- [25] Gingras, A.C., Gstaiger, M., Raught, B., and Aebersold, R. (2007). Analysis of protein complexes using mass spectrometry. *Nat Rev Mol Cell Biol* 8, 645-654.
- [26] Mellacheruvu, D., Wright, Z., Couzens, A.L., Lambert, J.P., St-Denis, N.A., Li, T., Miteva, Y.V., Hauri, S., Sardi, M.E., Low, T.Y., *et al.* (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat Methods* 10, 730-736.
- [27] Bindea, G., Galon, J., and Mlecnik, B. (2013). CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics* 29, 661-663.
- [28] Lopes, C.T., Franz, M., Kazi, F., Donaldson, S.L., Morris, Q., and Bader, G.D. (2010). Cytoscape Web: an interactive web-based network browser. *Bioinformatics* 26, 2347-2348.
- [29] Calkhoven, C.F., Muller, C., and Leutz, A. (2002). Translational control of gene expression and disease. *Trends Mol Med* 8, 577-583.
- [30] Descombes, P., and Schibler, U. (1991). A liver-enriched transcriptional activator protein, LAP, and a transcriptional inhibitory protein, LIP, are translated from the same mRNA. *Cell* 67, 569-579.
- [31] Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* 17, 1030-1032.
- [32] Giambruno, R., Grebien, F., Stukalov, A., Knoll, C., Planyavsky, M., Rudashevskaya, E.L., Colinge, J., Superti-Furga, G., and Bennett, K.L. (2013). Affinity purification strategies for proteomic analysis of transcription factor complexes. *J Proteome Res* 12, 4018-4027.
- [33] Steinberg, X. P., Hepp, M. I., Fernández García, Y., Suganuma, T., Swanson, S. K., Washburn, M., Workman, J.L., and Gutiérrez, J.L. (2012). Human CCAAT/Enhancer-Binding Protein β Interacts with Chromatin Remodeling Complexes of the Imitation Switch Subfamily. *Biochemistry* 51, 952-962.
- [34] Siersbæk, R., Rabiee, A., Nielsen, R., Sidoli, S., Traynor, S., Loft, A., Poulsen, L.L.C., Rogowska-Wrzesinska, A., Jensen, O.N., and Mandrup, S. (2014). Transcription Factor Cooperativity in Early Adipogenic Hotspots and Super-Enhancers. *Cell Reports* 7, 1443-1455.

- [35] Yao, N., Coryell, L., Zhang, D., Georgescu, R.E., Finkelstein, J., Coman, M.M., Hingorani, M.M., and O'Donnell, M. (2003). Replication factor C clamp loader subunit arrangement within the circular pentamer and its attachment points to proliferating cell nuclear antigen. *J Biol Chem* 278, 50744-50753.
- [36] Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., *et al.* (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25, 25-29.
- [37] Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2013). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* 42, D199-205.
- [38] Grebien, F., Vedadi, M., Getlik, M., Giamb Bruno, R., Grover, A., Avellino, R., Skucha, A., Vittori, S., Kuznetsova, E., Smil, D., *et al.* (2015). Pharmacological targeting of the Wdr5-MLL interaction in C/EBPalpha N-terminal leukemia. *Nat Chem Biol* 11, 571-578.
- [39] Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B., *et al.* (2006). Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440, 631-636.
- [40] Wang, N.D., Finegold, M.J., Bradley, A., Ou, C.N., Abdelsayed, S.V., Wilde, M.D., Taylor, L.R., Wilson, D.R., and Darlington, G.J. (1995). Impaired energy homeostasis in C/EBP alpha knockout mice. *Science* 269, 1108-1112.
- [41] Guillaumond, F., Boyer, B., Becquet, D., Guillen, S., Kuhn, L., Garin, J., Belghazi, M., Bosler, O., Franc, J.L., and Francois-Bellan, A.M. (2011). Chromatin remodeling as a mechanism for circadian prolactin transcription: rhythmic NONO and SFPQ recruitment to HLTF. *FASEB J* 25, 2740-2756.
- [42] Kuwahara, S., Ikei, A., Taguchi, Y., Tabuchi, Y., Fujimoto, N., Obinata, M., Uesugi, S., and Kurihara, Y. (2006). PSPC1, NONO, and SFPQ are expressed in mouse Sertoli cells and may function as coregulators of androgen receptor-mediated transcription. *Biol Reprod* 75, 352-359.
- [43] Huen, J., Kakiyama, Y., Ugwu, F., Cheung, K.L., Ortega, J., and Houry, W.A. (2010). Rvb1-Rvb2: essential ATP-dependent helicases for critical complexes. *Biochem Cell Biol* 88, 29-40.
- [44] Liang, S., and Lutz, C.S. (2006). p54nrb is a component of the snRNP-free U1A (SF-A) complex that promotes pre-mRNA cleavage during polyadenylation. *RNA* 12, 111-121.
- [45] Bereshchenko, O. Mancini, E., Luciani, L., Gambardella, A., Riccardi, C., and Nerlov, C. Pontin is essential for murine hematopoietic stem cell survival (2012). *Haematologica* 97, 1291-1294.
- [46] Tarangelo, A., Lo, N., Teng, R., Kim, E., Le, L., Watson, D., Furth, EE., Raman, P., Ehmer, U., Viatour, P. (2015). Recruitment of Pontin/Reptin by E2f1 amplifies E2f transcriptional response during cancer progression. *Nat Commun.* 7, 10028
- [47] Hsu, T.Y., Simon, L.M., Neill, N.J., Marcotte, R., Sayad, A., Bland, C.S., Echeverria, G.V., Sun, T., Kurley, S.J., Tyagi, S., *et al.* (2015). The spliceosome is a therapeutic vulnerability in MYC-driven cancer. *Nature* 525, 384-388.
- [48] Koh, C.M., Bezzi, M., Low, D.H., Ang, W.X., Teo, S.X., Gay, F.P., Al-Haddawi, M., Tan, S.Y., Osato, M., Sabo, A., *et al.* (2015). MYC regulates the core pre-mRNA splicing machinery as an essential step in lymphomagenesis. *Nature* 523, 96-100.

[49] Hong, S., Park, S.J., Kong, H.J., Shuman, J.D., and Cheong, J. (2001). Functional interaction of bZIP proteins and the large subunit of replication factor C in liver and adipose cells. *J Biol Chem* 276, 28098-28105.

[50] Karmakar, S., Mahajan, M.C., Schulz, V., Boyapaty, G., and Weissman, S.M. (2010). A multiprotein complex necessary for both transcription and DNA replication at the beta-globin locus. *EMBO J* 29, 3260-3271.

[51] Zhou, J., Ermakova, O.V., Riblet, R., Birshtein, B.K., and Schildkraut, C.L. (2002). Replication and subnuclear location dynamics of the immunoglobulin heavy-chain locus in B-lineage cells. *Mol Cell Biol* 22, 4876-4889.

[52] Demczuk, A., Gauthier, M.G., Veras, I., Kosiyatrakul, S., Schildkraut, C.L., Busslinger, M., Bechhoefer, J., and Norio, P. (2012). Regulation of DNA replication within the immunoglobulin heavy-chain locus during B cell commitment. *PLoS Biol* 10, e1001360.

[53] Lafrance-Vanasse, J., Williams, G.J., and Tainer, J.A. (2015). Envisioning the dynamics and flexibility of Mre11-Rad50-Nbs1 complex to decipher its roles in DNA replication and repair. *Prog Biophys Mol Biol* 117, 182-193.

[54] Chapman, J.R., Taylor, M.R., and Boulton, S.J. (2012). Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell* 47, 497-510.

[55] Willis, D.M., Loewy, A.P., Charlton-Kachigian, N., Shao, J.S., Ornitz, D.M., and Towler, D.A. (2002). Regulation of osteocalcin gene expression by a novel Ku antigen transcription factor complex. *J Biol Chem* 277, 37280-37291.

[56] Singh, I., Ozturk, N., Cordero, J., Mehta, A., Hasan, D., Cosentino, C., Sebastian, C., Kruger, M., Looso, M., Carraro, G., *et al.* (2015). High mobility group protein-mediated transcription requires DNA damage marker gamma-H2AX. *Cell Res* 25, 837-850.

[57] Liu, X.S., Chandramouly, G., Rass, E., Guan, Y., Wang, G., Hobbs, R.M., Rajendran, A., Xie, A., Shah, J.V., Davis, A.J., *et al.* (2015). LRF maintains genome integrity by regulating the non-homologous end joining pathway of DNA repair. *Nat Commun* 6, 8325.

[58] Tsukada, J., Yoshida, Y., Kominato, Y., and Auron, P.E. (2011). The CCAAT/enhancer (C/EBP) family of basic-leucine zipper (bZIP) transcription factors is a multifaceted highly-regulated system for gene regulation. *Cytokine* 54, 6-19.

[59] van der Krieken, S.E., Popeijus, H.E., Mensink, R.P., and Plat, J. (2015). CCAAT/enhancer binding protein beta in relation to ER stress, inflammation, and metabolic disturbances. *Biomed Res Int* 2015, 324815.