Protein Interaction Screen on Peptide Matrix (PRISMA) reveals interaction footprints and PTM-dependent interactome of intrinsically disordered C/EBPβ

Gunnar Dittmar, Daniel Perez Hernandez, Elisabeth Kowenz-Leutz, Marieluise Kirchner, Günther Kahlert, Radoslaw Wesolowski, Katharina Baum, Maria Knoblich, Maria Hofstätter, Arnaud Muller, Jana Wolf, Ulf Reimer, Achim Leutz

PII: S2589-0042(19)30060-4
DOI: https://doi.org/10.1016/j.isci.2019.02.026
Reference: ISCI 326

To appear in: ISCIENCE

Received Date: 28 August 2017
Revised Date: 20 January 2019
Accepted Date: 23 February 2019


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Features of intrinsically disordered C/EBPβ

... etc.

... decompose C/EBPβ into 203 spot synthesized, matrix immobilized, tiling 14aa peptides (duplicates)

... incubate matrix with nuclear extract & wash briefly

... isolate peptide spots with bound interactome

... etc.

data acquisition & bioinformatics yields C/EBPβ interactome footprints

Short Linear Motifs (SLiMs)

Intrinsic Disordered Regions (IDRs)

Post Translational Modifications (PTM)
Protein Interaction Screen on Peptide Matrix (PRISMA) reveals interaction footprints and PTM-dependent interactome of intrinsically disordered C/EBPβ

Gunnar Dittmar†,1,2,3, Daniel Perez Hernandez*2,3, Elisabeth Kowenz-Leutz*2, Marieluise Kirchner*2,3, Günther Kahlert2, Radoslaw Wesolowski2, Katharina Baum2, Maria Knoblich2, Maria Hofstätter2, Arnaud Muller1, Jana Wolf2, Ulf Reimer4 and Achim Leutz†,2,5,6

†Proteome and Genome research laboratory, Luxembourg Institute of Health, 1a Rue Thomas Edison, 1445 Strassen, Luxembourg.

2Max Delbrück Center for Molecular Medicine, Robert-Roessle Strasse 10, 13125 Berlin, Germany.

3BIH Core Facility Proteomics, Robert-Roessle Strasse 10, 10125 Berlin, Germany.

4JPT Peptide Technologies GmbH, Volmerstrasse 5, 12489 Berlin, Germany.

5Humboldt-University of Berlin, Institute of Biology, 10115 Berlin, Germany.

6Lead Contact: Achim Leutz, aleutz@mdc-berlin.de, ORCID: 0000-0001-8259-927X

†corresponding authors; aleutz@mdc-berlin.de; gunnar.dittmar@lih.lu

*these authors contributed equally
Summary

CCAAT enhancer binding protein beta (C/EBPβ) is a pioneer transcription factor that specifies cell differentiation. C/EBPβ is intrinsically unstructured, a molecular feature common to many proteins involved in signal processing and epigenetics. The structure of C/EBPβ differs depending on alternative translation initiation and multiple post-translational modifications (PTM). Mutation of distinct PTM sites in C/EBPβ alters protein interactions and cell differentiation, suggesting a C/EBPβ PTM indexing code determines epigenetic outcomes. Herein, we systematically explored the interactome of C/EBPβ using an array technique based on spot-synthesized C/EBPβ-derived linear tiling peptides with and without PTM, combined with mass spectrometric proteomic analysis of protein interactions (PRISMA). We identified interaction footprints of ~1300 proteins in nuclear extracts, many with chromatin modifying, remodeling and RNA processing functions. The results suggest C/EBPβ acts as a multi-tasking molecular switchboard, integrating signal-dependent modifications and structural plasticity to orchestrate interactions with numerous protein complexes directing cell fate and function.

Keywords: C/EBPβ, PRISMA, intrinsically disordered protein, post-translational modification, mass spectrometry.
Introduction

CCAAT enhancer binding proteins (C/EBPα, β, δ, ε) are basic leucine zipper transcription factors that regulate chromatin structure and gene expression by dimerization and binding to cis-regulatory, palindromic 5′ATTGC•GCAAT3′, or quasi-palindromic DNA sites in gene enhancers and promoters. Prototypic C/EBPβ is widely expressed, highly regulated at the post-transcriptional level, and integrated in many signaling events communicating extracellular cues to epigenetic changes, examples of which include adipogenesis, hematopoiesis, innate immunity, female fertility, skin function, apoptosis, and cellular senescence (Nerlov, 2008; Rodier and Campisi, 2011; Tsukada et al., 2011).

In early hematopoiesis and adipogenesis, C/EBPβ acts as a pioneering factor that orchestrates complex steps in cell fate commitment (Kajimura et al., 2009; Lichtinger et al., 2012; Muller et al., 1995; Ness et al., 1993; Sierrsbaek et al., 2011). C/EBPβ communicates with numerous other transcription factors, co-factors, histone modifiers, and chromatin remodeling complexes to alter the susceptibility of chromatin to the gene regulatory machinery. C/EBPβ induces lymphoid-myeloid trans-differentiation and accelerates acquisition of the induced pluripotent state by the Yamanaka set of reprogramming transcription factors (Di Stefano et al., 2016; Kowenz-Leutz and Leutz, 1999; Siolversa et al., 2013; Xie et al., 2004). The chromatin and gene regulatory functionality of C/EBPβ is linked to distinct regions and post-translational modifications (PTM) that suspend auto-inhibition, direct the activity of C/EBPβ and regulate recruitment of chromatin remodelers and writers of histone modifications (Kowenz-Leutz and Leutz, 1999; Kowenz-Leutz et al., 2010; Kowenz-Leutz et al., 1994; Lee et al., 2010b; Mo et al., 2004; Pless et al., 2008; Sierrsbaek et al., 2011).

C/EBPβ functions are controlled by extracellular signaling cascades involving receptor tyrosine kinases, cytokine receptors, Ras GTPases, MAP kinases, and cAMP and SMAD signaling, and by more complex conditions such as metabolic adaptation,
inflammation, senescence or stress responses (Nerlov, 2008; Rodier and Campisi, 2011; Tsukada et al., 2011). The complexity and diversity of C/EBPβ activities in various cell lineages raises the question of how a single transcription factor can participate in a multitude of regulatory events. Previous research suggested that the combinatorial outcome of post-transcriptional and post-translational modifications in conjunction with intrinsic structural plasticity enables the C/EBPβ protein to adopt a plethora of context- and signal-dependent states that facilitate a variety of interactions (Leutz et al., 2011; Nerlov, 2008; Tsukada et al., 2011).

Post-transcriptional regulation of C/EBPβ generates three isoforms (LAP*, LAP, LIP) by alternative translation initiation of the single exon C/EBPβ transcript. Since consecutive C/EBPβ start sites are positioned in the same reading frame, the isoforms vary in their gene regulatory N-terminal extensions but retain the same C-terminal dimerization and DNA-binding bZip domain (Wethmar et al., 2010). The diversity of C/EBPβ isoforms is further increased by numerous PTM of amino acid side chains; in addition to phosphorylation of serine, threonine, and tyrosine residues, lysine acetylation and methylation also occurs, as does methylation of arginine. Enzymes responsible for PTM of C/EBPβ include CARM1/PRMT4, G9A/EHMT2, and CREBBP/KAT3A, all of which also serve as epigenetic histone code writers. The decoration of C/EBPβ by PTM alters its capacity to engage in protein-protein interactions (PPI) and to direct cell fate, suggesting that the signal-dependent C/EBPβ modification index reflects integration of various upstream signaling events to adjust its interactome and to determine its gene regulatory and epigenetic capacity (Leutz et al., 2011). The combination of translational and post-translational modifications may thus encrypt the dynamic interactome, with C/EBPβ as the keystone for a wide range of functional outcomes (Lee et al., 2010b; Sebastian et al., 2005; Sterneck et al., 1997; Stoilova et al., 2013).
The C-terminal third of C/EBPα, β, δ, ε contains highly conserved DNA binding and basic leucine zipper domains (bZip) that may dimerize within an extended trans-regulatory bZip network including C/EBP, AP-1, and ATF transcription factors. The N-terminal two-thirds of the C/EBP primary structure is predicted as an intrinsically disordered region (IDR) (Miller, 2006). Phylogenetic analysis nevertheless suggests that the C/EBP N-terminus also contains several highly conserved short peptide regions (CRs) that are enriched in amino acids with hydrophobic and bulky side chains. These CRs are discontinuous and separated by less conserved and family-specific regions of low complexity (LCRs) characterized by a predominance of small and polar amino acids (Leutz et al., 2011; Tsukada et al., 2011). Experimental studies involving a large number of deletions and CR/IDR shuffling mutants suggested a highly modular, context-dependent functionality of N-terminal C/EBPβ CRs (Kowenz-Leutz et al., 1994; Lee et al., 2010b; Leutz et al., 2011; Williams et al., 1995).

Screening for interaction partners using N-terminally-derived C/EBPβ peptides further supports the notion that many C/EBPβ interactions may occur in a modular and dynamic fashion that rely on molecular recognition features (MoRFs) and short linear peptide motifs (SLiMs) in combination with adjacent PTM (Dunker et al., 1998; Leutz et al., 2011; Tompa et al., 2014; van der Lee et al., 2014; Wright and Dyson, 1999).

Based on previous observations of the modular structure and biological functionality of C/EBPβ peptide regions, we developed a novel technique to systematically explore the C/EBPβ interactome. Briefly, C/EBPβ was deconstructed into small linear peptides of 14 amino acids immobilized on a solid-phase matrix. The starting point for each peptide was shifted by four amino acids, creating a tiled peptide array covering the entire amino acid sequence of C/EBPβ. PTM containing peptides were included according to recorded modifications and their position in the primary C/EBPβ structure. The matrix was incubated with nuclear extracts and protein enrichment on individual peptide spots was identified and quantified by mass spectrometry. The assay, termed PRotein Interaction Screen on peptide Matrices (PRISMA), revealed hundreds of C/EBPβ- and PTM-specific protein interactions.
that left footprints on C/EBPβ-derived peptides. Based on comparison with other affinity enrichment approaches, 45 protein complexes were predicted and several novel interactions with proteins and complexes were experimentally confirmed.
Results

The C/EBPβ peptide matrix

The occurrence of novel PTMs on endogenous C/EBPβ was investigated by mass spectrometry of C/EBPβ immunoprecipitates derived from the human anaplastic lymphoma cell line SU-DHL1 that critically depends for growth and survival on C/EBPβ (Anastasov et al., 2010; Jundt et al., 2005). Over 90 PTMs were identified which, combined with published data, suggested that more than 130 PTMs may occur on this protein, as summarized in Figure 1A (Table S1. C/EBPβ PTMs. Related to Figure 1A). To systematically explore the C/EBPβ interactome and its PTM specific regulation, we developed the PRISMA technology, which is based on a solid matrix consisting of immobilized peptides spanning the entire primary structure of rat C/EBPβ (297 amino acids), as schematically depicted in Figure 1B and detailed in Table S2. PRISMA C/EBPβ Peptides. Related to Figure 1B. To cover all linear binding regions of the entire CEBPβ protein sequence, tiling peptides of 14 amino acids in length with an offset of mostly four amino acids were spot-synthesized on a cellulose acetate matrix using Fmoc synthesis. Since PTMs were previously shown to impact on CEBPβ protein interactions and functionality, PTM peptides with S/T/Y-phosphorylation, K-acetylation, K-, R-methylation, and R-citrullination were included in the screen matrix. In total, the solid matrix contained 203 immobilized peptides, covering known and potential post-translational side chain modifications.

The C/EBPβ PRISMA screen

To examine the linear CEBPβ interactome, two replicates of the peptide matrix were incubated with nuclear extracts of HeLa cells. HeLa cells express CEBPβ and commercially available nuclear extracts have been used successfully by many research groups for the purification of biologically active proteins and protein complexes involved in gene regulation.
(Figure 1B). Individual peptide spots were excised and bound proteins proteolytically digested and analyzed by high-resolution mass spectrometry. In total, 406 analytical mass spectrometric 1 h runs were performed (approximately 17 days of measurement) and spectra were interpreted automatically using the MaxQuant software package.

Enrichment efficiency by PRISMA was examined by comparing the PRISMA intensity distribution with the total proteome of the nuclear extract. Approximately 5100 proteins were identified and quantified in the nuclear cell extract using the intensity-based absolute quantification (iBaq) method (Schwanhausser et al., 2011). The copy number of proteins in the nuclear cell extract and the number of proteins bound in the PRISMA screen spans six orders of magnitude, suggesting that the identified binders are not biased towards highly abundant proteins (Smits et al., 2013) (Figure S1. Comparison of enrichment by PRISMA and physico-chemical properties of PRISMA peptides. Related to Figure 1 and Figure 2). Furthermore, the distribution of peptide/protein interactions was not attributed to physico-chemical parameters of the peptides, as evidenced by comparison of the accumulated interaction intensities with peptide hydrophobicity (gravy index) and isoelectric points, which were determined for all non-modified PRISMA peptides. These results suggest that the PRISMA peptides on the matrix retained their specific protein-binding properties.

Data processing of C/EBPβ peptide binding proteins

An initial inspection of the two replicate datasets revealed signal intensity variation of the interacting proteins. The two datasets were therefore integrated to increase robustness, as outlined in Figure 1C and as explained in the Materials and Methods. The main source of variation arose from proteins that were identified only once on different peptide spots. This led to single (low confidence) and double (high confidence) identification categories for each peptide. The signal intensity for each protein was then normalized between 0 and 1 across
all 203 matrix peptides. Individual peptides displayed large differences in binding partner profiles, further indicating the specificity of the interactions, because random binding would be expected to result in a more equal distribution of interacting proteins. Further analysis of the data showed that ~25% of the identified proteins bound to multiple C/EBPβ peptides across the array with low but varying intensity. While these proteins may promiscuously bind to many distant peptides, some sections of the array showed much higher signals. In order to minimize noise from background binding, signal intensities for each protein were filtered for binding above 90% of the protein’s signal distribution (outlier filtering), removing all signals below this threshold. Another filtering criterion for discriminating the most robust interactors was based on the consecutive binding behavior of tiling peptides of the C/EBPβ primary structure. The rational here is that by shifting the sequence of tiling peptides by four amino acids, some of the SLiMs and fractions thereof are included in more than one peptide. This may generate maximal binding signals for peptides containing optimal SLiM and adjacent supporting amino acids, and attenuate signals from neighboring peptides in which the particular SLiM is shifted or only partially included. We used this predicted binding behavior to stringently filter the dataset further to remove all proteins that failed the consecutive binding criterion (Figure 1D, E; see also Materials and Methods). In total, 2363 interacting proteins were identified (Table S3. The PRISMA C/EBPβ Interactome. Related to Figure 1D ), of which 1384 proteins were detected in both replicates, and 1302 proteins fulfilled the consecutive binding criterion and were defined as C/EBPβ core interactions (βCI; Figure 1E).

Validation of the PRISMA-derived dataset

To assess the biological significance of data derived by PRISMA, we compared PRISMA data (SET1, SET2, βCI) to previously identified C/EBPβ interactome data, as shown in Figure 1F (Siersbaek et al., 2011; Steinberg et al., 2012). In addition, we performed pull-
down experiments using C/EBPβ antibodies and SU-DHL1 cells that express a high amount of CEBPβ and analyzed the CEBPβ interactome by mass spectrometry (Table S4. C/EBPβ interactome in SU-DHL cells. Related to Figure 1F,G). In total, 1369 proteins were significantly enriched in SU-DHL1 C/EBPβ samples compared with control samples using a false discovery rate (FDR) cutoff of 5%. The list of SU-DHL1 C/EBPβ interactors was included to extend the existing CEBPβ interaction datasets. The PRISMA core interaction data covered between 38% and 59% of the affinity purification-based datasets (see supplemental information). Affinity purification-based datasets were then combined and the overlap with PRISMA data was determined on the basis of their UniProt identifier entries, as shown in Figure 1G (see Materials and Methods). In total, 47% of data in all three sets were also found in the PRISMA core interactions, and 64% of PRISMA core interactions were also found in at least one of the other datasets. To estimate the FDR of the C/EBPβ interactors detected by PRISMA, we employed the method proposed previously (D’Haeseleer and Church, 2004), which relies on comparing the intersections of protein interaction datasets to approximate the number of false-positive PPIs. We obtained FDRs of 11.2% and 13.9% for proteins detected in PRISMA replicates 1 and 2, respectively (SET1 and SET2 in Figure 1D and E; see Supplementary Information for details). FDRs were reduced to values below 4% when applying the filtering step leading to PRISMA core interactions. These results suggest that PRISMA data depict strong overlap and extensive coverage of the interactome related to native C/EBPβ. We conclude that the PRISMA method successfully extends the interactome data and serves as a resource for locating interaction footprints on C/EBPs.

High-resolution C/EBPβ interactome footprints

The global protein interaction profile obtained from the C/EBPβ peptide matrix is depicted as a non-hierarchically clustered heat map in Figure 2A. The numerical distribution of proteins identified by individual peptides is shown in the upper part of Figure 2A. Peptides
representing the DNA binding region (DB) and the C-terminal part of the leucine zipper (LZ) exhibit the highest number of protein interactions, yet locally-enriched binding hot-spots were also found with peptides from the N-terminal part of C/EBPβ.

Clusters of protein-peptide interactions tended to colocalize with regions predicted to undergo disorder-to-order transition during interaction with binding partners and coincide with the profile of conserved C/EBPβ regions (CR1–7), interaction and MoRF predictions (Disfani et al., 2012; Meszaros et al., 2009), as shown in Figure 2B. Predicted intrinsically disordered regions in the transactivating domain (CR1–4), IDR2 (between CR2 and CR3), and in the regulatory domain between CR7 and the basic/acidic region (IDR7) also displayed large numbers of interaction partners.

Next, PRISMA-derived data were compared with previously mapped C/EBPβ-interacting proteins. As shown in Figure 2C, both co-activator acetyltransferases, CBP and P300 (KAT3A/KAT3B) generated highly similar interaction footprints in the transactivation regions CR4 (strongest binding) and CR3 (additional binding), with some residual binding in CR2. Both C/EBPβ and C/EBPδ have been shown to interact with CBP/p300, and the main interaction region encompasses the CBP Taz2 domain as well as CR3 and CR4 in both C/EBPs. Removal of CR3 or CR4, or mutation of a critical tyrosine residue in CR3 or the DLF motif in CR4 all abrogated interaction with Taz2 and subsequent transcriptional co-activation (Kovacs et al., 2003; Schwartz et al., 2003). Previously published crystallographic data revealed that CR3/CR4 regions in C/EBPε may adopt an L-shaped α-helical structure that folds into the p300 Taz2 domain (Bhaumik et al., 2014), in agreement with the CBP/P300 protein footprints revealed by PRISMA (Figure 2C).

The multi-subunit Mediator (MED) complex (Conaway and Conaway, 2013; Jeronimo and Robert, 2017) has also been reported to interact with C/EBPβ (Mo et al., 2004), and several MED components were identified in a previously published C/EBPβ interaction dataset (Siersbaek et al., 2014), as well as in the SU-DHL1 interactome presented here. As
shown in Figure 2C, PRISMA revealed 20 components of the multi-subunit Mediator (MED) complex that predominantly interact with CR2-, CR3-, and CR4-derived peptides from C/EBPβ TAD.

CR1 of the LAP* C/EBPβ isoform was reported to specify conjugation by SUMO3 (Eaton and Sealy, 2003) and in accordance, PRISMA showed interaction between SUMO3 and CR1. CR1 also interacts with the Brg1/SMARCA4 ATPase of the SWI/SNF/BAF complex. Moreover, the Brg1-CR1 interaction is sensitive to methylation of arginine 3 (Kowenz-Leutz et al., 2010). Consistent with these results, PRISMA revealed interaction of Brg1 with the unmodified CR1 peptide but not with the methylated peptide. Moreover, interactions with 10 additional protein components of BAF-SWI/SNF type protein complexes were detected with various peptides of the C/EBPβ transactivation domain and the bZip domain, indicating multiple interactions between C/EBPβ and chromatin modifying complexes, as previously suggested (Kowenz-Leutz and Leutz, 1999).

The landscape of CEBPβ protein interactions

PRISMA data (SET1 and SET2) were inspected for enrichment of gene ontology (GO) terms, protein families and protein domains. As shown in Table 1, sorting of C/EBPβ interaction hits based on the quantity of terms resulted in exceedingly low Benjamini-Hochberg corrected FDRs (cut-off at a Benjamini-Hochberg corrected p-value of 0.01) (Finn et al., 2017; Finn et al., 2016; Gene Ontology, 2015). Strongly enriched terms included ‘nucleic acid binding’ (745 hits), ‘gene expression’ (611 hits), ‘protein binding’ (534 hits), ‘regulation of transcription’ (301 hits), ‘cell cycle’ (211 hits), ‘RNA splicing’ (172 hits), and several terms involving transcription, chromatin binding and remodeling. Gene set enrichment analysis showed that approximately 25% of PRISMA replicate validated proteins (267) fell into categories that involve GO terms related to RNA processing (FDR 5.63e-128).
RNA binding proteins that contain a RNP-1 RNA binding domain fall into three classes that preferentially interact with the C/EBPβ bZip domain and the adjacent Fork and BA motifs, while a third group interacts with the CR2 region. A similar binding pattern was found for DEAD-box helicases, a group of enzymes involved in ATP-driven conformational adjustment of ribonucleoprotein assembly.

Protein domain enrichment using InterPro and PFAM databases revealed nucleotide binding proteins (102 hits), RNA binding proteins (80 hits), armadillo, WD40, histone fold, and DEAD box domain proteins as major C/EBPβ interaction partners. The number of domain hits obtained using standard algorithms to search protein domain databases may be underrepresented, because manual curation of PRISMA data increased the number of hits for WD40 domain proteins (IPR017986: WD40 repeat, region) from 32 to 40, and the number for DEAD box proteins (IPR011545: DNA/RNA helicase; DEAD/DEAH box type, N-terminal) from 20 to 24.

The C/EBPβ structure is predicted to fold back onto itself, permitting intramolecular signaling, or to stretch out to allow contact with several interaction partners simultaneously (Kowenz-Leutz et al., 1994; Lee et al., 2010a; Lynch et al., 2011). A strong indication of the involvement of multiple contacts with interaction partners through adjacent or distant C/EBPβ regions is shown in Figure S2A (Internal C/EBPβ connections and GO term distribution. Related to Figure 2) that lists multivalent interactions, including potential interactions with different proteins of the same complex.

Discrete parts of the C/EBPβ primary structure have previously been assigned to different biochemical or cellular functions. Protein binding data from different C/EBPβ regions was extracted and GO term analysis performed for each of the regions to map structure-function relationships, as shown in Figure S2B. Internal C/EBPβ connections and GO term distribution. Related to Figure 2). While some functional attributions displayed partition over several regions of C/EBPβ (regulation of gene expression, RNA
splicing), others are more localized to distinct regions. For example, basic and DNA binding regions shows strong enrichment for ‘regulation of mitotic cell cycle processes’ or ‘nuclear export’ whereas ‘nuclear import’ is associated with CR5, CR7, IR7 and LZ. While the C-terminal region of C/EBPβ is involved in the majority of GO terms, transcription factor co-activator functions are mainly associated with N-terminal activation of CR2, CR3, CR4 and IR7.

**C/EBPβ-interacting protein complexes**

Next, PRISMA data were analyzed for potential enrichment of large protein complexes interacting with CEBPβ. In addition to anticipated C/EBPβ interacting complexes such as the basic RNA polymerase II transcription machinery or Mediator, previously unknown interrelations became apparent. These included potential complexes involved in the RNA transcript processing machinery responsible for transcript capping, splicing, termination, and polyadenylation. Interactions included 3'end processing, cleavage and polyadenylating factor (CPSF1, 3, 7, 30, and 100), and several components of the Integrator complex involved in snRNA expression and the RNA exosome (EXOSC2, 4, 6, 7, 8, 9, and 10). Furthermore, many components of the nuclear pore and associated adapter complexes were identified. In addition, components of the transcript export THO-TREX complex, including THO1, 2, 3, 5, 6, 7, Aly, DDX39, and the THO-TREX-associated mRNA export factors Nxf1-Nxt1 were identified, along with a large number of heterogeneous nuclear ribonucleoproteins that may couple transcript splicing, maturation, and the formation of messenger ribonucleoprotein particles. Taken together, these data imply the existence of a previously unknown connection between CEBPβ and several steps involved in transcript generation, processing, maturation, and transport (Muller-McNicoll and Neugebauer, 2013; Wickramasinghe and Laskey, 2015).
Proteins co-occurring in both the PRISMA and SU-DHL1 proteomic data were then systematically explored to identify soluble protein complexes listed in the CORUM database (2017-03-06) using the g:Profiler bioinformatics toolkit (Reimand et al., 2016; Ruepp et al., 2010). A list of 1432 human protein complexes built from 2678 proteins (single UniProt identifier; ID) was derived after removal of redundant and non-human complex entries. Proteins listed in PRISMA replicates (SET1+SET2) and protein interactions derived from SU-DHL1 IP-mass spectrometry data were matched by UniProt and gene name ID using Perseus version 1.5.2.4 (Tyanova et al., 2016). Redundant gene names and isoform entries were merged into a single UniProt ID for each protein. Altogether, 816 proteins of the PRISMA-derived dataset and 490 proteins of the IP SU-DHL1 dataset were included in any of the 1432 CORUM complexes. The 1432 CORUM complexes were then ranked by a combination of two criteria: (i) the percentage of proteins in complexes obtained by PRISMA, and (ii) deviation from randomness of the coverage of the complexes obtained by PRISMA and SU-DHL1 (p-values with CORUM background, see Supplementary Information for details). Considering only complexes sharing at least one protein with the PRISMA dataset and which have at least three overlaps with PRISMA and SU-DHL1 data resulted in a ranked list of 417 candidate complexes. Whenever possible, complexes were grouped into categories indicating functional connections, such as DNA repair, nuclear pore or centromere, in addition to well-characterized multi-subunit complexes, such as Mediator, SWI/SNF, MLL, NuRD, and others. As shown in Figure 3A, 45 multi-protein complexes and categories were extracted, and normalized protein interaction values at any of the 203 C/EBPβ peptides were summed over all instances of the category in the list and depicted as a bar-plot to visualize the distribution of interaction sites of different categories or complexes in relation to the C/EBPβ primary structure. These 45 entities were composed of 30 complexes or categories representing the full upper quartile of the ranked complex list (i.e. up to rank 104, Table S5, with minor exceptions, Figure S3), together with 15 lower-ranking functional counterparts.
Figure 3B shows prominent node-link diagrams of a selection of 14 multi-protein complexes among the highest-ranking categories that share many proteins identified by both PRISMA and SU-DHL1 immunoprecipitation methods.

Validation of region- and PTM-specific C/EBPβ interactions

Next, conventional protein pull-down, immunoprecipitation, and immunoblotting analysis was used in combination with CEBPβ mutants to validate data from the PRISMA-derived interactome (Figure 4). According to PRISMA, the nuclear factor kappa B subunit RelA, the signal transducer and activator of transcription STAT3 and the nuclear pore complex protein NUP50/NPAP60L predominantly interacted with peptides located in CR7. Bacterially expressed GST-C/EBPβ constructs were probed with HEK293 extracts to examine RelA and NUP50/NPAP60L interactions. As shown in Figure 4B and 4D (lower panel) affinity capture with GST-CEBPβ constructs or co-immunoprecipitation of STAT3 with the C/EBPβ mutant protein lacking CR7 and subsequent immunoblotting confirmed CR7 as the interaction site for all three proteins.

The histone acetyltransferase GCN5 of the SAGA complex and the DNA methyltransferase-associated protein DMAP1 of the NuA4/TIP60 acetyltransferase complex are involved in a wide variety of developmental and genome regulatory activities, including transcription, DNA repair, DNA methylation, and chromatin remodeling (Mohan et al., 2006; Weake and Workman, 2012). DMAP1 and GCN5 engage in major binding to CR3 and CR4, respectively, and minor binding elsewhere in C/EBPβ. Various C-terminally FLAG-tagged C/EBPβ mutant constructs lacking CR1, CR2 (5D43), the entire TAD (5D141), CR4 (LAP*Δ4), CR7 (LAP*Δ7), or both (LAP*Δ4,7) were expressed in HEK293 cells to examine selective and multi-site interaction with resident DMAP1/GCN5 or their respective complexes. As shown in Figure 4C and D, pull-down and immunoblotting revealed that
removal of the DMAP1 minor binding site in CR2 partially abrogated the interaction, and removal of binding to CR3 completely abrogated the interaction. Similarly, removal of the binding site for GCN5 located in CR4 strongly affected GCN5 association and deletion of the binding site in CR7 entirely abolished binding to C/EBPβ.

The impact of PTMs on C/EBPβ-mediated interactions

Detection of PTM-dependent protein interactions in linear C/EBPβ peptides was a major objective for the development of the PRISMA screening method. Protein binding to each PRISMA peptide and its PTM-modified versions were therefore examined. The binding signal for each protein was normalized against the signal from the corresponding unmodified peptide to compare enhanced or reduced binding. Interacting proteins were then clustered, as shown in Figure S4. PTM-dependent binding across C/EBPβ. Related to Figure 5. Depending on PTM, as exemplified in Figure 5A, interactions fell into four categories; (i) generally repressed binding (ii) PTM-independent binding, (iii) PTM-specific binding, and (iv) generally enhanced binding. Most of the identified binding partners were indifferent to the type of PTM and were classified as binding proteins that may recognize parts of the peptide sequences. Generally enhanced binding may in part relate to the fact that many PTMs increase hydrophobicity and may thus enhance interactions non-specifically. The most interesting binding partners are expected to be found in the last two categories and represent proteins that are disturbed by any PTM in the peptide or represent interaction partners that were attracted or repelled by a particular PTM. Examples of the latter are the NuA4/TIP60 complex components DMAP1 and YEATS4, both binding to CR3, or RelA and the transducin-like enhancer protein TLE3 binding to CR7, which depends on the methylation status of arginine residues that were subsequently examined.
The transactivation region CR3 (residues 53–68: AIGEHERAIDFSPYLE) contains an arginine residue that is conserved in C/EBPβ but not in C/EBPα, δ, ε and was previously found to be methylated (Leutz et al., 2011). The PRISMA data suggested that DMAP1 and YEATS4 binding to CR3 depended on methylation of R59. Another interaction hot-spot was mapped to the start site of the highly conserved, alternatively initiated LIP C/EBPβ isoform that also represents part of the regulatory region CR7 (residues 158–170: FPFALRAYLGYQAT). The respective arginine residues in both chicken (R60, R193) and rat (R58, R162) were previously found to be methylated. Mutation of the equivalent amino acid residue to alanine/leucine in chicken C/EBPβ strongly altered transcriptional activity, suggesting that the methylation status of the conserved arginine in CR7 may be critical (Leutz et al., 2011). WT C/EBPβ, methylation-defective R60A and R193A, and methylation mimicking R60L and R193L constructs were subjected to co-immunoprecipitation to compare alterations in binding, as shown in Figure 5B. As suggested by PRISMA, interaction of all four proteins with C/EBPβ was sensitive to the amino acid side chain configuration at the respective arginine positions, demonstrating both amino acid and PTM specificity, in addition to evolutionary conservation of chicken and rat C/EBPβ interactions. Binding of YEATS4 or DMAP1 to C/EBPβ was enhanced or tolerated by the R60L exchange but not the R60A exchange, respectively, confirming the side chain specificity and preference for increased side chain hydrophobicity. RelA binding to the R193A mutant was slightly enhanced in comparison to WT or reduced by the R193L mutant, respectively, confirming avoidance of the methylated, more hydrophobic PRISMA peptide. By contrast, the TLE3 co-repressor (Agarwal et al., 2015) favored binding to the R193L mutant, suggesting that the increase in hydrophobicity by methylation of the arginine side chain but not the positive charge was important for interaction with TLE3. Pull-down of TLE3 by immuno-precipitation confirmed preferential binding of C/EBPβR193L, as shown in Figure 5C. Next, we performed methylation assays with the protein arginine methyltransferase PRMT4/CARM1 and C/EBPβ peptides, as shown in Figure 5D. C/EBPβR193 was identified
as a methylation target in CR7, in addition to the previously described R3 in CR1 (Kowenz-Leutz et al 2010). Co-immunoprecipitation of TLE3 together with WT C/EBPβ was dependent on co-expression of PRMT4/CARM1 (Figure 5E). Importantly, and irrespective of co-expressed PRMT4/CARM1, the methylation-defective C/EBPβR193A mutant failed to bind TLE3, confirming C/EBPβ CR7 arginine side chain methylation as a prerequisite for binding to TLE3.

Both C/EBPβ and TLE3 had been described as important regulators in adipogenesis (Villanueva et al., 2013). As shown in Figure 5F, ectopic expression of C/EBPβ in 3T3L1 cells induced partial adipogenic differentiation in the absence of the in vitro differentiation cocktail (IBMX, Insulin, Dexamethasone). Fat cell differentiation was evident by up-regulation of the adipogenic genes ADIPOQ, CFD, FABP4 (Figure 5F, bar graph panel) and enhanced Oil-red-O staining (below) in local adipogenic nests. Adipogenic gene expression of ADIPOQ and FABP4 was enhanced by the C/EBPβR193L mutant and TLE3 cooperated with C/EBPβ WT and more strongly with C/EBPβR193L but not with the C/EBPβR193A mutant to activate ADIPOQ, CFD, and FABP4. We conclude that PRISMA may uncover functional C/EBPβ PTM-regulated protein-protein interactions to adjust the biology of C/EBPβ.

Novel connections between C/EBPβ, transcription elongation, MLL, and NuRD

Of particular interest was the observation that many proteins forming part of the machinery involved in RNA polymerase II (RNAPII) pausing and elongation were included in the PRISMA data. Components of the Trithorax MLL/Set1/COMPAS histone H3 lysine 4 (H3K4) methyltransferase complexes (Shilatifard, 2012) were also represented in the dataset, raising the possibility of a connection between C/EBPβ, enhancer/promoter binding, and regulation of RNAPII processivity. Many components implicated in both processes were
found, including the MLL complex components CHD3 and 4 (Mi-2), ASH2, Dpy30, PTIP, HCF-1, the general transcriptional elongation factor B (TFIIS) and associated elongin A/B/C factors (Aso et al., 1995), the P-TEFb components cyclinK/T1, cdk9, and additional regulatory components including LARP7, HEXIM1, SPT5 (DSIF), and the chromatin adaptor bromodomain factor Brd4 and components of the super elongation complex (SEC) including AF9, PCAF1, and AF4 (AFF1) (Luo et al., 2012). AFF1 is a central SEC scaffold component, and AF9 and ENL are highly similar YEATS domain family members that compete for binding to AFF1. The N-terminal YEATS domain of AF9 and ENL both bind to the PAF complex to connect SEC to RNAPII on chromatin templates. The CDK9/CYCT1 P-TEFb complex is required for rapid transcriptional induction, phosphorylation of the C-terminal domain of RNAPII, and engagement of BRD4, which also interacts with H3K9ac. P-TEFb is also associated with a 7SK snRNA subcomplex that contains the regulatory components LARP7, HEXIM1, and SPT5 (DSIF) and connects to PAFc in the dynamic transcription elongation machinery (He et al., 2011; Luo et al., 2012; Peterlin and Price, 2006). In addition, the histone chaperone FACT complex facilitates passage of the transcription apparatus through chromatin and is thought to restore the chromatin structure and the epigenetic state during transcription, replication and repair (Hammond et al., 2017; Hondele and Ladurner, 2013; Orphanides et al., 1998).

Immunoprecipitation and immunoblotting were performed in order to validate several of these novel connections. As shown in Figure 6, multiple MLL, FACT, and SEC components were all co-immunoprecipitated with C/EBPβ, confirming the predictive capacity of PRISMA. It is also important to note that AFF1, AF9, and ENL are among the most frequent oncogenic fusion partners with the MLL gene product that transforms early hematopoietic progenitors and causes childhood leukemia by short-circuiting enhancer and promoter activation and transcriptional elongation checkpoint controls (Krivtsov and Armstrong, 2007; Shilatifard, 2012; Slany, 2009; Smith et al., 2011).
The Mi2/Nucleosome Remodeling Deacetylase (NuRD) corepressor complex is widely expressed and maintains chromatin in a closed state via ATP-dependent chromatin remodelling and histone tail deacetylation. Some of the NuRD components (CHD3/4(Mi2), HDAC1, RBBP4/7, MTA1,2, GATAD2A/p66α, MBD2, MBD3) identified by PRISMA, SU-DHL1 IP and other approaches (Steinberg et al 2012; Siersbaek et al 2014) are shared with the SIN3A and CoREST repression complexes (Gregoretti et al 2004), and all were implicated by CORUM analysis (Figure 3, Figure S3, Table S5). Two major NuRD complexes contain either the unique methyl-CpG binding proteins MBD3 or MBD2, the latter of which may also associate with the WDR77(MEP50)/PRMT5 sub-module (Le Guezennec et al., 2006; Zhang and Yinghua, 2011). Immunoprecipitation of either MBD2 or MBD3 revealed association of C/EBPβ, Mi2, and MTA with both MBD components, whereas only MBD2 co-precipitated with the WDR77(MEP50)/PRMT5 sub-module, suggesting that C/EBPβ may associate with both types of NuRD complexes.

Discussion

In conclusion, the PRISMA technique reveals a comprehensive and PTM-dependent interactome of the disordered C/EBPβ transcription factor hub and permits the detection of footprints of protein interactions and protein complexes. The results presented in this manuscript suggest that C/EBPβ is involved at the nexus between enhancer and promoter regulation, in many aspects of pre- and posttranscriptional control, including initiation and pausing control, splicing and exosomal RNA degradation, polyadenylation, RNA maturation and in nuclear export. Previously, we described an array peptide screening system (APS) for PTM-dependent protein interactions that relied on fluorophore-based detection of peptide interactions with bacterially expressed human proteins immobilized on macroarrays. Although it is difficult to compare APS with PRISMA because of major differences in peptide choice, peptide length, biochemical conditions, local concentration of interaction partners,
signal to noise ratio, or complexity of the protein expression library and requirement of a renaturation step, 35 common APS/PRISMA interaction partners, including Brg1/SMARCA4, were found with a N-terminal APS-peptide and a comparable section covered in PRISMA that contained CR1 (Pless et al., 2011).

The high local concentration of peptides in matrix spots may limit free diffusion and overcome the dissociation problem of low affinity interactions (Ruthenburg et al., 2007). These features permit the detection of weak interactions and interactions with rare protein partners. The interaction screen developed herein is augmented by inclusion of amino acids with modified side chains to enable detection of PTM-dependent interactions on a global scale. Together, the interaction footprints and PTM-dependent data may help in the rational design of mutants to further explore the functional C/EBPβ interactome, and aid in the development of pharmacological inhibitors of protein interactions. The PRISMA workflow is applicable to other regulatory proteins and particularly to transcriptional regulators with a high degree of intrinsic disorder such as Myc- or various Hox proteins. PRISMA may help to discern highly complex and dynamic transcription factor functions that arise through IDR- and PTM-regulated interactions.

PRISMA was developed for analysis of the interactome of proteins with a high degree of intrinsic disorder. Such proteins lack a defined 3D structure, and protein interactions and complex formation are based on SLiMs, PTMs, and intrinsic flexibility. IDRs in conjunction with PTMs may facilitate structural transitions in partner protein exchanges (Wright and Dyson, 2015). Comprehension of the dynamics and context-dependent interactions between structurally flexible SLiMs and various partner proteins is an emerging hallmark of signal transmission and key to understand the regulation of chromatin structure and gene expression (Minde et al., 2017; Tompa et al., 2014; van der Lee et al., 2014; Wright and Dyson, 2015). SLiM interactions often occur with high specificity but low affinity to enable multiple contacts of a dynamic nature, such as occurring during phase separation.
Deciphering the functionality of SLiM/MoRF-type protein regions and PTMs in homeostasis and disease can be challenging when using full-length proteins due to functional redundancy and compensatory effects. A variant of PRISMA has meanwhile been employed to detect dysregulated interactions with disease causing mutations occurring in IDRs (Meyer et al., 2018). PRISMA data presented here thus provide a repository of high molecular resolution C/EBPβ interactions and serve as a basis to explore gene regulatory and PTM-modulated C/EBPβ functions.

Limitations of the study
Protein interactions can be mediated by binding domains or by smaller interaction motifs. One key aspect of the study is the identification of the interactome of intrinsically disordered regions and SLiMs and binding regulation by PTMs. As the PRISMA technique is based on the use of peptides, the detection of interactions is limited to those that are based on structurally variable SLiMs, as binding directly depends on the amino acid sequence properties. The high local concentrations of the peptides on the membrane permits the detection of low-affinity interactors and leaves components of their corresponding protein complexes intact under mild washing conditions. For interactions that are based on recognition of three-dimensional structures, however, the method will probably not be applicable.

Acknowledgements
We would like to thank Manuel Baesler for generating a web tool (Peptide2Protein) for visualization of PRISMA interactions, Andreas Ladurner, LMU Munich, for FACT expression plasmid constructs, Bernd Lüscher and Juliane Lüscher-Firzlaff, RWTH, Aachen, for MLL complex expression plasmids, and Peter Tontonoz, HHMI, UCLA for TLE3 constructs. Parts
of the work were supported by the CPIL-PEARL grant to GD and DFG Grants, LE770/4-2, CRC167B11 to AL.

Author contributions


Declaration of interests

The authors declare no competing interest

References


Figure 1. Outline of PRISMA screen and comparison of data

A. Schematic representation of C/EBPβ and the distribution of known post-translational modifications (PTMs). Conserved regions (CRs) are depicted in color, while intrinsic disordered regions (IDR) are shown in grey. B. Schematic description of the workflow of
protein binding and data acquisition. C. Workflow of data and proteomic analysis. D. Overlap between the two replicates (SET1, SET2) of the PRISMA screen. E. Number of proteins in the two PRISMA datasets that show consecutive peptide binding (core interactions, dark amber). F. Overlap of three affinity purification-based datasets of C/EBPβ interactors as described in the literature and combined with data obtained from a proteomic interaction screen in SU-DHL1 cells. Overlaps were determined using IP SU-DHL1 as reference dataset and thus the numbers add up to the size of this set only (see Material and Methods). G. Overlap of the PRISMA-derived C/EBPβ interactor datasets from E (union of SET1 and SET2, light amber) with core interactions (dark amber) from the union of the three datasets from F (blue). Overlaps are given using the PRISMA-derived data as reference datasets. The overlap count using the union of the datasets from F as a reference dataset is denoted in brackets.

Figure 2. Distribution of protein interactions discovered by PRISMA

A. Heatmap of the normalized intensities of interaction partners. The bar-graph (top) shows the distribution of accumulated protein interactions by normalized binding intensities for each peptide. The tiled peptides of C/EBPβ (below) are plotted on the horizontal axis, and proteins are shown on the vertical axis following non-supervised hierarchical clustering. B. Prediction of C/EBPβ interaction regions using ‘Anchor’ (blue) and ‘MoRF’ (magenta) bioinformatics tools and prediction of intrinsic disorder (grey) in relation to the schematic representation of C/EBPβ and the primary structure (bottom of A). Different conserved regions are colored as shown in Figure 1A. C. Heatmaps of selected complexes.

Figure 3. PRISMA-based prediction of protein complexes interacting with C/EBPβ
A. Bar-graphs of accumulated normalized intensities for 45 selected complexes. Based on CORUM database protein complex annotations, the corresponding normalized intensities of proteins identified by PRISMA were extracted and plotted. B. Network representation of 14 potential complexes identified. Nodes are color-coded according to their detection in PRISMA and SU-DHL1 IP experiments (yellow = detected in IP and PRISMA, green = detected in IP, dark blue = detected in the PRISMA core set, light blue = detected in PRISMA SET1 and 2).

Figure 4. Validation of site-specific interactions with C/EBPβ

A. Accumulated binding intensities of interaction partners (indicated on the left) from the PRISMA screen. Bargraphs indicate interactions according to the peptide position in C/EBPβ, as schematically shown underneath. B. Top: Color-coded scheme of the C/EBPβ protein with dashed lines to aid comparison of constructs shown in B, C, and D. Top, right: Immunoblots (right) showing interaction of RelA and Nup50/NPAP60L of HEK293T cell lysates with bacterially expressed GST-C/EBPβ constructs. Integrity of GST-C/EBPβ constructs was examined by immunoblotting (as indicated). GST constructs were probed with anti-GST and a polyclonal C/EBPβ antibody that preferentially, but not exclusively, recognizes CR7 as a major epitope. The most intense signals of RelA and Nup50/NPAP60L are associated with the presence of the CR7 region of C/EBPβ. A, Antibody; IP, immunoprecipitation; IB, Immunoblotting. C. Co-immunoprecipitation of FLAG-tagged LAP* C/EBPβ or three N-terminal C/EBPβ deletion mutants, followed by immunoblot detection of co-precipitated DMAP1 from HEKT cells. D. Immunoprecipitation of HEKT cell expressing LAP* and three different internal deletion mutants of C/EBPβ followed by immunoblot detection of GCN5 and Stat3.
Figure 5. PTM-specific C/EBPβ interactions.

A. PTM-dependent protein binding to FPFALRAYLGYQAT peptides of C/EBPβ CR7. Normalized binding intensity relative to the unmodified peptide. Four binding categories are considered: proteins with reduced binding by any PTM (repressed, blue, far left); proteins not responding to modifications on C/EBPβ (independent, yellow, middle), proteins with PTM-specific binding (regulated, yellow-red, small), and enhanced binding by any PTM (enhanced, red, far right). B. Immunoprecipitation (IP) from transfected HEK293T cell lysates: FLAG tagged C/EBPβ wild-type (WT) or mutants, as indicated, abrogating (C/EBPβ R60A; R193A) or mimicking (C/EBPβ R60L, R193L) methylation. Co-immunoprecipitation (co-IP) of endogenous YEATS4 and DMAP1 detected by immunoblotting. YEATS4 and DMAP1 bind to C/EBPβ R60L and failed to bind to the R60A C/EBPβ mutant. RelA favors WT or C/EBPβ R193A binding over R193L. TLE3 preferentially binds to C/EBPβ R193L. Underneath: Expression controls of endogenous YEATS4, DMAP1, RelA, and TLE3. C. Co-IP of WT and mutant C/EBPβ, as indicated, with HA-tagged TLE3. C/EBPβ R193L but not WT or R193A co-IPs with TLE3. D. In vitro methylation of C/EBPβ peptides by CARM1. Scheme of C/EBPβ with approximate positions of peptides and R-residues indicated by red bars and dots, R residue sequence positions underneath. Bar graphs show peptide specific incorporation of methyl-group from donor S-adenosyl-L-[methyl-3H] methionine. E. Co-IP of C/EBPβ with TLE3 depends on CARM1. Co-expression of WT C/EBPβ (R) or mutants (R193A, R193L), TLE3 and CARM1 are indicated on the top. WT C/EBPβ but not R193A co-IPs with TLE3 in the presence of CARM1, whereas the C/EBPβ R193L mutation supersedes CARM1 requirement. F. RT PCR of ADIPOQ, CFD, and FABP4 expression (arbitrary units: target gene to internal 36B4 expression control) induced by C/EBPβ WT, R193A, R193L or vector control in the absence or presence of TLE3 in NIH 3T3 L1 cells without adipogenic differentiation cocktail (upper panel; n=3, ANOVA; *P≤0.05, **P≤0.01, ***P≤0.001; ns: not significant). Oil-red O staining of stably transfected NIH 3T3 L1 cells with C/EBPβ
constructs, vector control and TLE3, as indicated, 3 weeks post-confluency (middle panel, representative images), protein expression controls (lower panel).

Figure 6. C/EBPβ-interacting multi-protein complexes in HEK293T cells.

A. Co-immunoprecipitation of FLAG-tagged C/EBPβ and the MLL/Ash2/Wdr5 complex. B. Left panel: FLAG-tagged C/EBPβ was co-expressed with GFP-tagged Spt16 and SSRP1. Right panel: Immunoprecipitation of Flag tagged C/EBPβ co-precipitates both endogenous FACT complex subunits. C. Interaction with the super elongation complex. Immunoprecipitation of human C/EBPβ (C/EBPβ) pulls down the entire super-elongation complex. D. Interaction with the NuRD complex. The integral subunits of the NuRD complex MBD2 (HA-tagged) and MBD3 (GFDP-tagged) were pulled down by immunoprecipitation. Detection by western blotting shows that C/EBPβ interacts with both subunits.

Main Table Title

Table 1: C/EBPβ interactome enriched gene ontology (GO) terms

Supplemental Figures Titles and Legends

Supplemental Figure 1: Comparison of enrichment by PRISMA and physico-chemical properties of PRISMA peptides. Related to Figures 1 and 2. A. Distribution of the copy number. Estimations based on the iBaq quantification for the PRISMA screen and the nuclear extracts. Both, nuclear extracts and PRISMA binders are in the same copy number range. B. Correlation of the accumulated intensity of the PRISMA binders to different peptide properties. C. Binding of proteins in the PRISMA screen in comparison to various calculated...
peptide properties. The top panel shows the accumulated binding intensity of the PRISMA 
binders, while the other panels show the different calculated peptide properties for the 
peptides listed below.

**Supplemental Figure 2: Internal C/EBPβ connections and GO term distribution.**

**Related to Figure 2.**

**A.** Co-binding of interaction partners to different C/EBPβ regions. The number of interaction 
partners was calculated that bind to two different regions, as indicated on the abscissa. Dot 
sizes represent the number of binding partners in the C/EBPβ regions, while the width of the 
arcs represent the number of interaction partners found in the connected regions. **B.**

Enrichment analysis of GO-terms for binding partners in the different regions of C/EBPβ.

GO-terms were selected for DNA related processes (pink), nuclear import/export (orange), 
chromatin (green), RNA splicing (red) and transcription (blue).

**Supplemental Figure 3: Potential C/EBPβ interacting protein complexes. Related to**

**Figure 3.**

Complex coverage and p-values for each of the 1432 complexes as used for the ranking of 
CORUM complexes. Each dot corresponds to one complex. Red symbols show complexes 
which have less than three protein hits in PRISMA and SU-DHL1 data sets. Circles 
represent complexes ranking within the upper quartile (dark grey) (104 complexes, dashed 
line between upper quartile and lower ranks), triangles encode lower ranking complexes.

The 104 highest ranking complexes are shown in the close-up on the right. Highest ranking 
complexes for each of the 14 categories from Fig. 4B are indicated in black; other complexes 
which are not included in Fig. 4A, in grey.

**Supplemental Figure 4: PTM-dependent binding across C/EBPβ. Related to Figure 5.**

The positions of the PTM-modified peptides are shown on the schematic presentation of the
primary C/EBPβ sequence on top. For each peptide, the relative binding to the PTM-modified peptides was calculated in relation to the unmodified peptide. Heatmaps display the relative binding of the proteins on the y-axis in relation to different PTM-carrying peptides, indicated below the heat maps. Red indicates enhanced binding and blue shows reduced binding. Note that the Figure has infinite zoom function for detailed inspection of data.

Supplemental Excel Table Titles

Supplemental Table 1: C/EBPβ PTMs, Related to Figure 1A

Supplemental Table 2: PRISMA C/EBPβ Peptides, Related to Figure 1B

Supplemental Table 3: The PRISMA C/EBPβ Interactome, Related to Figure 1D

Supplemental Table 4: C/EBPβ interactome in SU-DHL cells, Related to Figure 1F and 1G

Supplemental Table 5: PRISMA deduced C/EBPβ complexes, Related to Figure 3
<table>
<thead>
<tr>
<th>enrichment score</th>
<th>pathway / domain description</th>
<th>counts in PRISMA</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO: Biological Processes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:0010467</td>
<td>gene expression</td>
<td>611</td>
<td>4.19e-103</td>
</tr>
<tr>
<td>GO:0006355</td>
<td>regulation of transcription, DNA-templated</td>
<td>301</td>
<td>1.27e-07</td>
</tr>
<tr>
<td>GO:0006396</td>
<td>RNA processing</td>
<td>267</td>
<td>5.63e-128</td>
</tr>
<tr>
<td>GO:0007049</td>
<td>cell cycle</td>
<td>211</td>
<td>3.06e-31</td>
</tr>
<tr>
<td>GO:0008380</td>
<td>RNA splicing</td>
<td>172</td>
<td>7.14e-106</td>
</tr>
<tr>
<td>GO:0051276</td>
<td>chromosome organization</td>
<td>169</td>
<td>1.16e-34</td>
</tr>
<tr>
<td>GO:0006461</td>
<td>protein complex assembly</td>
<td>128</td>
<td>4.17e-12</td>
</tr>
<tr>
<td>GO:0006568</td>
<td>chromatin modification</td>
<td>89</td>
<td>2.50e-13</td>
</tr>
<tr>
<td>GO:0006281</td>
<td>DNA repair</td>
<td>86</td>
<td>9.46e-38</td>
</tr>
<tr>
<td>GO: Molecular Function</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:0003676</td>
<td>nucleic acid binding</td>
<td>745</td>
<td>3.8e-199</td>
</tr>
<tr>
<td>GO:0003723</td>
<td>RNA binding</td>
<td>607</td>
<td>0E+00</td>
</tr>
<tr>
<td>GO:0005515</td>
<td>protein binding</td>
<td>534</td>
<td>1.92e-40</td>
</tr>
<tr>
<td>GO:0003677</td>
<td>DNA binding</td>
<td>269</td>
<td>6.43e-17</td>
</tr>
<tr>
<td>GO:0005524</td>
<td>ATP binding</td>
<td>216</td>
<td>2.82e-24</td>
</tr>
<tr>
<td>GO:0003682</td>
<td>chromatin binding</td>
<td>94</td>
<td>4.43e-18</td>
</tr>
<tr>
<td>GO:0003712</td>
<td>transcription cofactor activity</td>
<td>81</td>
<td>2.27e-09</td>
</tr>
<tr>
<td>GO:0008134</td>
<td>transcription factor binding</td>
<td>66</td>
<td>5.86e-09</td>
</tr>
<tr>
<td>GO:0019900</td>
<td>kinase binding</td>
<td>61</td>
<td>3.30e-06</td>
</tr>
<tr>
<td>GO:0004693</td>
<td>cyclin-dependent protein serine/threonine kinase activity</td>
<td>17</td>
<td>6.20e-10</td>
</tr>
<tr>
<td>GO: Cellular Component</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:0000785</td>
<td>chromatin</td>
<td>103</td>
<td>6.04e-31</td>
</tr>
<tr>
<td>GO:0005681</td>
<td>spliceosomal complex</td>
<td>85</td>
<td>5.10e-56</td>
</tr>
<tr>
<td>GO:0005643</td>
<td>nuclear pore</td>
<td>32</td>
<td>9.77e-19</td>
</tr>
<tr>
<td>GO:0030880</td>
<td>RNA polymerase complex</td>
<td>30</td>
<td>2.98e-10</td>
</tr>
<tr>
<td>GO:0070603</td>
<td>SWI/SNF superfamily-type complex</td>
<td>28</td>
<td>9.63e-14</td>
</tr>
<tr>
<td>GO:0034708</td>
<td>methyltransferase complex</td>
<td>24</td>
<td>4.86e-09</td>
</tr>
<tr>
<td>GO:0016592</td>
<td>mediator complex</td>
<td>20</td>
<td>3.98e-13</td>
</tr>
<tr>
<td>GO:0000178</td>
<td>exosome (RNase complex)</td>
<td>12</td>
<td>4.26e-09</td>
</tr>
<tr>
<td>GO:0016581</td>
<td>NuRD complex</td>
<td>11</td>
<td>1.50e-09</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------------------</td>
<td>-----</td>
<td>----------</td>
</tr>
<tr>
<td>GO:0016593</td>
<td>Cdc73/Paf1 complex</td>
<td>6</td>
<td>1.01e-05</td>
</tr>
</tbody>
</table>

**Interpro Domains**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Score</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPR027417</td>
<td>P-loop containing nucleoside triphosphate hydrolase</td>
<td>102</td>
<td>9.71e-30</td>
</tr>
<tr>
<td>IPR000504</td>
<td>RNA recognition motif domain</td>
<td>80</td>
<td>8.77e-53</td>
</tr>
<tr>
<td>IPR005225</td>
<td>Small GTP-binding protein domain</td>
<td>41</td>
<td>2.11e-13</td>
</tr>
<tr>
<td>IPR016024</td>
<td>Armadillo-type fold</td>
<td>34</td>
<td>1.69e-11</td>
</tr>
<tr>
<td>IPR017986</td>
<td>WD40-repeat-containing domain</td>
<td>32</td>
<td>6.78e-10</td>
</tr>
<tr>
<td>IPR009072</td>
<td>Histone-repeat</td>
<td>27</td>
<td>3.89e-09</td>
</tr>
<tr>
<td>IPR001650</td>
<td>Helicase, C-terminal</td>
<td>23</td>
<td>1.94e-14</td>
</tr>
<tr>
<td>IPR011989</td>
<td>Armadillo-like helical</td>
<td>21</td>
<td>5.11e-06</td>
</tr>
<tr>
<td>IPR011545</td>
<td>DEAD/DEAH box helicase domain</td>
<td>20</td>
<td>2.11e-13</td>
</tr>
<tr>
<td>IPR002041</td>
<td>Ran GTPase</td>
<td>20</td>
<td>5.34e-12</td>
</tr>
</tbody>
</table>

**PFAM Domains**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Score</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF00076</td>
<td>RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)</td>
<td>72</td>
<td>4.32e-49</td>
</tr>
<tr>
<td>PF00071</td>
<td>Ras family</td>
<td>31</td>
<td>8.17e-09</td>
</tr>
<tr>
<td>PF00400</td>
<td>WD domain, G-beta repeat</td>
<td>29</td>
<td>1.38e-10</td>
</tr>
<tr>
<td>PF00271</td>
<td>Helicase conserved C-terminal domain</td>
<td>23</td>
<td>2.27e-14</td>
</tr>
<tr>
<td>PF00125</td>
<td>Core histone H2A/H2B/H3/H4</td>
<td>23</td>
<td>1.89e-09</td>
</tr>
<tr>
<td>PF00270</td>
<td>DEAD/DEAH box helicase</td>
<td>20</td>
<td>2.63e-13</td>
</tr>
<tr>
<td>PF00013</td>
<td>KH domain</td>
<td>10</td>
<td>1.36e-06</td>
</tr>
<tr>
<td>PF00244</td>
<td>14-3-3 protein</td>
<td>6</td>
<td>2.32e-04</td>
</tr>
<tr>
<td>PF01423</td>
<td>LSM domain</td>
<td>9</td>
<td>2.32e-04</td>
</tr>
<tr>
<td>PF00538</td>
<td>linker histone H1 and H5 family</td>
<td>5</td>
<td>4.50e-04</td>
</tr>
</tbody>
</table>
Figure 1, Dittmar et al.
Figure 2; Dittmar et al.
Figure 3; Dittmar et al.
Figure 4; Dittmar et al.
Figure 5, Dittmar et al.
Figure 6; Dittmar et al.
Highlights

- PRISMA yields a high resolution SLiM and PTM based C/EBPβ interactome
- Disclosure of intrinsically disordered protein hub functions and interaction networks
- New connections between C/EBPβ and RNA processing, transcription elongation, MLL, NuRD
- C/EBPβ interaction with TLE3 is regulated by arginine methylation