The human oncoprotein PRAME unveiled:
interaction networks and genomic targets

Adalberto Costessi

To the inspiring and dedicated teachers
I had the opportunity to meet
and to the memory of Gabriella Greblo.

"A teacher affects eternity;
he can never tell where his influence stops."
Henry Adams

The human oncoprotein PRAME unveiled

interaction networks and genomic targets

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,
volgens besluit van het college van decanen
in het openbaar te verdedigen op maandag 21 januari 2013
om 10.30 uur precies

door Adalberto Costessi
geboren op 23 april 1979
te Trieste (Italië)

The human oncoprotein PRAME unveiled

interaction networks and genomic targets

DOCTORAL THESIS

to obtain the degree of doctor
from the Radboud University Nijmegen
on the authority of the Rector Magnificus prof. dr. S.C.J.J. Kortmann,
according to the decision of the Council of Deans
to be defended in public on Monday January 21, 2013
at 10.30 hours

by Adalberto Costessi
born on April 23, 1979
in Trieste (Italy)

Supervisor:
Prof. dr. H.G. Stunnenberg

Doctoral thesis committee:
Prof. dr. G. Pruijn (chairman)
Prof. dr. C.P. Verrijzer (Erasmus University Rotterdam)
Dr. D. Libri (Centre de Genetique Moleculaire, CNRS, Gif sur Yvette, France)

Table of contents

AIM

The human protein PRAME (preferentially expressed antigen of melanoma) is frequently overexpressed in tumours. Importantly, high expression levels are associated to poor clinical outcome in melanoma, neuroblastoma, and breast cancer. However, its function and molecular mechanisms in normal and cancer cells are not yet well understood and the cellular pathways affected by PRAME are still elusive. Although PRAME has been indicated as a promising potential target for cancer therapy and anti-tumour immunization protocols, efforts to develop strategies to interfere with the oncogenic activity of PRAME are hindered by a lack of information about its function.

The aim of this thesis was to investigate the molecular function of human PRAME using cell line model systems and a powerful combination of proteomic and genomic techniques. We developed innovative biochemical approaches to dissect the protein-protein interaction network of PRAME, and we setup chromatin immunoprecipitation assays coupled to next-generation sequencing technologies to reveal binding profile of PRAME to the human genome in cancer cells.

OUTLINE

Chapter 1 provides a general introduction and an overview of the knowledge available on PRAME and the PRAME-like gene family. This chapter introduces also the proteomic and genomic techniques developed and applied in this thesis.

Chapter 2 describes the setup and application of epitope-tagging affinity purification and mass spectrometry analysis to the identification of the protein interactors of PRAME.

In *chapter 3* we extend the results of chapter 2 and characterize PRAME as a component of Cullin2-based ubiquitin ligases. Furthermore, we present the first genome-wide ChIP-seq profiling of PRAME that associates this oncoprotein to the NFY signaling pathway.

Chapter 4 presents the results of improved protein-complex purification protocols and analysis of the PRAME interactome to a deeper level. We characterize direct interactions between PRAME and the human EKC/KEOPS complex and we show that EKC subunits associate with PRAME binding sites on chromatin.

Chapter 5 reports the development of rabbit antibodies against PRAME and their validation for use in several assays. We further discuss the experimental strategies that can be undertaken to generate genome-wide ChIP-seq profiles of a protein of interest, including epitope-tagging approaches. Finally, we investigated the reported function of PRAME as dominant repressor of retinoic acid signaling using proliferation assays and *PRAME* knockdown systems.

Chapter 6 summarizes the results of the previous chapters and discusses the implications of this thesis for further research.

In the *Appendix*, the work performed for this thesis is presented in a chronological context, as a reference for similar projects in the future.

LIST OF ABBREVIATIONS AND GENE SYMBOLS

Ab, antibody

ChIP, chromatin immunoprecipitation

CRL, Cullin-RING ligases

CUL2, cullin 2

EKC, Endopeptidase-like Kinase Chromatin-associated

ELOB/TCEB2, elongin B/transcription elongation factor B (SIII), polypeptide 2

ELOC/TCEB1, elongin C/transcription elongation factor B (SIII), polypeptide 1

IP, immunoprecipitation

KEOPS, kinase, putative endopeptidase and other proteins of small size

LAGE3, L antigen family, member 3

MS, mass spectrometry

NFY, nuclear transcription factor Y

OSGEP, O-sialoglycoprotein endopeptidase

PRAME, preferentially expressed antigen in melanoma

RBX1, ring-box 1

RING, really interesting new gene

TP53RK, TP53 regulating kinase

TPRKB, TP53RK binding protein

WB, western blot

GENERAL INTRO– DUCTION

Cancer is a general term that refers to more than hundred different diseases that are all characterised by uncontrolled cell proliferation. Collectively, cancers constitute the second leading cause of human mortality worldwide (WHO, 2012). Cancers are mostly classified based on the tissue of origin and histological analysis of tissue sections or cells, based on which diagnosis and prognosis are made. Some cancers are also tested for specific molecular markers of known prognostic value, i.e. ERalpha and HER2 for breast cancer. The advent of array-based techniques, and more recently high-throughout sequencing, to quantitatively measure the levels of messenger RNA transcripts has substantially increased the power to sub-classify cancers based on gene expression profiles (McDermott *et al*, 2011). Importantly, specific gene expression signatures have been characterised that are predictive of prognosis and can therefore guide therapeutic interventions, like the Mammaprint 70-gene profile for breast cancer (McDermott *et al*, 2011; van 't Veer *et al*, 2002).

The human oncoprotein PRAME

The human gene *PRAME* (preferentially expressed antigen of melanoma) is frequently overexpressed in tumours and constitutes a molecular marker of poor clinical outcome for several cancers like melanoma, neuroblastoma and breast cancer (Kilpinen *et al*, 2008; Oberthuer *et al*, 2004; Epping & Bernards, 2006; Doolan *et al*, 2008; Haqq *et al*, 2005). The discovery of the *PRAME* gene was reported in 1998 by the Coulie lab (Ikeda *et al*, 1997). The authors identified *PRAME* as the gene encoding an antigenic peptide recognised by tumour-specific cytolytic T-lymphocytes isolated from a melanoma patient. The gene encodes a putative protein of 509 amino acids with unknown function. In healthy adult tissues, *PRAME* is expressed at high levels only in the testis, while low levels were detected in the adrenals, ovary, endometrium, and very low levels in skin, brain, breast, and kidney (Ikeda *et al*, 1997). However, *PRAME* is frequently overexpressed in a wide range of solid and haematological tumours including melanoma, head and neck carcinoma, ovarian cancer, breast carcinoma, lung carcinoma, renal carcinoma, neuroblastoma, and various acute and chronic haematological malignancies (Ikeda *et al*, 1997; Kilpinen *et al*, 2008).

Based on its peculiar expression pattern, PRAME belongs to the category of cancer-testis antigens, similarly to the MAGE, BAGE and GAGE gene families. The expression in healthy tissues other than testis, however,

differentiates PRAME from other cancer-testis antigens, whose expression is restricted to normal testis. Since the testis does not bear HLA molecules, it is invisible to the immune system and high expression of PRAME (and other CTA proteins) in this tissue does not trigger an immune response. It has been suggested that the low expression of PRAME in other normal tissues is insufficient for recognition by cytolytic T-lymphocytes, and only upon malignant overexpression would the immune system be able to mount an anti-PRAME immune response (Ikeda *et al*, 1997; van Baren *et al*, 1998).

PRAME expression in cancer

A large number of studies investigated the transcript levels of *PRAME* in cohorts of cancer patients. While normal skin and benign nevi were found negative, high *PRAME* expression was found to correlate with the stage of melanoma lesions (Haqq *et al*, 2005), and a similar trend was reported for neuroblastoma: 93% of primary lesions and 100% of advanced cases overexpressed PRAME (Oberthuer *et al*, 2004). On the contrary, *PRAME* levels were lower in neuroblastoma cases characterised by a good prognosis and a high frequency of spontaneous regression (Fischer *et al*, 2006).

Importantly, significant associations were found between higher *PRAME* expression levels and poor clinical outcome (i.e. shorter overall and disease-free survival) in neuroblastoma (Oberthuer *et al*, 2004), as well as serous ovarian adenocarcinoma patients (Partheen *et al*, 2008). High *PRAME* mRNA levels were also found to be an independent prognostic factor of poor clinical outcome in breast cancer by two independent studies (Doolan *et al*, 2008; Epping *et al*, 2008).

PRAME transcripts were detected in about 30-40% of acute and chronic leukaemias, and hypomethylation of the *PRAME* locus seems to have a role in this process (Roman-Gomez *et al*, 2007; Schenk *et al*, 2007). Studies on patients affected by BCR-ABL-positive chronic myeloid leukaemia (Ph+ CML) indicated that *PRAME* was activated in the advanced stages and blast crisis (Radich *et al*, 2006; van Baren *et al*, 1998). The fact that *PRAME* was not expressed in the early stages, when the BCR-ABL translocation is already present, prompted the authors to suggest that the BCR-ABL translocation is not sufficient to activate *PRAME*, and the later activation could be due to a secondary correlation with the progression of the disease. However, two studies reported that *PRAME* expression could be induced by transfection of BCR-ABL in cell line models (De Carvalho *et al*, 2011; Watari *et al*, 2000).

Therefore, it cannot be excluded that additional (still BCR-ABL-mediated) signals are necessary for full activation of *PRAME* in CML blasts.

Although a growing body of evidence indicates that PRAME contributes markedly to the oncogenic phenotype of several cancers, some studies have shown that *PRAME* expression can be associated also with leukaemia subtypes characterised by a favourable prognosis. In particular, high *PRAME* levels were reported in patients with the genomic translocations t(8;21) and t(15;17), which result in the expression of the AML1-ETO and PML-RARA fusion proteins, respectively (van Baren *et al*, 1998). A larger recent study on patients carrying the PML-RARA translocation reported that high *PRAME* levels at diagnosis were associated with a better clinical outcome (Santamaría *et al*, 2008).

Finally, several studies indicated that *PRAME* expression levels are a useful marker to monitor the presence of cancer cells in the body (also called "minimal residual disease") in patients during and after chemotherapy (Steinbach *et al*, 2002; Santamaría *et al*, 2008).

Molecular studies of PRAME

There are indications that epigenetic processes are involved in the regulation of *PRAME* expression. DNA hypomethylation of two different regions of the *PRAME* locus was reported to correlate with expression in cell lines as well as in the progression of CML. Notably, hypomethylation induced by 5-aza-dC treatment resulted in *PRAME* upregulation in several cell lines, suggesting that all the factors required for *PRAME* expression were present in the cells, and that DNA methylation mediates the silencing of this locus (Ikeda *et al*, 1997; Roman-Gomez *et al*, 2007; Schenk *et al*, 2007).

Five different transcript variants of *PRAME* have been described (Schenk *et al*, 2007). These variants show two different transcription start sites (positions +1 and +151), but they all encode the same amino acid sequence. Regulation of expression of the different variants has not yet been reported, and their functional relevance is unclear (Schenk *et al*, 2007). Recent analysis of small non-coding RNAs revealed that PRAME can be targeted by the microRNA miR-211, and downregulation of miR-211 correlates with upregulation of PRAME in melanoma cell lines (Sakurai *et al*, 2011).

Although numerous studies reported expression patterns of *PRAME* in human malignancies, the molecular and cellular functions of PRAME are not yet clear. In a functional genetic screen, ectopic expression of PRAME

was found to repress the growth inhibitory effects of histone deacetylase inhibitors (HDACi) (Epping *et al*, 2007). HDACi induce proliferation arrest and apoptosis in cancer cells but not in normal healthy cells; therefore, they are promising compounds for use in the treatment of cancer (Kim & Bae, 2011). Ectopic expression of human PRAME in mouse embryonic fibroblasts (MEFs) prevented growth arrest and apoptosis upon treatment with the HDACi PXD101, while PRAME overexpression did not affect proliferation in standard culturing conditions (Epping *et al*, 2007). Luciferase reporter assays indicated that PRAME could repress retinoic acid signaling induced by PXD101, while PRAME silencing in A375 melanoma cells sensitized the cells to PXD101 treatment *in vitro* as well as in a xenograft model (Epping *et al*, 2007).

The authors reported in a separate publication that PRAME interacts with the retinoic acid receptor alpha (RARA) and suggested that PRAME mediates repression of retinoic acid signaling via recruitment of the transcriptional repressor EZH2 to RARA target genes (Epping *et al*, 2005). In particular, the authors reported in vitro protein-protein interactions between PRAME and the ligand binding domain of RARA fused to GST, while a mammalian two-hybrid system detected an interaction between the ligand binding domain of RARA and the C-terminus of PRAME (aa. 416-509). Co-immunoprecipitation experiments further suggested that this interaction requires the C-terminal LRELL amino acid motif in PRAME and the presence of the ligand all-trans retinoic acid (atRA). Luciferase reporter assays indicated that expression of PRAME represses retinoic acid signaling and that this effect is mediated by protein-protein interactions with the transcriptional repressor EZH2. A ChIP assay suggested that PRAME associates with the RARB2 promoter, a known target of retinoic acid signaling. Finally, PRAME silencing in A375 melanoma cells was reported to restore retinoic acid signaling and inhibit retinoic acid induced growth arrest both in proliferation assays in vitro and in a xenograft model in vivo (Epping *et al*, 2005). Contrary to the model put forward by the Bernards lab, Wadelin *et al*. could not reproduce the interactions between PRAME and the ligand binding domain of RARA in GST pulldown and yeast two-hybrid assays (Wadelin *et al*, 2010). Furthermore, studies on primary AML and breast cancer samples indicated that PRAME expression did not associate with a down-regulation of retinoic acid signaling (Epping *et al*, 2008; Steinbach *et al*, 2007).

Notably, it was reported that overexpression of SOX9 in melanoma cell lines lead to increased sensitivity to retinoic acid and at the same time

downregulation of PRAME (Passeron *et al*, 2009). However, the authors did not provide a causative link between these two findings. Therefore, it is still unclear whether the downregulation of PRAME plays any role in the increased sensitivity to retinoic acid caused by SOX9 overexpression, or whether it is an unlinked event.

In conclusion, a putative role for PRAME in the retinoic acid signaling pathway remains unclear, and additional studies are required to shed light on these mechanisms.

PRAME in leukaemia

The molecular functions of PRAME in leukaemia cells are also controversial, since conflicting reports suggested that PRAME might induce caspase-independent cell death (Tajeddine *et al*, 2005), or rather promote cell survival repressing pro-apoptotic genes (Tanaka *et al*, 2011).

Tajeddine *et al.* reported that PRAME causes decreased proliferation and caspase-independent cell death in CHO-K1, HeLa and KG1 cell lines. Using a microarray-based approach, transfection of PRAME in leukaemia KG1 cells was found to downregulate expression of the *HSP27*, *CDKN1A/p21* and *S100A4* genes, which are inhibitors of apoptosis. However, these associations were not reported in a microarray gene expression analysis of paediatric AML samples with high versus low *PRAME* expression (Goellner *et al*, 2006). Moreover, an independent study on paediatric AML samples by real-time PCR found that high levels of PRAME were not associated with down-regulation of *HSP27*, *S1000A*, *CDKN1A/p21* or *RARB* (Steinbach *et al*, 2007).

Tajeddine *et al.* also reported that downregulation of PRAME by siRNA increased proliferation and tumorigenicity of K562 leukaemia cells in a xenograft nude mice model. On the contrary, Tanaka *et al.* showed that siRNA transfection in the same cell line suppressed proliferation in vitro and promoted apoptosis. Another study in agreement with Tajeddine *et al.* found that silencing of PRAME promoted cell proliferation of K562 cells; however, the same effect was obtained upon overexpression of PRAME in HL60 cells (Oehler *et al*, 2009). Moreover, PRAME silencing on K562 cells expressing increased numbers of RARA receptors did not show any difference in proliferation in the same assays (Oehler *et al*, 2009). It is possible that the functional relevance of PRAME differs between different cell types and cancers, and further analyses are required in order to understand the mechanisms and pathways in which PRAME is involved.

Evolution of the PRAME gene family

PRAME is the founding member of a large family of genes with a peculiar evolutionary history. In human, the *PRAME* gene family is comprised of the single *PRAME* gene on chromosome 22, and a cluster of 22 *PRAME*-like genes and 10 pseudogenes on chromosome 1 (Birtle *et al*, 2005).

Molecular evolution analyses established that the *PRAME*-like gene cluster originated from a translocation dating shortly before the divergence between rodents and primates (~85-95 million years ago). In fact, synthenic regions of rodent and primate genomes contain *PRAME*-like homologues, while dog and chicken genomes do not (Birtle *et al*, 2005). After the first translocation event, the mouse and primate loci underwent an independent and very rapid expansion by "birth-and-death" processes, i.e. gene duplication and pseudogene creation. This process explains the finding that mouse *Prame*-like genes are monophyletic, similarly to the human *PRAME*-like genes. In other words, while ~80% of mouse genes have a single human orthologue, there is no pair of mouse and human *PRAME* genes with a simple 1:1 orthology. This evidently hinders the development of mouse model systems for the study of human *PRAME* genes.

Furthermore, independent duplication events have been detected in all branches of the primate lineage (Rhesus Macaque Genome Sequencing and Analysis Consortium *et al*, 2007). Extensive duplications occurred early in the primate evolution as well as in recent chimpanzee evolution, while the macaque branch seems to have been less dynamic (Figure 1). The chimpanzee and hominin branches separated around 6-7 million years ago, and two independent large segmental duplications occurred in the terminal human branch in the past 3 million years. This implies that a number of human *PRAME*-like genes do not have a 1:1 orthology even with chimpanzee *PRAME* genes.

Besides the extraordinarily rapid and independent expansion of PRAME genes in the rodent and the primate lineages, there is also evidence for copy number variation in the *PRAME*-like gene cluster among human populations (Birtle *et al*, 2005). This indicates that there have been additional recent duplications that are not yet fixed in the human population, nor captured in the human genome reference sequence.

Evidence has also been found for positive selection of amino acid residues in PRAME (Birtle *et al*, 2005; Rhesus Macaque Genome Sequencing and Analysis Consortium *et al*, 2007). Homology modelling suggested that the

positively-selected residues identified are likely to cluster on an accessible surface of PRAME proteins, underscoring a potential role in binding interactions. Moreover, since positive selection is a common feature of genes involved in immunity and reproduction, it was suggested that PRAME genes might have roles in these processes (Birtle *et al*, 2005; Wadelin *et al*, 2010). Given its peculiar characteristics, it was suggested that the PRAME gene family has played a key role in species evolution and that the differences between human and chimpanzee PRAME genes may have contributed to the functional divergence along the hominin lineage (Birtle *et al*, 2005).

While the evolution of the PRAME-like gene cluster has been studied in detail, the origin of the *PRAME* gene on chr22 is unclear. Inspection of the orthologous genomic regions in the mouse, rat and chimp genomes, did not identify *PRAME* ortho-logues, and gaps in the genome builds have not been reported that could accommodate a missing *PRAME* gene (Birtle *et al*, 2005).
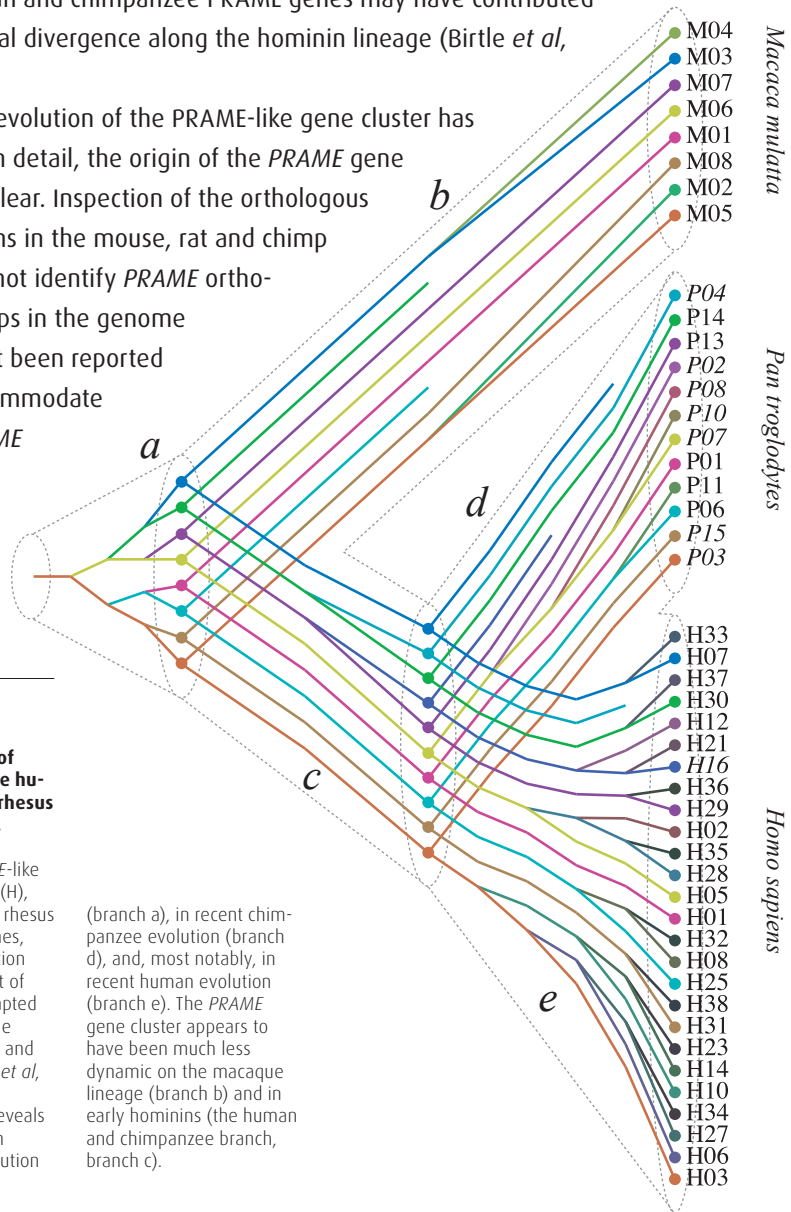
FIGURE 1

Phylogenetic tree of *PRAME* genes in the human-chimpanzee-rhesus macaque lineages.
Maximum-likelihood phylogeny for *PRAME*-like genes in the human (H), chimpanzee (P), and rhesus macaque (M) genomes, showing the duplication history in the context of the species tree (adapted from Rhesus Macaque Genome Sequencing and Analysis Consortium *et al*, 2007).
This reconstruction reveals extensive duplication early in primate evolution (branch a), in recent chimpanzee evolution (branch d), and, most notably, in recent human evolution (branch e). The *PRAME* gene cluster appears to have been much less dynamic on the macaque lineage (branch b) and in early hominins (the human and chimpanzee branch, branch c).

PRAME-like genes and early development

Despite the extraordinary evolutionary history of *PRAME* genes, their biological and molecular function is still poorly understood. While no reports have been published yet on human *PRAME*-like genes, a number of studies in the mouse identified roles for this gene family in oogenesis (Dadé *et al*, 2003), spermatogenesis (Wang *et al*, 2001) and stemness (Bortvin *et al*, 2003; Cinelli *et al*, 2008).

A number of mouse *PRAME*-like genes were first identified as genes with specific expression in the oocyte, and therefore they were first called "oogenesin" (Dadé *et al*, 2003; Minami *et al*, 2003). Analysis of their predicted amino acid sequence suggested that the encoded proteins belong to the leucine-rich repeat (LRR) superfamily and fold in a non-globular horse-shoe structure (Dadé *et al,* 2003).

Four PRAME-like genes were found to play roles in pluripotency and development of mouse embryos cloned by nuclear transfer (Bortvin *et al*, 2003). In a study aimed at identifying genes involved in establishing or maintaining developmental pluripotency, the authors performed an in silico analysis to identify genes with an expression pattern similar to that of Oct4 (Bortvin *et al*, 2003). OCT4 is a transcription factor expressed in pluripotent cells of the early embryo and adult germ cells that is known to play an essential role in the control of developmental pluripotency. Ten genes were found that satisfied the search criteria, among which *Pramel4*, *Pramel5*, *Pramel6*, *Pramel7*. Analysis of expression of these genes established that they are active in pluripotent embryonic tissues (blastocyst, morula and 4-cell embryo), while they are not expressed in differentiated somatic tissues. Importantly, analysis of their activity in mouse embryos cloned from somatic or embryonic stem (ES) cells revealed that successful reactivation of the full set of *Oct4*-genes correlated with successful development of the cloned embryos to term. The authors concluded that although the biological functions of these genes have yet to be elucidated, they serve as informative markers for pluripotent cell types and cloned embryos (Bortvin *et al*, 2003).

Another study fond that the MBD3 component of the nucleosome re-modelling and histone deacetylation (NuRD) complex mediates silencing of *Pramel6* and *Pramel7* expression in mouse embryonic stem cells (Kaji *et al*, 2006). These genes were found to be expressed at very low levels in wild type ES cells, while knockout of *Mbd3* resulted in elevated transcript levels as well as increased H3K9 and H4 acetylation levels at their genomic loci,

but not at the silent *Pramel5* promoter. This correlated with LIF-independent maintenance of pluripotency in culture and a defect in lineage commitment of *Mbd3-/-* ES cells. Although a causative role for *Pramel6* and *Pramel7* in these processes was not addressed, expression analysis suggested that *Pramel6* and *Pramel7* contribute in deciding the fate of ES cells (Kaji *et al*, 2006).

An independent study confirmed that *Pramel6* and *Pramel7* are prevalently expressed in embryonic pluripotent cells, and further determined that whereas *Pramel6* is typically expressed in all cells of the morula and blastocyst, *Pramel7* is expressed only in the inner part of the morula and in the inner cell mass of the blastocyst, similarly to *Nanog*, a factor known to be involved in pluripotency (Cinelli *et al*, 2008). Notably, ectopic overexpression of the STAT3 transcription factor resulted in overexpression of *Pramel6* and *Pramel7* and allowed the derivation of pluripotent embryonic stem cell lines from mouse strains that are otherwise "non permissive" to this procedure. Remarkably, overexpression of *Pramel7* in ES cells was sufficient to sustain maintenance of pluripotency when cells were cultured in the absence of LIF, indicating an important role for this protein in maintaining ES cells in a pluripotent state (Cinelli *et al*, 2008).

Ubiquitin pathways

Modification of proteins by the addition of ubiquitin (a highly conserved 76-amino-acid polypeptide) or ubiquitin-like proteins has emerged as a major mechanism by which virtually all events in cells can be regulated, including cell cycle progression, gene transcription, DNA repair, protein turn-over, receptor trafficking, and apoptosis (Deshaies & Joazeiro, 2009; Budhidarmo *et al*, 2012).

Ubiquitination (also known as ubiquitylation) is best-known for targeting proteins for degradation by the 26S proteasome, but other functions include modulation of protein interactions and alteration of subcellular distribution.

Members of ubiquitin pathways play also important roles in cancer and a number of them have been identified as either oncogenes (i.e. *MDM2*) or tumour suppressor genes (i.e. *BRCA1* and *VHL*), and the first pharmacological inhibitors of the proteasome have entered the clinic as anti-cancer drugs, and significant efforts are undertaken in this field of drug discovery (Joazeiro *et al*, 2006).

Attachment of ubiquitin to target proteins is performed by a cascade of three enzymes (E1-E2-E3) (Deshaies & Joazeiro, 2009; Metzger *et al*, 2012

and references therein). The first step is the activation of ubiquitin by one predominant ubiquitin-activating enzyme E1. Activated ubiquitin is then transferred from E1 to one of the 30-40 mammalian ubiquitin-conjugating enzymes (UBCs or E2s), and eventually to the final substrates that are bound specifically to one of the several hundreds different ubiquitin ligases (E3s). Substrates can be modified by a single ubiquitin on one or more sites (monoubiquitination and multi-monoubiquitination, respectively) or can be tagged with chains of ubiquitin (polyubiquitination). Chain formation can occur through the seven internal lysine residues on ubiquitin (K6, K11, K27, K29, K33, K48, and K63) or its N-terminus (M1). Ubiquitin chains are recognized by specialised ubiquitin-binding domains (UBDs) that form transient, noncovalent interactions either with ubiquitin moieties or with the linkage region in their chains and determine the function of the modification (Ikeda *et al*, 2010). K48-linked polyubiquitin chains were originally described as the signal that targets substrates for proteasomal degradation, while atypical K11, K63, and M1-linked chains were associated to DNA repair, cell cycle progression and inflammation. However, recent reports have challenged the notion that specific chain topologies regulate specific cellular processes, and it still unclear how a specific type of linkage is generated and how this is translated to the fate of the modified protein (Ikeda *et al*, 2010).

The ubiquitination cascades result in the formation of a covalent bond between the C-terminus of ubiquitin and, most frequently, the epsilon-amino group of a lysine side chain in the substrate. The ubiquitination of the N-terminus of proteins or other amino acids like threonine, serine, or cysteine has also been described. E3-mediated attachment of ubiquitin to substrates is highly regulated in response to cellular cues and one substrate can be also ubiquitinated by different ubiquitin ligases while, on the other hand, one ligase can target several proteins. Moreover, E3s themselves can undergo both self-ubiquitination and be ubiquitinated by heterologous ligases. Examples are the RING finger 1B component of the polycomb repressive complex 1 (PRC1), which self-ubiquitinates to increase its activity, while MDM2 self-ubiquitination targets itself and the p53 proto-oncogene for proteasome degradation (Weissman *et al*, 2011).

Since E3s dictate the specificity toward a variety of substrates and at the same time interact with an E2, they are the most structurally complex and diverse enzymes in the pathway (Zimmerman *et al*, 2010). E3 ligases can be classified into two main structural classes that bear either a HECT (ho-

mologous to E6-AP carboxy terminus) domain or a RING (really interesting new gene) domain, although a few have a U-box, which is structurally and functionally similar to the RING domain (Budhidarmo *et al*, 2012). HECT-containing E3 ligases, of which there are ~20 in humans, form an intermediate conjugate with the ubiquitin received from a charged E2 before transferring it to the target. In contrast, the more prevalent (~500) RING-containing E3 ligases bring together an ubiquitin-charged E2 and a substrate protein and facilitate the transfer of ubiquitin directly from the charged E2 to the substrate (Deshaies & Joazeiro, 2009).

Cullin-RING E3 ubiquitin ligases

Among the E3 RING ligases, the structurally-related multisubunit cullin-RING ubiquitin ligases (CRLs) are particularly significant (Petroski & Deshaies, 2005; Sarikas *et al*, 2011). This is a diverse group of E3 complexes characterized by containing both a cullin family protein and a RING finger protein (RBX1 or RBX2). The human genome encodes for six cullin proteins, CUL1, CUL2, CUL3, CUL4A, CUL4B, and CUL5, and two additional related proteins, PARC and CUL7, which are considered atypical cullins.

Extensive genome-wide analysis of the cullin family has suggested that three ancestral cullin genes, termed 'Culα', 'Culβ' and 'Culγ', appeared in early eukaryotic evolution, from which the cullin genes evolved after the split of the unikonts (which include animals and fungi) and the bikonts (which include plants) (Marín, 2009). The human *CUL1, CUL2, CUL5, CUL7* and *PARC* genes were derived from one common ancestral gene (Culα), and the available data indicate that the emergence of multiple Culα genes in animals has been accompanied by a diversification of their protein partners and complex subunits. The *CUL3* and *CUL4A/4B* genes evolved from two distinct ancestors, the Culβ and Culγ gene, respectively (Marín, 2009).

CRLs are the most abundant E3s as it is predicted that about 400-500 different E3 are encoded in the human genome. Their diversity is due to the fact that many alternative complexes can be generated in a combinatorial way: multiple related substrate receptor proteins may substitute for each other to form similar but functionally distinct complexes.

The roles of the cullin protein in this type of complex are quite well understood. Structural data indicate that cullins have an overall elongated crescent-shaped structure consisting of a stalk-like N-terminal domain (NTD), and a globular C-terminal domain (CTD) (Zimmerman *et al*, 2010) (Figure

2). The NTD binds adaptor proteins and substrate receptor subunits, which confers substrate specificity, while the CTD binds the RING-finger containing protein RBX1 or RBX2, which recruits the ubiquitin-conjugating enzyme E2. In this context, cullins act as backbones that facilitate ubiquitination by correctly positioning the RING finger protein and the substrate (Marín, 2009; Zimmerman *et al*, 2010; Petroski & Deshaies, 2005) (Figure 2).

Cullin-RING ligases can be classified in a number of main CRL classes based on the different adaptors, substrate receptors, and RING-finger proteins used (Marín, 2009; Petroski & Deshaies, 2005):

1. Cullin1/RBX1/SKP1/F-box-receptor CRLs (historically known as SCF complexes) that contain SKP1 or related proteins as adaptors and proteins with F-box motifs as substrate receptors;

FIGURE 2   27

Structural model of cullin-RING (CRL) ubiquitin ligases.
The cullin-based E3 ubiquitin ligases position a substrate in close proximity to the RING-bound E2 ubiquitin-conjugating enzyme (not shown), which catalyzes the transfer of ubiquitin to the substrate. (A) Schematic model of the CRL modular architecture (adapted from Sarikas *et al*, 2011). (B-E) Structural models of CRLs with the scaffold cullin proteins, RING-box proteins, adaptor proteins and substrate receptors (from Zimmerman *et al*, 2010).

2. Cullin2/RBX1/ElonginBC/VHL-box-receptor and Cullin5/RBX2/Elongin-BC/SOCS-box-receptor CRLs (BC-box CRLs), which contain two Elongin proteins as adaptors and BC-box containing proteins as substrate receptors (Petroski & Deshaies, 2005; Okumura *et al*, 2012; Mahrour *et al*, 2008);

3. Cullin3/RBX1/BTB-receptor CRLs, characterized by lacking additional adaptors and containing proteins with BTB (broad complex/tram-track/bric-a-brac) domains as substrate receptors, directly bound to the cullins;

4. Cullin4/RBX1/DDB1/WD40-proteins CRLs, which contain DDB1 as adaptor and proteins with WD40 domains as substrate receptors, also called DCAFs (DDB1–CUL4A-associated factors).

Despite sharing the identical adaptor proteins Elongin B and Elongin C, CUL2 and CUL5 direct the assembly of E3 ligases with distinct substrate receptor proteins: CUL2 with von Hippel-Lindau (VHL) or related BC-box proteins, and CUL5 with suppressors of cytokine signaling (SOCS)-box proteins. Sequence comparison of the VHL and SOCS-box proteins revealed that they each include a degenerate, ~12 amino acid sequence motif with consensus sequence [S,T,P]Lxxx[C,S,A]xxxΦ (Mahrour *et al*, 2008; Kamura *et al*, 2004). This motif, which is referred to as the BC-box, is required for binding to the Elongin BC complex. Downstream of the BC-box, a Cul2- or Cul5-box is present, which mediates the specific interaction to the corresponding cullin scaffold (Mahrour *et al*, 2008; Kamura *et al*, 2004). The most studied substrate of EloBC-cullin ligases is the von Hippel-Lindau (VHL) protein, the product of the *VHL* tumour suppressor gene. More than half of sporadic clear cell renal carcinomas contain inactivating mutations of VHL, and the VHL complex has ubiquitin ligase activity and targets the hypoxia-inducible factor-α (HIF-α) family of transcription factors (HIF-1–3α) for proteasomal degradation. At normal oxygen levels, proline residues within the oxygen-dependent degradation domain of HIF-α are hydroxylated and recognized by pVHL for polyubiquitination and proteasome-mediated degradation (Okumura *et al*, 2012).

The ligase activity of CRLs was found to be modulated by the covalent conjugation of an ubiquitin-like molecule, Nedd8 (Deshaies & Joazeiro, 2009). This modification, termed neddylation, activates the E3 ligase activity by promoting substrate polyubiquitination, likely via the induction of extensive conformational changes. Neddylation is a reversible modification that can

be enzymatically removed by the COP9 complex, and cullins are thought to undergo dynamic regulation by neddylation-deneddylation cycles. Another level of regulation of CRLs activity is performed by the CAND1 protein (Cullin-Associated and Neddylation-Dissociated-1), which inhibits the E3 ligase activity of CRLs by binding to both the C-terminus as well as the N-terminus of all cullins in their un-neddylated forms (Deshaies & Joazeiro, 2009).

The EKC/KEOPS complex

Two independent laboratories have identified an ancient and highly conserved multiprotein complex named KEOPS (Downey *et al*, 2006) or EKC (Kisseleva-Romanova *et al*, 2006), starting from genetic clues in yeast. This complex has orthologues from Archaea to Eukarya and so far it has been implicated in telomeres maintenance, transcriptional regulation, and t^{6}A modification of tRNAs (Kisseleva-Romanova *et al*, 2006; Downey *et al*, 2006; Daugeron *et al*, 2011; Yacoubi *et al*, 2011; Srinivasan *et al*, 2011). Yeast EKC comprises four subunits which are also conserved in the human genome (human orthologues are indicated in brackets): Pcc1p (LAGE3, also known as ESO3), the ATPase Kae1p (OSGEP), the kinase Bud32p (TP53RK, also known as PRPK), and Cgi121p (TPRKB). In addition, yeast EKC also includes Gon7p (also known as Pcc2p), which appears to be fungi-specific (Kisseleva-Romanova *et al*, 2006).

The laboratory of Domenico Libri isolated Pcc1, a protein of small size with no predicted functional domains in a search for suppressors of a splicing defect and demonstrated that its function is not related to splicing. Their genetic and biochemical experiments led to the description of the five-subunit EKC complex. Furthermore, the authors showed that Pcc1p is required for efficient transcription of several yeast genes induced by pheromones or galactose, and that deletion of Pcc1 inhibited the recruitment of the SAGA and Mediator co-activators (Kisseleva-Romanova *et al*, 2006).

A parallel genome-wide screen in yeast identified the cgi121 null mutant as a suppressor of the telomere-capping defect of cdc13-1 cells (Downey *et al*, 2006). Deletion of CGI121 or other EKC subunits resulted in short telomeres in an otherwise wild-type context. A detailed analysis of three-dimensional crystal structures of archaeal EKC subunits revealed that the complex has a linear topology Pcc1-Kae1-Bud32-Cgi121, and Pcc1 is an obligate dimer that can support dimerization of the full complex in a V shape (Figure 3). Dimerization of the Pcc1p and Kae1p subunits was found to be necessary for
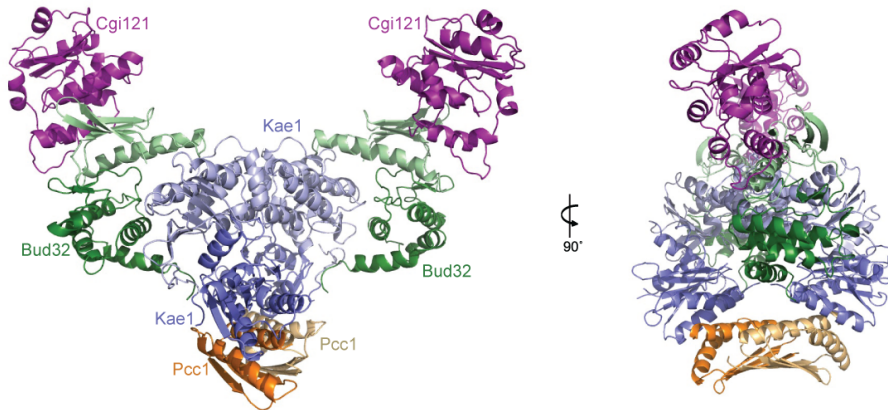
FIGURE 3 **Composite model of the archaeal EKC/KEOPS complex in dimeric form (from Mao et al, 2008).**

cell viability (Mao *et al*, 2008), and an intact interaction surface between Kae1p and Bud32p was required for both the transcriptional and telomere maintenance functions (Hecker *et al*, 2008).

Intriguingly, the OSGEP subunit is also present in bacteria (YgjD) and eukaryotic genomes express an OSGEP paralogue (Qri7/OSGEPL1) that localizes to mitochondria (Reinders *et al*, 2006). Comparative genomic studies identified OSGEP as one of the very few genes present in all genomes sequenced so far (Galperin, 2008), suggesting an extremely conserved function. Very recently, several groups have reported a crucial role for the YgjD/Kae1/OSGEP protein family in the biosynthesis of N6-threonylcarbamoyl adenosine (t⁶A) (Daugeron *et al*, 2011; Yacoubi *et al*, 2011; Srinivasan *et al*, 2011): a universal modification at position 37 of tRNAs decoding ANN codons, which is required for accurate translation of messenger RNAs (Yarian *et al*, 2002).

Human LAGE3 belongs to the NY-ESO gene family together with the closely related LAGE1 and LAGE2 (Alpen *et al*, 2002), and all three genes are clustered in the same region on chromosome X. While LAGE3 is ubiquitously expressed, LAGE1 and LAGE2 are cancer-testis antigens with high expression in healthy testis and upregulation in a number of cancer tissues, similar to PRAME. Notably, expression of the N-terminal domain of LAGE3 fused to the C-terminus of Pcc1p could partially rescue the growth defects of a Pcc1 N-terminal mutant yeast strain, while full length LAGE3 could not substitute for Pcc1 (Kisseleva-Romanova *et al*, 2006)

Human TP53RK was first cloned by cDNA subtraction as an interleukin-2 upregulated gene in cytotoxic T-cells, and it was shown to phosphorylate the human oncoprotein p53 at Serine 15 (Abe *et al*, 2001). Recombinant TP53RK

30

expressed in bacteria was found catalytically inactive, but incubation with COS-7 cell lysates was sufficient to activate the enzymatic activity, suggesting the need for activators and potentially protein interactors. The kinase Akt/PKB was found to play a role in the activation of TP53RK through phosphorylation of the Ser250 residue (Facchin *et al*, 2007). Underscoring a high level of functional homology, TP53RK could partially rescue the growth phenotype of Bud32 deletion in yeast, and Bud32 could interact with and phosphorylate human p53 in vitro (Facchin *et al*, 2003), although yeast does not possess a bone fide p53 homologue.

Human TPRKB was first isolated by a two-hybrid screen using TP53RK as a bait, and was suggested to inhibit the binding of TP53RK to p53 (Miyoshi *et al*, 2003).

Although the structures of several EKC subunits have been characterised and phenotypes caused by mutations or depletion of EKC have been described in yeast, the precise molecular functions and mechanisms of EKC in Archaea as well as Eukarya are still unclear.

Purification and identification of protein complexes

Recent advances in genome sequencing technologies have resulted in hundreds of genomes being completely sequenced, and this number is rapidly growing (http://www.ensemblgenomes.org/). In parallel, computational algorithms have been developed that can efficiently predict and annotate genes, thereby generating comprehensive lists of predicted proteins for many species (the "proteomes"). However, such lists are not sufficient to understand cellular processes, which are dependent on the coordinated action of multiprotein complexes and dynamic signaling networks of interacting proteins. Therefore, experimental approaches have been developed to dissect protein-protein interaction networks (the "interactomes") in order to fully understand the cellular machinery (Mering *et al*, 2002; Gingras *et al*, 2007).

A powerful technique to identify protein-protein interactors is the purification of entire protein complexes by epitope-tagging and affinity-based methods coupled to mass spectrometry identification. This approach has been successfully applied at a high-throughput scale for the dissection of the yeast and human interactomes (Ewing *et al*, 2007; Gavin *et al*, 2002), and also to the study of specific proteins of interests (Le Guezennec *et al*, 2006).

Epitope tags are short amino acid sequences that can be recognised by

specific immunoprecipitation-grade antibodies (i.e. FLAG, HA, MYC) or other affinity reagents (i.e. calmodulin-binding columns, streptavidin beads). The first step entails cloning the coding sequence of the target protein in frame with sequences coding for one or multiple tags in a suitable expression vector, often retroviral. Subsequently, viral particles are produced to transduce the target cells and polyclonal or monoclonal cultures with stable expression of the tagged protein are selected. The tagged protein of interest is purified from protein extracts together with its protein partners, which are subsequently identified by mass spectrometry (Figure 4). The list of interactors identified can be further validated by epitope-tagging of these proteins in a so-called "complex walking" approach. Importantly, proteins can take part to multiple complexes at the same time, but this cannot be inferred from a simple list of interactors. However, several lists of interactors generated by complex walking can be integrated to generate an interaction network. In this way, distinct complexes can be dissected and shared or specific subunits among these complexes can be identified.

FIGURE 4
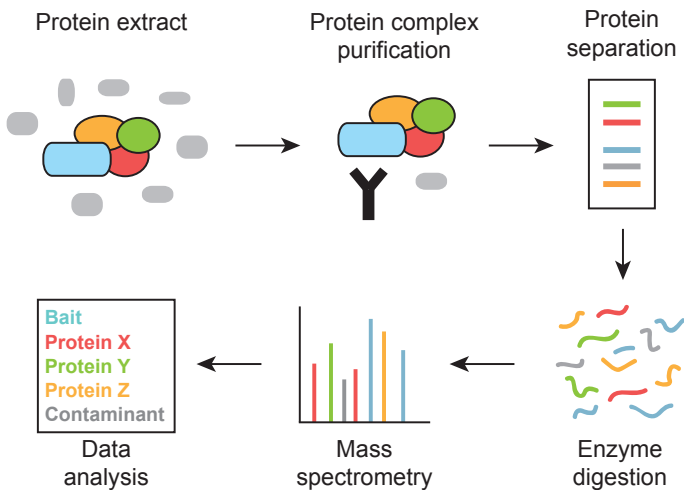
Discovery of protein interactors by complex purification and mass spectrometry.
The main steps of biochemical protein-complex purifications and subsequent identification of the purified proteins by mass spectrometry are illustrated. First, a protein extract is made from a suitable cell system and antibodies (or other affinity reagents) are used to enrich for the protein complex of interest (in colours) and wash away un-related proteins (grey circles). The proteins purified are separated by electrophoresis and subjected to enzymatic digestion with trypsin. The peptides obtained are identified by mass spectrometry, and data analysis generates a list of proteins specifically present in the original sample.

Importantly, this approach allows to deduce the function of unknown or predicted proteins, since proteins that interact with one another or are part of the same complex are often involved in the same cellular processes. If an uncharacterised target protein interacts specifically with a protein of known function, such function can be generally extended to the target protein of interest and follow-up experiments can be designed based on this data.

Chromatin immunoprecipitation and high-throughput sequencing

Chromatin immunoprecipitation (ChIP) is a powerful technique to study protein-DNA interactions in the context of chromatin *in vivo*. Genome-wide profiling of protein-DNA interactions (ChIP-seq) together with quantitative measurements of transcriptomes (mRNA-seq) are increasingly used to decipher the regulatory networks that underlie biological processes in health and disease (Park, 2009).

In a ChIP assay, proteins are covalently crosslinked to their binding sites on chromatin by incubation of the cells or tissues with formaldehyde (Figure 5). Crosslinked chromatin is then extracted from the cell nuclei and fragmented by sonication or enzymatic digestion into pieces of a few hundred base pairs. The fragmented chromatin is used as input in an immunoprecipitation reaction using antibodies against a target protein, i.e. a transcription factor, histone, or polymerase. In this way, the DNA associated to the protein of interest is enriched and purified. The presence of genomic DNA regions of interest (binding sites) can be determined by real-time quantitative PCR assays using specific primers. However, this approach is inherently limited to a priori knowledge or hypotheses on putative binding sites for the protein studied.

The introduction of microarrays (ChIP-chip) and more recently next-generation sequencing approaches (ChIP-seq) has enabled the discovery of unknown binding sites and the rapid and efficient mapping of protein-DNA interactions at the genomic scale in the same experiment. For ChIP-seq experiments on the Illumina sequencing platform, the DNA obtained from a ChIP experiment is subject to an adaptor ligation step, which adds specific sequences to both ends of the DNA molecules. After a PCR amplification step, the "sequencing library" is obtained. The library is hybridized on a glass surface via sequences present in the adaptors, and each molecule is subsequently amplified by isothermal bridge amplification, generating clusters of about 1000 identical molecules. The clustered flowcell is then
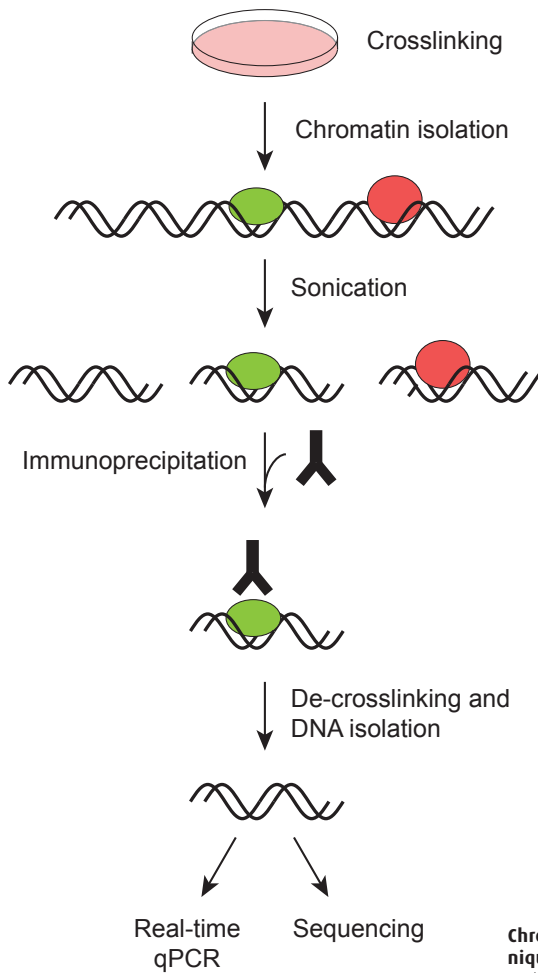
Chromatin immunoprecipitation technique (ChIP). The different steps involved in the ChIP technique are illustrated (see main text for a detailed description).

FIGURE 5

installed on the sequencer (Genome Analyzer or HiSeq) and subject to a sequence-by-synthesis protocol, generating millions of sequence reads in one experiment within a few days.

The ChIP-seq technique is characterized by a number of key features: first of all, ChIP-seq has single-nucleotide resolution, which allows the precise determination of a binding site and a more accurate discovery of relevant sequence motifs. Very low amounts of ChIP DNA (10-50 ng) and a limited number of PCR amplification cycles are required. Although the PCR can introduce some GC-bias, the background levels of a ChIP-seq experiment are significantly lower than ChIP-chip, and the dynamic range is theoretically unlimited. Taken together, these factors result in a very high sensitivity for this technique, which allows the identification of binding sites with moder-

ate enrichment that would not be detected with an array approach. This is not only relevant for the study of low-affinity binding sites, but also for the profiling of proteins that do not bind DNA directly but rather via protein-protein interactions and of proteins that establish only weak interactions with chromatin. The genomic coverage in ChIP-Seq is only limited by alignability of the sequence reads to the reference genome and not by probe design and cross-hybridization issues. If needed, the specificity of the alignments can be further increased by generating longer or paired-end reads. Moreover, ChIP-seq data can be reanalysed at a later stage if a new genome build becomes available. Finally, the steady drop in the costs of Next-Generation sequencing together with recent protocols allowing barcoding and multiplexing of samples have made ChIP-seq experiments more cost-effective, especially for higher eukaryotes and large genomes.

A critical aspect of ChIP-seq experiments is that a suitable ChIP-grade antibody against the protein of interest is needed. Nowadays, thousands of antibodies are available from many different commercial suppliers, but only a limited number of them works in native immunoprecipitation assays, and even fewer work in ChIP. Therefore, antibodies must be thoroughly tested before use in order to assess their specificity (cross-reaction with other proteins containing a similar epitope is a common problem) and sensitivity (signal vs. background in ChIP experiment). Therefore, proper negative controls are crucial in order to draw valid conclusions. Several different types of negative controls are being used in the literature, but not all of them have the same power. Some researchers compare the ChIP-qPCR recoveries of the antibody of interest with an unrelated antibody on the same genomic locus. While this type of negative control might establish a background level related to the primers used in qPCR, it does not address the specificity (i.e. cross-reactivity) of the antibody of interest or the background levels related to the antibody itself. Primers designed on genomic loci that are not bound by the protein of interest can better indicate background levels related to the antibody. Other useful controls to test the specificity of the antibody are experiments performed on cells that lack the protein of interest, for instance upon knockout or knockdown. Similarly, ChIP assays can be performed on epitope-tagged proteins in a cell line expressing the tagged protein while the wild type cell line is used as negative control. Alternatively, serum or antibodies harvested before the immunization protocol against the protein of interest (preimmune serum) can establish if the signal detected in ChIP is an aspecific signal relat-

ed to the animal used for the immunization, or rather a specific signal caused by antibodies raised in the immunization. A positive result can arguably represent a signal related to the protein of interest, but cross reactivity with proteins containing the same epitope cannot be formally excluded.

Next-generation sequencing approaches have been successfully applied also to genome-wide expression profiling (mRNA-seq) (Pepke *et al*, 2009). Also in this case, the larger dynamic range, higher sensitivity and the digital nature of the mRNA-seq data have constituted distinct advantages compared to array-based approaches. The increasing use and applications of these sequencing technologies and the large-scale data collection should form the basis for truly integrative systems biology analyses. Developments of the computation aspects will allow researchers to combine and correlate hundreds or thousands of protein binding profiles, transcriptomes, and protein-protein interaction datasets, with the final goal to decipher molecular regulatory pathways and eventually address important biological questions.

REFERENCES

Abe Y, Matsumoto S, Wei S, Nezu K, Miyoshi A, Kito K, Ueda N, Shigemoto K, Hitsumoto Y, Nikawa J & Enomoto Y (2001) Cloning and characterization of a p53-related protein kinase expressed in interleukin-2-activated cytotoxic T-cells, epithelial tumor cell lines, and the testes. *J Biol Chem* **276**: 44003–44011

Alpen B, Güre AO, Scanlan MJ, Old LJ & Chen Y-T (2002) A new member of the NY-ESO-1 gene family is ubiquitously expressed in somatic tissues and evolutionarily conserved. *Gene* **297**: 141–149

Birtle Z, Goodstadt L & Ponting C (2005) Duplication and positive selection among hominin-specific PRAME genes. *BMC Genomics* **6**: 120

Bortvin A, Eggan K, Skaletsky H, Akutsu H, Berry DL, Yanagimachi R, Page DC & Jaenisch R (2003) Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. *Development* **130**: 1673–1680

Budhidarmo R, Nakatani Y & Day CL (2012) RINGs hold the key to ubiquitin transfer. *Trends Biochem Sci* **37**: 58–65

Cinelli P, Casanova EA, Uhlig S, Lochmatter P, Matsuda T, Yokota T, Rülicke T, Ledermann B & Bürki K (2008) Expression profiling in transgenic FVB/N embryonic stem cells overexpressing STAT3. B*MC Developmental Biology 2008 8:57* **8**: 57

Dadé S, Callebaut I, Mermillod P & Monget P (2003) Identification of a new expanding family of genes characterized by atypical LRR domains. Localization of a cluster preferentially expressed in oocyte. *FEBS Lett* **555**: 533–538

Daugeron M-C, Lenstra TL, Frizzarin M, Yacoubi El B, Liu X, Baudin-Baillieu A, Lijnzaad P, Decourty L, Saveanu C, Jacquier A, Holstege FCP, de Crécy-Lagard V, van Tilbeurgh H & Libri D (2011) Gcn4 misregulation reveals a direct role for the evolutionary conserved EKC/KEOPS in the t6A modification of tRNAs. *Nucleic Acids Res* **39**: 6148–6160

De Carvalho DD, Binato R, Pereira WO, Leroy JMG, Colassanti MD, Proto-Siqueira R, Bueno-Da-Silva AEB, Zago MA, Zanichelli MA, Abdelhay E, Castro FA, Jacysyn JF & Amarante-Mendes GP (2011) BCR-ABL-mediated upregulation of PRAME is responsible for knocking down TRAIL in CML patients. *Oncogene* **30**: 223–233

Deshaies RJ & Joazeiro CAP (2009) RING domain E3 ubiquitin ligases. Annu. Rev. *Biochem*. **78**: 399–434

Doolan P, Clynes M, Kennedy S, Mehta JP, Crown J & O'Driscoll L (2008) Prevalence and prognostic and predictive relevance of PRAME in breast cancer. *Breast Cancer Res Treat* **109**: 359–365

Downey M, Houlsworth R, Maringele L, Rollie A, Brehme M, Galicia S, Guillard S, Partington M, Zubko MK, Krogan NJ, Emili A, Greenblatt JF, Harrington L, Lydall D & Durocher D (2006) A genome-wide screen identifies the evolutionarily conserved KEOPS complex as a telomere regulator. *Cell* **124**: 1155–1168

Epping MT & Bernards R (2006) A causal role for the human tumor antigen preferentially expressed antigen of melanoma in cancer. *Cancer Res* **66**: 10639–10642

Epping MT, Hart AAM, Glas AM, Krijgsman O & Bernards R (2008) PRAME expression and clinical outcome of breast cancer. *Br J Cancer* **99**: 398–403

Epping MT, Wang L, Edel MJ, Carlée L, Hernandez M & Bernards R (2005) The human tumor antigen PRAME is a dominant repressor of retinoic acid receptor signaling. *Cell* **122**: 835–847

Epping MT, Wang L, Plumb JA, Lieb M, Gronemeyer H, Brown R & Bernards R (2007) A functional genetic screen identifies retinoic acid signaling as a target of histone deacetylase inhibitors. *Proc Natl Acad Sci USA* **104**: 17777–17782

Ewing RM, Chu P, Elisma F, Li H, Taylor P, Climie S, McBroom-Cerajewski L, Robinson

MD, O'Connor L, Li M, Taylor R, Dharsee M, Ho Y, Heilbut A, Moore L, Zhang S, Ornatsky O, Bukhman YV, Ethier M, Sheng Y, *et al* (2007) Large-scale mapping of human protein-protein interactions by mass spectrometry. *Mol Syst Biol* **3**: 89 Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=17353931&retmode=ref&cmd=prlinks

Facchin S, Lopreiato R, Ruzzene M, Marin O, Sartori G, Götz C, Montenarh M, Carignani G & Pinna LA (2003) Functional homology between yeast piD261/Bud32 and human PRPK: both phosphorylate p53 and PRPK partially complements piD261/Bud32 deficiency. *FEBS Lett* **549**: 63–66

Facchin S, Ruzzene M, Peggion C, Sartori G, Carignani G, Marin O, Brustolon F, Lopreiato R & Pinna LA (2007) Phosphorylation and activation of the atypical kinase p53-related protein kinase (PRPK) by Akt/PKB. *Cell Mol Life Sci* **64**: 2680–2689

Fischer M, Oberthuer A, Brors B, Kahlert Y, Skowron M, Voth H, Warnat P, Ernestus K, Hero B & Berthold F (2006) Differential expression of neuronal genes defines subtypes of disseminated neuroblastoma with favorable and unfavorable outcome. *Clin Cancer Res* **12**: 5118–5128

Galperin MY (2008) Social bacteria and asocial eukaryotes. *Environ Microbiol* **10**: 281–288

Gavin A-C, Bösche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon A-M, Cruciat C-M, Remor M, Höfert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, *et al* (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**: 141–147

Gingras A-C, Gstaiger M, Raught B & Aebersold R (2007) Analysis of protein complexes using mass spectrometry. *Nat Rev Mol Cell Biol* **8**: 645–654

Goellner S, Steinbach D, Schenk T, Gruhn B, Zintl F, Ramsay E & Saluz HP (2006) Childhood acute myelogenous leukaemia: association between PRAME, apoptosis- and MDR-related gene expression. *Eur J Cancer* **42**: 2807–2814

Haqq C, Nosrati M, Sudilovsky D, Crothers J, Khodabakhsh D, Pulliam BL, Federman S, Miller JR, Allen RE, Singer MI, Leong SPL, Ljung B-M, Sagebiel RW & Kashani-Sabet M (2005) The gene expression signatures of melanoma progression. *Proc Natl Acad Sci USA* **102**: 6092–6097

Hecker A, Lopreiato R, Graille M, Collinet B, Forterre P, Libri D & van Tilbeurgh H (2008) Structure of the archaeal Kae1/Bud32 fusion protein MJ1130: a model for the eukaryotic EKC/KEOPS subcomplex. *EMBO J*

Ikeda F, Crosetto N & Dikic I (2010) What determines the specificity and outcomes of ubiquitin signaling? *Cell* **143**: 677–681

Ikeda H, Lethé B, Lehmann F, van Baren N, Baurain JF, de Smet C, Chambost H, Vitale M, Moretta A, Boon T & Coulie PG (1997) Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory receptor. *Immunity* **6**: 199–208

Joazeiro CAP, Anderson KC & Hunter T (2006) Proteasome inhibitor drugs on the rise. In pp 7840–7842.

Kaji K, Caballero IM, MacLeod R, Nichols J, Wilson VA & Hendrich B (2006) The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. *Nat Cell Biol* **8**: 285–292

Kamura T, Maenaka K, Kotoshiba S, Matsumoto M, Kohda D, Conaway RC, Conaway JW & Nakayama KI (2004) VHL-box and SOCS-box domains determine binding specificity for Cul2-Rbx1 and Cul5-Rbx2 modules of ubiquitin ligases. *Genes Dev* **18**: 3055–3065

Kilpinen S, Autio R, Ojala K, Iljin K, Bucher E, Sara H, Pisto T, Saarela M, Skotheim RI, Björkman M, Mpindi J-P, Haapa-Paananen S, Vainio P, Edgren H, Wolf M, Astola J, Nees M, Hautaniemi S & Kallioniemi O (2008) Systematic bioinformatic analysis of expression levels of 17,330 human genes across 9,783 samples from 175 types of healthy and pathological tissues. *Genome Biol* **9**: R139

Kim H-J & Bae S-C (2011) Histone deacetylase inhibitors: molecular mechanisms of action and clinical trials as anti-cancer drugs. *Am J Transl Res* **3**: 166–179

Kisseleva-Romanova E, Lopreiato R, Baudin-Baillieu A, Rousselle J-C, Ilan L, Hofmann K, Namane A, Mann C & Libri D (2006) Yeast homolog of a cancer-testis antigen defines a new transcription complex. *EMBO J* **25**: 3576–3585

Le Guezennec X, Vermeulen M, Brinkman AB, Hoeijmakers WAM, Cohen A, Lasonder E & Stunnenberg HG (2006) MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. *Mol Cell Biol* **26**: 843–851

Mahrour N, Redwine WB, Florens L, Swanson SK, Martin-Brown S, Bradford WD, Staehling-Hampton K, Washburn MP, Conaway RC & Conaway JW (2008) Characterization of Cullin-box sequences that direct recruitment of Cul2-Rbx1 and Cul5-Rbx2 modules to Elongin BC-based ubiquitin ligases. *J Biol Chem* **283**: 8005–8013

Mao DYL, Neculai D, Downey M, Orlicky S, Haffani YZ, Ceccarelli DF, Ho JSL, Szilard RK, Zhang W, Ho CS, Wan L, Fares C, Rumpel S, Kurinov I, Arrowsmith CH, Durocher D & Sicheri F (2008) Atomic structure of the KEOPS complex: an ancient protein kinase-containing molecular machine. *Mol Cell* **32**: 259–275

Marín I (2009) Diversification of the cullin family. *BMC Evolutionary Biology 2009 9:267* **9**: 267

McDermott U, Downing JR & Stratton MR (2011) Genomics and the continuum of cancer care. *N. Engl. J. Med.* **364**: 340–350

Mering von C, Krause R, Snel B, Cornell M, Oliver SG, Fields S & Bork P (2002) Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* **417**: 399–403

Metzger MB, Hristova VA & Weissman AM (2012) HECT and RING finger families of E3 ubiquitin ligases at a glance. *Journal of Cell Science* **125**: 531–537

Minami N, Aizawa A, Ihara R, Miyamoto M, Ohashi A & Imai H (2003) Oogenesin is a novel mouse protein expressed in oocytes and early cleavage-stage embryos. *Biol Reprod* **69**: 1736–1742

Miyoshi A, Kito K, Aramoto T, Abe Y, Kobayashi N & Ueda N (2003) Identification of CGI-121, a novel PRPK (p53-related protein kinase)-binding protein. *Biochemical and Biophysical Research Communications* **303**: 399–405

Oberthuer A, Hero B, Spitz R & Berthold F (2004) The tumor-associated antigen PRAME is universally expressed in high-stage neuroblastoma and associated with poor outcome. *Clinical Cancer Research*

Oehler VG, Guthrie KA, Cummings CL, Sabo K, Wood BL, Gooley T, Yang T, Epping MT, Shou Y, Pogosova-Agadjanyan E, Ladne P, Stirewalt DL, Abkowitz JL & Radich JP (2009) The preferentially expressed antigen in melanoma (PRAME) inhibits myeloid differentiation in normal hematopoietic and leukemic progenitor cells. *Blood* **114**: 3299–3308

Okumura F, Matsuzaki M, Nakatsukasa K & Kamura T (2012) The Role of Elongin BC-Containing Ubiquitin Ligases. *Front Oncol* **2**: 10

Park PJ (2009) ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet* **10**: 669–680

Partheen K, Levan K, Osterberg L, Claesson I, Fallenius G, Sundfeldt K & Horvath G (2008) Four potential biomarkers as prognostic factors in stage III serous ovarian adenocarcinomas. *Int J Cancer* **123**: 2130–2137

Passeron T, Valencia JC, Namiki T, Vieira WD, Passeron H, Miyamura Y & Hearing VJ (2009) Upregulation of SOX9 inhibits the growth of human and mouse melanomas and restores their sensitivity to retinoic acid. *J. Clin. Invest.* **119**: 954–963

Pepke S, Wold B & Mortazavi A (2009) Computation for ChIP-seq and RNA-seq studies. *Nat Meth* **6**: S22–32

Petroski MD & Deshaies RJ (2005) Function and regulation of cullin-RING ubiquitin ligases. *Nat Rev Mol Cell Biol* **6**: 9–20 Available at: http://www.nature.com.proxy.ubn.ru.nl:8080/nrm/journal/v6/n1/extref/nrm1547-s1.pdf

Radich JP, Dai H, Mao M, Oehler V, Schelter J, Druker B, Sawyers C, Shah N, Stock W, Willman CL, Friend S & Linsley PS (2006) Gene expression changes associated with progression and response in chronic myeloid leukemia. *Proc Natl Acad Sci USA* **103**: 2794–2799

Reinders J, Zahedi RP, Pfanner N, Meisinger C & Sickmann A (2006) Toward the complete yeast mitochondrial proteome: multidimensional separation techniques for mitochondrial proteomics. *J Proteome Res* **5**: 1543–1554

Rhesus Macaque Genome Sequencing and Analysis Consortium, Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER, Remington KA, Strausberg RL, Venter JC, Wilson RK, Batzer MA, Bustamante CD, Eichler EE, Hahn MW, Hardison RC, Makova KD, Miller W, Milosavljevic A, Palermo RE, *et al* (2007) Evolutionary and Biomedical Insights from the Rhesus Macaque Genome. *Science* **316**: 222–234

Roman-Gomez J, Jimenez-Velasco A, Agirre X, Castillejo JA, Navarro G, Jose-Eneriz ES, Garate L, Cordeu L, Cervantes F, Prosper F, Heiniger A & Torres A (2007) Epigenetic regulation of PRAME gene in chronic myeloid leukemia. *Leukemia Research* **31**: 1521–1528

Sakurai E, Maesawa C, Shibazaki M, Yasuhira S, Oikawa H, Sato M, Tsunoda K, Ishikawa Y, Watanabe A, Takahashi K, Akasaka T & Masuda T (2011) Downregulation of microRNA-211 is involved in expression of preferentially expressed antigen of melanoma in melanoma cells. *Int J Oncol* **39**: 665–672

Santamaría C, Chillón MC, García-Sanz R, Balanzategui A, Sarasquete ME, Alcoceba M, Ramos F, Bernal T, Queizán JA, Peñarrubia MJ, Giraldo P, San Miguel JF & Gonzalez M (2008) The relevance of preferentially expressed antigen of melanoma (PRAME) as a marker of disease activity and prognosis in acute promyelocytic leukemia. *Haematologica* **93**: 1797–1805

Sarikas A, Hartmann T & Pan Z-Q (2011) The cullin protein family. Genome Biol 12: 220

Schenk T, Stengel S, Goellner S, Steinbach D & Saluz HP (2007) Hypomethylation of PRAME is responsible for its aberrant overexpression in human malignancies. *Genes Chromosomes Cancer* **46**: 796–804

Srinivasan M, Mehta P, Yu Y, Prugar E, Koonin EV, Karzai AW & Sternglanz R (2011) The highly conserved KEOPS/EKC complex is essential for a universal tRNA modification, t6A. *EMBO J* **30**: 873–881

Steinbach D, Hermann J, Viehmann S, Zintl F & Gruhn B (2002) Clinical implications of PRAME gene expression in childhood acute myeloid leukemia. *Cancer Genet Cytogenet* **133**: 118–123

Steinbach D, Pfaffendorf N, Wittig S & Gruhn B (2007) PRAME expression is not associated with down-regulation of retinoic acid signaling in primary acute myeloid leukemia. *Cancer Genet Cytogenet* **177**: 51–54

Tajeddine N, Gala J-L, Louis M, Van Schoor M, Tombal B & Gailly P (2005) Tumor-associated antigen preferentially expressed antigen of melanoma (PRAME) induces caspase-independent cell death in vitro and reduces tumorigenicity in vivo. *Cancer Res* **65**: 7348–7355

Tanaka N, Wang Y-H, Shiseki M, Takanashi M & Motoji T (2011) Inhibition of PRAME expression causes cell cycle arrest and apoptosis in leukemic cells. *Leuk Res* **35**: 1219–1225

van Baren N, Chambost H, Ferrant A, Michaux L, Ikeda H, Millard I, Olive D, Boon T & Coulie PG (1998) PRAME, a gene encoding an antigen recognized on a human melanoma by cytolytic T cells, is expressed in acute leukaemia cells. *Br J Haematol* **102**: 1376–1379

van t Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AAM, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R & Friend SH (2002) Gene expression profiling predicts clinical outcome of breast

cancer. *Nature* **415**: 530–536

Wadelin F, Fulton J, McEwan PA, Spriggs KA, Emsley J & Heery DM (2010) Leucine-rich repeat protein PRAME: expression, potential functions and clinical implications for leukaemia. *Mol. Cancer* **9**: 226

Wang PJ, McCarrey JR, Yang F & Page DC (2001) An abundance of X-linked genes expressed in spermatogonia. *Nat Genet* **27**: 422–426

Watari K, Tojo A, Nagamura-Inoue T, Nagamura F, Takeshita A, Fukushima T, Motoji T, Tani K & Asano S (2000) Identification of a melanoma antigen, PRAME, as a BCR/ABL-inducible gene. *FEBS Lett* **466**: 367–371

Weissman AM, Shabek N & Ciechanover A (2011) The predator becomes the prey: regulating the ubiquitin system by ubiquitylation and degradation. *Nat Rev Mol Cell Biol* **12**: 605–620

WHO World Health Statistics 2012 World Health Organization Available at: http://www.who.int/healthinfo/EN_WHS2012_Full.pdf

Yacoubi El B, Hatin I, Deutsch C, Kahveci T, Rousset J-P, Iwata-Reuyl D, Murzin AG & de Crécy-Lagard V (2011) A role for the universal Kae1/Qri7/YgjD (COG0533) family in tRNA modification. *EMBO J* **30**: 882–893

Yarian C, Townsend H, Czestkowski W, Sochacka E, Malkiewicz AJ, Guenther R, Miskiewicz A & Agris PF (2002) Accurate translation of the genetic code depends on tRNA modified nucleosides. *J Biol Chem* **277**: 16391–16395

Zimmerman ES, Schulman BA & Zheng N (2010) Structural assembly of cullin-RING ubiquitin ligase complexes. *Current opinion in structural biology* **20**: 714–721

CHAPTER 2

IDENTIFI-
CATION
of protein interactors of
the human oncoprotein
PRAME by epitope-tag
affinity purifications

Adalberto Costessi, Esther Tijchon, Pascal W. Jansen, Hendrik G. Stunnenberg

Department of Molecular Biology, Nijmegen Centre for Molecular Life Sciences, Radboud University, Nijmegen, The Netherlands

ABSTRACT

Preferentially expressed antigen of melanoma (PRAME) is frequently overexpressed in tumors and high *PRAME* levels correlate with poor clinical outcome of several cancers. However, the molecular mechanisms and pathways in which PRAME acts are still unclear. In order to investigate the molecular function(s) of PRAME, we set out to identify its protein interactors. To this end, we developed a biochemical protein-complex purification approach using epitope-tagged versions of PRAME and immunoprecipitation assays coupled to mass spectrometry analysis. In this chapter, we report the setup and optimization of these techniques, which led us to identify specific interactions between PRAME and Cullin2-based E3 ubiquitin ligases.

44

INTRODUCTION

The human protein PRAME (Preferentially expressed antigen in melanoma) is frequently overexpressed in cancer, and high expression levels often correlate with poor clinical outcome (Ikeda *et al*, 1997; Kilpinen *et al*, 2008; Oberthuer *et al*, 2004; Haqq *et al*, 2005). *PRAME* is the founding member of a large family of *PRAME*-like genes, characterized by a recent evolutionary history and a high degree of conservation (Birtle *et al*, 2005; Consortium *et al*, 2007).

Computational analysis of the human genome identified 22 *PRAME*-like genes and 10 pseudogenes (Birtle *et al*, 2005). Analysis of the primary amino acid sequence of PRAME revealed that it is a leucine-rich repeat protein (Epping *et al*, 2005; Birtle *et al*, 2005) and it was suggested that PRAME could fold into a horse-shoe structure (Birtle *et al*, 2005; Wadelin *et al*, 2010). However, defined functional motifs could not be identified and the molecular function(s) of PRAME and PRAME-like proteins are still unclear.

An experimental approach to gain insight into the function of an unknown protein is to identify its protein-protein interactors, since proteins that interact with one another or are part of the same complex are generally involved in the same cellular processes (Phizicky *et al*, 2003). If the interactors identified have a known function, this function can often be extended to the unknown protein and specific follow-up experiments can be designed. If the interactors identified are also uncharacterized proteins, inspection of their amino acid sequences might reveal functional domains or motifs that can eventually give clues to putative functions. A powerful approach to determine protein interactors and entire protein complexes involves the stable expression of epitope-tagged versions of the protein of interest in a suitable cell system using retroviral delivery. Subsequently, the protein and its interacting partners are biochemically purified using antibodies or affinity-based reagents which specifically and efficiently bind to the epitope tags, and the proteins are identified by mass spectrometry and database searches (Aebersold & Mann, 2003; Rigaut *et al*, 1999).

In this chapter, we describe the application of epitope-tagging and mass spectrometry analysis to the unbiased identification of the protein interactors of PRAME. We successfully setup a single-step protein complex purification protocol and discovered that the human oncoprotein PRAME establishes interactions with Cullin2-based E3 ubiquitin ligases.

RESULTS AND DISCUSSION

The ProtA-TEV-3xMyc epitope tag in FM6 cells

Since PRAME is overexpressed in the vast majority of human melanomas, we set out to identify its protein interactors applying a tap-tag (tandem affinity purification) approach in the human melanoma cell line FM6 (Rigaut *et al*, 1999). A modified pZOME retroviral system was used to generate FM6-pZ3XN-PRAME cell lines stably expressing PRAME fused to a combination of N-terminal tags: 2xProtA-2xTEV-3xMYC (ProtA-TEV-3xMYC-PRAME; 72 kDa) (Fig. 1A). A similar tag with a single MYC epitope was successfully used for the characterization of proteins complexes involved in transcriptional regulation (Le Guezennec *et al*, 2006).

The purification strategy includes a first incubation of protein extracts with crosslinked IgG beads for which the protein A domains have a high affinity (Fig. 1B). After incubation, several washes are performed to discard unbound proteins, and recombinant TEV (Tobacco Etch Virus) protease is added to cleave the tags at its recognition site. This proteolytic cleavage releases the bait protein and its interacting partners in solution. The last step is a MYC immunoprecipitation to further purify the complex.

Western blot of protein extracts from FM6-pZ3XN-PRAME cells showed a prominent band at the expected size, and a lower band with a significantly lower intensity (Fig. 1C). These extracts were used to optimize the conditions for TEV cleavage. The amounts of TEV required and the incubation temperature were tested in preliminary experiments (not shown). Western blot analysis of whole cell extracts after incubation with TEV showed that cleavage was most efficient after overnight incubation at 4°C (Fig. 1D, lanes 7 and 8).

Next, a small-scale purification experiment was performed to test the efficiency of IgG binding and TEV cleavage. Figure 1E shows that after incubation of protein extract with IgG beads, about 50% of ProtA-TEV-3xMYC-PRAME was bound to IgG beads (lane 4), while about 50% was still present in the supernatant (lane 2). When TEV protease was added to the beads and incubated overnight, however, almost no cleavage was detected in the eluates (lanes 6 and 9) and the intact protein was still bound to the beads. Independent experiments confirmed that TEV could efficiently cleave its recognition site in ProtA-TEV-3xMYC-PRAME from soluble protein extracts, but not when the tagged protein was bound to IgG beads.

A possible explanation for the inefficient cleavage is that the 27 kDa TEV protease might not be able to efficiently access its target sites when ProtA-TEV-3xMYC-PRAME is immobilized on IgG beads. In conclusion, although this epitope tag had

Figure 1

47
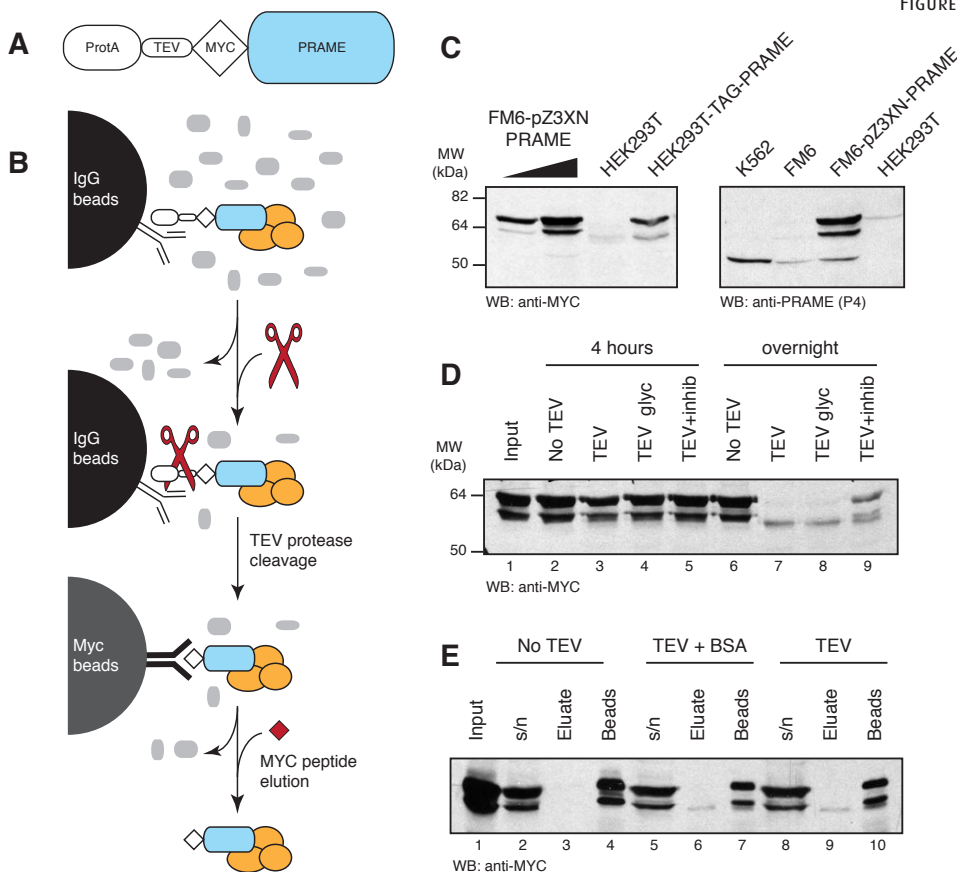
TAP approach in FM6-pZ3XN-PRAME cells.

(A) Schematic representation of ProtA-TEV-3xMYC-PRAME.
(B) Overview of the tandem affinity purification (TAP) procedure. The target protein is light blue, its protein interactors orange. The first purification is based on binding of the proteinA domains of the tagged protein to IgG beads. Most unbound proteins (in grey) are washed off and the target protein complex is released by TEV protease cleavage. The complex is further purified using a second purification using beads with Myc antibodies and elution with molar excess of MYC peptide (red square).
(C) Protein levels of PRAME in cell lines. Western blot analysis of endogenous and epitope-tagged PRAME. Left panel: FM6-pZ3XN-PRAME stable cell lines and HEK293T cells transiently transfected with pcDNA5-StrepII-2xTEV-MYC-3xHA-PRAME (TAG-PRAME,

69 kDa) are analyzed with MYC antibody. Right panel: whole cell extracts from the cell lines indicated are probed with anti-PRAME affinity-purifed antibody P4 (1:1000 dilution). Equal amounts of extracts were loaded for all samples, and this was checked by Ponceaus S staining of the membranes.
(D) TEV cleavage on protein extracts. Whole cell extracts from FM6-pZ3XN-PRAME cells were subjected to TEV cleavage at 4°C for 4 hours or overnight. For each reaction, 18 µl of whole cell extracts were mixed with 59 µl IPP150/0,1 buffer and 1 µl of TEV protease (preparation batch 14-2), or 2 µl of the same protease preparation diluted 1:1 in glycerol (TEV glyc), or 1 µl buffer (no TEV control). The effect of protease inhibitors was tested by incubating extracts with 1 µl TEV and 1x protease inhibitors cocktail. After incubation, 14 µl laemli buffer 4x was added and the sample boiled to stop the reaction; 25 µl of all reactions were

analyzed by western blot with MYC antibody.
(E) TEV cleavage of epitope tagged PRAME bound on IgG beads is not efficient. IgG beads were incubated for 1h with PBS (TEV) or PBS with 1% BSA (No TEV, and TEV+BSA). Beads were subsequently washed with IPP150/0,1 buffer without protein inhibitors; 15 µl of beads were mixed with 100 µl IPP150/0,1 and 50 µl of FM6-pZ3XN-PRAME whole cell extract, and rocked for 2h at 4 degrees. The supernatant was collected, and the beads washed with 500 µl of IPP150/0,1. Elution was performed by adding 100 µl of IPP150/0,1 and 2 µl of TEV protease with rocking at 4 degrees overnight. The supernatant (TEV eluate) was collected, and beads washed and boiled in 100 µl laemli buffer. 20% fractions of all samples were analyzed by western blot with MYC antibody.

Figure 2

A

LZRS-StrepII-TEV-MYC-3xHA-PRAME in suspension cells.

(A) Map of the LZRS-ZeoR retroviral plasmid. The multi cloning site (MCS) is indicated in the square. The sequence coding for TAG-PRAME was cloned as a BamHI-XhoI fragment.

(B) Nucleotide and amino acid sequences of the StrepII-2xTEV-MYC-3xHA tag.

(C) Western blot analysis of single cell clones of HeLaS3-TAG-PRAME and K562-TAG-PRAME cells with HA.Y11 antibody against the HA TAG. Arrows indicate the two main bands for TAG-PRAME; asterisks indicate background bands. Clone numbers are indicated; WT, wild type cells (negative control); poly, polyclonal culture; ctrl, HEK293T cells transiently transfected with pcDNA5-TAG-PRAME. Equal amounts of whole cell extracts were loaded and checked by Ponceau S staining of the membrane (not shown).

(D) Gene expression analysis of *PRAME* by real-time RT-PCR. Left panel, cDNA prepared from the cell lines indicated was analyzed with two different primers for *PRAME* and the results normalized against *GAPDH*. Right panel, independent experiment with normalization against a pool of reference genes (*GAPDH, ACTB, GUSB*). HeLaS3 cells growing in suspension or adhesion (plate) were tested.

48

B

C

D

been successfully applied to the purification of protein complexes (Le Guezennec *et al*, 2006), it could not be applied to the analysis of PRAME protein interactors.

The StrepII-2xTEV-MYC-3xHA epitope tag in K562 cells

We next tested a different N-terminal combinatorial tag: StrepII-2xTEV-MYC-3xHA (further referred to as TAG) that could allow the use of several combinations of two-step purification strategies (Le Guezennec *et al*, 2006). Stable cell lines expressing StrepII-2xTEV-MYC-3xHA-PRAME (TAG-PRAME, 69 kDa) were generated using the retroviral LZRS-ZeoR vector in the human leukaemia K562 cells, which express PRAME at high levels, and the cervical cancer HeLaS3 cells, which express PRAME at very low to undetectable levels (Fig. 2A and B).

Single cell colonies were selected and tested for expression: K562-TAG-PRAME clone 12 and HeLaS3-TAG-PRAME clone 3 were chosen for further experiments (Fig. 2C). Gene expression analysis by real-time RT-PCR indicated that *TAG-PRAME* was expressed at the same levels as endogenous *PRAME* in K562-TAG-PRAME clone 12 cells, and about 3-fold higher in HeLaS3-TAG-PRAME clone 3 cells (Fig. 2D).

Preliminary results indicated that large amounts of protein extracts would be needed for purifications of sufficient quantities of PRAME and associated proteins. We therefore set up protocols for growth and harvest of suspension cells from multiple 5-liter spinner flasks in parallel. For this purpose, we optimized a nuclear/cytoplasmic protein extraction protocol for both K562 and HeLaS3 cells and we assessed the quality of all extractions by western blot analyses for the nuclear protein HDAC2 and the cytoplasmic marker α-tubulin. Representative experiments are presented in Figure 3A. TAG-PRAME was detected in both the cytoplasmic and nuclear fractions, similarly to the endogenous PRAME protein (not shown).

Notably, cytoplasmic extracts yielded on average between 3 and 4-fold more proteins than nuclear extracts (in mass). Therefore, for quality control of the protein extractions, equivalent fractions in volume - and not the same amount of micro-grams - of nuclear and cytoplasmic extracts were analyzed by western blot. In this way, it is possible to directly compare the two fractions and assess the potential contamination of the cytoplasmic extract by nuclear proteins and vice versa.

Two-step TAG-PRAME purifications

At the time of our study, it was generally assumed that a single immuno-precipitation (IP) step does not sufficiently enrich for the bait, because too many aspecific interactors are still present in the eluates. This would result in a mixture

Tandem affinity purification approach from K562-TAG-PRAME cells.

(A) Western blot analysis of nuclear and cytoplasmic protein extracts from K562 cell lines.

(B-D) Western blot analyses of pilot experiments for TAG-PRAME complex purification with a two-step IP approach; 3% of each sample were analyzed, unless indicated otherwise. All membranes were probed with mouse HA.Y11 antibody.

FIGURE 3

of proteins with a too high complexity to allow the successful identification of bait and interacting proteins by mass spectrometry. Therefore, most studies used a tandem affinity purification protocol in order to increase the purity of the protein complex of interest (Rigaut *et al*, 1999). In line with these studies, we performed several small-scale pilot experiments with tandem IPs, each one with two rounds of peptide elution. The two eluates from the first IP were pooled and used as input for a second IP with a different antibody. We tested three antibodies in different combinations: mouse MYC produced in house from hybridoma cultures, rabbit HA

from Abcam (ab9110), and the PRAME affinity-purified antibody P4. Representative experiments are shown in Figure 3B-D: MYC-HA, HA-MYC, and P4-MYC purifications.

In order to assess the efficiency of the protocol, equivalent fractions for all steps were analysed by western blot: input (nuclear extract), supernatant after incubation of the extract with beads-Ab, first and second peptide eluates, and beads after elution. Comparison of the protein levels in input and supernatant gives information on the efficiency of the binding of the tagged protein to the beads, while comparison of eluates and beads gives information on the efficiency of elution.

Lanes 1 to 5 (Fig. 3B-D) show that TAG-PRAME could be efficiently bound and peptide-eluted with any of the three antibodies used in the first IP. Notably, later experiments revealed that peptide elution from the HA-ab9110 antibodies was only possible when using a specific lot of the antibody, and later lots did not allow peptide elution any more.

Although the first IP steps were efficient for all three antibodies, the final yields after the second purifications were considered too low to proceed with a large-scale purification (compare final eluates in lanes 7 with input levels in lanes 1). With this low efficiency, too large amounts of protein extracts and antibodies would be required to obtain enough purified proteins for detection by mass spectrometry.

Single-step TAG-PRAME purification and mass spectrometry

Based on the previous results, we set out to experimentally test whether a single IP purification approach could be applied to the identification of PRAME protein partners. In order to reduce the background levels, we used relatively stringent buffers to wash the antibody-beads after incubation with the protein extracts and before peptide elution. Figure 4A shows that increasing the salt concentration of the wash buffer from 150 to 300 mM reduced very significantly the number of background interactions detected by SDS-PAGE and silver stain of peptide eluates from a MYC IP on either K562-TAG-PRAME or K562 wild type extracts. A salt concentration of 500 mM NaCl did not improve the results further, and we therefore chose the 300 mM concentration for follow up experiments. Notably, we did not identify clear bands specific for the TAG-PRAME eluates, consistent with the notion that TAG-PRAME is present at very low abundance in such eluates as compared to background proteins.

Next, we performed a large-scale immunoprecipitation experiment with crosslinked MYC beads and 10 ml of nuclear extracts from K562-TAG-PRAME cells,
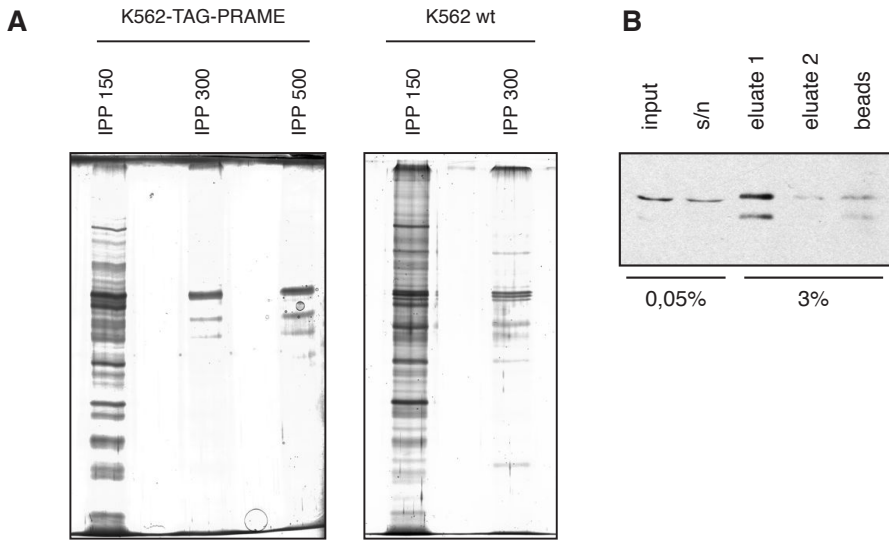
Single-step immunopre-cipitations from K562 cells.

(A) Background levels in single-step IP protocol are decreased by 300 mM NaCl buffer. MYC IPs were performed with 300 µl nuclear extracts from the cell lines indicated and 20 µl MYC-beads. After incubation, the beads were washed with buffers of different salt concentration, as indicated. Peptide eluates were resolved by SDS-PAGE and protein bands visualized by silver staining.
(B) Western blot analysis of large scale MYC IP from K562-TAG-PRAME nuclear extracts. The fraction of samples analyzed is indicated. The mouse HA.Y11 antibody was used.
(C) Silver staining of eluates from large scale MYC IP. Both first and second eluates from large scale IPs on the nuclear extracts indicated were resolved on SDS-PAGE and proteins visualized by silver staining. Molecular weights (kDa) of the marker are indicated on the left.
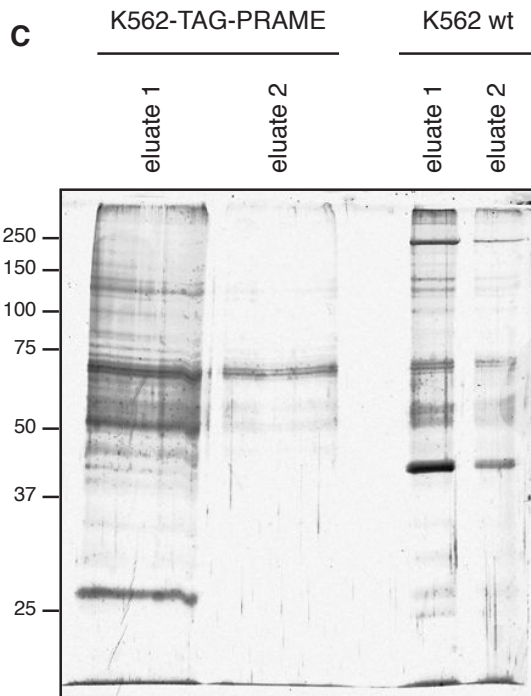
FIGURE 4

corresponding to about 6 litres of culture and $6x10^9$ cells. Immunoprecipitation from 4 ml of nuclear extracts of parental K562 cells (about 2,5 litre culture and $2,5x10^9$ cells) was performed as negative control. Western blotting of an aliquot of the peptide eluates confirmed the efficiency of the IP protocol (Fig. 4B). We resolved the eluates on SDS-PAGE and visualized the proteins by silver staining (Fig. 4C). Several bands were identified, the most intense being present in both the TAG-PRAME and the negative control samples, representing background interactions with beads or antibodies. However, a number of fainter bands were specifically present in the TAG-PRAME eluates and not in the negative control. The entire gel lanes were excised in multiple fractions and processed for mass spectrometry analysis.

The peptides identified were used in a MASCOT database search and the proteins found in the TAG-PRAME and WT eluates were compared to generate a list of proteins specific for the TAG-PRAME sample (Table 1). PRAME was successfully identified with 7 different peptides, establishing that a single IP/MS approach can be successfully applied to the purification of PRAME protein complexes.

Remarkably, peptides were found in the TAG-PRAME eluates that matched all known subunits of Cullin2 E3 ubiquitin ligases (Petroski & Deshaies, 2005): Cullin2 (CUL2), Transcription elongation factor B polypeptide 1 (TCEB1, also called Elongin C), Transcription elongation factor B polypeptide 2 (TCEB2, also called Elongin B), and RING-box protein 1 (RBX1), suggesting a functional link between PRAME and this complex. These interactions were confirmed by western blot of independent immunoprecipitation experiments and were further characterized in other chapters of this thesis.

Taken together, our results show that a specific epitope tag may not always be useful for the purification of a protein of interest. The ProtA-2xTEV-MYC tag had been previously used for the purification of protein complexes (Le Guezennec *et al*, 2006), but in the case of PRAME the TEV cleavage step proved too inefficient. On the contrary, we successfully used the N-terminal StrepII-2xTEV-MYC-3xHA tag and identified PRAME interactors using a single-step purification approach with crosslinked MYC beads and peptide elution. We identified specific interactions between PRAME and all the known subunits of Cullin2-based E3 ubiquitin ligases, suggesting a functional link between ubiquitination pathways and the oncogenic properties of PRAME.

IPI ID	Mass (kDa)	Description	N. peptides
IPI00014311	87,6	Cullin-2	13
IPI00003362	72,5	HSPA5 protein	12
IPI00300341	12,6	Transcription elongation factor B polypeptide 1	7
IPI00019282	58,7	**Melanoma antigen preferentially expressed in tumors**	7
IPI00013890	27,9	Isoform 1 of 14-3-3 protein sigma	6
IPI00021447	58,3	Alpha-amylase 2B precursor	6
IPI00012382	31,1	U1 small nuclear ribonucleoprotein A	5
IPI00026670	13,2	Transcription elongation factor B polypeptide 2	5
IPI00644708	43,8	TIA-1 related protein isoform 2	5
IPI00017963	13,6	Small nuclear ribonucleoprotein Sm D2	5
IPI00013881	49,5	heterogeneous nuclear ribonucleoprotein H1	5
IPI00015809	36,9	Probable O-sialoglycoprotein endopeptidase	4
IPI00294997	42,9	Nucleoside diphosphate kinase 7	4
IPI00027285	24,8	Isoform SM-B' of Small nuclear ribonucleoprotein-associated proteins B and B'	4
IPI00419373	39,8	Isoform 1 of Heterogeneous nuclear ribonucleoprotein A3	4
IPI00029266	10,9	Small nuclear ribonucleoprotein E	3
IPI00000874	22,3	Peroxiredoxin-1	3
IPI00005615	41,9	Nucleolysin TIAR	3
IPI00152871	65,3	Leucine-rich repeat-containing protein 15 precursor	3
IPI00219483	50,6	Isoform 2 of U1 small nuclear ribonucleoprotein 70 kDa	3
IPI00033025	51,1	Isoform 1 of Septin-7	3
IPI00179964	57,4	Isoform 1 of Polypyrimidine tract-binding protein 1	3
IPI00216402	15,5	Histone H3.1t	3
IPI00081836	13,8	Histone H2A type 1-H	3
IPI00013885	27,9	Caspase-14 precursor	3
IPI00006662	21,5	Apolipoprotein D precursor	3
IPI00013296	17,6	40S ribosomal protein S18	3
IPI00013396	17,6	U1 small nuclear ribonucleoprotein C	2
IPI00302850	13,3	Small nuclear ribonucleoprotein Sm D1	2
IPI00016572	8,5	Small nuclear ribonucleoprotein G	2
IPI00010740	76,2	Isoform Long of Splicing factor, proline- and glutamine-rich	2
IPI00291398	43,3	Isoform Long of Nucleolysin TIA-1 isoform p40	2
IPI00171611	15,3	Histone H3.2	2
IPI00152785	13,8	Histone H2B type 1-O	2
IPI00217466	22,2	Histone H1.3	2
IPI00003881	46,0	heterogeneous nuclear ribonucleoprotein F	2
IPI00002966	95,1	Heat shock 70 kDa protein 4	2
IPI00020487	14,2	Extracellular glycoprotein lacritin precursor	2
IPI00013031	22,5	Dual specificity protein phosphatase 14	2
IPI00062206	35,5	DDX39 protein	2
IPI00742905	147,9	ATP-dependent RNA helicase A	2
IPI00217030	29,7	40S ribosomal protein S4, X isoform	2
IPI00003949	17,2	Ubiquitin-conjugating enzyme E2 N	1
IPI00003386	12,7	RING-box protein 1	1
IPI00032314	14,9	L antigen family member 3	1

TABLE 1 — **PRAME putative protein interactors from the first large scale MYC IP.** The tables lists all proteins identified with at least 2 peptides and a selection of proteins identified with only 1 peptide.

MATERIALS AND METHODS

Cell culture and plasmids

HEK293T and FM6 cells were cultured in DMEM, K562 and HeLaS3 in RPMI medium (Gibco, Invitrogen) at 37°C in 5% CO_2. Both media were supplemented with 5% Glutamax, 10% fetal bovine serum, 100 U/ml of penicillin, 100 μg/ml of streptomycin (Gibco, Invitrogen). Stable cell lines were generated by retroviral transduction, as described (Costessi *et al*, 2011). Generation of plasmids pcDNA5-TAG-PRAME and LZRS(Zeo)-TAG-PRAME, and preparation of nuclear/cytoplasmic extracts are described elsewhere (Costessi *et al*, 2011).

Antibodies and western blot

Mouse monoclonal MYC (9E11) were produced in-house from hybridoma cultures; the PRAME antibody P4 was generated by peptide immunization of rabbits and affinity-purification. The following commercial antibodies were used: rabbit HA (Abcam ab9110), mouse monoclonal HA.Y11 (Covance MMS-101P; 1:3000 dilution for WB), alpha-tubulin (Santa Cruz B-7, sc-5286; 1:2000 dilution for WB), HDAC2 (Biomol SA-402; 1:3000-1:5000). Proteins were separated on conventional 10% SDS-PAGE gels and western blots were visualized by ECL (GE Healthcare) or Odyssey (LiCor). All antibodies were diluted in PBS/T with 5% skimmed milk for western blot analyses.

Tandem affinity purifications of StrepII-2xTEV-MYC-3xHA-PRAME

Pilot purifications with two tandem IP steps were performed on nuclear extracts of K562-TAG-PRAME cl.12 cells. For the first IP reaction, 300 μl nuclear extract were mixed with 700 μl IPP150/0,5 buffer (300 mM NaCl, 10 mM Tris-HCl pH 8, 0,5 mM EDTA, 0,5% NP-40) supplemented with protease and phosphatase inhibitors (1 mM PMSF, 1 mM Na_3VO_4, 1 mM NaF, 1x Protease Inhibitor Cocktail, Roche), and either 50 μl of PRAME P4 antibody, 9 μl HA antibody (Abcam ab9110) or 100 μl of crosslinked MYC-beads. After either 2 hours or overnight incubation, 100 μl of ProtG beads (Sigma) were added to P4 and HA IPs and incubated for 2 hours. Subsequently, the reactions were spun down and supernatant (containing unbound proteins) collected. The beads were washed 3 times with IPP150/0,5, and subjected to 2 round of

peptide elution with 500 μl elution buffer (IPP150/0,5 with inhibitors and 100 μg/ml peptide) for 30 minutes in a thermoshaker at 28°C and 1000 rpm. After elution, the beads (containing bound proteins that were not eluted) were boiled in 250 μl laemli buffer. The two eluates from the first IP were pooled and used as input to the second IP, using half the amount of antibody and beads: either 50 μl MYC-beads, 4,5 μl HA antibody, or 25 μl P4 antibody. Incubation and elutions were performed as in the first IP step with 100 μl of elution buffer.

Large scale single-step MYC immunoprecipitation

Large scale IP was performed by mixing 10 ml nuclear extracts from K562-TAG-PRAME cells (from about 6 liter cultures) with 50 μl crosslinked MYC-beads, and 4 ml nuclear extracts from K562 wild type cells with 15 μl MYC-beads. After overnight incubation at 4°C in rotation, the beads were spun down and washed 3 times with 5 ml IPP300/0,5, and once with IPP150/0,5. Two consecutive peptide elutions were performed with either 100 μl (TAG-PRAME) or 30 μl (wild type) elution buffer, as described above.

Silver staining

The peptide eluates were resolved by SDS-PAGE and the proteins were fixed by incubating the gel in fixation solution (40:10:50 v/v/v ethanol: acetic acid: water) for 20-30 minutes at room temperature and gentle rocking. The gel was wash twice with washing solution (30% v/v ethanol in water) on a shaking platform for 20 minutes, rinsed twice with water, and incubated with water for 20 minutes on a shaking platform.

The gel was incubated in sensitizing solution (0,02% sodium thiosulfate) for 1-2 minutes, rinsed 3 times with water, and incubated for 30 minutes in chilled 0,1% $AgNO_3$ solution at 4°C. The gel was quickly rinsed twice with water and incubated in developing solution (0,05% formaldehyde in 2% sodium carbonate). When a sufficient degree of staining was obtained the reaction was stopped with 5% acetic acid. The gel was washed several times in 1% acetic acid and stored in water.

Preparation of samples for mass spectrometry

The lanes from the SDS-PAGE gel corresponding to the first eluates were excised with a clean scalpel, chopped into small pieces and transferred into 1,5 ml tubes. The gel particles were equilibrated for 5 minutes with 100 µl ABC (50 mM ammonium bicarbonate) and centrifuged to discard the liquid. Then, 100 µl acetronitrile was added and incubated for 5 minutes to shrink the gel particles. The ABC and acetonitrile steps were repeated twice, or until the particles turned white. The gel pieces were swelled in 100 µl reduction buffer (10mM dithiothreitol in 50mM NH_4HCO_3) and incubated for 20 minutes at 56°C to reduce the protein. All liquid was discarded and the gel pieces were shrunk in 100 µl acetonitrile. The acetonitrile was replaced by 100 µl alkylation buffer (55 mM iodoacetamide in 50 mM NH_4HCO_3) and incubated for 20 minutes at room temperature in the dark. The gel particles were washed with 100 µl ABC (50mM Ammonium Bi Carbonate) for 5 minutes and ABC was replaced by 100µl acetonitrile; these steps were repeated twice. Gel pieces were rehydrated at room temperature in digestion buffer (50 mM NH_4HCO_3) containing 12,5 ng/µl trypsin (Promega). After 30 minutes, 10-50 µl ABC was added to cover the gel pieces and samples were incubated overnight at 37°C. Peptides were extracted from the gel matrix by adding 100 µl of 2% TFA (trifluoroacetic acid) for 10 minutes at room temperature; gel particles were spun down and the supernatant recovered. The peptide extraction was repeated a second time. Then, 100 µl acetonitrile was added for 5-10 minutes at room temperature to shrink the gel pieces, 100 µl 2% TFA was added for 5-10 minutes and supernatant was recovered after centrifugation. All samples were concentrated in a vacuum centrifuge (Savant) to a maximum of 200 µl, which was transferred to a stagetip. The peptides were eluted in 2-4 µl buffer B (80% acetonitrile, 1% TFA) and diluted 10 times in buffer A (3% acetonitrile, 0,5% acetic acid, 1% TFA). Mass spectrometry analysis was performed as previously described (Olsen *et al*, 2004).

ACKNOWLEDGEMENTS

REFERENCES

Aebersold R & Mann M (2003) Mass spectrometry-based proteomics. *Nature* **422**: 198–207

Birtle Z, Goodstadt L & Ponting C (2005) Duplication and positive selection among hominin-specific PRAME genes. *BMC Genomics* **6**: 120

Consortium RMGSAA, Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER, Remington KA, Strausberg RL, Venter JC, Wilson RK, Batzer MA, Bustamante CD, Eichler EE, Hahn MW, Hardison RC, Makova KD, Miller W, Milosavljevic A, Palermo RE, *et al* (2007) Suppl: Evolutionary and biomedical insights from the rhesus macaque genome. *Science* **316**: 222–234

Costessi A, Mahrour N, Tijchon E, Stunnenberg R, Stoel MA, Jansen PW, Sela D, Martin-Brown S, Washburn MP, Florens L, Conaway JW, Conaway RC & Stunnenberg HG (2011) The tumour antigen PRAME is a subunit of a Cul2 ubiquitin ligase and associates with active NFY promoters. *EMBO J* **30**: 3786–3798

Epping MT, Wang L, Edel MJ, Carlée L, Hernandez M & Bernards R (2005) The human tumor antigen PRAME is a dominant repressor of retinoic acid receptor signaling. *Cell* **122**: 835–847

Haqq C, Nosrati M, Sudilovsky D, Crothers J, Khodabakhsh D, Pulliam BL, Federman S, Miller JR, Allen RE, Singer MI, Leong SPL, Ljung B-M, Sagebiel RW & Kashani-Sabet M (2005) The gene expression signatures of melanoma progression. *Proc Natl Acad Sci USA* **102**: 6092–6097

Ikeda H, Lethé B, Lehmann F, van Baren N, Baurain JF, de Smet C, Chambost H, Vitale M, Moretta A, Boon T & Coulie PG (1997) Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory receptor. *Immunity* **6**: 199–208

Kilpinen S, Autio R, Ojala K, Iljin K, Bucher E, Sara H, Pisto T, Saarela M, Skotheim RI, Björkman M, Mpindi J-P, Haapa-Paananen S, Vainio P, Edgren H, Wolf M, Astola J, Nees M, Hautaniemi S & Kallioniemi O (2008) Systematic bioinformatic analysis of expression levels of 17,330 human genes across 9,783 samples from 175 types of healthy and pathological tissues. *Genome Biol* **9**: R139

Le Guezennec X, Vermeulen M, Brinkman AB, Hoeijmakers WAM, Cohen A, Lasonder E & Stunnenberg HG (2006) MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. *Mol Cell Biol* **26**: 843–851

Oberthuer A, Hero B, Spitz R & Berthold F (2004) The tumor-associated antigen PRAME is universally expressed in high-stage neuroblastoma and associated with poor outcome. *Clinical Cancer Research*

Olsen JV, Ong S-E & Mann M (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol Cell Proteomics* **3**: 608–614

Petroski MD & Deshaies RJ (2005) Function and regulation of cullin-RING ubiquitin ligases. *Nat Rev Mol Cell Biol* **6**: 9–20

Phizicky E, Bastiaens PIH, Zhu H, Snyder M & Fields S (2003) Protein analysis on a proteomic scale. *Nature* **422**: 208–215

Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M & Séraphin B (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**: 1030–1032

Wadelin F, Fulton J, McEwan PA, Spriggs KA, Emsley J & Heery DM (2010) Leucine-rich repeat protein PRAME: expression, potential functions and clinical implications for leukaemia. *Mol. Cancer* **9**: 226

CHAPTER 3

THE TUMOR ANTIGEN PRAME

is a subunit of a Cul2 ubiquitin ligase and associates with active NFY promoters

Adalberto Costessi[1][*], Nawel Mahrour[2][*], Esther Tijchon[1], Rieka Stunnenberg[1], Marieke A. Stoel[1], Pascal W. Jansen[1], Dotan Sela[2], Skylar Martin-Brown[2], Michael P. Washburn[2,3], Laurence Florens[2], Joan W. Conaway[2,4], Ronald C. Conaway[2,4], Hendrik G. Stunnenberg[1]

[1] Department of Molecular Biology, Nijmegen Centre for Molecular Life Sciences, Radboud University, 6500 HB, Nijmegen, The Netherlands
[2] Stowers Institute for Medical Research, 1000 E 50th Street, Kansas City, MO 64110, USA
[3] Department of Pathology and Laboratory Medicine, Kansas University Medical Center, Kansas City, KS 66160, USA
[4] Department of Biochemistry and Molecular Biology, Kansas University Medical Center, Kansas City, KS 66160, USA
[*] equal contribution

ABSTRACT

The human tumor antigen PRAME (Preferentially expressed antigen of melanoma) is frequently overexpressed in tumors. High *PRAME* levels correlate with poor clinical outcome of several cancers, but the mechanisms by which PRAME could be involved in tumorigenesis remain largely elusive. We applied protein-complex purification strategies and identified PRAME as a substrate recognition subunit of a Cullin2-based E3 ubiquitin ligase. PRAME can be recruited to DNA in vitro, and genome-wide chromatin immunoprecipitation experiments revealed that PRAME is specifically enriched at transcriptionally active promoters that are also bound by NF-Y and at enhancers. Our results are consistent with a role for the PRAME ubiquitin ligase complex in NF-Y-mediated transcriptional regulation.

INTRODUCTION

PRAME is the founding member of a large family of PRAME genes (Birtle *et al*, 2005). PRAME genes have been implicated in oogenesis (Dadé *et al*, 2003), spermatogenesis (Wang *et al*, 2001) and stemness (Bortvin *et al*, 2003; Cinelli *et al*, 2008), but their function is poorly understood. Human PRAME was initially identified as the protein triggering an autologous anti-tumor immune response in a melanoma patient (Ikeda *et al*, 1997). Since then, many reports have shown that PRAME is not expressed in most normal adult tissues, while it is overexpressed in a wide variety of solid and hematological malignancies (Kilpinen *et al*, 2008). Although not prognostic for all cancers (Steinbach *et al*, 2002; Tajeddine *et al*, 2005; van Baren *et al*, 1998), high PRAME expression has been found to correlate with the stage of melanoma lesions (Haqq *et al*, 2005) and with shorter overall survival of neuroblastoma (Oberthuer *et al*, 2004) and serous ovarian adeno-carcinoma patients (Partheen *et al*, 2008). High PRAME mRNA levels were also shown to be an independent prognostic factor of poor clinical outcome in breast cancer (Doolan *et al*, 2007; Epping *et al*, 2008). In addition, PRAME is significantly activated in BCR-ABL-positive chronic myeloid leukemia (Ph+ CML) patients at the onset of advanced phase and blast crisis (Radich *et al*, 2006), and hypomethylation of the PRAME locus seems to play a role in this process (Roman-Gomez *et al*, 2007; Schenk *et al*, 2007). A growing body of evidence indicates that PRAME contributes markedly to the oncogenic phenotype of several cancers. PRAME has been recently implicated in retinoic acid signaling and polycomb-mediated transcriptional repression in melanoma cells (Epping *et al*, 2005); however, these mechanisms do not seem to play a role in breast cancer (Epping *et al*, 2008) and acute myeloid leukemia (Steinbach *et al*, 2007). While evidence that expression of PRAME is restricted mainly to tumor cells has designated it as a promising therapeutic target, efforts to develop strategies to interfere with the oncogenic activity of PRAME are hindered by a lack of information about its function(s) (Paydas, 2008).

In this study, we combined biochemical and genomics approaches to identify PRAME as a subunit of Cullin2-based E3 ubiquitin ligases. In addition, we show that PRAME is recruited in cells to epigenetically and transcriptionally active promoter regions bound by NFY, a transcription factor that is essential for early embryonic development (Bhattacharya *et al*, 2003) and has been implicated in the maintenance of the high proliferative capacity of ES cells (Grskovic *et al*, 2007) as well as in the inhibition of differentiation (Gurtner *et al*, 2003).

RESULTS

## PRAME associates with Cullin2-based E3 ubiquitin ligases

To investigate the molecular function of human PRAME, we used a proteomic approach to identify PRAME-interacting proteins. To this end, we stably expressed epitope-tagged versions of PRAME suitable for protein purifications in two cellular backgrounds: human K562 leukemia cells, which endogenously express detectable levels of PRAME, and human embryonic kidney HEK293 cells, which lack detectable endogenous PRAME (Fig. 1A). When analyzed by size exclusion chromatography, both endogenous and epitope-tagged PRAME from K562 cells showed a broad elution profile, with a substantial fraction of the protein eluting earlier than expected of free PRAME, which has a predicted molecular mass of 58 kDa (Fig. 1B). This observation suggests that PRAME may be part of one or more protein complexes.

We purified PRAME-associated proteins from K562-TAG-PRAME or HEK293-HA-PRAME cells by anti-epitope immunoaffinity chromatography and analyzed them by mass spectrometry. As shown in Tables 1 and 2, eluates from K562-TAG-PRAME and HEK293-HA-PRAME included a collection of proteins not found in control samples. Among the most highly represented were Cullin family member 2 (Cul2), Elongins B (TCEB2) and C (TCEB1), and RING finger protein Rbx1, all of which are subunits of the large family of multisubunit, Cullin-based, ubiquitin ligases. In these ubiquitin ligases, the Elongin BC dimer acts as an adaptor that links an Elongin BC-box containing substrate recognition subunit, or BC-box protein, to heterodimeric modules composed of either Cullin 2 and Rbx1, or Cullin 5 and Rbx2 (Kamura *et al*, 2004; Kamura *et al*, 1998; Lonergan *et al*, 1998). The RING finger domains of the related Rbx1 and Rbx2 proteins in turn recruit and activate an E2 ubiquitin conjugating enzyme for ubiquitination of substrates (Kamura *et al*, 1999; Petroski & Deshaies, 2005).

The interaction of TAG-PRAME with ubiquitin ligase subunits was confirmed by western blotting of TAG-eluates. As shown in Figure 1C, these fractions contained PRAME, Elongins B and C, and two forms of Cul2 that most likely correspond to unmodified and a more slowly migrating Neddylated-Cul2, a hallmark of active ubiquitin ligases. Supporting the physiological relevance of these interactions, endogenous PRAME, Cul2, Elongin B, and Elongin C were all detected in anti-PRAME immunoprecipitates from nuclear extracts of parental K562 cells (Fig. 1D). Notably, while PRAME was clearly

FIGURE 1
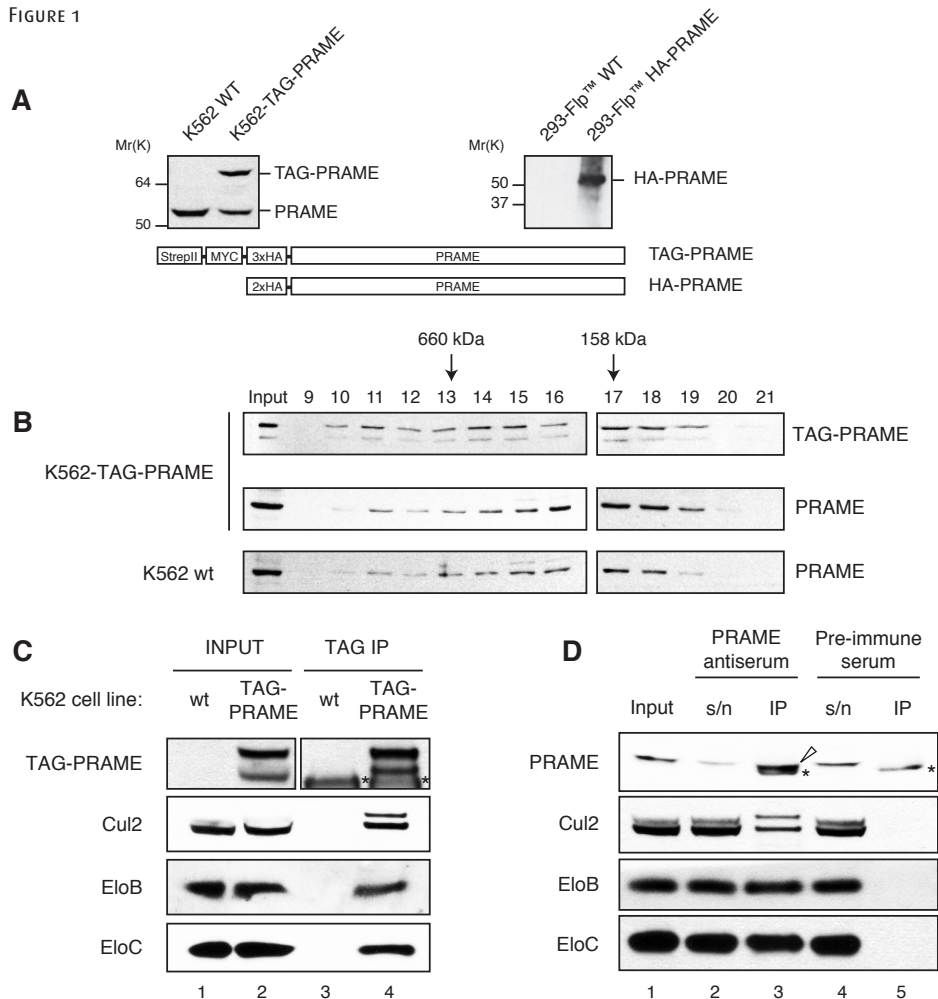
PRAME assembles with Cullin2 ubiquitin ligases.

(A) Ectopic expression of tagged versions of PRAME suitable for purifications. Whole-cell extracts of K562 cells or HA-immunoprecipitates from 293 cells were analyzed by western blot with anti-PRAME antibodies. K562 cells expressing StrepII-MYC-3xHA-PRAME (TAG-PRAME) at near-endogenous levels were selected, and 293 Flp-In cells stably expressing HA-PRAME were generated using the Invitrogen Flp-In system.

(B) Superose 6 gel filtration analysis of nuclear extracts from K562-TAG-PRAME and K562 wild type cells. Western blot was performed with anti-HA or affinity-purified anti-PRAME antibodies. Both TAG-PRAME and endogenous PRAME partially elute in high-molecular weight fractions. The numbers represent eluates fractions. A less prominent and faster migrating band for TAG-PRAME was detected, that most likely corresponds to a C-terminal truncation.

(C) TAG-PRAME co-immunoprecipitates Cullin2-EloBC complex. Nuclear extracts from K562 wild type and K562-TAG-PRAME cells were subjected to anti-TAG immunoprecipitation and western blotting as indicated; 0.8% of input and 33% of IP were loaded. A more slowly migrating form of Cullin2 was enriched after IP; this band is consistent with being post-translationally neddylated Cullin2. Asterisks indicate antibody heavy chains.

(D) Endogenous PRAME interacts with Cullin2-EloBC complex. Nuclear extracts from K562 wild type cells were subjected to immunoprecipitations with anti-PRAME or pre-immune sera and western blotting as indicated; 1% of input, 1% of supernatant and 20% of IP were loaded. PRAME was detected using a light-chain specific secondary antibody. PRAME was readily depleted from the extract after incubation with the specific antibodies (compare lanes 1, 2 and 4). Asterisks indicate protein A that dissociated from the beads after elution; the open arrow indicates PRAME; s/n, supernatant; IP, immunoprecipitates.

TABLE 1   **PRAME-associated ubiquitin ligase subunits identified by Mass Spectrometry of purified complexes from K562 cells.**

Protein name	Protein ID	TAG-PRAME (emPAI)	TAG-PRAME (pept.)	Control (pept.)
Elongin C	IPI00300341.5	4.6	6	1
Elongin B	IPI00026670.3	4.2	5	0
Rbx1	IPI00003386.3	0.8	1	0
PRAME	IPI00019282.1	0.7	6	0
CUL2	IPI00014311.4	0.7	12	0

Pept., number of peptides.

TABLE 2   **PRAME-associated ubiquitin ligase subunits identified by mass spectrometry of purified complexes from HEK 293 cells expressing epitope-tagged wild type or mutant PRAME proteins.**

64

Protein name	Gen Bank acc. no.	PRAME WT				PRAME BC-box mutant				PRAME Cul2-box mutant				Control
		Pept.	SpC	Cov. (%)	NSAF x1000	Pept.	SpC	Cov. (%)	NSAF x1000	Pept.	SpC	Cov. (%)	NSAF x1000	Pept.
PRAME	NP_996839	11	255	24.9	22	11	275	23	29.5	15	267	27.1	17.3	0
Elongin B	NP_009039	9	235	48.3	87.6	2	3	21.2	1.4	4	26	31.4	7.3	1
Elongin C	NP_005639	6	160	50.9	62.8	0	0	0	0	4	13	38.4	3.8	0
CUL2	NP_003582	14	455	18.4	26.9	0	0	0	0	0	0	0	0	0
NEDD8	NP_006147	2	9	17.3	4.9	0	0	0	0	0	0	0	0	0

Pept., number of peptides; SpC., spectral counts; Cov., sequence coverage; NSAF, Normalized Spectral Abundance Factor.

depleted from the extract after incubation with the specific antibodies (Fig. 1D, compare lanes 1, 2, and 4), Cul2, Elongin B, and Elongin C were not, consistent with their function as components of the many Cullin-based ubiquitin ligases present in cells (Kamura *et al*, 1998; Mahrour *et al*, 2008).

PRAME is a BC-box protein

Although we identified all of the other subunits of Cul2-based ubiquitin ligases as PRAME-interacting proteins, we did not detect any of the previously identified BC-box proteins in our fractions. BC-box proteins are characterized by a sequence motif of about 10 amino acids, called the BC-box, that binds to Elongin C. In addition, a downstream Cullin-box specifies binding to either Cul2 or Cul5 (Kamura *et al*, 2004; Mahrour *et al*, 2008). Multiple se-

FIGURE 2
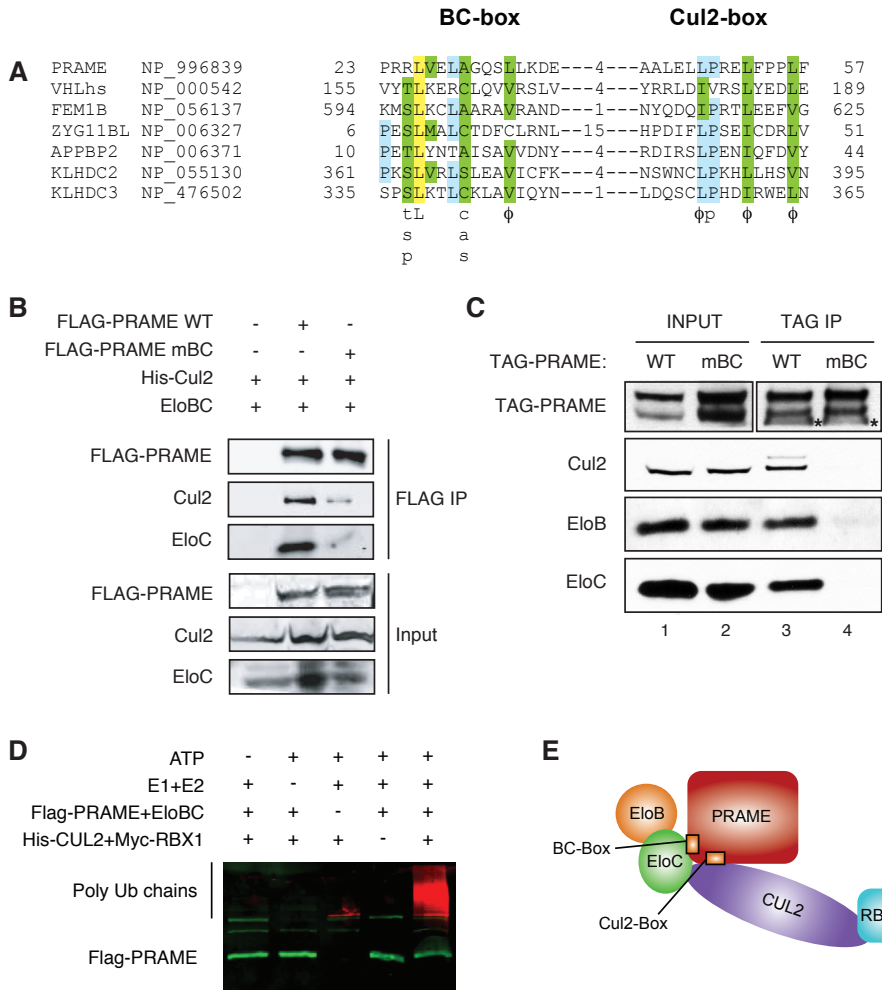
PRAME is a BC-box substrate receptor for Cullin2-EloBC ligases.
(A) Multiple sequence alignment of known Elongin BC-associating proteins with PRAME. Sequences were aligned using the blosum62mt2 alignment matrix with the AlignX program of the Vector NTI software package. Color coding is based on the AlignX default similarity table. Identical amino acids are highlighted in yellow, very similar ones in blue, and similar amino acids are shown in green. The canonical BC-box and Cul2-box sequences are indicated below the alignment. (B) PRAME has a functional BC-box. SF21 cells were co-infected with the baculoviruses as indicated. Total cell lysates and anti-FLAG immunoprecipitates were analyzed by western blotting. IP, immunoprecipitation; mBC, BC-box mutant. (C) BC-box defective PRAME lacks interaction with endogenous Cul2-EloBC ligases. Immunoblot analysis of 12CA5 HA immunoprecipitates from nuclear extracts of K562 cells stably expressing TAG-PRAME wild type or TAG-PRAME BC-box mutant (PRAME mBC). Asterisks indicate antibody heavy chains. (D) The Cul2-PRAME ligase has in vitro ubiquitination activity. Epitope-tagged subassemblies of the PRAME Cul2-based ubiquitin ligase complex were purified from insect cells and used in an in vitro ubiquitin assay (as described in Methods). The results were analyzed by western blot using rabbit anti-Flag and mouse anti-ubiquitin antibodies. (E) Model summarizing Cul2-EloBC-PRAME interactions.

quence alignments of PRAME with several known BC-box proteins identified putative BC- and Cul2-boxes near the PRAME N-terminus (Fig. 2A), raising the possibility that PRAME might function as a substrate recognition subunit of a Cul2-based ubiquitin ligase.

To assess the functionality of the predicted motifs, we introduced point mutations in key residues of the putative BC-box (L26P/A30F) or the putative Cul2-box (L48A/P49A). We observed that the BC-box mutation greatly reduced the binding of PRAME to Elongin C and Cul2 in baculovirus-infected insect cells (Fig. 2B). We next generated K562 and HEK293 cells stably expressing epitope-tagged PRAME carrying the BC-box or Cul2-box mutations and performed immunoprecipitations from nuclear extracts. In these cells, both mutations disrupted the interaction of PRAME with endogenous Cullin2 complex components (Fig. 2C and Table 2), demonstrating the functionality of the motifs.

In order to test the biochemical activity of the Cullin2-PRAME complex, we performed in vitro ubiquitination assays using purified recombinant proteins. As shown in Fig. 2D, we observed significant synthesis of poly-ubiquitin chains only when all the subunits were combined in the presence of ATP, confirming that the Cullin2-PRAME complex is a catalytically active ubiquitin ligase.

Taken together, these results indicate that PRAME's BC- and Cul2-boxes are essential for the assembly of a PRAME-containing Elongin BC ubiquitin ligase, and they identify PRAME as a BC-box substrate recognition subunit of a Cul2-based ubiquitin ligase (Fig. 2E).

PRAME associates with chromatin

Cell fractionation and cesium chloride gradient purification of cross-linked chromatin from K562 cells showed that PRAME associates with chromatin (data not shown). In addition, we observe that PRAME can be recruited to DNA in vitro. Consistent with the absence of recognizable DNA binding domains in PRAME and other PRAME complex subunits, purified PRAME complexes did not bind directly to DNA immobilized on agarose beads. When incubated with bead-bound DNA in the presence of nuclear extract, however, PRAME complexes were retained on the DNA beads, suggesting that activity(s) in the extract are needed to bridge PRAME's interaction with DNA (Suppl. Fig. 1A).

To investigate the chromatin-binding profile of PRAME, we performed ChIP assays followed by deep sequencing (ChIP-seq) using anti-PRAME

Figure 3
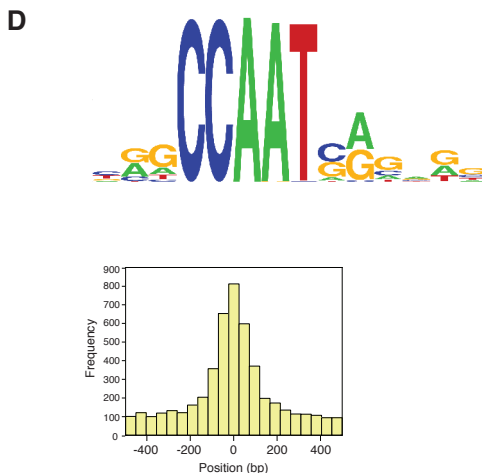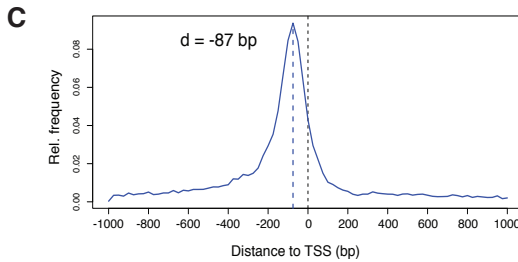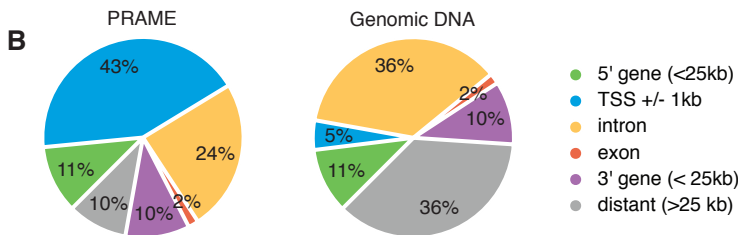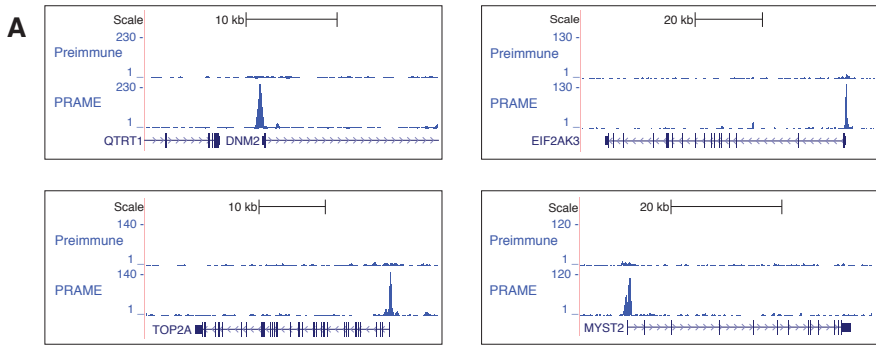
Genome-wide profiling of PRAME association with chromatin.

(A) Overview of PRAME binding sites. UCSC genome browser views of ChIP-seq profiles generated with preimmune (negative control) and PRAME antibodies in K562 cells. Specific PRAME peaks are detected at the promoters of the *DNM2*, *EIF2AK3*, *TOP2A*, and *MYST2* genes.

(B) Genomic distribution of PRAME binding sites (left) compared to the distribution of genomic DNA (right). The majority of sites (43%) are located within ±1kb from a transcription start site.

(C) Promoter binding of PRAME occurs at the TSS. Frequency plot of the distances of PRAME binding sites from TSS shows that 83% of PRAME promoter sites are located between -400 bp and +100 bp, with a peak at -87bp.

(D) Motif analysis identifies a CCAAT motif under PRAME promoter peaks. Top, logo of the Jaspar motif matching the most enriched sequence identified by GimmeMotifs. Bottom, the positional preference plot represents the frequency of the locations of the CCAAT motifs with respect to the summits of PRAME promoter summits, and indicate that the CCAAT motif is enriched at the center of PRAME binding regions.

67

antibodies or pre-immune serum as control for the specificity of the signals obtained. Overlapping sequence tags were joined and displayed as tag-density profiles in the UCSC genome browser. Strikingly, inspection of the profiles revealed that the promoter regions of many protein-coding genes had specific enrichment for PRAME, as exemplified by the screenshots in Figure 3A. Validation experiments by ChIP-qPCR confirmed the high quality of the sequencing data (Suppl. Fig. 1B). Furthermore, ChIP-seq with stably expressed, epitope-tagged PRAME resulted in very similar if not identical patterns of PRAME enrichment (data not shown). To substantiate our finding, we used the MACS algorithm (Zhang *et al*, 2008) at a false discovery rate of 10^{-6} and identified a total of 14132 PRAME binding sites. Location analysis showed that 43% of the PRAME sites are closer than 1 kb upstream or downstream of a transcription start site (TSS) (Fig. 3B). Importantly, promoters were 9-fold enriched as compared to the background distribution of genomic DNA.

68

Recently, Heintzman *et al.* identified enhancer elements in K562 cells based on ChIP-on-chip histone modification signatures and binding of the enhancer-associated acetyltransferase p300 (Heintzman *et al*, 2009). Most enhancers were cell type specific, and they were enriched in genomic domains containing genes specifically expressed in K562 cells. When we compared the binding profile of PRAME to the predicted enhancers we found a significant overlap between promoter-distal PRAME binding sites and enhancer regions: 44% of the distal PRAME sites overlapped predicted enhancers, whereas only 4.1% of random genomic sites did (11-fold enrichment). Moreover, similarly to the annotated enhancers, the distal PRAME sites were not randomly distributed in the genome, but clustered in the vicinity of PRAME-bound promoters (Suppl. Fig. 1C), suggesting possible promoter-enhancer interactions.

PRAME associates with NF-Y promoters

We next analyzed all PRAME binding sites falling within 1 kb of promoters. When we plotted their distance to the closest TSS, we found that 83% of promoter-proximal sites bound by PRAME are concentrated in a narrow window between -400 bp and +100 bp from TSS, peaking around -87 bp relative to the TSS.

Promoters are characterized by the presence of transcription factor binding sites that critically contribute to the regulation of gene expression. Since neither PRAME, nor other PRAME complex subunits contain recognizable DNA-binding domains, we reasoned that other transcription factor(s) might mediate the recruitment of PRAME to these regions. We therefore performed

de novo motif discovery with GimmeMotifs (van Heeringen & Veenstra, 2010) using the DNA sequences underlying the promoter regions bound by PRAME. This analysis identified a CCAAT sequence as the top-scoring motif, which stood out with highly significant p-values ($p<10^{-20}$) and the highest enrichment and MNCP scores as compared to a pool of random promoters (Fig. 3D and Suppl. Fig. 2). The CCAAT sequence can be bound by the Nuclear Transcription Factor Y (NFY), a heterotrimeric transcription factor ubiquitously expressed in cycling cells (Maity & de Crombrugghe, 1998). The NFYA subunit harbors the DNA-binding domain that recognizes the CCAAT sequence, while NFYB and NFYC are characterized by histone-like folds (Maity & de Crombrugghe, 1998). Similar *de novo* motif analysis performed on the distal PRAME binding sites identified two significant motifs: a GATA/Evi1 motif, and a Fos motif.

ChIP-seq experiments with antibodies against the NFYA and NFYB subunits revealed that PRAME sites at promoters are indeed bound by NFY in vivo. Screenshots of the sequencing data show clear enrichment of NFY at the representative PRAME sites DNM2 and EIF2AK3, as well as PRAME enrichment at the classical NFY promoters TOP2A and CDC25C (Fig. 4A). Consecutive ChIP experiments (ChIP-re-ChIP) using crosslinked NFYA antibodies detected PRAME and NFY concomitantly on the same regions, indicating that these proteins are simultaneously localized at the same promoter regions (Fig. 4B).

Using the MACS algorithm we identified a total of 28844 regions that were bound by NFYA and/or NFYB. The distance distributions of both NFYA and NFYB binding sites from TSSs showed a distinct peak around -78/-84 bp very similar to PRAME (Fig. 4C). We next counted the number of sequence tags within NFY-binding sites and performed a regression curve analysis (Suppl. Fig. 3). The binding of NFYA and NFYB were very closely correlated at the vast majority of the sites, as expected for stoichiometric binding and in agreement with in vitro experiments showing that all subunits are required for efficient DNA binding (Sinha *et al*, 1995).

Globally, 71% of all PRAME binding sites overlapped NFY regions. Importantly, virtually all PRAME-bound promoters were also bound by NFY, corroborating and extending the relevance of the CCAAT motif (Fig. 4D). In contrast, a sizeable number of NFY promoters did not contain PRAME peaks (Fig. 4D). We performed functional annotation of the associated genes and found that NFY-only genes as well as NFY genes with PRAME enrichment are associated with metabolic and cancer-related pathways. Furthermore, NFY-only genes are specifically enriched for immunological processes and intercellular junc-

FIGURE 4
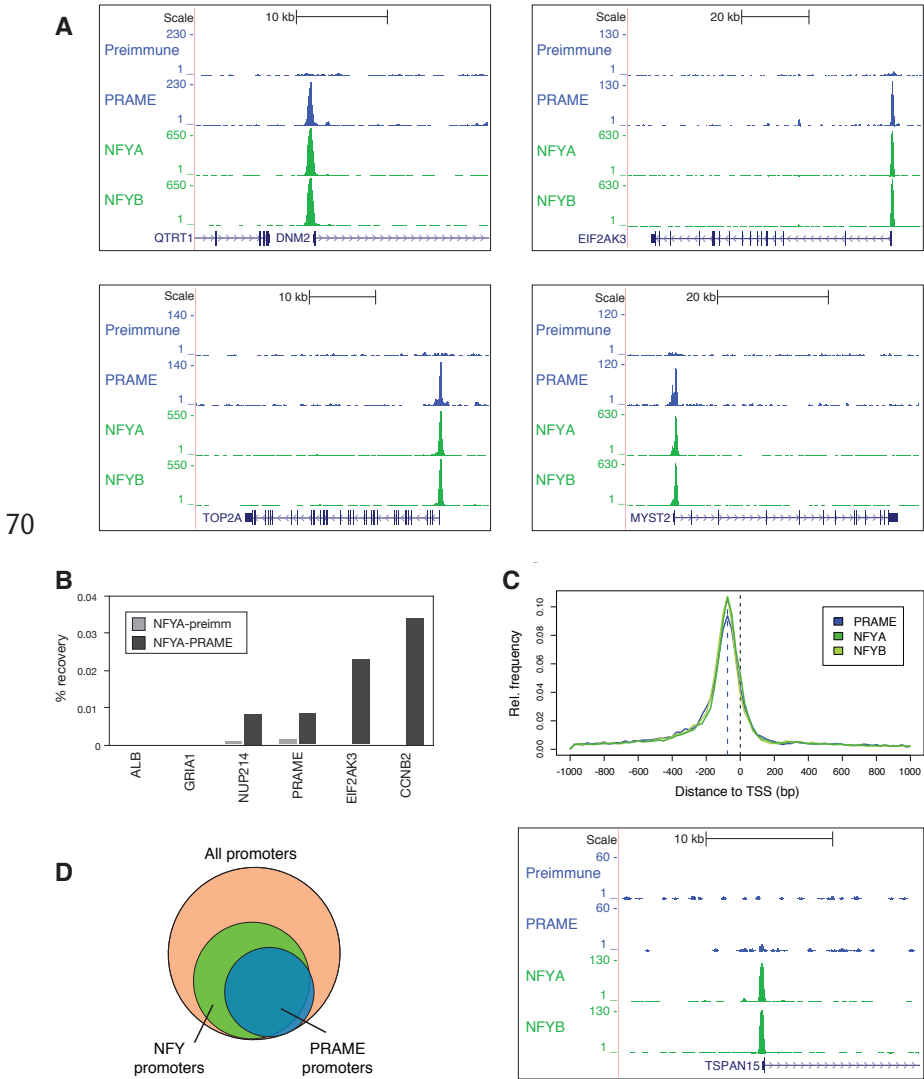
70

PRAME associates with NFY-bound promoters genome-wide.
(A) Overview of NFY binding to PRAME loci. UCSC genome browser views show clear binding of PRAME, NFYA, and NFYB at the promoters of the *DNM2*, *EIF2AK3*, *TOP2A*, and *MYST2* genes.
(B) Sequential ChIP (ChIP-re-ChIP) experiments show co-occupancy of PRAME and NFY. The first ChIP was performed with covalently crosslinked NFYA-beads, followed by either anti-PRAME serum (NFYA-PRAME) or preimmune serum as negative control (NFYA-preimmune).
(C) Promoter binding of NFY and PRAME occurs at the TSS. Frequency plot as in Fig. 3C shows very similar distributions of distances from TSS for NFY and PRAME.
(D) PRAME associates with a subset of NFY-promoters genome-wide. Venn diagram of the overlap between NFY-bound (green) and PRAME-bound (blue) promoters shows that PRAME promoters overlap almost completely with NFY promoters. On the contrary, a minority of NFY promoters are not bound by PRAME. An example is the *TSPAN15* gene (right).

tions, while NFY genes with PRAME binding are significantly associated with numerous transcription-related pathways and cell cycle regulation (Supplementary Table 1).

PRAME is specifically recruited to transcriptionally active NFY promoters

Post-translational modifications of the histone proteins, like lysine methylations and acetylations, are implicated in the regulation of gene transcription. Therefore, we used histone marks to assess the transcriptional status of PRAME loci. For this purpose we processed ChIP-seq datasets generated by the ENCODE project for K562 cells (see methods), and we quantitated the signals at all NFY promoters. We found that the active histone marks H3K9ac and H3K4me3 showed a strong correlation with the binding of PRAME (Fig. 5A). In contrast, NFY-only promoters had near-background levels of these marks.

To investigate the transcriptional status of the associated genes, we quantitated the occupancy in gene bodies of RNA Polymerase II and of the histone marks H3K36me3, which has been linked to transcriptional elongation (Krogan *et al*, 2003), and H3K27me3, which has been linked to repression (Barski *et al*, 2007). As shown in Figure 5B, NFY genes without PRAME binding displayed low levels of PolII and H3K36me3, and high levels of the repressive H3K27me3 mark. In contrast, PRAME-bound genes had significantly higher occupancy of both PolII and H3K36me3 (Fig. 5B and D, and Suppl. Fig. 4), suggesting that PRAME is associated specifically with transcriptionally active genes. To substantiate these findings, we analyzed gene expression levels by RNA-seq in our K562 cells. In agreement with the ChIP-seq data, we detected few, if any, transcripts from NFY-only genes, while PRAME-bound NFY genes were expressed at high levels (Fig. 5C).

We next tested if components of the Cullin2 complex are present together with PRAME on chromatin. In the absence of ChIP-grade antibodies, we established stable K562 cell lines expressing Elongin C tagged with a double TY1 epitope (TTE-EloC), for which a ChIP-grade antibody was validated in our lab. ChIP-qPCR experiments showed specific binding of EloC at all the PRAME promoters tested (Fig. 5E). As a control, we performed similar experiments with TTE-PRAME.

Taken together, our data supports a model in which a Cul2/EloBC/PRAME complex is specifically bound to the subset of epigenetically and transcriptionally active NFY promoters and nearby enhancer regions.

71

FIGURE 5

72
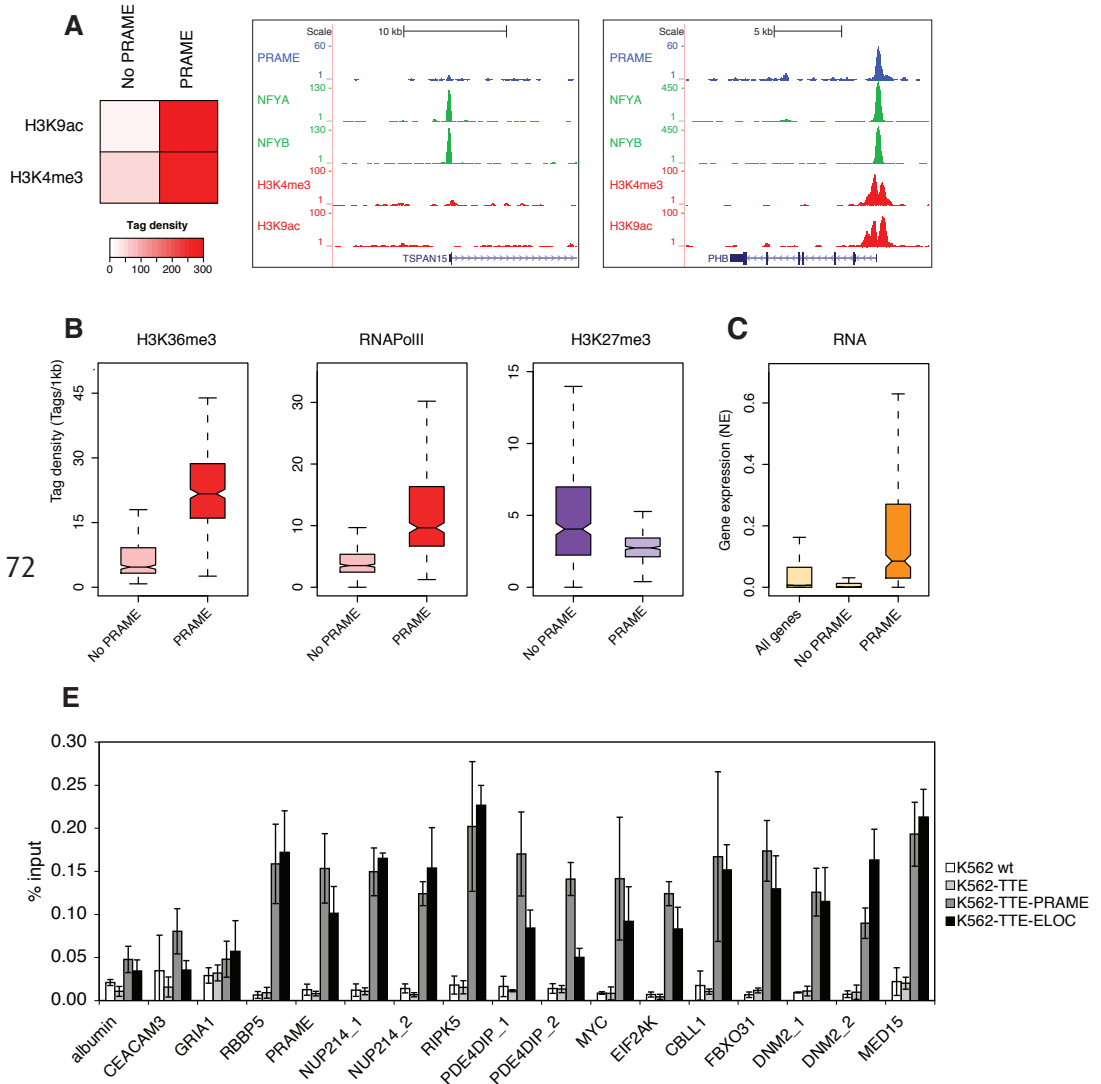
PRAME discriminates between active and inactive NFY loci.
(A) PRAME binding positively correlates with H3K9ac and H3K4me3 modifications at NFY promoters. Left, heat map representing median tag densities (tags/kb) of H3K9ac and H3K4me3 at promoter-associated NFY binding sites without ("No PRAME") or with ("PRAME") PRAME binding. Higher levels of both histone marks are found if PRAME is present. Right, UCSC genome browser views of two representative genes: *TSPAN15* and *PHB*.
(B) PRAME binding of NFY promoters correlates with high H3K36me3 and RNA PolII, and with low H3K27me3 in the gene body. Box plots show the tag density (tags/kb) over the bodies of the genes associated with NFY promoters without or with PRAME binding.
(C) PRAME binding correlates with high levels of gene expression. Box plot of Normalized Expression values computed from RNA-seq experiments show that PRAME-bound NFY genes are expressed at higher levels than those without PRAME.
(D) Genome browser view of the genome-wide data combined. Two neighboring tail-to-tail NFY genes are shown: *SPRYD4* is bound by PRAME and is active, while *GLS2* is not bound by PRAME and is in a repressed state.
(E) Model summarizing the genome-wide data.
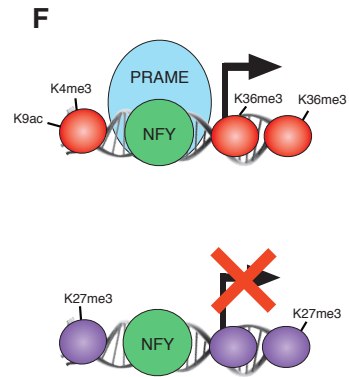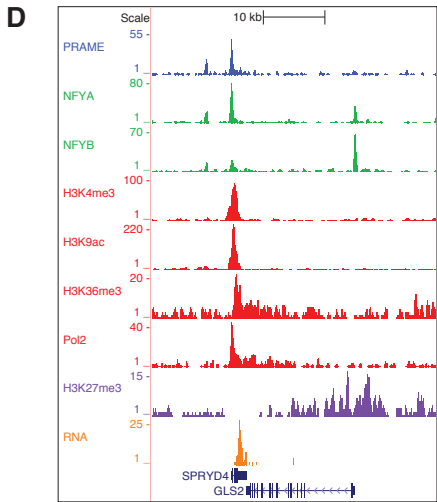
DISCUSSION

Our proteomic data reveals that the human oncoprotein PRAME is a BC-box substrate recognition subunit of a Cul2-based E3 ubiquitin ligase, suggesting a functional link between ubiquitination and the oncogenic properties of PRAME. Notably, the Von Hippel-Lindau tumor suppressor protein is one of the best-characterized BC-box proteins, and its tumor suppressor activity is due in part to its role as the substrate recognition subunit of a Cul2-based ubiquitin ligase (Kaelin, 2007). Intriguingly, all the annotated PRAME-like proteins display a complete conservation of the BC-box and Cul2-box at their N-terminus (Suppl. Fig. 5), suggesting that all members of the large PRAME protein family might function as substrate recognition subunits of Cul2-based ubiquitin ligases, possibly displaying different degrees of specificity and/or redundancy in substrate selection.

To date, ubiquitination has been demonstrated to play important roles in gene regulation, affecting several different steps in the transcriptional process (Collins & Tansey, 2006; Conaway *et al*, 2002). Many activators are subject to ubiquitin-mediated degradation at a rate that parallels their trans-activation abilities (Salghetti *et al*, 2001). Histones, transcription factors, and RNA polymerase II (RNAPII) are targets of ubiquitination, and ubiquitination

events are required for the initiation-competent RNAPII complex to exchange factors and become competent for cotranscriptional mRNA processing (Muratani *et al*, 2005). A recent study indicated that the stability of NFYA can be modulated by ubiquitination and degradation via the proteasome, and that degradation-resistant NFYA mutants increased cell proliferation (Manni *et al*, 2008). Interestingly, some of the lysine residues in NFYA that are target of ubiquitination could also be acetylated by p300, suggesting a regulatory competition between these post-translational modifications (Manni *et al*, 2008).

Our genome-wide analyses reveal that PRAME is strongly enriched at NFY promoters, and binding of PRAME predicts a permissive epigenetic landscape and transcriptional activity at these regions. Moreover, we found that PRAME is also significantly enriched at adjacent enhancer regions, suggesting that it could participate in promoter-enhancer interactions. Recently, Ceribelli *et al*. reported a ChIP-on-chip analysis of NFYB in HeLa cells with 2.2% coverage of the human genome. Similar to our K562 data, the majority of the NFY-bound promoters identified in HeLa cells were transcriptionally active and carried H3K9-14ac and/or H3K4me3 marks, while about 27% of NFY-bound promoters were devoid of these active marks and mostly silent. Further experiments implicated NFY not only in the expression of active genes, but also in the repression of silent ones, revealing an unexpected dual nature of the NFY complex (Ceribelli *et al*, 2008). The basis of this dual functionality, however, is still unclear, and genome-wide data suggest that the activity and inactivity of NFY promoters is determined by additional factors and mechanisms that are yet to be identified.

PRAME is the first factor reported to discriminate between active and inactive NFY loci. Our genome-wide analyses, summarized in Figure 5F, reveal that PRAME is specifically recruited to NFY loci associated with the promoters of genes that are highly transcribed and that have high levels of H3K9ac and H3K4me3 at their promoters and high H3K36me3 but low H3K27me3 levels in their gene bodies. This data suggest a role for PRAME in transcriptional regulation, and possibly in the switch between activity and inactivity of NFY loci. It is tempting to speculate that PRAME might mediate ubiquitination, and possibly degradation, of one or more of the factors that are involved in the regulation of NFY promoters. Despite evidence that PRAME and NFYA can co-occupy the same genetic loci (Fig. 4B), we have not been able to detect stable and direct interactions between PRAME and NFY subunits by coimmunoprecipitation following their expression in transfected mammalian cells or

in baculovirus-infected insect cells (data not shown). We have also observed that PRAME can be recruited in vitro to a DNA fragment containing the NFY-dependent HSPA5 promoter by a mechanism that depends on activity(s) present in a nuclear extract. It seems likely that PRAME-NFY interactions, if they exist, occur only on DNA and depend on additional, yet to be identified factors, since (i) purified NFY is not sufficient to recruit PRAME to HSPA5 promoter DNA, and (ii) supplementing nuclear extract with exogenously added NFY does not further enhance PRAME recruitment (Suppl. Figure 1A).

Interactions have been reported between NFY and the histone acetyltransferases GCN5 and PCAF (Currie, 1998), and single-gene studies showed that NFY can activate at least some of its target genes in cooperation with the acetyltransferase p300, which is able to acetylate histone tails as well as NFY subunits (Caretti *et al*, 2003; Gurtner *et al*, 2008; Li *et al*, 1998; Salsi *et al*, 2003). While these acetyltransferases constitute potential targets of PRAME, other unknown factors could also take part to these mechanisms. Notably, p300 is not only recruited to promoters, but it is also a hallmark of enhancer elements. Heintzman and coworkers used p300 ChIP-on-chip binding profiles to identify an enhancer-specific chromatin signature, which was successfully used to predict enhancer elements in several cell lines, including K562 (Heintzman *et al*, 2007). Significantly, we found that about half of the promoter-distal PRAME sites are localized at predicted enhancers, suggesting that PRAME might be implicated in promoter-enhancer crosstalk.

In an effort to gain insight into PRAME's potential role in transcription regulation, we knocked down PRAME with retroviral shRNA vectors and measured the mRNA levels of several target genes. Despite a 70% decrease of PRAME transcripts, we did not observe reproducible changes in the transcript levels of the target genes tested (data not shown). This is perhaps not surprising given that PRAME is a member of a large family of closely related proteins that could have overlapping or redundant functions in cultured cells. Further, the fact that the PRAME gene family underwent an extraordinary expansion specifically in the rodents and primates (Birtle *et al*, 2005; Consortium *et al*, 2007) suggests that they are more likely to fine-tune transcriptional regulatory processes rather than to be essential for transcription. Regrettably, the lack of a one-to-one orthology between the human and mouse PRAME genes (Birtle *et al*, 2005) has thus far precluded the development of suitable knock-out mouse models for studying PRAME function.

Notably, it is becoming clear that many proteins associate with the promoter regions of genes, but their depletion does not seem to cause significant or widespread transcriptional defects. For example, the yeast Set1 histone methyl transferase is widely associated with active promoters, but deletion of the Set1 gene causes expression changes of 1.5-fold or more in only about twenty genes (Miller *et al*, 2001). Similarly, the Jmjd3 histone demethylase was recently shown to associate with the promoters of LPS-induced genes, but the expression of most of them was unaffected by knock-out of Jmjd3 (De Santa *et al*, 2009). These observations are consistent with the model that some factors are not strictly required for the bulk of transcription, but instead have ancillary or regulatory functions that become apparent only in very specific conditions or subpopulations of cells.

While numerous reports established a role for PRAME in human cancers, several studies indicated that PRAME-like genes have roles in the early stages of spermatogenesis (Wang *et al*, 2001) and oogenesis (Dadé *et al*, 2003), as well as in embryonic development and embryonic stem cells (Bortvin *et al*, 2003). Recently, overexpression of Pramel7 in E14 embryonic stem cells was reported to maintain a pluripotent state in the absence of the anti-differentiation factor LIF (Cinelli *et al*, 2008). These findings are consistent with the notion that stem cells and neoplastic tissues share many properties, and several oncogenic pathways were recently found to also regulate self-renewal mechanisms in stem cells (Reya *et al*, 2001). Interestingly, recent studies suggest important roles for NFY in stem cell biology. Knock-out mice demonstrated that NFYA is essential for early mouse development (Bhattacharya *et al*, 2003), and NFYA was found to promote the self-renewal of hematopoietic stem cells (Zhu *et al*, 2005). Moreover, a systematic analysis of cis-regulatory elements identified the CCAAT box and NFY as crucial regulators of both mouse and human stem cells-specific genes, and NFY was involved in the maintenance of the proliferative capacity of ES cells (Grskovic *et al*, 2007).

Taken together, our data unambiguously shows that the PRAME oncoprotein is part of a Cullin2-based ubiquitin ligase complex and that PRAME localizes to active NFY promoters, paving the road for understanding the functions of PRAME in normal cells and in human malignancies. Future experiments will address the role of the PRAME ubiquitin ligase at TSSs, its role in transcription of NFY-regulated genes, and its contribution to the oncogenic properties of PRAME.

MATERIALS AND METHODS

Cell culture and stable cell lines

HEK293 cells were cultured in DMEM and K562 in RPMI medium (Gibco, Invitrogen) at 37°C in 5% CO2. Both media were supplemented with 5% Glutamax, 10% fetal bovine serum, 100 U/ml of penicillin, 100 μg/ml of streptomycin (Gibco, Invitrogen). Stable HEK293-HA-PRAME cells were generated using pcDNA5/FRT/HA-PRAME and the Flp-in system (Invitrogen). K562 stable cell lines were generated with retroviral constructs according to the following 5-days protocol. On the first day, HEK293T cells at 50-60% confluence in a 6-cm dish were transfected using Lipofectamine2000 (Invitrogen) and 1,8 μg of MD-MLV, 300 ng VSV-G, 3.6 μg retroviral construct, 100 ng pGFP. After 24 h (day 2), the virus-containing supernatant was harvested, centrifuged for 10 min at 3000 rpm and used for a first round of transduction of 10^6 K562 cells in the presence of 8 μg/ml of Polybrene (Sigma) for 24 h. A second round of transduction was performed on day 3, and on day 4 the target cells were transferred to fresh RPMI and selection was started on day 5 with either 6-10 μg/ml of puromycin (Sigma) or 1-2 mg/ml Zeocin (Invitrogen).

Antibodies and western blot

PRAME antisera and affinity-purified antibodies were generated by immunizing rabbits with the previously described peptide FPEPEAAQPMTKKRKVDG (Epping *et al*, 2005) (V27 serum). Mouse monoclonal HA-12CA5 and Myc-9E11 were produced in-house from hybridoma cultures. Commercial antibodies used were: mouse monoclonal FLAG-M2 (Sigma), c-Myc monoclonal (Roche Applied Science), rabbit HA (Bethyl Laboratories, or Abcam ab9110), mouse monoclonal HA.11 (Covance MMS-101P), rabbit Cul2 (Zymed Laboratories 51-1890, mouse ElonginC (BD Transduction Laboratories 610760), goat ElonginB (Santa Cruz P-16, sc-23407), alpha-tubulin (Santa Cruz B-7, sc-5286), HDAC2 (Biomol SA-402).

Proteins were separated by conventional SDS-PAGE or NuPAGE 4-12% Bis-Tris gels (Invitrogen) run with MES buffer. Western blots were visualized by ECL (GE Healthcare) or Odyssey (LiCor).

Plasmids and cloning

Full-length human PRAME was PCR-amplified and cloned into NotI and XhoI sites of pcDNA5/FRT/TO/StrepII-2xTEV-Myc-3xHA (Le Guezennec *et al*, 2006) to generate pcDNA5-TAG-PRAME. A BamHI-XhoI fragment encoding TAG-PRAME was further subcloned into LZRS(Zeo) to generate the retroviral expression vector LZRS(Zeo)-TAG-PRAME. Full-length human PRAME (ATCC) was subcloned with N-terminal HA epitope tags into pcDNA5/FRT (Invitrogen). Mutagenesis to generate PRAME mutants was performed using the QuikChange II XL site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions. To create the pTTE retroviral vector for ChIP, the ProtA-2xTEV-Myc cassette of the pZXN plasmid (Le Guezennec *et al*, 2006) was removed by digestion with BamHI and SalI. The resulting backbone was ligated to overlapping synthetic oligos with BamHI and XhoI overhangs, encoding two consecutive TY1 epitopes, a single ERalpha epitope and an extended MCS. PRAME was cloned using the EcoRI site to generate pTTE-PRAME plasmid; EloC was cloned using BglII and XhoI to generate pTTE-EloC plasmid. All constructs were checked by sequencing.

Gel filtration analysis

Size exclusion chromatography was performed with a Superose 6 column on a SMART system (Amersham Pharmacia Biotech) with buffer containing 300 mM NaCl, 10% glycerol, 20 mM TRIS, 0,2 mM EDTA, 1 mM DTT and 1 mM PMSF. 50 μl nuclear extract were separated in 100 μl fractions and analyzed by SDS–PAGE and western blot.

Purification of protein complexes from K562 cells

K562 cells were grown in spinner flasks and harvested at a confluence of about 10^6 cells/ml. For each protein extraction, about 10^{10} cells were centrifuged for 15 minutes at 1600 rpm and washed with PBS. Cell pellets were resuspended in 2,5 volumes of hypotonic buffer T1 (10 mM Hepes pH 7,9, 10 mM KCl, 0,1 mM $MgCl_2$, 0,1 mM EDTA pH 8, 1 mM DTT, 10% glycerol) supplemented with phosphatase and protease inhibitors (1 mM $NaVO_4$, 1 mM NaF, 1 mM PMSF, Protease inhibitor cocktail (Roche)). After 15 minutes of incubation on ice, the cell membranes were disrupted

with 20 strokes of Pestle A, or until 80-90% of the cells were stained by trypan blue. After centrifugation for 20 minutes at 3000 rpm at 4ºC, the supernatant was harvested and regarded as the cytoplasmic fraction. The pellet containing the nuclei was resuspended in 1 volume of T2 buffer (20 mM Hepes pH 7,9, 420 mM NaCl, 1,5 mM MgCl$_2$, 0,1 mM EDTA pH 8, 1 mM DTT, 10% glycerol) supplemented with inhibitors, and incubated in rotation for 1 hour at 4ºC to extract the nuclear proteins. The suspension was cleared by ultracentrifugation for 30 minutes at 100000 g at 4ºC (Sorvall RC M150 GX). The supernatant was the nuclear fraction. The quality of the cytoplasmic/nuclear extractions was determined by western blot of alpha tubulin (cytoplasmic marker) and HDAC2 (nuclear marker).

About 5-10 ml (60-120 mg) of nuclear extracts were incubated overnight at 4°C with anti-Myc (9E11) antibodies covalently crosslinked to protA-sepharose CL-4B (GE Healthcare) (see below). The following day, beads were washed three times with 10 ml buffer IPP300/0,5 (300 mM NaCl, 10 mM Tris-HCl pH 8, 0,5 mM EDTA and 0,5% NP-40), and the protein complexes were eluted with 20 µg/ml of Myc peptide in a thermoshaker for 30 minutes at 28°C.

Crosslinking of antibody to beads

ProtA-sepharose CL-4B (GE Healthcare) or protG-sepharose (Sigma) beads were incubated with antibody in PBS at room temperature for 1-2 h, then washed once with 10 volumes of sodium borate solution (0,2 M, pH 9). Chemical cross-linking of the antibodies to the beads was performed resuspending the beads in a sodium borate solution containing 5.2 mg/ml DMP (dimethyl-pimelimidate, Sigma) and incubating for 30 minutes at room temperature with gentle rotation. The beads were washed once in 10 volumes ethanolamine (0,2 M, pH 8), incubated with ethanolamine for 2 h at room temperature, washed twice with glycine (0,1 M, pH 3) to remove un-crosslinked antibodies, and eventually with PBS to neutralize the pH.

Mass Spectrometry and emPAI calculation for K562 complexes

The eluates of large scale immunoprecipitations from K562 cells were concentrated by speedvac and run shortly (~1 cm) on SDS-PAGE gels to remove detergents and the excess of peptide. The gel lane was then fixed, cut in small pieces, reduced and alkylated. Proteins were digested overnight with trypsin (Promega) and eluted from the gel with trifluoroacetic acid. Peptides were sequenced using a nano-high-pressure liquid chromatography Agilent 1100 nanoflow system connected online to a 7-Tesla linear quadrupole ion trap-Fourier transform (FT) mass spectrometer (Thermo Electron, Bremen, Germany) essentially as described (Olsen *et al*, 2004). Proteins were identified using the Mascot search algorithm (Matrix Science) with IPI human database v3.37, which contains 69 164 unique entries. First ranked peptides (Mascot peptide scores ≥20) were parsed from Mascot database search html files with MSQuant (http://www.msquant.sourceforge.net) to generate unique first ranked peptide lists that were further filtered for absolute calibrated mass error <5 and peptide delta score >5. To remove redundancy at the protein level and to uniquely assign peptides to one IPI Db entry, the peptides were remapped to the IPI human Db using the program Protein Coverage Summarizer (http://ncrr.pnl.gov/software/) as described (Lasonder *et al*, 2008). To determine the protein abundance, we computed the exponentially modified Protein Abundance Index (emPAI) values for our data (Ishihama *et al*, 2005; Lasonder *et al*, 2008); emPAI values for all proteins were calculated as $10^{PAI}-1$ (PAI = observed peptides/observable peptides). The number of observable peptides was calculated by *in silico* tryptic digestion of the IPI human Db v3.37 using the program Protein Digestion Simulator (http://ncrr.pnl.gov/software/).

Purification of protein complexes from 293 cells and MudPIT analysis

To prepare protein samples for mass spectrometry, 293-HA-PRAME and control wild type cells were grown to 70–80% confluence in 10–20 roller bottles then washed, lysed and subjected to anti-HA agarose immunoaffinity chromatography as described (Mahrour *et al*, 2008). TCA-precipitated proteins were urea-denatured, reduced, alkylated and digested with endoproteinase Lys-C (Roche) followed by modified trypsin (Roche) as described (Florens & Washburn, 2006; Washburn *et al*, 2001). Peptide mixtures were loaded onto 100 µm fused silica microcapillary columns packed with 5-µm C$_{18}$ reverse phase (Aqua, Phenomenex), strong cation

exchange particles (Partisphere SCX, Whatman), and reverse phase. Loaded microcapillary columns were placed in-line with a Quaternary Agilent 1100 series HPLC pump and a LTQ ion trap mass spectrometer equipped with a nano-LC electrospray ionization source (ThermoFinnigan). Fully automated 10-step MudPIT runs were carried out on the electrosprayed peptides, as described (Florens & Washburn, 2006). Tandem mass (MS/MS) spectra were interpreted using SEQUEST against a database of 61 430 sequences, consisting of 37742 human proteins (downloaded from NCBI on 2008-03-04), 177 usual contaminants (such as human keratins, IgGs, and proteolytic enzymes), and to estimate false discovery rates, 30716 randomized amino acid sequences derived from each non-redundant protein entry. Peptide/spectrum matches were sorted and selected using DTASelect with the following criteria set: spectra/peptide matches were only retained if they had a DeltCn of at least 0,08, and minimum XCorr of 1,8 for singly-, 2,0 for doubly-, and 3,0 for triply-charged spectra. In addition, peptides had to be fully-tryptic and at least 7 amino acids long. Combining all runs, proteins had to be detected by at least 2 such peptides, or 1 peptide with 2 independent spectra. Under these criteria the final false discovery rate at the peptide level is 0,151 ± 0,009. Peptide hits from multiple runs were compared using CONTRAST. To estimate relative protein levels, we took advantage of evidence that Normalized Spectral Abundance Factors (NSAFs) vary with the relative abundance of proteins in a sample. NSAFs were calculated for each detected protein, using the formula:

$$(NSAF)_k = \frac{(SpC/Length)_k}{\sum_{i=1}^{N}(SpC/Length)_i}$$

where SpC = spectral count, L = protein length in amino acids, and i = all proteins detected in the MudPIT runs, as described (Florens et al, 2006; Paoletti et al, 2006).

Expression of recombinant proteins in Sf21 insect cells

cDNAs encoding wild type or mutant PRAME, Cul2, Elongins B and C were subcloned into pBAcPAK8 and recombinant baculoviruses were generated. SF21 cells were cultured at 27°C in SF-900 medium supplemented with 10% fetal bovine serum. SF21 cells were co-infected with the recombinant baculoviruses indicated in the figures. 50 hours after infection, cells were collected, lysed, and processed as described (Mahrour et al, 2008).

In vitro ubiquitination assays

The PRAME-Elongin BC complex and the His-T7-Xpress-Cullin2/Myc-RBX1 complex were obtained by co-infecting Sf21 cells with baculoviruses encoding for each individual protein and purifying through Flag tag or nickel-agarose chromatography, respectively. Different combinations of these complexes were then incubated for 1 h at 37°C with ~50 nM Uba1, ~0,2 μM UbcH5a, and 0,25 mg/ml ubiquitin, in 20 μl reactions containing 50 mM Tris-HCl, pH 7,6, 50 mM NaCl, 5 mM MgCl$_2$, 0.5 mM EDTA (pH 7,9), 5% glycerol, 2 mM ATP, 0.5 mM dithiothreitol, 0.3 units/ml pyrophosphatase, 60 mM creatine phosphate, and 0,3 mg/ml creatine phosphokinase. N-terminally (His)6-tagged mouse E1 was expressed in Sf21 insect cells and purified by nickel-agarose chromatography as described. Human UbcH5a with an N-terminal 6-histidine tag and a C-terminal Flag tag were expressed in E. coli strain BL21(DE3) and purified by nickel agarose.

ChIP and qPCR

Chromatin harvests, ChIPs, and qPCR analyses were performed as previously described (Denissov et al, 2007) with the following minor adjustments. Cells were crosslinked for 15 min in 1% formaldehyde, they were resuspended in ChIP incubation buffer at a concentration of 16 millions/ml and sonicated using a Bioruptor sonicator (Diagenode) for 15 min with settings: high power, 30 s ON, 30 s OFF. For every ChIP reaction, 200 μl of sonicated chromatin were used. The immunoprecipitated chromatin was eluted with 200 μl of elution buffer at room temperature for 20 minutes and DNA was purified with QIAquick PCR purification kit (QIAGEN). For ChIP-seq, DNA eluted from 5-7 ChIP reactions was pooled and the DNA concentration was measured with a Qubit fluorometer and Quant-iT dsDNA HS Assay (Invitrogen).

The following antibodies were used for ChIP: preimmune serum (preV27, "no antibody" control) and PRAME antiserum (V27); NFYA (Santa Cruz H-209, sc-10779); NFYB (pAb-007-100) and BB2 (against TY1 epitope) were from Diagenode.

For sequential ChIP experiments (ChIP-re-ChIP), in order to avoid antibody carry-over, the first ChIP was performed with antibody covalently crosslinked to ProteinA/G-Sepharose beads (Santa Cruz) with DMS as described above. Immunoprecipitated chromatin was then eluted in a smaller volume and further processed as a normal ChIP with PRAME immune serum or the preimmune serum as negative control. Real-time qPCR was performed using the SYBR Green mix (Biorad) with the MyIQ thermocycler (Bio-Rad). Specific primers for qPCR (Biolegio) were designed as previously described (Denissov *et al*, 2007) and are listed in Supplementary Table 2. Primers amplifying exon 12 of *albumin* were used as negative control.

RNA-seq library preparation and data analysis

Total RNA was extracted from K562 cells with the RNeasy kit and on-column DNase treatment (Qiagen). 100 μg of total RNA were subjected to two rounds of poly(A) selection using the Oligotex mRNA Mini Kit (Qiagen) and the concentration of poly(A)+ RNA was measured with a Qubit fluorometer (Invitrogen). 200 ng of purified poly(A)+ RNA were diluted in 160 μl RNAse-free water and fragmented by addition of 40 μl 5x fragmentation buffer (200 mM Tris acetate pH 8,2, 500 mM potassium acetate and 150 mM magnesium acetate) and incubation at 94°C for 120 seconds. After purification with RNeasy Minelute Kit (Qiagen), the quality of fragmented RNA was analyzed on an Experion (Bio-Rad). Double stranded cDNA was synthesized from the fragmented RNA with SuperscriptIII (Invitrogen) using random hexamers, and then used for Illumina sample prepping and sequencing (see below).

A total of 7 246 489 RNA-seq reads were uniquely mapped to hg18 and used for bioinformatical analysis. Normalized gene expression values (NE) for RefSeq genes were computed with Genomatix Region Miner tools.

Illumina high-throughput sequencing

A starting amount of 10-20 ng of ChIP DNA or double stranded cDNA was used for sample prepping according to standard protocols. Briefly, end repair was performed with T4 DNA ligase, Klenow DNA polymerase and T4 PNK; a 3′ protruding adenosine was added using Klenow DNA polymerase, and T-overhang adaptors were ligated. DNA molecules with a size of 270-300 bp were selected from gel, purified and amplified with 14 cycles of PCR. Quality controls were performed by qPCR of DNA before and after sample prepping, and by analysis of the PCR products with an Experion (Bio-Rad). Cluster generation and sequencing-by-synthesis (35bp) were performed using the Illumina Genome Analyzer II according to standard protocols (Illumina). The image files were processed to extract DNA sequence data, and sequences were mapped to the human hg18 reference genome using the eland program allowing at most one base mismatch. Sequence reads were filtered based on quality controls and sequences that could not be mapped uniquely on the reference genome were discarded. For visualization purposes, ChIP-seq reads were directionally extended to the length of the fragments used for sequencing and the number of overlapping sequence reads was determined for each base pair in the genome, averaged over a 10 bp window, and visualized in the University of California Santa Cruz genome browser (http://genome.ucsc.edu). ChIP-seq and RNA-seq data are published in the GEO database (GSE26439), and in the published history "PRAME" in Galaxy (http://main.g2.bx.psu.edu/).

ChIP-seq analysis and peak detection

Bioinformatical analysis was performed using a combination of in-house developed scripts, a Genomatix Genome Analyzer server (www.genomatix.de), and Galaxy tools (http://main.g2.bx.psu.edu/).

ChIP-seq experiments generated the following number of uniquely-mapped reads: 5,5 millions for PRAME, 5,6 millions for preimmune (control), 8,9 millions for NFYA, and 8,8 millions for NFYB. In order to compensate for the small difference in sequencing depth between the PRAME and preimmune datasets, the number of tags of preimmune ChIP-seq was equalized to the number of tags of PRAME ChIP-seq by uniform removal of tags as described (Nielsen *et al*, 2008). PRAME binding sites (peaks) were subsequently identified with the MACS peak calling algorithm (Zhang *et al*, 2008) using the two-sample method with a p-value (e) of 10^{-6}, which compared the PRAME profile to the preimmune (control) profile. NFYA and NFYB binding sites were determined with MACS using the standard

method with extended lambda (10 kb, 50 kb, 100 kb) and p-value (e) of 10^{-6}. MACS identified a total of 141 32 PRAME peaks, 25 151 NFYA peaks, and 16350 NFYB peaks. NFYA and NFYB peaks were combined in a total of 28 844 NFY binding sites that were tested for tag density. The number of sequence reads were counted in 500bp regions centered on the summits of the 28844 NFY binding sites (with the summit being the site of highest enrichment as determined by MACS), and the data were visualized in dot plots using R.

Location analysis and motif discovery

To analyze the genomic distribution of the binding sites identified, we first determined the fraction of peak summits localized within a 2 kb window centered around the TSS of primary transcripts (TSS ± 1 kb) using the Genomatix Genome Inspector. The genomic distribution of the binding sites located outside promoter regions ("distal sites") relative to genes was analyzed with PinkThing (http://pinkthing.cmbi.ru.nl/). The background distribution of genomic DNA was determined applying the same approach to pools of random genomic sites (n > 33 000).

The distances of promoter-associated binding sites from TSSs were computed with the Genomatix Genome Inspector tool, and plots were generated with R.

The localization of PRAME at enhancer elements was determined by overlaps between PRAME peaks summits and enhancer regions recently predicted for K562 cells (Heintzman *et al*, 2009). The genomic coordinates of the published enhancer regions were converted from hg17 to hg18 with the Galaxy Lift-over function, resulting in 24 545 enhancers.

De novo motif discovery was performed using the GimmeMotifs pipeline (van Heeringen & Veenstra, 2010), which incorporates a number of widely-used motif analysis tools, including MDmodule, MEME, Weeder, MotifSampler, BioProspector. Based on the distribution of PRAME binding sites, we performed motif searches on the DNA sequences underlying PRAME summits located within -400 bp and +100 bp of a primary TSS (n=4981). A CCAAT motif scored the highest values of enrichment and MNCP scores, as compared to pools of background sequences with the same genomic distribution.

ENCODE datasets

We retrieved and analyzed publicly-available ChIP-seq datasets of H3K9ac, H3K4me3, H3K36me3, H3K27me3, and RNA-PolII generated from K562 cells by the "ENCODE Chromatin Group at Broad Institute and Massachusetts General Hospital" (http://hgdownload.cse.ucsc.edu/goldenPath/hg18/encodeDCC/wgEncodeBroadChipSeq/).

Starting from 27 903 RefSeq genes with a RNA-seq expression value, we identified 7 214 genes longer than 1 kb with a NFY peak summit in the proximal promoter region. Next, we quantified the signals of PRAME in a 500 bp window under the NFY binding site, H3K9ac and H3K4me3 in a 3 kb window, H3K36me3 and H3K27 in the gene body, and RNA-PolII in the gene body excluding the first 500 bp.

Immobilized template assay method

A 5'-biotinylated DNA fragment of ~670 bp containing nucleotides -280 to +150 (relative to the annotated TSS) of the human HSPA5 promoter downstream of ~240 of DNA derived from the pGEM-T Easy Vector (Promega) was immobilized on streptavidin-coupled magnetic beads (Dynabeads M-280 Streptavidin, Invitrogen) following the manufacturer's instructions. This DNA fragment contains 3 ER stress elements (ERSE), each of which includes contains binding sites for NFY and ATF6. Reactions containing different combinations of Hela nuclear extract, Flag-PRAME complex (~50ng), and NFY (~100ng) were incubated with ~200ng of immobilized DNA for 30 minutes at 30ºC in buffer containing 50mM NaCl, 40mM Hepes-NaOH pH 7.9, 15% Glycerol, 2mM $MgCl_2$, 0.15mM EDTA, 1mM DTT, 0.02% NP-40, and 0.1mg/ml BSA. The magnetic beads were washed three times in the same buffer without BSA. Bound proteins were eluted by boiling the beads in 1X SDS loading buffer and analyzed by western blotting. NFY was purified by anti-Flag immunoaffinity chromatography from Sf21 cells coinfected with baculoviruses encoding Flag- NFYA, Myc-NFYB, and HA-NFYC. Flag-PRAME and PRAME-associated proteins (including Elongins B and C, Cul2, and Rbx1) were purified using anti-Flag agarose from nuclear extracts from 293-FRT cells stably expressing Flag-PRAME. NE, HeLa nuclear extract.

ACKNOWLEDGEMENTS

REFERENCES

Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. *Cell* **129**: 823-837

Bhattacharya A, Deng JM, Zhang Z, Behringer R, de Crombrugghe B, Maity SN (2003) The B subunit of the CCAAT box binding transcription factor complex (CBF/NF-Y) is essential for early mouse development and cell proliferation. *Cancer Res* **63**: 8167-8172

Birtle Z, Goodstadt L, Ponting C (2005) Duplication and positive selection among hominin-specific PRAME genes. *BMC Genomics* **6**: 120

Bolognese F, Wasner M, Dohna CL, Gurtner A, Ronchi A, Muller H, Manni I, Mossner J, Piaggio G, Mantovani R, Engeland K (1999) The cyclin B2 promoter depends on NF-Y, a trimer whose CCAAT-binding activity is cell-cycle regulated. *Oncogene* **18**: 1845-1853

Bortvin A, Eggan K, Skaletsky H, Akutsu H, Berry DL, Yanagimachi R, Page DC, Jaenisch R (2003) Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. *Development* **130**: 1673-1680

Caretti G, Salsi V, Vecchi C, Imbriano C, Mantovani R (2003) Dynamic recruitment of NF-Y and histone acetyltransferases on cell-cycle promoters. *J Biol Chem* **278**: 30435-30440

Ceribelli M, Dolfini D, Merico D, Gatta R, Viganò AM, Pavesi G, Mantovani R (2008) The histone-like NF-Y is a bifunctional transcription factor. *Mol Cell Biol* **28**: 2047-2058

Cinelli P, Casanova EA, Uhlig S, Lochmatter P, Matsuda T, Yokota T, Rülicke T, Ledermann B, Bürki K (2008) Expression profiling in transgenic FVB/N embryonic stem cells overexpressing STAT3. *BMC Developmental Biology 2008 8:57* **8**: 57

Collins GA, Tansey WP (2006) The proteasome: a utility tool for transcription? *Curr Opin Genet Dev* **16**: 197-202

Conaway RC, Brower CS, Conaway JW (2002) Emerging roles of ubiquitin in transcription regulation. *Science* **296**: 1254-1258

Consortium RMGSaA, Consortium RMGSaA, Gibbs R, Rogers J, Katze M, Bumgarner R, Weinstock G, Mardis E, Remington K, Strausberg R, Venter J, Wilson R, Batzer M, Bustamante C, Eichler E, Hahn M, Hardison R, Makova K, Miller W, Milosavljevic A *et al* (2007) Evolutionary and Biomedical Insights from the Rhesus Macaque Genome. *Science* **316**: 222

Currie RA (1998) NF-Y is associated with the histone acetyltransferases GCN5 and P/CAF. *J Biol Chem* **273**: 1430-1434

Dadé S, Callebaut I, Mermillod P, Monget P (2003) Identification of a new expanding family of genes characterized by atypical LRR domains. Localization of a cluster preferentially expressed in oocyte. *FEBS Letters* **555**: 533-538

De Santa F, Narang V, Yap ZH, Tusi BK, Burgold T, Austenaa L, Bucci G, Caganova M, Notarbartolo S, Casola S, Testa G, Sung W-K, Wei C-L, Natoli G (2009) Jmjd3 contributes to the control of gene expression in LPS-activated macrophages. *The EMBO Journal* **28**: 3341-3352

Denissov S, van Driel M, Voit R, Hekkelman M, Hulsen T, Hernandez N, Grummt I, Wehrens R, Stunnenberg H (2007) Identification of novel functional TBP-binding sites and general factor repertoires. *EMBO J* **26**: 944-954

Doolan, Clynes, Kennedy, Mehta, Crown, O'driscoll (2007) Prevalence and prognostic and predictive relevance of PRAME in breast cancer. *Breast Cancer Res Treat*

Epping MT, Hart AAM, Glas AM, Krijgsman O, Bernards R (2008) PRAME expression and clinical outcome of breast cancer. *Br J Cancer* **99**: 398-403

Epping MT, Wang L, Edel MJ, Carlée L, Hernandez M, Bernards R (2005) The human tumor antigen PRAME is a dominant repressor of retinoic acid receptor signaling. *Cell* **122**: 835-847

Florens L, Carozza M, Swanson S, Fournier M, Coleman M, Workman J, Washburn M (2006) Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. *Methods* **40**: 303-311

Florens L, Washburn M (2006) Proteomic analysis by multidimensional protein identification technology. *Methods Mol Biol* **328**: 159-175

Grskovic M, Chaivorapol C, Gaspar-Maia A, Li H, Ramalho-Santos M (2007) Systematic identification of cis-regulatory sequences active in mouse and human embryonic stem cells. *PLoS Genet* **3**: e145

Gurtner A, Fuschi P, Magi F, Colussi C, Gaetano C, Dobbelstein M, Sacchi A, Piaggio G (2008) NF-Y dependent epigenetic modifications discriminate between proliferating and postmitotic tissue. *PLoS ONE* **3**: e2047

Gurtner A, Manni I, Fuschi P, Mantovani R, Guadagni F, Sacchi A, Piaggio G (2003) Requirement for down-regulation of the CCAAT-binding activity of the NF-Y transcription factor during skeletal muscle differentiation. *Molecular Biology of the Cell* **14**: 2706-2715

Haqq C, Nosrati M, Sudilovsky D, Crothers J, Khodabakhsh D, Pulliam BL, Federman S, Miller JR, Allen RE, Singer MI, Leong SPL, Ljung B-M, Sagebiel RW, Kashani-Sabet M (2005) The gene expression signatures of melanoma progression. *Proc Natl Acad Sci USA* **102**: 6092-6097

Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, Ching KA, Antosiewicz-Bourget JE, Liu H, Zhang X, Green RD, Lobanenkov VV, Stewart R, Thomson JA, Crawford GE, Kellis M *et al* (2009) Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**: 108-112

Heintzman ND, Stuart RK, Hon GC, Fu Y, Ching CW, Hawkins RD, Barrera LO, Calcar Sv, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39**: 311-318

Ikeda H, Lethé B, Lehmann F, van Baren N, Baurain JF, de Smet C, Chambost H, Vitale M, Moretta A, Boon T, Coulie PG (1997) Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory receptor. *Immunity* **6**: 199-208

Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* **4**: 1265-1272

Kaelin WG (2007) Von Hippel-Lindau disease. Annual review of pathology 2: 145-173

Kamura T, Koepp DM, Conrad MN, Skowyra D, Moreland RJ, Iliopoulos O, Lane WS, Kaelin WG, Elledge SJ, Conaway RC, Harper JW, Conaway JW (1999) Rbx1, a component of the VHL tumor suppressor complex and SCF ubiquitin ligase. *Science* **284**: 657-661

Kamura T, Maenaka K, Kotoshiba S, Matsumoto M, Kohda D, Conaway RC, Conaway JW, Nakayama KI (2004) VHL-box and SOCS-box domains determine binding specificity for Cul2-Rbx1 and Cul5-Rbx2 modules of ubiquitin ligases. *Genes Dev* **18**: 3055-3065

Kamura T, Sato S, Haque D, Liu L, Kaelin WG, Conaway RC, Conaway JW (1998) The Elongin BC complex interacts with the conserved SOCS-box motif present in members of the SOCS, ras, WD-40 repeat, and ankyrin repeat families. *Genes Dev* **12**: 3872-3881

Kilpinen S, Autio R, Ojala K, Iljin K, Bucher E, Sara H, Pisto T, Saarela M, Skotheim RI, Björkman M, Mpindi J-P, Haapa-Paananen S, Vainio P, Edgren H, Wolf M, Astola J, Nees M, Hautaniemi S, Kallioniemi O (2008) Systematic bioinformatic analysis of expression levels of 17,330 human genes across 9,783 samples from 175 types of healthy and pathological tissues. *Genome Biol* **9**: R139

Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, Shilatifard A, Buratowski S, Greenblatt J (2003) Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol* **23**: 4207-4218

Lasonder E, Janse CJ, van Gemert G-J, Mair GR, Vermunt AMW, Douradinha BG, van Noort V, Huynen MA, Luty AJF, Kroeze H, Khan SM, Sauerwein RW, Waters AP, Mann M, Stunnenberg HG (2008) Proteomic profiling of Plasmodium sporozoite maturation identifies new proteins essential for parasite development and infectivity. *PLoS Pathog* **4**: e1000195

Le Guezennec X, Vermeulen M, Brinkman AB, Hoeijmakers WAM, Cohen A, Lasonder E, Stunnenberg HG (2006) MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. *Mol Cell Biol* **26**: 843-851

Li Q, Herrler M, Landsberger N, Kaludov N, Ogryzko VV, Nakatani Y, Wolffe AP (1998) Xenopus NF-Y pre-sets chromatin to potentiate p300 and acetylation-responsive transcription from the Xenopus hsp70 promoter in vivo. *The EMBO Journal* **17**: 6300-6315

Lonergan KM, Iliopoulos O, Ohh M, Kamura T, Conaway RC, Conaway JW, Kaelin WG (1998) Regulation of hypoxia-inducible mRNAs by the von Hippel-Lindau tumor suppressor protein requires binding to complexes containing elongins B/C and Cul2. *Mol Cell Biol* **18**: 732-741

Mahrour N, Redwine WB, Florens L, Swanson SK, Martin-Brown S, Bradford WD, Staehling-Hampton K, Washburn MP, Conaway RC, Conaway JW (2008) Characterization of Cullin-box sequences that direct recruitment of Cul2-Rbx1 and Cul5-Rbx2 modules to Elongin BC-based ubiquitin ligases. *J Biol Chem* **283**: 8005-8013

Maity SN, de Crombrugghe B (1998) Role of the CCAAT-binding protein CBF/NF-Y in transcription. *Trends Biochem Sci* **23**: 174-178

Manni I, Caretti G, Artuso S, Gurtner A, Emiliozzi V, Sacchi A, Mantovani R, Piaggio G (2008) Posttranslational regulation of NF-YA modulates NF-Y transcriptional activity. *Molecular Biology of the Cell* **19**: 5203-5213

Miller T, Krogan NJ, Dover J, Erdjument-Bromage H, Tempst P, Johnston M, Greenblatt JF, Shilatifard A (2001) COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc Natl Acad Sci USA* **98**: 12902-12907

Muratani M, Kung C, Shokat KM, Tansey WP (2005) The F box protein Dsg1/Mdm30 is a transcriptional coactivator that stimulates Gal4 turnover and cotranscriptional mRNA processing. *Cell* **120**: 887-899

Nielsen R, Pedersen TA, Hagenbeek D, Moulos P, Siersbaek R, Megens E, Denissov S, Børgesen M, Francoijs K-J, Mandrup S, Stunnenberg HG (2008) Genome-wide profiling of PPARgamma:RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. *Genes Dev* **22**: 2953-2967

Oberthuer A, Hero B, Spitz R, Berthold F, Fischer M (2004) The tumor-associated antigen PRAME is universally expressed in high-stage neuroblastoma and associated with poor outcome. *Clin Cancer Res* **10**: 4307-4313

Olsen JV, Ong S-E, Mann M (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol Cell Proteomics* **3**: 608-614

Paoletti A, Parmely T, Tomomori-Sato C, Sato S, Zhu D, Conaway R, Conaway J, Florens L, Washburn M (2006) Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc Natl Acad Sci USA* **103**: 18928-18933

Partheen K, Levan K, Osterberg L, Claesson I, Fallenius G, Sundfeldt K, Horvath G (2008) Four potential biomarkers as prognostic factors in stage III serous ovarian adenocarcinomas. *Int J Cancer* **123**: 2130-2137

Paydas S (2008) Is everything known in all faces of iceberg in PRAME? *Leukemia Research* **32**: 1356-1357

Petroski MD, Deshaies RJ (2005) Function and regulation of cullin-RING ubiquitin ligases. *Nat Rev Mol Cell Biol* **6**: 9-20

Radich JP, Dai H, Mao M, Oehler V, Schelter J, Druker B, Sawyers C, Shah N, Stock W, Willman CL, Friend S, Linsley PS (2006) Gene expression changes associated with progression and response in chronic myeloid leukemia. *Proc Natl Acad Sci USA* **103**: 2794-2799

Reya T, Morrison SJ, Clarke MF, Weissman IL (2001) Stem cells, cancer, and cancer stem cells. *Nature* **414**: 105-111

Roman-Gomez J, Jimenez-Velasco A, Agirre X, Castillejo JA, Navarro G, Jose-Eneriz ES, Garate L, Cordeu L, Cervantes F, Prosper F, Heiniger A, Torres A (2007) Epigenetic regulation of PRAME gene in chronic myeloid leukemia. *Leukemia Research* **31**: 1521-1528

Salghetti SE, Caudy AA, Chenoweth JG, Tansey WP (2001) Regulation of transcriptional activation domain function by ubiquitin. *Science* **293**: 1651-1653

Salsi V, Caretti G, Wasner M, Reinhard W, Haugwitz U, Engeland K, Mantovani R (2003) Interactions between p300 and multiple NF-Y trimers govern cyclin B2 promoter function. *J Biol Chem* **278**: 6642-6650

Schenk T, Stengel S, Goellner S, Steinbach D, Saluz HP (2007) Hypomethylation of PRAME is responsible for its aberrant overexpression in human malignancies. *Genes Chromosomes Cancer* **46**: 796-804

Sinha S, Maity SN, Lu J, de Crombrugghe B (1995) Recombinant rat CBF-C, the third subunit of CBF/NFY, allows formation of a protein-DNA complex with CBF-A and CBF-B and with yeast HAP2 and HAP3. *Proc Natl Acad Sci USA* **92**: 1624-1628

Steinbach D, Hermann J, Viehmann S, Zintl F, Gruhn B (2002) Clinical implications of PRAME gene expression in childhood acute myeloid leukemia. *Cancer Genet Cytogenet* **133**: 118-123

Steinbach D, Pfaffendorf N, Wittig S, Gruhn B (2007) PRAME expression is not associated with down-regulation of retinoic acid signaling in primary acute myeloid leukemia. *Cancer Genet Cytogenet* **177**: 51-54

Tajeddine N, Gala J-L, Louis M, Van Schoor M, Tombal B, Gailly P (2005) Tumor-associated antigen preferentially expressed antigen of melanoma (PRAME) induces caspase-independent cell death in vitro and reduces tumorigenicity in vivo. *Cancer Res* **65**: 7348-7355

van Baren N, Chambost H, Ferrant A, Michaux L, Ikeda H, Millard I, Olive D, Boon T, Coulie PG (1998) PRAME, a gene encoding an antigen recognized on a human melanoma by cytolytic T cells, is expressed in acute leukaemia cells. *Br J Haematol* **102**: 1376-1379

van Heeringen SJ, Veenstra GJC (2010) GimmeMotifs: a de novo motif prediction pipeline for ChIP-sequencing experiments. Bioinformatics (Oxford, England)

Wang PJ, McCarrey JR, Yang F, Page DC (2001) An abundance of X-linked genes expressed in spermatogonia. *Nat Genet* **27**: 422-426

Washburn M, Wolters D, Yates J (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* **19**: 242-247

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137

Zhu J, Zhang Y, Joe GJ, Pompetti R, Emerson SG (2005) NF-Ya activates multiple hematopoietic stem cell (HSC) regulatory genes and promotes HSC self-renewal. *Proc Natl Acad Sci USA* **102**: 11728-11733

SUPPLEMENTARY INFORMATION
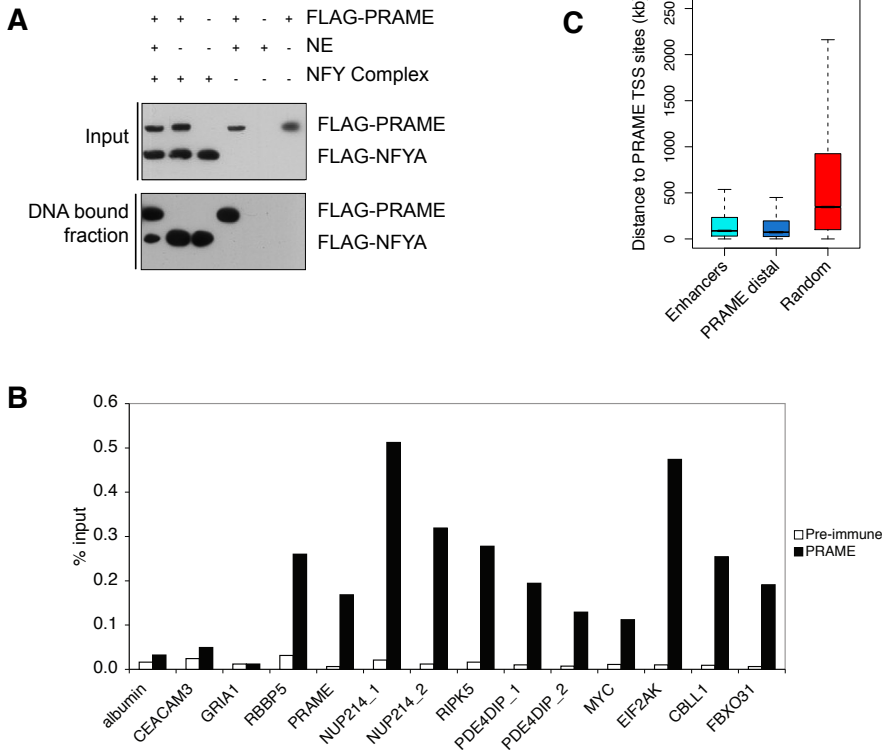
PRAME binding to DNA, ChIP-seq validation and analysis.
(A) PRAME binds a DNA fragment containing the NFY-dependent HSPA5 promoter by a mechanism that depends on activity(s) present in Hela nuclear extract. A DNA fragment containing the HSPA5 promoter was immobilized on Streptavidin-coupled magnetic beads and incubated with various combinations of Flag-PRAME complex, HeLa nuclear extract and recombinant NFY expressed in and purified from insect cells coinfected with

baculovirus vectors encoding Flag- NFYA, Myc-NFYB and HA-NFYC. After washing, bound proteins were eluted and analyzed by western blotting.
(B) Validation of PRAME binding sites by ChIP-qPCR. ChIP assays were performed with preimmune serum or anti-PRAME serum in K562 cells. The qPCR primers used targeted regions in the proximal promoters of the genes indicated, with the exception of albumin (exon 12). Albumin, CEACAM3 and GRIA1 are used as negative controls. For some

genes, two independent primer pairs were used.
(C) Distal PRAME binding sites are clustered in the vicinity of promoter-associated PRAME binding sites. The box plots clearly show that the PRAME binding sites outside promoters (n=8095) are not randomly distributed in the genome, but are instead significantly clustered close to promoter-associated PRAME binding sites (median distance: 73 kb vs. 346 kb for random sites), similarly to annotated enhancer regions.

SUPPLEMENTARY FIGURE 2

Motif found   **JASPAR match**   **Significance and enrichment vs. background**   **Positional preference plot**

De novo motif discovery at PRAME binding sites. Significant sequence motifs identified at PRAME binding sites at TSS. "Motif found" is the motif identified in the de novo discovery; "Jaspar match" is the annotated motif present in the Jaspar database that is most similar to the discovered motif (a p-value for the similarity test is shown); "significance and enrichment" report statistics of the enrichment of the discovered motif in the input sequences versus two different sets of backgrounds; the "positional preference plot" shows the frequency of the location of the discovered motif relative to the summit of PRAME binding sites.

A

B

NFY

- 5' (1-25kb)
- TSS +/- 1kb
- intron
- exon
- 3 ' (< 25kb)
- distant (>25 kb)

89

C

NFY ChIP-seq data analysis.
(A) Correlation between NFYA and NFYB binding detected by ChIP-seq. NFYA and NFYB tag densities were determined in 500 bp regions centered around the summits of NFYA and/or NFYB binding sites (n=28844). The data is displayed after sorting on NFYA (left) or NFYB (right) tag densities. The plots show that NFYA and NFYB signals are very well correlated for the majority of the binding sites. A subpopulation of binding sites shows high signals for NFYA, but low for NFYB (~5-10% of the sites, depending on the threshold applied), which might reflect dynamic in vivo binding of NFY subunits in the context of chromatin. PRAME binding sites didn't show a bias towards or against these regions. (B) Genomic distribution of NFY binding sites.
(C) NFYA, NFYB, and PRAME binding occur at the same site, and not generally within promoters. Each NFYA promoter summit is centred on the zero, and the distances to NFYB (green) or PRAME (blue) summits plotted as frequency density.

90

UCSC genome browser views of all genome-wide datasets combined.
Overviews of the active NFY loci TOP2A and STAT2, and of the inactive NFY loci TSPAN15 and SLC4A8.

Supplementary figure 5

BC-box **Cul2-box**

```
              20                        40                        60
 PRAME  MERRRLWGSIQSRYISMSVWTSPRRLVELAGQSLLKDEALAIAALELLPRELFPPLFMAA 60
homo_24  ----------------MSLQSPS-RLLELAGQSLLRNQFLTIFILDELPREVFPLMFMEA 43
 homo_2  ----------------MSIQAPP-RLLELAGQSLLRDQALSISAMEELPRVLYLPLFMEA 43
homo_31  ----------------MSLQSPS-RLLELAGQSLLRNQFLTIFILDELPREVFPLMFMEA 43
homo_11  ----------------MSLQSPS-RLLELAGQSLLRNQFLTIFILDELPREVFPLMFMEA 43
 homo_7  ----------------MSLQAPS-RLLELAGQSLLRNQFLTIFTLDELPREVFPLMFMEA 43
homo_22  ----------------MSIQAPP-RLLELAGQSLLRDQALSISAMEELPRVLYLPLFMEA 43
homo_29  ----------------MSIQAPP-RLLELAGQSLLRDQALSISAMEELPRVLYLPLFMEA 43
 homo_5  ----------------MSIQAPP-RLLELAGQSLLRDQALSISAMEELPRVLYLPLFREA 43
 homo_8  ----------------MSIRAPP-RLLELARQRLLRDQALAISTMEELPRELFPTLFMEA 43
homo_26  ----------------MSIRAPP-RLLELARQRLLRDQALAISTMEELPRELFPTLFMEA 43
homo_19  ----------------MSIRAPP-RLLELARQRLLRDQALAISTMEELPRELFPTLFMEA 43
 homo_1  ----------------MSLQAPP-RLLELAEQSLLRDRALAIPTLEELPRELFPPLFMEA 43
homo_21  ----------------MSIRTPP-RLLELAGRSLLRDQALAMSTLEELPTELFPPLFMEA 43
homo_28  ----------------MSIRTPP-RLLELAGRSLLRDQALAMSTLEELPTELFPPLFMEA 43
 homo_6  ----------------MSIWTPP-RLLELAGRSLLRDQALAMSTLEELPTELFPPLFMEA 43
homo_10  ----------------MSIRTPP-RLLELAGRSLLRDQALAMSTLEELPTELFPPLFMEA 43
homo_17  ----------------MSIRTPP-RLLELAGRSLLRDQALAMSTLEELPTELFPPLFMEA 43
homo_25  ----------------MSIRTPP-RLLELAGRSLLRDEALAISTLEELPTELFPPLFMEA 43
homo_32  ----------------MSIRTPP-RLLELAGRSLLRDEALAISTLEELPTELFPPLFMEA 43
homo_23  ----------------MSFQAP-RRLLELAGQSLLRDQALAISVLDELPRELFPRLFVEA 43
homo_30  ----------------MSFQAP-RRLLELAGQSLLRDQALAISVLDELPRELFPRLFVEA 43
homo_15  ----------------MSLQAP-RRLLELAGQSLLGDQALAISILDELPRELFPPLFVEA 43
homo_12  ----------------MSLQAP-RRLLELAGQSLLGDQALAISILDELPRELFPPLFVEA 43
 homo_3  ----------------MSIRIPP-RLLELAGXSLLRDQALAVSTLEELPTELFPPLFMEA 43
homo_13  ----------------MSIQAPP-RLLELAGQSLLRDKPLAISALEELPRELYLPLFLEA 43
homo_14  ----------------MSLQTPS-RLLELAGQSLMRKQFLTIFTDXLPREVSPLMFMEA 43
homo_16  ----------------MSLQSPS-RLLELAGQSLLRNQFLTIFILDELPREVFPLMFMEA 43
 homo_4  ----------------MSLQTPS-RLLELAGQSLLRNQFLTIFTLDELPREVSP-LFMEA 42
homo_27  ----------------MSIQAPP-RLLELAGQSLLRNQFLTIFTLDELPREVFPLMFMEA 43
homo_20  ----------------MSLQAPS-RLLELAGQSLLRNQFLTIFTLDELPREVFPLMFMEA 43
```

Complete conservation of the BC-box and Cul2-box in the PRAME-like protein family.

All human PRAME-like proteins share a conserved BC-box and Cul2-box at their N-terminus. The amino acid sequence of human PRAME was aligned with ClustalX to the previously published multiple align- ment of the human PRAME-like proteins encoded on human chromosome 1 [Birtle et al. 2005]. The alignment was edited with CLC Sequence Viewer to indicate dissimilar residues in grey, and to identical in yellow the identical residues within the BC- and Cul2-boxes.

Supplementary table 1   **Functional analysis of NFY genes.** E., enrichment; adj. P, adjusted P-value.

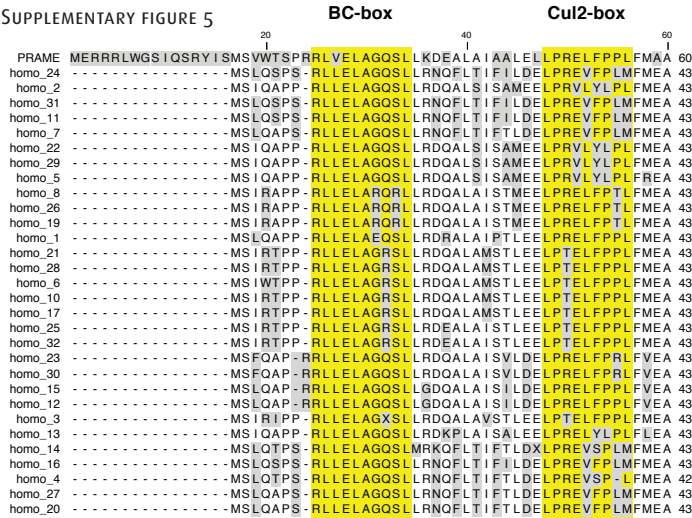| NFY and PRAME	E	Adj.P
Spliceosome	9.7	2.3E-18
Ribosome	10.2	6.5E-14
Metabolic pathways	2.5	2.1E-09
Huntington's disease	4.8	9.3E-08
Aminoacyl-tRNA biosynthesis	10.4	6.5E-07
Terpenoid backbone biosynthesis	17.0	1.3E-05
Endocytosis	3.9	3.8E-05
Oxidative phosphorylation	4.1	2.0E-04
Proteasome	7.1	2.0E-04
Oocyte meiosis	4.5	2.0E-04
Progesterone-mediated oocyte maturation	5.0	4.0E-04
Cell cycle	4.0	5.0E-04
Nucleotide excision repair	6.8	6.0E-04
MAPK signaling pathway	2.7	1.3E-03
Alzheimer's disease	3.3	1.3E-03
Mismatch repair	9.3	1.3E-03
DNA replication	7.1	1.3E-03
Parkinson's disease	3.5	1.9E-03
Homologous recombination	7.6	2.4E-03
RNA degradation	5.1	2.7E-03
Acute myeloid leukemia	5.0	2.7E-03
Ubiquitin mediated proteolysis	3.1	8.3E-03
Valine, leucine and isoleucine biosynthesis	11.6	9.4E-03
Vibrio cholerae infection	4.6	9.5E-03
Fc gamma R-mediated phagocytosis	3.5	9.6E-03
Selenoamino acid metabolism	6.6	1.3E-02
Purine metabolism	2.8	1.3E-02
NOD-like receptor signaling pathway	4.1	1.3E-02
RNA polymerase	5.9	1.7E-02
ErbB signaling pathway	3.4	1.7E-02
Glyoxylate and dicarboxylate metabolism	8.5	1.8E-02
Prostate cancer	3.4	1.8E-02
Pathways in cancer	2.1	1.9E-02
p53 signaling pathway	3.7	1.9E-02
Long-term depression	3.7	2.0E-02
Steroid biosynthesis	7.5	2.2E-02
mTOR signaling pathway	4.1	2.3E-02
Vitamin B6 metabolism	14.2	2.3E-02
Pyrimidine metabolism	3.0	2.5E-02

NFY-only	E	Adj.P
Metabolic pathways	3.0	1.7E-11
Pathways in cancer	3.8	6.6E-06
Wnt signaling pathway	5.0	1.0E-04
Hedgehog signaling pathway	8.2	2.0E-04
Basal cell carcinoma	7.3	1.2E-03
Melanogenesis	5.1	1.6E-03
Axon guidance	4.5	1.6E-03
Glycerolipid metabolism	7.7	1.6E-03
Antigen processing and presentation	5.2	2.8E-03
Arginine and proline metabolism	6.4	3.8E-03
Adherens junction	5.2	4.6E-03
Cell adhesion molecules (CAMs)	3.9	5.3E-03
MAPK signaling pathway	2.8	8.8E-03
PPAR signaling pathway	5.0	1.2E-02
Fc gamma R-mediated phagocytosis	4.2	1.2E-02
Glycerophospholipid metabolism	4.9	1.2E-02
Asthma	7.7	1.3E-02
Alanine, aspartate and glutamate metabolism	7.4	1.4E-02
Citrate cycle (TCA cycle)	7.2	1.4E-02
Endometrial cancer	5.5	1.4E-02
Vibrio cholerae infection	5.1	1.6E-02
Insulin signaling pathway	3.4	1.6E-02
Fructose and mannose metabolism	6.8	1.6E-02
Renin-angiotensin system	10.2	1.6E-02
Arachidonic acid metabolism	5.0	1.7E-02
Allograft rejection	6.1	2.1E-02
Leukocyte transendothelial migration	3.4	2.1E-02
Regulation of actin cytoskeleton	2.7	2.1E-02
Prostate cancer	3.9	2.1E-02
Gap junction	3.8	2.1E-02
Graft-versus-host disease	5.5	2.5E-02
Epithelial cell signaling in Helicobacter pylori infection	4.2	2.5E-02

alb_F	Albumin (negative ctrl)	TGAAACATACGTTCCCAAAGAGTTT
alb_R	Albumin (negative ctrl)	CTCTCCTTCTCAGAAAGTGGCATAT
#72_F	CEACAM3 (negative ctrl)	TCTCTGTGACCTCCTTGCTG
#72_R	CEACAM3 (negative ctrl)	GGTCCCTGAACCTGGTGTTA
#6_F	RBBP5	GCCCAAATATATGCCTGACC
#6_R	RBBP5	TTCCCGGAGCAAATAATCTT
#14_F	NUP214_1	CAAGCTTGACGTCACCAAAG
#14_R	NUP214_1	AGCAAACTTCCTCCCCTCAG
#15_F	NUP214_2	GCCACGCCACTACTCAGC
#15_R	NUP214_2	CGTCAAGCTTGGTCAGAGGA
#16_F	PRAME	GGGGAGCTGTACCCTGAAG
#16_R	PRAME	CTGGGAGGAAGTGGGTTTT
#23_F	RIPK5	AAGGATCGCTGGAGGACAC
#23_R	RIPK5	TTCTCCAGCCTTGAGGTTTT
#24_F	PDE4DIP_1	GACTCTGCTGTCACCCTCCT
#24_R	PDE4DIP_1	TCTTGGGGGTCAGTCTCTCT
#25_F	PDE4DIP_2	AAGTGGTGTAACCCGACTCC
#25_R	PDE4DIP_2	TCTCTCTTTTCCGAGAAGCTG
#46_F	MYC	AGGGATCGCGCTGAGTATAA
#46_R	MYC	CTGCCTCTCGCTGGAATTA
#47_F	EIF2AK	CGCTAACTGCCTCTTGATTG
#47_R	EIF2AK	GCACTTGCCTCCCACTCT
#48_F	CBLL1	TGAAGCACATGAGGAGAAGC
#48_R	CBLL1	GGGACTCAGGCATAGTTC
#49_F	FBXO31	AGAAAGTGGTGATGCCGAGT
#49_R	FBXO31	CAGCGCGTCAGTCTCATAAT
#92_F	CCBN2	CCGGCTGTTGTGACAATCAAA
#92_R	CCBN2	CAAATACCGCGTCGCTTGC

THE HUMAN EKC/KEOPS COMPLEX

Is Recruited to Cullin2 Ubiquitin Ligases by the Human Tumour Antigen PRAME

Adalberto Costessi[1], Nawel Mahrour[2], Vikram Sharma[1], Rieka Stunnenberg[1], Marieke A Stoel[1], Esther Tijchon[1], Joan W Conaway[2,3], Ronald C Conaway[2,3], Hendrik G Stunnenberg[1]

[1] Department of Molecular Biology, Nijmegen Centre for Molecular Life Sciences, Radboud University, 6500 HB, Nijmegen, The Netherlands
[2] Stowers Institute for Medical Research, 1000 E 50th Street, Kansas City, MO 64110, USA
[3] Department of Biochemistry and Molecular Biology, Kansas University Medical Center, Kansas City, KS 66160, USA

ABSTRACT

The human tumour antigen PRAME (preferentially expressed antigen in melanoma) is frequently overexpressed during oncogenesis, and high *PRAME* levels are associated with poor clinical outcome in a variety of cancers. However, the molecular pathways in which PRAME is implicated are not well understood. We recently characterized PRAME as a BC-box subunit of a Cullin2-based E3 ubiquitin ligase. In this study, we mined the PRAME interactome to a deeper level and identified specific interactions with OSGEP and LAGE3, which are human orthologues of the ancient EKC/KEOPS complex. By characterizing biochemically the human EKC complex and its interactions with PRAME, we show that PRAME recruits a Cul2 ubiquitin ligase to EKC. Moreover, EKC subunits associate with PRAME target sites on chromatin. Our data reveal a novel link between the oncoprotein PRAME and the conserved EKC complex and support a role for both complexes in the same pathways.

INTRODUCTION

The human oncoprotein PRAME (preferentially expressed antigen in mela-noma) was first identified and cloned as the antigen responsible for an anti-tu-mour immune response in a melanoma patient (Ikeda *et al*, 1997). Follow-up experiments revealed that *PRAME* is expressed at low levels in few normal adult tissues like adrenals, ovaries, and endometrium, and at high levels only in the testis (Ikeda *et al*, 1997; Kilpinen *et al*, 2008). However, overexpression of *PRAME* is frequently found in a wide variety of human cancers, including acute and chronic haematological tumours, synovial sarcoma, lung, breast, and renal carcinoma (Allander *et al*, 2002; Ikeda *et al*, 1997). Importantly, high *PRAME* levels were found to correlate with advanced stages of disease in melanoma (Haqq *et al*, 2005), neuroblastoma (Oberthuer *et al*, 2004), serous ovarian adenocarcinoma (Partheen *et al*, 2008), and chronic myeloid leukae-mia (Radich *et al*, 2006), and to constitute an independent prognostic factor of poor clinical outcome in breast cancer (Doolan *et al*, 2008; Epping *et al*, 2008). In contrast, high levels of *PRAME* were found to correlate with good prognosis in leukaemia cases carrying the t(15;17) PML-RAR translocation (acute promye-locitic leukaemia) (Santamaría *et al*, 2008).

Although these findings suggested a role for PRAME in human malignan-cies, the detailed molecular mechanisms and pathways involved are not yet clear. PRAME was reported to repress retinoic acid signaling in melanoma cell lines (Epping *et al*, 2005), but this was not confirmed for breast cancer or leu-kaemia cases (Steinbach *et al*, 2007; Epping *et al*, 2008). Conflicting reports on leukaemia cells suggested that PRAME might induce caspase-independent cell death (Tajeddine *et al*, 2005), or repress apoptosis-related genes to pro-mote cell survival (Tanaka *et al*, 2011).

Recently, through biochemical characterization of PRAME-containing protein complexes, we established that this oncoprotein is a component of Cullin2-based E3 ubiquitin ligases and belongs to the family of BC-box proteins, associating PRAME to a clear biochemical activity and pathway (Costessi *et al*, 2011). PRAME establishes direct interactions with other ligase subunits through conserved N-terminal motifs: a BC-box (aa. 25-34) mediates interactions with the ElonginB-ElonginC heterodimer, and a downstream Cul2-box (aa. 48-56) mediates interactions with the Cullin2 scaffold protein. Genome-wide chromatin immunoprecipitation experiments further revealed that Cul2-PRAME ubiquitin ligases specifically associate with active promot-

ers regulated by the transcription factor NFY and with proximal enhancers (Costessi *et al*, 2011).

Two independent laboratories have identified an ancient and highly conserved multiprotein complex named KEOPS (Downey *et al*, 2006) or EKC (Kisseleva-Romanova *et al*, 2006), which has orthologues from Archaea to Eukarya and has been implicated in telomeres maintenance, transcriptional regulation, and t⁶A modification of tRNAs. Yeast EKC comprises four subunits which are also conserved in the human genome (human orthologues are indicated in brackets): Pcc1p (LAGE3, also known as ESO3), the ATPase Kae1p (OSGEP), the kinase Bud32p (TP53RK, also known as PRPK), and Cgi121p (TPRKB). In addition, yeast EKC also includes Gon7p (also known as Pcc2p), which appears to be fungi-specific (Kisseleva-Romanova *et al*, 2006).

Intriguingly, the OSGEP subunit is also present in bacteria (YgjD) and eukaryotic genomes express an OSGEP paralogue (Qri7/OSGEPL1) that localizes to mitochondria (Reinders *et al*, 2006). Comparative genomic studies identified OSGEP as one of the very few genes present in all genomes sequenced so far (Galperin, 2008), suggesting an extremely conserved function. Very recently, several research groups have reported a crucial role for the YgjD/Kae1/OSGEP protein family in the biosynthesis of N6-threonylcarbamoyl adenosine (t⁶A) (Daugeron *et al*, 2011; Yacoubi *et al*, 2011; Srinivasan *et al*, 2011): a universal modification at position 37 of tRNAs decoding ANN codons, which is required for accurate translation of messenger RNAs (Yarian *et al*, 2002).

Human LAGE3 belongs to the NY-ESO gene family together with the closely related LAGE1 and LAGE2 (Alpen *et al*, 2002), and all three genes are clustered in the same region on chromosome X. While LAGE3 is ubiquitously expressed, LAGE1 and LAGE2 are cancer-testis antigens with high expression in healthy testis and upregulation in a number of cancer tissues, similarly to PRAME.

The aim of the present study was to mine the protein-protein inter-actome of the PRAME oncoprotein to a deeper level by protein complex purification and mass spectrometry. Our experiments revealed that PRAME specifically interacts with the EKC complex in human cells and that it recruits Cul2-based E3 ubiquitin ligases to EKC. We further show that EKC subunits co-localize with PRAME at the promoter regions of transcriptionally active human genes, supporting a common functional role for these complexes on chromatin. Overall, our findings provide novel and important insights in the molecular pathways in which the oncoprotein PRAME is active.

RESULTS

PRAME interacts with the EKC/KEOPS complex

We have recently applied biochemical purification strategies and characterized PRAME as a component of Cullin2 E3 ubiquitin ligases (Costessi *et al*, 2011). Here, we improved the immunoaffinity purification protocols and mined the PRAME interactome to a deeper level (see materials and methods). We isolated epitope-tagged PRAME (StrepII-Myc-3xHA-PRAME, referred to as TAG-PRAME) from K562 cells, which express high levels of endogenous PRAME, and from HeLaS3 cells, which express endogenous PRAME at very low to undetectable levels. As expected, mass spectrometry analysis of HA immunoprecipitates detected Cullin2 ligase components with the highest emPAI: Cullin2, Elongin B, Elongin C, and Rbx1 (Table 1). Notably, western blot analysis showed that Cullin2 and Elongin C were expressed to similar levels in nuclear extracts from the two cell lines (not shown).

Furthermore, a number of additional proteins were specifically detected in eluates of K562-TAG-PRAME and HeLaS3-TAG-PRAME cells, but not in the control parental cell lines. These included OSGEP (O-sialoglycoprotein endopeptidase) and LAGE3 (L antigen family member 3), which are the human orthologues of the Kae1 and Pcc1 subunits of the recently described EKC complex (Table 1). OSGEP and LAGE3 were present with a similar abundance in TAG-PRAME eluates from both K562 and HeLaS3 cell lines, consistent with similar expression levels. However, in these eluates we did not identify peptides matching TP53RK or TPRKB, which are the predicted human orthologues of the EKC subunits Bud32 and Cgi121, respectively.

Other potential interactors identified in PRAME eluates were BAG2 and UBL4A (Table 1). BAG2 (BAG family molecular chaperone regulator 2) is a nucleotide exchange factor for the chaperone protein Hsp70 that contributes to the processing of misfolded proteins (Xu *et al*, 2008). Similarly, UBL4A (Ubiquitin-like protein 4A) is part of a chaperone protein complex that promotes correct membrane targeting of tail-anchored proteins (Wang *et al*, 2011; Mariappan *et al*, 2010). Peptides for both BAG2 and UBL4A were identified in all anti-HA eluates from K562-TAG-PRAME cells, but were not detected in the anti-MYC eluates. Additional experiments are therefore needed to establish whether these proteins are real interactors of PRAME or rather false positives. For the purpose of this study, we further concentrated on EKC subunits.

TABLE 1

Protein ID	Protein description	AVG emPAI	K562 cells									HeLaS3 cells			
			PRAME MYC 1	PRAME MYC 2	PRAME HA 1	PRAME HA 2	PRAME HA 3	Ctrl MYC 1	Ctrl MYC 2	Ctrl HA 1	Ctrl HA 2	PRAME nucl	PRAME cyto	Ctrl nucl	Ctrl cyto
IPI00026670.3	**TCEB2 Transcription elongation factor B polypeptide 2**	7.00	5	4	8	7	6	0	2	1	0	6	8	2	2
IPI00300341.5	**TCEB1 Transcription elongation factor B polypeptide 1**	5.01	6	6	4	8	6	1	0	1	0	7	7	4	1
IPI00019282.1	**PRAME Melanoma antigen preferentially expressed in tumors**	4.01	6	12	27	18	12	0	0	0	0	18	30	0	0
IPI00014311.4	**CUL2 Cullin-2**	2.05	12	22	34	31	23	0	6	2	0	18	30	0	0
IPI00015809.1	OSGEP Probable O-sialoglycoprotein endopeptidase	0.91	4	3	6	7	6	0	0	0	1	1	6	0	0
IPI00032314.2	LAGE3 L antigen family member 3	0.89	1	0	3	3	3	0	0	0	0	1	4	0	0
IPI00000643.1	BAG2 BAG family molecular chaperone regulator 2	0.77	0	0	9	5	1	0	0	1	0	5	6	4	0
IPI00003386.3	**RBX1 RING-box protein 1**	0.62	1	0	1	1	1	0	1	0	0	0	0	0	0
IPI00395674.1	SNRPB Isoform SM-B of Small nuclear ribonucleoprotein-associated proteins B and B'	0.59	4	5	1	1	2	0	4	0	0	1	4	0	0
IPI00221089.5	RPS13 40S ribosomal protein S13	0.58	0	3	3	1	4	0	4	1	0	3	4	7	0
IPI00005658.3	UBL4A Ubiquitin-like protein 4A	0.57	0	0	5	2	3	0	0	2	0	7	4	2	0
IPI00020042.2	PSMC4 Isoform 1 of 26S protease regulatory subunit 6B	0.54	0	0	13	1	1	0	0	13	1	2	0	0	0
IPI00383296.5	HNRPM Isoform 2 of Heterogeneous nuclear ribonucleoprotein M	0.48	0	16	11	11	6	0	12	1	0	1	0	0	0
IPI00003881.5	HNRPF Heterogeneous nuclear ribonucleoprotein F	0.48	2	7	6	1	1	0	7	0	0	6	0	0	1
IPI00219153.4	RPL22 60S ribosomal protein L22	0.47	1	1	1	1	1	0	1	0	0	0	1	1	0
IPI00017381.1	RFC4 Replication factor C subunit 4	0.46	0	4	9	3	6	0	0	7	1	11	1	0	3
IPI00477313.3	HNRNPC Isoform C2 of Heterogeneous nuclear ribonucleoproteins C1/C2	0.45	2	6	5	2	3	0	1	3	0	0	1	2	0
IPI00456631.5	AOF2 Isoform 1 of Lysine-specific histone demethylase 1	0.38	0	4	14	12	1	0	0	1	0	6	0	3	0
IPI00554723.5	RPL10 60S ribosomal protein L10	0.34	0	0	4	2	2	0	0	3	0	2	11	4	0
IPI00013881.6	HNRPH1 Heterogeneous nuclear ribonucleoprotein H	0.32	3	5	4	3	1	1	3	0	0	6	0	3	1

PRAME-interacting proteins identified by Mass Spectrometry of purified complexes from K562 and HeLaS3 cells. PRAME-associating proteins were purified from nuclear extracts of K562-TAG-PRAME by anti-MYC or anti-HA immunoprecipitation and from HeLaS3-TAG-PRAME cells by anti-HA immunoprecipitation and subjected to mass spectrometry. Immunoprecipitations from wild type K562 and HeLaS cells were used as a control for immunospecificity of the purifications. Components of Cullin2 ligases and EKC complex are indicated in bold.

The specificity of the interactions between PRAME and EKC subunits was confirmed by transient transfections in HEK293T cells and coimmunoprecipitation experiments of TAG-PRAME with OSGEP and LAGE3 fused to a TTE (TY1-TY1-ER) epitope tag (Fig. 1A). We also used a baculovirus expression system to co-express FLAG-PRAME with different combinations of the four human EKC subunits in Sf21 insect cells. Consistent with the transient transfection experiments, we observed that PRAME could bind to OSGEP and LAGE3 independent of TP53RK and TPRKB. Interestingly, although both OSGEP and LAGE3

FIGURE 1

PRAME interacts with OSGEP and LAGE3.
(A) Coimmunoprecipitation assays verify the interaction of PRAME with the EKC subunits OSGEP and LAGE3. Constructs expressing the indicated proteins were transiently transfected in 293T cells. Anti-TAG immunoprecipitations were performed with rabbit HA antibody (Abcam). TAG-PRAME was detected with monoclonal anti-HA (Covance) and TTE-tagged proteins with monoclonal BB2 (Diagenode) which recognizes the TY1 tag. The lower panel shows a scheme of the tagged proteins used (tags and coding sequences are not on scale).
(B-C) PRAME directly interacts with EKC complex through OSGEP and LAGE3. SF21 cells were co-infected with baculoviruses expressing the indicated proteins. Anti-FLAG immunoprecipitations were performed, and total cell lysates and eluates were analyzed by western blotting. Panels showing anti-HA immunoblots are from the same exposure of the same blot, as are the panels showing anti-cMyc immunoblots. Asterisks indicate light chains of the antibodies used in the immunoprecipitation.
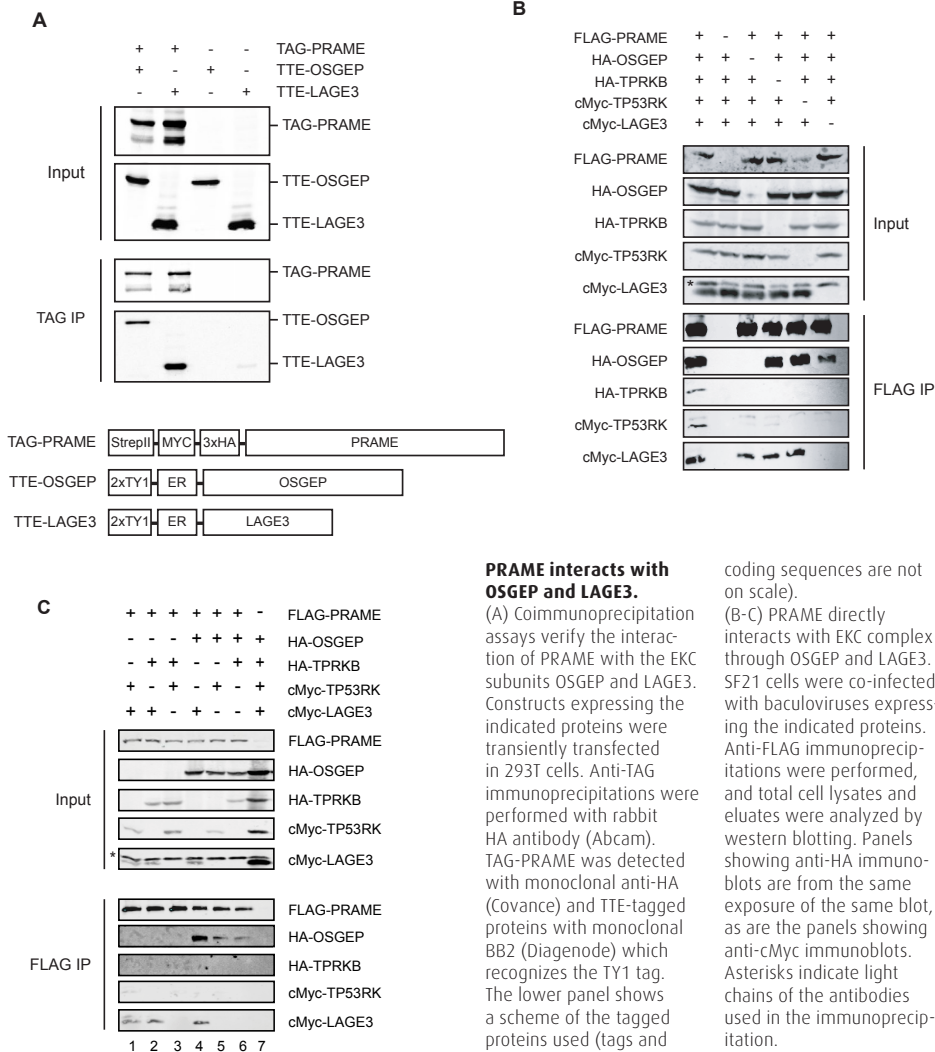
TABLE 2

Protein ID	Protein description	K562 LAGE3	K562 OSGEP	K562 TP53RK	K562 TPRKB	HeLa LAGE3	K562 ctrl 1	K562 ctrl 2	K562 ctrl 3	HeLa ctrl	K562 LAGE3 emPAI	K562 OSGEP emPAI	K562 TP53RK emPAI	K562 TPRKB emPAI	HeLa LAGE3 emPAI
IPI00032314.2	**LAGE3 L antigen family member 3**	6	8	8	6	7	0	0	0	0	4.62	9.00	9.00	4.62	6.50
IPI00015809.1	**OSGEP Probable O-sialoglycoprotein endopeptidase**	12	17	17	12	21	0	1	0	0	3.28	6.85	6.85	3.28	11.74
IPI00290305.3	**TP53RK TP53-regulating kinase**	8	10	15	9	7	0	0	0	0	2.73	4.18	10.79	3.39	2.16
IPI00301432.3	**TPRKB Isoform 1 of TP53RK-binding protein**	5	11	15	8	6	0	0	0	0	2.16	11.59	30.62	5.31	2.98
IPI00026670.3	**TCEB2 Transcription elongation factor B polypeptide 2**	3	8	1	2	2	1	0	4	3	1.68	12.90	0.39	0.93	0.93
IPI00014311.4	**CUL2 Cullin-2**	21	0	1	2	0	2	0	8	0	1.49	0.00	0.04	0.00	0.00
IPI00019282.1	**PRAME Melanoma antigen preferentially expressed in tumors**	6	18	3	0	0	0	0	0	0	0.74	4.25	0.32	0.00	0.00
IPI00646167.2	C14orf142 hypothetical protein LOC84520	1	2	1	0	3	0	0	0	0	0.59	1.51	0.59	0.00	2.98
IPI00007797.3	FABP5;FABP5L7 Fatty acid-binding protein, epidermal	2	0	0	1	0	0	0	2	0	0.59	0.00	0.00	0.26	0.00
IPI00005658.3	UBL4A Ubiquitin-like protein 4A	2	8	6	3	5	2	0	2	2	0.47	3.64	2.16	0.78	1.61
IPI00019912.3	HSD17B4 Peroxisomal multifunctional enzyme type 2	7	0	10	3	0	2	0	0	0	0.44	0.00	0.69	0.17	0.00
IPI00395865.4	RBBP7 Histone-binding protein RBBP7	3	6	1	2	8	2	2	3	0	0.44	1.07	0.13	0.27	1.64
IPI00304903.4	SPRR1B Cornifin-B	1	0	0	0	0	1	0	0	0	0.39	0.00	0.00	0.00	0.00
IPI00742943.1	NUP43 Nucleoporin Nup43	2	0	0	0	1	0	0	0	0	0.36	0.00	0.00	0.00	0.00
IPI00300341.5	**TCEB1 Transcription elongation factor B polypeptide 1**	1	6	3	1	1	1	0	2	1	0.33	4.62	1.37	0.33	0.17
IPI00293434.2	SRP14 Signal recognition particle 14 kDa protein	1	2	3	0	1	0	0	2	0	0.33	0.78	1.37	0.00	0.00
IPI00446354.1	- CDNA FLJ41805 fis, clone NOVAR2000962	1	0	0	0	0	0	0	0	0	0.33	0.00	0.00	0.00	0.00
IPI00020025.3	SH3PXD2B CDNA FLJ20831 fis, clone ADKA03080	1	0	0	0	0	0	0	0	0	0.23	0.00	0.00	0.00	0.00
IPI00027626.3	CCT6A T-complex protein 1 subunit zeta	3	24	11	8	8	0	1	0	3	0.22	3.85	1.06	0.69	0.69
IPI00869118.2	- Anti-(ED-B) scFv (Fragment)	1	0	1	1	0	0	1	1	0	0.21	0.00	0.21	0.21	0.00
IPI00386765.2	PDE4D Isoform 7 of cAMP-specific 3',5'-cyclic phosphodiesterase 4D	1	0	0	1	0	0	0	0	0	0.21	0.00	0.00	0.21	0.00
IPI00007423.1	ANP32B Isoform 1 of Acidic leucine-rich nuclear phosphoprotein 32 family member B	1	0	6	2	0	0	0	4	0	0.19	0.00	1.89	0.43	0.00

Human EKC-interacting proteins identified by Mass Spectrometry of purified complexes from nuclear extracts of K562 and HeLaS3 cells. Proteins interacting with each of human EKC subunits were identified by anti-HA immunoprecipitations and mass spectrometry on the cell lines indicated. Components of Cullin2 ligases and EKC complex are indicated in bold.

could bind PRAME independent of the other (Fig. 1B, lanes 3 and 6), LAGE3 appears to enhance the PRAME-OSGEP interaction, since substantially more OS-GEP bound to PRAME when LAGE3 was also expressed (Fig 1B, compare lanes 1, 4, and 5 to lane 6, and Fig. 1C, compare lane 4 to lanes 5 and 6). Finally, we observed small amounts of TP53RK and TPRKB copurifying with FLAG-PRAME dependent on both OSGEP and LAGE3 (Fig. 1B, lanes 1, 3, and 6).

Taken together, our results identified stable interactions between PRAME and human orthologues of the EKC/KEOPS complex.

EKC purifications from K562 cells

We next studied the expression patterns of EKC, Cullin2 ligase subunits and *PRAME* in normal and cancer cells by querying the GeneSapiens gene expression database (Kilpinen *et al*, 2008). This database contains mRNA expression data of most human genes across almost 10000 array experiments from 175 healthy and pathological tissues. Consistent with the literature, we found that PRAME is expressed only in very few normal adult tissues while it is upregulated in several cancers (Suppl. Fig. S1). On the contrary, *LAGE3*, *OSGEP*, *TP53RK*, *TPRKB*, and *CUL2* seem to be expressed in most if not all tissues (Suppl. Figures S1-S4).

To explore further the protein-protein interactions between PRAME and EKC, we generated K562 cell lines stably expressing epitope-tagged versions of each of the four human predicted human EKC subunits and performed protein complex purifications. As shown in Table 2, each EKC subunit co-purified with the other three subunits of the complex, indicating that a complete EKC complex is present in human cells.

Furthermore, TAG-OSGEP and TAG-LAGE3 (but not TAG-TP53RK or TAG-TPRKB) copurified with significant amounts of endogenous PRAME, consistent with TAG-PRAME purifications. Remarkably, TAG-OSGEP and TAG-LAGE3 associated also with Cullin2 ligase components (Table 2). The few peptides for Cul2-PRAME components identified in TAG-TP53RK and TAG-TPRKB eluates were not supported by western blot of the same samples (not shown).

Interestingly, peptides matching the putative protein C14orf142 were specifically identified in eluates from OSGEP, LAGE3 and TP53RK (Table 2). Similarly to LAGE3 and TPRKB, C14orf142 is predicted to be a small protein without clear structural domains. It is tempting to speculate that this protein might constitute a novel EKC interactor in human cells.

Although we did not identify peptides matching Cul2 by mass spectrometry of TAG-OSGEP eluates, we did detect small amounts of Cul2 by western

FIGURE 2
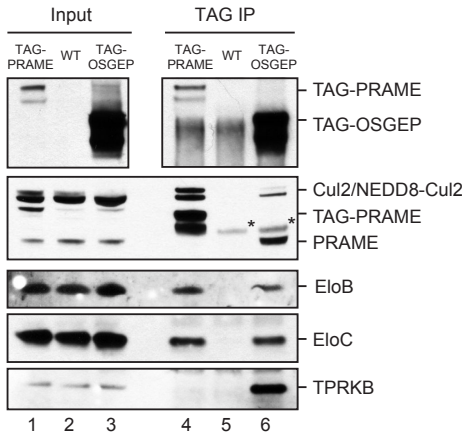
OSGEP interacts with PRAME and Cul2 ubiquitin complex components.
OSGEP interacts with PRAME and Cul2-Elo-BC ligases. Immunoblot analysis of TAG-PRAME and TAG-OSGEP protein complexes purified from K562 cells to verify the mass spectrometry data. Mock purification was performed on wild type cells. 0.8% of input and 33% of IP were separated on NuPage 4-12% gels. Tagged proteins were detected with mouse HA antibody (Covance, top panel); endogenous PRAME and TAG-PRAME were detected in the second panel with affinity-purified PRAME antibody after staining for Cul2; the other proteins were detected as indicated. Asterisks indicate protein A that dissociated from the beads after elution.
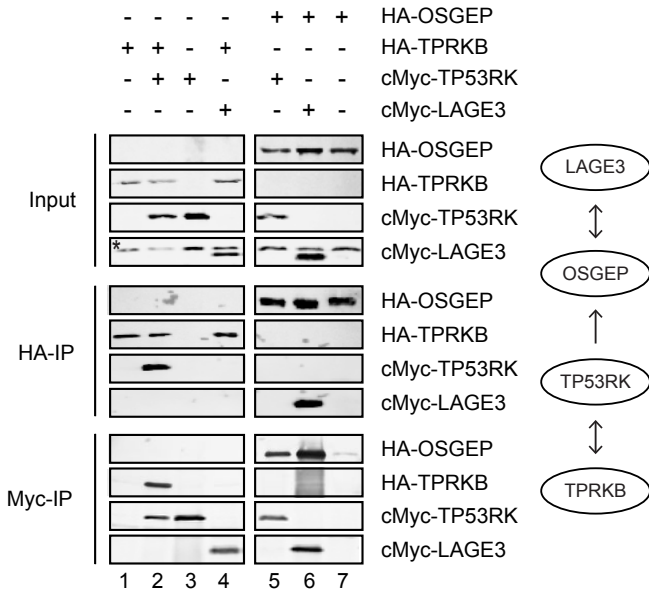
blotting of the same samples (Fig. 2, lane 6). Intriguingly, the level of Cul2 detected in TAG-OSGEP eluates was substantially lower than in TAG-PRAME eluates (Fig. 2, lanes 4 and 6), while Elongins B and C were present in similar amounts in all purifications. These findings might indicate that OSGEP establishes less stable interactions with Cullin2 as compared to PRAME. Consistent with the mass spectrometry results, TAG-PRAME did not copurify with TPRKB (Fig. 2, lane 4), while a clear band was detected in TAG-OSGEP eluates.

Our proteomic experiments show for the first time that a complete EKC complex is present in human cells, and that the OSGEP and LAGE3 subunits are present in a complex together with PRAME and Cullin2 ligase components.

Topology of the human EKC complex

In order to study the topology of the interactions within the human EKC complex, we expressed pairs of EKC subunits in the baculovirus system and performed immunoprecipitations (Fig. 3). As shown in figure 3, we observed reciprocal coimmunoprecipitation of HA-LAGE3 and c-Myc-OSGEP and of HA-TPRKB and c-Myc TP53RK, respectively. In addition, c-Myc TP53RK copurified with HA-OSGEP upon anti-c-Myc immunopurification, although this interaction was not detected in the reciprocoal experiment, raising the possibility that the c-Myc epitope was occluded and/or that binding of the antibody to the epitope disrupts the interaction. Taken together, these experiments established that human EKC has the linear configuration LAGE3-OSGEP-TP53RK-TPRKB, consistent with the three-dimensional structures of archael EKC orthologues that were recently reported (Mao *et al*, 2008).
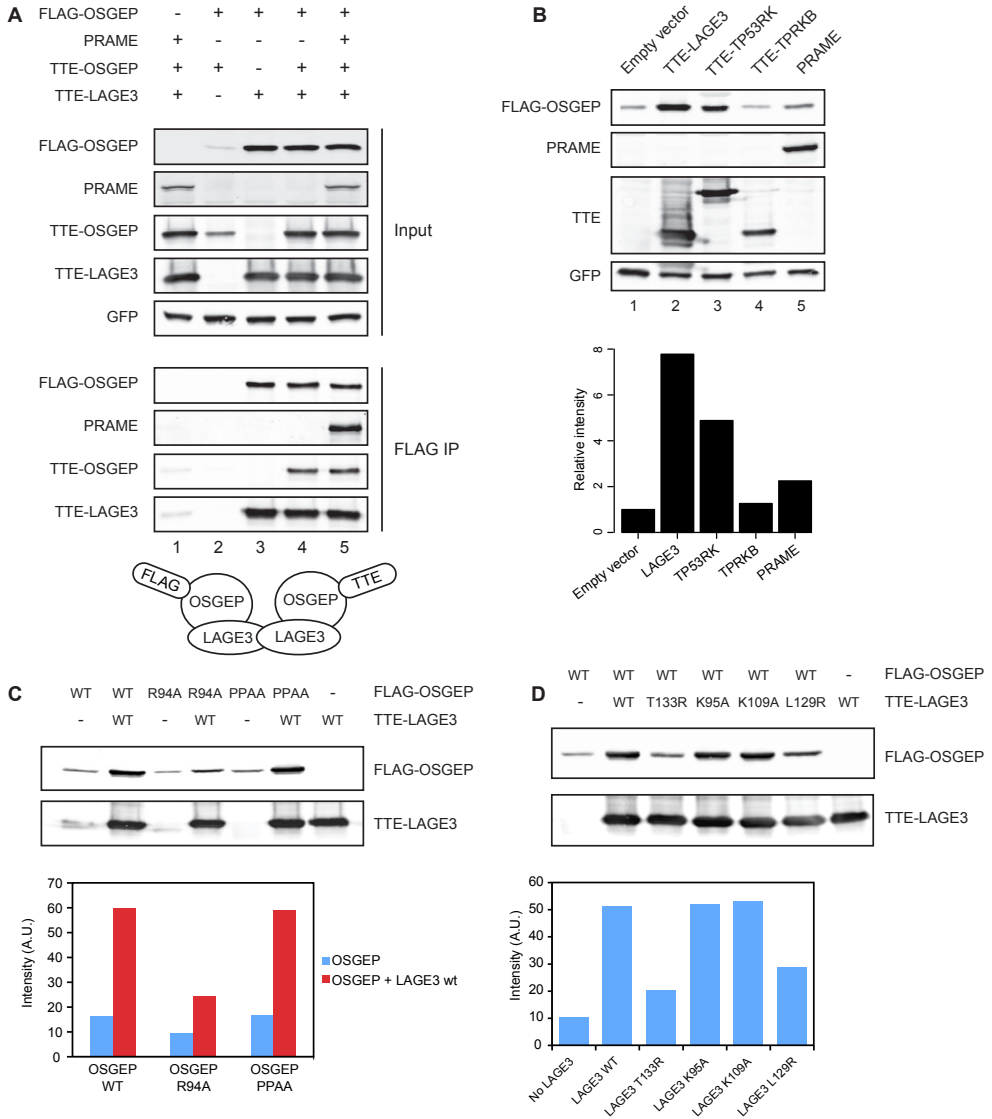
FIGURE 3

	1	2	3	4	5	6	7	
HA-OSGEP	-	-	-	-	+	+	+	
HA-TPRKB	+	+	-	+	-	-	-	
cMyc-TP53RK	-	+	+	-	+	-	-	
cMyc-LAGE3	-	-	-	+	-	+	-	

Input
HA-OSGEP
HA-TPRKB
cMyc-TP53RK
*cMyc-LAGE3

HA-IP
HA-OSGEP
HA-TPRKB
cMyc-TP53RK
cMyc-LAGE3

Myc-IP
HA-OSGEP
HA-TPRKB
cMyc-TP53RK
cMyc-LAGE3

1 2 3 4 5 6 7

LAGE3
↕
OSGEP
↑
TP53RK
↕
TPRKB

Organization of human EKC complex. SF21 cells were co-infected with baculoviruses expressing the indicated proteins. Anti-FLAG immunoprecipitations were performed, and total cell lysates and eluates were analyzed by western blotting. Interactions detected are summarized in the diagram on the right. Panels showing anti-HA immunoblots are from the same exposure of the same blot, as are the panels showing anti-cMyc immunoblots. Asterisk indicates light chains of the antibodies used in the immunoprecipitation.

Interestingly, Mao *et al.* reported that archael Pcc1, the orthologue of human LAGE3, forms highly interdigitated homodimers and hypothesized that dimerization of Pcc1 might support dimerization of the full EKC complex. Moreover, the crystal structures indicated that one moiety of Pcc1 could bind only one moiety of Kae1. We tested the dimerization properties of human EKC using combinations of different epitope tags in transient transfection experiments. As shown in figure 4A, FLAG-OSGEP coimmunoprecipitated TTE-OSGEP when LAGE3 was cotransfected (lane 4), establishing that multiple OSGEP moieties are present in the same biochemical complex. Importantly, the concomitant expression of PRAME did not affect these dimerization properties (Fig. 4A, compare lanes 4 and 5). Taking into consideration the structural information available for archaeal EKC, our data suggest the formation of a tetrameric LAGE3-OSGEP assembly (Fig. 4A).

In the same experiments, we noticed that OSGEP levels were significantly lower in the absence of LAGE3 (Fig. 4A, compare lanes 2 and 4), suggesting that the stability of OSGEP could be critically dependent on interactions with other proteins. To address this hypothesis, we transfected FLAG-OSGEP together with each of the other EKC subunits or PRAME. Fig. 4B shows that LAGE3, TP53RK and PRAME could all enhance the stability of OSGEP. TPRKB did not alter OSGEP levels, consistent with the absence of a direct interaction between these two proteins. Interestingly, LAGE3 exerted the highest positive

Figure 4

A

FLAG-OSGEP	-	+	+	+	+
PRAME	+	-	-	-	+
TTE-OSGEP	+	+	-	+	+
TTE-LAGE3	+	-	+	+	+

FLAG-OSGEP
PRAME
TTE-OSGEP — Input
TTE-LAGE3
GFP

FLAG-OSGEP
PRAME
TTE-OSGEP — FLAG IP
TTE-LAGE3

1 2 3 4 5

106

B

Empty vector | TTE-LAGE3 | TTE-TP53RK | TTE-TPRKB | PRAME

FLAG-OSGEP
PRAME
TTE
GFP

1 2 3 4 5

Relative intensity

Empty vector, LAGE3, TP53RK, TPRKB, PRAME

C

	WT	WT	R94A	R94A	PPAA	PPAA	-	FLAG-OSGEP
	-	WT	-	WT	-	WT	WT	TTE-LAGE3

FLAG-OSGEP
TTE-LAGE3

Intensity (A.U.)
■ OSGEP ■ OSGEP + LAGE3 wt
OSGEP WT, OSGEP R94A, OSGEP PPAA

D

WT	WT	WT	WT	WT	WT	-	FLAG-OSGEP
-	WT	T133R	K95A	K109A	L129R	WT	TTE-LAGE3

FLAG-OSGEP
TTE-LAGE3

Intensity (A.U.)
No LAGE3, LAGE3 WT, LAGE3 T133R, LAGE3 K95A, LAGE3 K109A, LAGE3 L129R

OSGEP protein levels are modulated by protein-protein interactions.

(A) Multiple OSGEP moieties are present in the same complex. Constructs expressing the indicated proteins were transiently transfected in 293T cells. Total cell lysates and FLAG eluates were analyzed by western blotting. A plasmid expressing GFP was cotransfected to control for the transfection efficiency. (B) OSGEP levels are affected by protein-protein interactions. 293T cells were transfected with constructs expressing FLAG-OSGEP and the proteins indicated. The bar graph shows FLAG-OSGEP levels quantified by immunoblot with the Odyssey system (values are the average of two independent experiments). A plasmid expressing GFP was cotransfected to control for the transfection efficiency. (C) and (D) LAGE3-OSGEP interface mutants decrease OSGEP protein levels. Transient transfections in 293T cells with the mutant constructs indicated and immunoblot by Odyssey of total cell lysates. Graphs report intensities of FLAG-OSGEP quantified with Odyssey (A.U., arbitrary units).

effect, which correlates with the possibility to form tetrameric LAGE3-OSGEP complexes, while TP53RK can bind only one moiety of OSGEP. The finding that PRAME had a relatively modest effect on OSGEP levels might reflect a less stable interaction between OSGEP and PRAME, than between OSGEP and TP53RK.

We investigated further the effect of LAGE3 on the protein levels of OSGEP in transient transfection experiments with mutant constructs. Based on the crystallographic structures of archael EKC (Mao *et al*, 2008), we introduced point mutations in aminoacid residues predicted to be involved in the protein-protein interactions between OSGEP and LAGE3: R94A in OSGEP, and L129R or T133R in LAGE3. Additionally, we generated OSGEP mutant P35A/P36A (referred to as PPAA) and LAGE3 mutants K95A and K109A, which are not predicted to contribute to the OSGEP-LAGE3 interaction. As shown in figure 4C, the OSGEP mutant R94A showed lower steady-state levels, which could not be enhanced by co-expression of wild-type LAGE3, while OSGEP PPAA behaved like the wild type. Similarly, levels of wild-type OSGEP were not enhanced upon co-expression of either T133R or L129R LAGE3 mutants, while the LAGE3 lysine mutants behaved the same as wild type LAGE3.

Taken together, our data suggests (i) that the human EKC complex is organized in the same way as its archaeal counterpart, (ii) that dimerization of LAGE3 can mediate formation of higher order structures containing at least two molecules of OSGEP, (iii) that formation of these higher order structures is independent of PRAME, and (iv) that the protein levels of OSGEP can be significantly modulated by protein-protein interactions.

PRAME bridges Cullin2 ligases to the EKC complex

We next studied the role of the Cullin2 complex in the association of PRAME with EKC. Transient transfection experiments showed that the interaction of PRAME with OSGEP and LAGE3 does not require an intact BC-box, since a BC-box defective PRAME (PRAME M2) coimmunoprecipitated OSGEP and LAGE3 as efficiently as wild type PRAME (Fig. 5A).

Next, we purified TAG-LAGE3 protein complexes from HeLaS3 cells, which express very low to undetectable levels of endogenous PRAME. As expected, all EKC subunits copurified with TAG-LAGE3, while no peptides were identified for PRAME (Table 2). Importantly, components of Cullin2 ligases did not co-purify with LAGE3 in these cells. The abundance of EKC subunits was similar in the eluates from K562-TAG-LAGE3 and HeLaS3-TAG-LAGE3 cells, consistent with similar protein levels in the two cell lines.
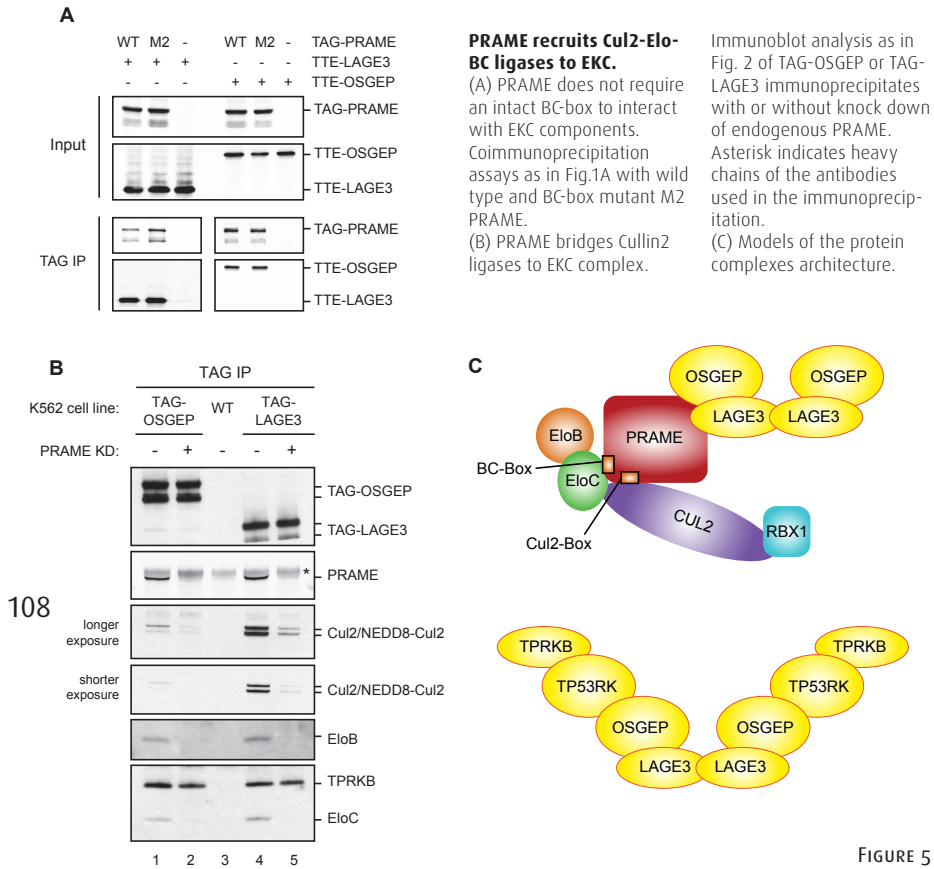
A

	WT	M2	-	WT	M2	-	TAG-PRAME
	+	+	+	-	-	-	TTE-LAGE3
	-	-	-	+	+	+	TTE-OSGEP

Input

TAG-PRAME
TTE-OSGEP
TTE-LAGE3

TAG IP

TAG-PRAME
TTE-OSGEP
TTE-LAGE3

PRAME recruits Cul2-Elo-BC ligases to EKC.
(A) PRAME does not require an intact BC-box to interact with EKC components. Coimmunoprecipitation assays as in Fig.1A with wild type and BC-box mutant M2 PRAME.
(B) PRAME bridges Cullin2 ligases to EKC complex.

Immunoblot analysis as in Fig. 2 of TAG-OSGEP or TAG-LAGE3 immunoprecipitates with or without knock down of endogenous PRAME. Asterisk indicates heavy chains of the antibodies used in the immunoprecip-itation.
(C) Models of the protein complexes architecture.

B

TAG IP

K562 cell line: TAG-OSGEP | WT | TAG-LAGE3

PRAME KD: - + - - +

TAG-OSGEP
TAG-LAGE3
PRAME *
longer exposure — Cul2/NEDD8-Cul2
shorter exposure — Cul2/NEDD8-Cul2
EloB
TPRKB
EloC

1 2 3 4 5

108

C

FIGURE 5

These results, together with the finding that TAG-OSGEP and TAG-LAGE3 associate with endogenous PRAME and Cullin2 ligase components in K562 cells, suggested that PRAME could be involved in recruiting a Cullin2 ligase to the EKC complex.

To test this hypothesis, we knocked down endogenous PRAME with stable transfection of retroviral constructs in K562-TAG-OSGEP and K562-TAG-LAGE3 cell lines and performed immunoprecipitations. Figure 5B clearly shows that downregulation of PRAME significantly reduced the interaction between each EKC subunit and Cul2, EloB and EloC, establishing that PRAME is required for these interactions. As expected, the interaction of TAG-OSGEP and TAG-LAGE3 with TPRKB was not affected by PRAME knockdown. Notably, we found that the level of Cul2 detected in TAG-OSGEP eluates was substan-tially lower than in TAG-LAGE3 eluates (Fig. 5B, lanes 1 and 4), despite similar amounts of EloB and EloC. These results are similar to the previous compari-son between TAG-OSGEP and TAG-PRAME eluates (Fig. 2, lanes 4 and 6).

Taken together, our data reveal that PRAME can bridge Cullin2 ligases to the EKC complex (Fig. 5C).

Cul2-PRAME ligases and EKC subunits co-localize to active promoters

The yeast orthologues of OSGEP (Kae1p) and LAGE3 (Pcc1p) have been reported to associate with the promoters and transcribed regions of several genes in a transcription-dependent manner (Kisseleva-Romanova *et al*, 2006). Genome-wide chromatin immunoprecipitation experiments revealed that PRAME and Cullin2 complex components are specifically enriched at transcriptionally active promoters that are also bound by the transcription factor NFY, and at nearby enhancers (Costessi *et al*, 2011).

In the absence of validated ChIP-grade antibodies against EKC subunits, we generated K562 cell lines stably expressing the proteins of interest fused to the TTE epitope tag (TY1-TY1-ER), which can be efficiently and specifically used in ChIP-qPCR experiments (Costessi *et al*, 2011). To determine whether human OSGEP and LAGE3 are present on chromatin, we performed ChIP assays with the BB2 antibody (which recognizes the TY1 tag) and tested a panel of PRAME-bound promoters. The results in figure 6 clearly show that all PRAME-bound promoters tested are efficiently recovered in both K562-TTE-OSGEP and K562-TTE-LAGE3 cells, but not in parental K562 cells or control K562-TTE cells.

Our genomic experiments extend the biochemical interactions to the chromatin environment, establishing that PRAME and EKC subunits associate with the same promoter regions.

109

FIGURE 6

OSGEP and LAGE3 are present at PRAME-bound promoters on chromatin.
ChIP-qPCR experiments using BB2 antibody were performed on K562 cells lines stably expressing TTE-OSGEP or TTE-LAGE3. As control for specificity, ChIPs were performed on the parental cells (wt) and on cells transduced with the vector expressing the tag only (K562-TTE) using a panel of primer sets for PRAME biding sites. Values are expressed as mean recoveries ± standard deviation from three independent experiments.
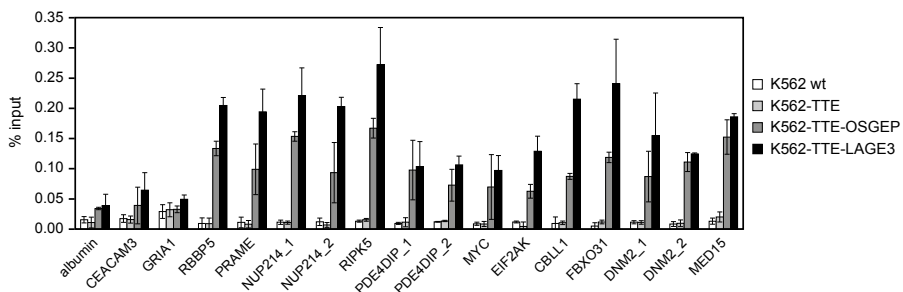
DISCUSSION

We recently reported the first purifications of PRAME-containing protein complexes by epitope-tagged immunoprecipitations and mass spectrometry, which revealed that PRAME is a BC-box substrate receptor component of Cullin2-based E3 ubiquitin ligases (Costessi *et al*, 2011). In the present study, we identified novel interactions of Cul2-PRAME ligases with the human EKC/KEOPS complex, we discovered that PRAME recruits ligase components to EKC, and we showed that these proteins associate with the same promoter regions, strongly suggesting a functional link between these complexes in the same pathway.

EKC is an ancient and extremely conserved complex with orthologues in Archaea and Eukarya. It has been implicated in telomere maintenance, transcriptional regulation, and t^{6}A modification of tRNAs in yeast (Kisseleva-Romanova *et al*, 2006; Downey *et al*, 2006; Daugeron *et al*, 2011; Yacoubi *et al*, 2011; Srinivasan *et al*, 2011). Our interaction data support a linear architecture LAGE3-OSGEP-TP53RK-TPRKB for the human EKC complex and dimerization around the LAGE3 subunit, with two OSGEP moieties per complex. Our results are fully consistent with recently reported crystal structures of archaeal EKC proteins (Mao *et al*, 2008).

Consistent with the notion that all subunits contribute to the activity of the complex, yeast EKC was reported to constitute a stable entity. Downey *et al*. reported that deletion of either Bud32p, Cgi121p, or Gon7p resulted in shorter telomeres. An independent study found that Pcc1p was required for efficient transcription of several yeast genes induced by pheromones or galactose, and that deletion of Pcc1 inhibited the recruitment of the SAGA and Mediator co-activators (Kisseleva-Romanova *et al*, 2006). Dimerization of the Pcc1p and Kae1p subunits was found to be necessary for cell viability (Mao *et al*, 2008), and an intact interaction surface between Kae1p and Bud32p was required for both the transcriptional and telomere maintenance functions (Hecker *et al*, 2008). Most recently, the yeast EKC complex was shown to be essential for a universal chemical modification of tRNAs, called threonyl carbamoyl adenosine (t^{6}A) (Srinivasan *et al*, 2011; Yacoubi *et al*, 2011; Daugeron *et al*, 2011). This modification occurs at position 37 of tRNAs decoding ANN codons, including the first ATG of all protein-coding mRNAs, and is required for efficient translation (Yarian *et al*, 2002). The t^{6}A modification activity was dependent on Pcc1p, the ATPase activity of Kae1p, and the kinase activity of

Bud32p, but not on Cgi121p (Srinivasan *et al*, 2011).

In our experiments, all human EKC orthologues co-purified together and with similar abundance both in K562 and HeLaS3 cells, demonstrating for the first time that a complete and stable EKC complex is present in human cells. However, we consistently found that PRAME complexes contained primarily OSGEP and LAGE3, and very low or undetectable levels of the other two subunits TP53RK and TPRKB. This might reflect a PRAME-mediated disruption of an otherwise stable EKC complex, or a preferential binding of PRAME to an already existing (and possibly dynamic) EKC submodule that was not previously detected in yeast. It is tempting to speculate that PRAME could mediate ubiquitination of EKC components via recruitment of the Cul2 E3 ligase complex. In an attempt to test this hypothesis, we performed ubiquitination assays upon transient transfections in HEK293 cells in presence or absence of the proteasome inhibitor MG132, and in vitro assays with purified proteins. As expected for BC-box proteins, we detected polyubiquitin chains associated with PRAME itself. However, the high background signals in these assays did not allow us to detect a putative ubiquitination of EKC subunits. Hence, our assays could not rule out that EKC might be ubiquitination targets of Cul2-PRAME, and more efforts are needed to test this hypothesis experimentally.

Interestingly, dynamic regulation of EKC components has been reported: the kinase activity of Bud32p was repressed by association with Kae1p (Hecker *et al*, 2008), but promoted by binding to Cgi121p (Mao *et al*, 2008). Moreover, we found that the protein levels of OSGEP can be significantly modulated by association with other proteins (Fig. 4), with LAGE3 exerting the most pronounced positive effect, followed by TP53RK and PRAME. Interestingly, the effect of LAGE3 correlates with the formation of tetrameric complexes in which an LAGE3 dimer binds two OSGEP moieties (Fig. 4A), which are likely not in direct contact with each other (Mao *et al*, 2008).

The results of our transient transfection experiments and those obtained using the baculovirus system suggested (i) that PRAME is able to interact independently with LAGE3 and OSGEP and (ii) that LAGE3 enhances the association of OSGEP with PRAME. Considering the dimerization properties of the LAGE3-OSGEP submodule, we therefore considered the possibility that multiple PRAME moieties might be present in the same protein complex. In this case, both epitope-tagged PRAME and endogenous PRAME should be present in the same protein complex in K562-TAG-PRAME cells, since the expression

level of the two proteins were similar. Contrary to this hypothesis, however, TAG-PRAME did not co-immunoprecipitate endogenous PRAME in our experiments (see Fig. 2 lane 4). Although we cannot formally exclude the possibility that TAG-PRAME might compete with the endogenous protein, our data are most consistent with the idea that there is only one PRAME molecule per complex. Notably, our data suggest that the interaction between PRAME and LAGE3-OSGEP is very stable, since it was efficiently detected in all assays and despite the long incubations times in our protein-complex purification protocols. However, we cannot exclude that high levels of PRAME might be required for the interactions to take place in cells. Taking into account the published crystal structures of archaeal EKC, we propose the models depicted in Figure 5C for the architecture of the human EKC complex and the relationship between EKC and Cul2-PRAME ligases. Further experiments are needed to confirm the stoichiometry of PRAME-EKC complexes, and to characterize the interaction surfaces involved.

Besides the functional experiments in yeast, no information is available yet about EKC functions in other organisms. At present, only studies focusing on single subunits of the human complex have been reported. LAGE3 appears to be ubiquitously expressed, while the two closely related LAGE1 and LAGE2 are cancer-testis antigens with high expression confined to healthy testis and cancer tissues (Alpen *et al*, 2002). Interestingly, expression of the N-terminal domain of LAGE3 fused to the C-terminus of Pcc1p could partially rescue the growth defects of a Pcc1 N-terminal mutant yeast strain, while full length *LAGE3* could not substitute for Pcc1 (Kisseleva-Romanova *et al*, 2006). The *OSGEP* gene was found expressed in all somatic tissues tested and it lies in a head-to-head orientation with the tightly regulated *APEX* gene, which encodes a multifunctional DNA repair enzyme (Ikeda *et al*, 2002). Analyses of their bidirectional promoter identified a CCAAT box crucial for *OSGEP* transcription. Futhermore, *OSGEP* could partially rescue the growth defects of the yeast *kae1Δ* mutant (Kisseleva-Romanova *et al*, 2006), underlining degrees of functional conservation. Human *TP53RK* was first cloned by cDNA subtraction as an interleukin-2 upregulated gene in cytotoxic T-cells, transcripts were detected in a number of normal tissues and cancer cell lines, and it was shown to phosphorylate the human oncoprotein p53 at Serine 15 (Abe *et al*, 2001). Recombinant TP53RK expressed in bacteria was found catalytically inactive, but incubation with COS-7 cell lysates was sufficient to activate the enzymatic activity, suggesting the need for activators and po-

tentially protein interactors. The kinase Akt/PKB was found to play a role in the activation of TP53RK through phosphorylation of the Ser250 residue (Facchin *et al*, 2007). Underscoring a high level of functional homology, TP53RK could partially rescue the growth phenotype of Bud32 deletion in yeast, and Bud32 could interact with and phosphorylate human p53 in vitro (Facchin *et al*, 2003), although yeast does not possess a *bona fide* p53 homologue. Human *TPRKB* was first isolated by a two-hybrid screen using TP53RK as a bait, and was suggested to inhibit the binding of TP53RK to p53 (Miyoshi *et al*, 2003). Expression of *TP53RK* was detected in the high-PRAME expressing cell line K562 as well as in the low-PRAME expressing cell line KG1 (Miyoshi *et al*, 2003). Gene expression data obtained from the GeneSapiens database further suggests that EKC subunits are expressed in most if not all human tissues (Suppl. Figures S2-S4).

Considering the extreme conservation of EKC components during evolution, it can be expected that at least some of the functions and pathways already identified for this complex in yeast could be valid for human cells as well. However, it cannot be excluded that human EKC acquired additional, yet to be identified functions during evolution. The absence of efficient knock-down or knock-out models have so far hindered advanced functional analysis of EKC in higher eukaryotes. Future studies will have to address these issues, and in particular to understand the connections between apparently different functions ascribed to this complex, like phosphorylation of the p53 oncogene, chemical modification of tRNAs, and physical association with promoter regions.

Kae1p and Pcc1p have been found to associate with active chromatin and to regulate expression of some of the yeast promoters to which they are recruited (Kisseleva-Romanova *et al*, 2006). A *pcc1-4* mutation prevented activation of *GAL1* by impairing recruitment of the SAGA histone acetyltransferase and mediator without affecting recruitment of the activator Gal4p. In contrast, while Pcc1p was recruited to the constitutively active *PMA1* and the heat-shock inducible *HSP104* genes in a transcription-dependent manner, transcription of these genes was not affected in the *pcc1-4* mutant.

Our ChIP assays revealed that LAGE3 and OSGEP are also present on chromatin and, in particular, that they colocalize with PRAME at many of its target sites, consistent with our protein-protein interaction data. Since neither PRAME nor EKC proteins possess known DNA-binding domains, it is

not clear whether they can bind DNA directly. Genome-wide ChIP-seq experiments have recently revealed that PRAME associates with transcriptionally active NFY-regulated promoters and nearby enhancers (Costessi *et al*, 2011). Therefore, the transcription factor NFY might play a direct or indirect role in the targeting of a Cul2-PRAME-EKC complex to chromatin. It is also not clear whether there is a hierarchy in the association of PRAME and EKC proteins with chromatin. In order to address this point, we established stable cell lines with downregulation of endogenous PRAME using the published retroviral pRetroSuper system in either K562 wild type or K562-TTE-LAGE3 cells and we performed ChIP experiments. Unfortunately, despite a 70% downregulation of PRAME mRNA levels, our preliminary results did not support a solid conclusion on the hierarchy of binding to chromatin for PRAME and EKC. More experiments, and likely in a different cell and knockdown system, are required to address this point.

Importantly, EKC proteins have an ancient evolutionary origin and seem to be ubiquitously expressed in adult tissues, while PRAME appeared more recently in evolution, and its expression seems to be tightly confined to particular developmental stages and tissues and to be particularly deregulated in oncogenesis. These characteristics are consistent with a model in which EKC can associate with chromatin independently of PRAME. Only when *PRAME* is expressed would a Cul2-PRAME ubiquitin ligase be recruited to chromatin by virtue of specific protein-protein interactions between LAGE3-OSGEP and PRAME.

In an attempt to study the putative function of EKC subunits on target gene transcription, we established K562 cells with stable downregulation of OSGEP or LAGE3 using pRetroSuper vectors and performed genome-wide mRNA-seq analyses. Despite a 50-80% knockdown efficiency for OSGEP and LAGE3, we did not detect genes with significant expression changes when compared to a GFP knockdown control. Our inconclusive data could of course be the result of inefficient downregulation of the target genes. Although the K562 cell line was successfully used to dissect the protein-protein interaction network, we cannot exclude that this cell line might not be suitable for the study of the finer putative chromatin functions and signaling of EKC and PRAME.

The development of ChIP-grade antibodies against EKC components will be a crucial step to address important questions as to whether PRAME target

sites are a subset of EKC sites (or viceversa), and to address the functional relationships between EKC, Cul2-PRAME ligases and NFY on transcriptional regulation of target genes, and potentially other pathways like the recently described modification of tRNAs.

Taken together, our data reveal a novel link between the human oncoprotein PRAME and the ancient EKC complex. Our data clearly show that PRAME can recruit Cullin2 ubiquitin ligases to EKC, and that these complexes associate with the same genomic regions, supporting a functional link in the same pathways. Our results add a new twist to both PRAME and EKC biology and provide an important contribution to understanding the molecular mechanisms in which these proteins are active, both in healthy and cancer cells.

MATERIALS AND METHODS

Cell culture and stable cell lines

HEK293T cells were cultured in DMEM and K562 and HeLaS3 in RPMI medium (Gibco, Invitrogen) at 37°C in 5% CO2; cell lines were obtained from ATCC. Both media were supplemented with 5% Glutamax, 10% fetal bovine serum, 100 U/ml of penicillin, 100 µg/ml of streptomycin (Gibco, Invitrogen). K562 and HeLaS3 stable cell lines were generated with retroviral constructs as described (Costessi *et al*, 2011). Transient transfections were performed with Lipofectamine2000 (Invitrogen).

Antibodies and western blot

Mouse monoclonal HA-12CA5 and Myc-9E11 were produced in-house from hybridoma cultures, PRAME antibodies were previously described (Costessi *et al*, 2011). Commercial antibodies used were: mouse monoclonal FLAG-M2 (Sigma), c-Myc monoclonal (Roche Applied Science), rabbit HA (Bethyl Laboratories), mouse monoclonal HA.11 (Covance MMS-101P), rabbit Cul2 (Zymed Laboratories 51-1890), mouse ElonginC (BD Transduction Laboratories 610760), goat ElonginB (Santa Cruz P-16, sc-23407), mouse TPRKB (Abcam ab68245), mouse BB2 against TY1 tag (Diagenode), GFP (Santa Cruz). Proteins were separated by conventional SDS-PAGE or NuPAGE 4-12% Bis-Tris gels (Invitrogen) run with MES buffer. Western blots were visualized by ECL (GE Healthcare) or Odyssey (LiCor) for quantification.

Plasmids and cloning

A modified version of the retroviral vector LZRS(Zeo) with an improved MCS was made by ligation of a synthetic oligo to generate LZRSn(Zeo). The sequence coding for the StrepII-TEV-MYC-3xHA tag was excised with BamHI and NotI from pcDNA5-TAG-PRAME and subcloned in LZRSn(Zeo) to generate LZRSn-TAG, where the EcoRI site is in the same frame as in the pTTE retroviral vector (Costessi *et al*, 2011).

The TTE cassette was subcloned from pTTE (Costessi *et al*, 2011) to pcDNA3.1(+) using BamHI and XhoI sites to generate pcDNA-TTE for transient transfections.

Full-length human OSGEP was amplified by PCR from cDNA prepared from K562 cells to generate a NotI-EcoRI-OSGEP-SalI fragment, which was cloned into pFLAG-CMV2 (pFLAG-OSGEP). OSGEP was subsequently subcloned as an EcoRI-SalI fragment into the EcoRI and XhoI sites of pTTE and pcDNA-TTE, to generate pTTE-OSGEP and pcDNA-TTE-OSGEP, respectively. OSGEP was cloned as a NotI-SalI fragment downstream the TAG in LZRSn(Zeo), generating LZRSn-TAG-OSGEP.

LAGE3 CDS was obtained from Invitrogen, PCR amplified to introduce EcoRI and XhoI sites, and subcloned in the same sites of pTTE, pcDNA-TTE, and LZRSn-TAG-LAGE3.

Full-length TP53RK/PRPK and TPRKB were amplified by PCR from cDNA of K562 cells and subcloned as EcoRI and XhoI fragments to generate all constructs used.

Mutagenesis to generate OSGEP and LAGE3 mutants was performed using the QuikChange II XL site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions.

All constructs were checked by sequencing.

Purification of protein complexes and mass spectrometry analysis

Protein complex purifications and mass spectrometry were performed from nuclear extracts (unless otherwise indicated) essentially as described (Costessi *et al*, 2011). For HeLaS3 cells the protein extraction protocol was performed with a tighter douncer (pestle B). The quality of cytoplasmic and nuclear extractions were tested by immunoblotting of equivalent volume fractions for alpha tubulin (cytoplasmic marker) and HDAC2 (nuclear marker).

For TAG-PRAME purifications with MYC antibody, about 10 ml of protein extracts (~100 mg) were incubated with 50µl (MYC1 experiment) or 600µl (MYC2 experiment) of crosslinked MYC beads and peptide elution was performed. For protein complex purifications with HA antibody, about 5-6 ml of protein extracts were incubated with 600µl HA-12CA5 crosslinked beads and acidic glycine (pH 2.9) was used for elution.

The data analysis steps of peptide-to-protein remapping and emPAI calculation were automated with a PERL script to generate output tables of unique protein IDs with peptide frequencies and emPAI values for each samples (see Supporting Information Protocol S1 and Analysis Script

S1). The protein list was manually filtered to discard proteins identified in the control samples (false positives).

Expression of recombinant proteins in Sf21 insect cells

cDNAs encoding wild type or mutant PRAME, OSGEP, LAGE3, TP53RK and TPRKB were subcloned into pBAcPAK 8 and recombinant baculoviruses were generated. SF21 cells were cultured at 27°C in SF-900 medium supplemented with 10% fetal bovine serum. SF21 cells were co-infected with the recombinant baculoviruses indicated in the figures. 50 hours after infection, cells were collected, lysed, and processed as described (Mahrour *et al,* 2008).

ChIP and qPCR

Chromatin harvests, ChIPs and qPCR analyses were performed as previously described (Costessi *et al,* 2011). The BB2 antibody recognizing the TY1 epitope (Diagenode) was used to ChIP TTE-tagged proteins from K562 stable cell lines.

ACKNOWLEDGMENTS

REFERENCES

Abe Y, Matsumoto S, Wei S, Nezu K, Miyoshi A, Kito K, Ueda N, Shigemoto K, Hitsumoto Y, Nikawa J & Enomoto Y (2001) Cloning and characterization of a p53-related protein kinase expressed in interleukin-2-activated cytotoxic T-cells, epithelial tumor cell lines, and the testes. *J Biol Chem* **276**: 44003–44011

Allander SV, Illei PB, Chen Y, Antonescu CR, Bittner M, Ladanyi M & Meltzer PS (2002) Expression profiling of synovial sarcoma by cDNA microarrays: association of ERBB2, IGFBP2, and ELF3 with epithelial differentiation. *Am J Pathol* **161**: 1587–1595

Alpen B, Güre AO, Scanlan MJ, Old LJ & Chen Y-T (2002) A new member of the NY-ESO-1 gene family is ubiquitously expressed in somatic tissues and evolutionarily conserved. *Gene* **297**: 141–149

Costessi A, Mahrour N, Tijchon E, Stunnenberg R, Stoel MA, Jansen PW, Sela D, Martin-Brown S, Washburn MP, Florens L, Conaway JW, Conaway RC & Stunnenberg HG (2011) The tumour antigen PRAME is a subunit of a Cul2 ubiquitin ligase and associates with active NFY promoters. *EMBO J* **30**: 3786–3798

Daugeron M-C, Lenstra TL, Frizzarin M, Yacoubi El B, Liu X, Baudin-Baillieu A, Lijnzaad P, Decourty L, Saveanu C, Jacquier A, Holstege FCP, de Crécy-Lagard V, van Tilbeurgh H & Libri D (2011) Gcn4 misregulation reveals a direct role for the evolutionary conserved EKC/KEOPS in the t6A modification of tRNAs. *Nucleic Acids Res* **39**: 6148–6160

Doolan P, Clynes M, Kennedy S, Mehta JP, Crown J & O'Driscoll L (2008) Prevalence and prognostic and predictive relevance of PRAME in breast cancer. *Breast Cancer Res Treat* **109**: 359–365

Downey M, Houlsworth R, Maringele L, Rollie A, Brehme M, Galicia S, Guillard S, Partington M, Zubko MK, Krogan NJ, Emili A, Greenblatt JF, Harrington L, Lydall D & Durocher D (2006) A genome-wide screen identifies the evolutionarily conserved KEOPS complex as a telomere regulator. *Cell* **124**: 1155–1168

Epping MT, Hart AAM, Glas AM, Krijgsman O & Bernards R (2008) PRAME expression and clinical outcome of breast cancer. *Br J Cancer* **99**: 398–403

Epping MT, Wang L, Edel MJ, Carlée L, Hernandez M & Bernards R (2005) The human tumor antigen PRAME is a dominant repressor of retinoic acid receptor signaling. *Cell* **122**: 835–847

Facchin S, Lopreiato R, Ruzzene M, Marin O, Sartori G, Götz C, Montenarh M, Carignani G & Pinna LA (2003) Functional homology between yeast piD261/Bud32 and human PRPK: both phosphorylate p53 and PRPK partially complements piD261/Bud32 deficiency. *FEBS Lett* **549**: 63–66

Facchin S, Ruzzene M, Peggion C, Sartori G, Carignani G, Marin O, Brustolon F, Lopreiato R & Pinna LA (2007) Phosphorylation and activation of the atypical kinase p53-related protein kinase (PRPK) by Akt/PKB. Cell Mol Life Sci 64: 2680–2689

Galperin MY (2008) Social bacteria and asocial eukaryotes. *Environ Microbiol* **10**: 281–288

Haqq C, Nosrati M, Sudilovsky D, Crothers J, Khodabakhsh D, Pulliam BL, Federman S, Miller JR, Allen RE, Singer MI, Leong SPL, Ljung B-M, Sagebiel RW & Kashani-Sabet M (2005) The gene expression signatures of melanoma progression. *Proc Natl Acad Sci USA* **102**: 6092–6097

Hecker A, Lopreiato R, Graille M, Collinet B, Forterre P, Libri D & van Tilbeurgh H (2008) Structure of the archaeal Kae1/Bud32 fusion protein MJ1130: a model for the eukaryotic EKC/KEOPS subcomplex. *EMBO J*

Ikeda H, Lethé B, Lehmann F, van Baren N, Baurain JF, de Smet C, Chambost H, Vitale M, Moretta A, Boon T & Coulie PG (1997) Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory

receptor. *Immunity* **6**: 199–208

Ikeda S, Ayabe H, Mori K, Seki Y & Seki S (2002) Identification of the functional elements in the bidirectional promoter of the mouse O-sialoglycoprotein endopeptidase and APEX nuclease genes. *Biochemical and Biophysical Research Communications* **296**: 785–791

Kilpinen S, Autio R, Ojala K, Iljin K, Bucher E, Sara H, Pisto T, Saarela M, Skotheim RI, Björkman M, Mpindi J-P, Haapa-Paananen S, Vainio P, Edgren H, Wolf M, Astola J, Nees M, Hautaniemi S & Kallioniemi O (2008) Systematic bioinformatic analysis of expression levels of 17,330 human genes across 9,783 samples from 175 types of healthy and pathological tissues. *Genome Biol* **9**: R139

Kisseleva-Romanova E, Lopreiato R, Baudin-Baillieu A, Rousselle J-C, Ilan L, Hofmann K, Namane A, Mann C & Libri D (2006) Yeast homolog of a cancer-testis antigen defines a new transcription complex. *EMBO J* **25**: 3576–3585

Mahrour N, Redwine WB, Florens L, Swanson SK, Martin-Brown S, Bradford WD, Staehling-Hampton K, Washburn MP, Conaway RC & Conaway JW (2008) Characterization of Cullin-box sequences that direct recruitment of Cul2-Rbx1 and Cul5-Rbx2 modules to Elongin BC-based ubiquitin ligases. *J Biol Chem* **283**: 8005–8013

Mao DYL, Neculai D, Downey M, Orlicky S, Haffani YZ, Ceccarelli DF, Ho JSL, Szilard RK, Zhang W, Ho CS, Wan L, Fares C, Rumpel S, Kurinov I, Arrowsmith CH, Durocher D & Sicheri F (2008) Atomic structure of the KEOPS complex: an ancient protein kinase-containing molecular machine. *Mol Cell* **32**: 259–275

Mariappan M, Li X, Stefanovic S, Sharma A, Mateja A, Keenan RJ & Hegde RS (2010) A ribosome-associating factor chaperones tail-anchored membrane proteins. *Nature* **466**: 1120–1124

Miyoshi A, Kito K, Aramoto T, Abe Y, Kobayashi N & Ueda N (2003) Identification of CGI-121, a novel PRPK (p53-related protein kinase)-binding protein. *Biochemical and Biophysical Research Communications* **303**: 399–405

Oberthuer A, Hero B, Spitz R & Berthold F (2004) The tumor-associated antigen PRAME is universally expressed in high-stage neuroblastoma and associated with poor outcome. *Clinical Cancer Research*

Partheen K, Levan K, Osterberg L, Claesson I, Fallenius G, Sundfeldt K & Horvath G (2008) Four potential biomarkers as prognostic factors in stage III serous ovarian adenocarcinomas. *Int J Cancer* **123**: 2130–2137

Radich JP, Dai H, Mao M, Oehler V, Schelter J, Druker B, Sawyers C, Shah N, Stock W, Willman CL, Friend S & Linsley PS (2006) Gene expression changes associated with progression and response in chronic myeloid leukemia. *Proc Natl Acad Sci USA* **103**: 2794–2799

Reinders J, Zahedi RP, Pfanner N, Meisinger C & Sickmann A (2006) Toward the complete yeast mitochondrial proteome: multidimensional separation techniques for mitochondrial proteomics. *J Proteome Res* **5**: 1543–1554

Santamaría C, Chillón MC, García-Sanz R, Balanzategui A, Sarasquete ME, Alcoceba M, Ramos F, Bernal T, Queizán JA, Peñarrubia MJ, Giraldo P, San Miguel JF & Gonzalez M (2008) The relevance of preferentially expressed antigen of melanoma (PRAME) as a marker of disease activity and prognosis in acute promyelocytic leukemia. *Haematologica* **93**: 1797–1805

Srinivasan M, Mehta P, Yu Y, Prugar E, Koonin EV, Karzai AW & Sternglanz R (2011) The highly conserved KEOPS/EKC complex is essential for a universal tRNA modification, t6A. *EMBO J* **30**: 873–881

Steinbach D, Pfaffendorf N, Wittig S & Gruhn B (2007) PRAME expression is not associated with down-regulation of retinoic acid signaling in primary acute myeloid leukemia. *Cancer Genet Cytogenet* **177**: 51–54

Tajeddine N, Gala J-L, Louis M, Van Schoor M, Tombal B & Gailly P (2005) Tumor-associated

antigen preferentially expressed antigen of melanoma (PRAME) induces caspase-independent cell death in vitro and reduces tumorigenicity in vivo. *Cancer Res* **65**: 7348–7355

Tanaka N, Wang Y-H, Shiseki M, Takanashi M & Motoji T (2011) Inhibition of PRAME expression causes cell cycle arrest and apoptosis in leukemic cells. *Leuk Res* **35**: 1219–1225

Wang Q, Liu Y, Soetandyo N, Baek K, Hegde R & Ye Y (2011) A ubiquitin ligase-associated chaperone holdase maintains polypeptides in soluble states for proteasome degradation. *Mol Cell* **42**: 758–770

Xu Z, Page RC, Gomes MM, Kohli E, Nix JC, Herr AB, Patterson C & Misra S (2008) Structural basis of nucleotide exchange and client binding by the Hsp70 cochaperone Bag2. Nat. Struct. *Mol. Biol.* **15**: 1309–1317

Yacoubi El B, Hatin I, Deutsch C, Kahveci T, Rousset J-P, Iwata-Reuyl D, Murzin AG & de Crécy-Lagard V (2011) A role for the universal Kae1/Qri7/YgjD (COG0533) family in tRNA modification. *EMBO J* **30**: 882–893

Yarian C, Townsend H, Czestkowski W, Sochacka E, Malkiewicz AJ, Guenther R, Miskiewicz A & Agris PF (2002) Accurate translation of the genetic code depends on tRNA modified nucleosides. *J Biol Chem* **277**: 16391–16395

121

SUPPLEMENTARY INFORMATION

Protocol S1.
Detailed protocol for comparative analyses of mass spectrometry data.
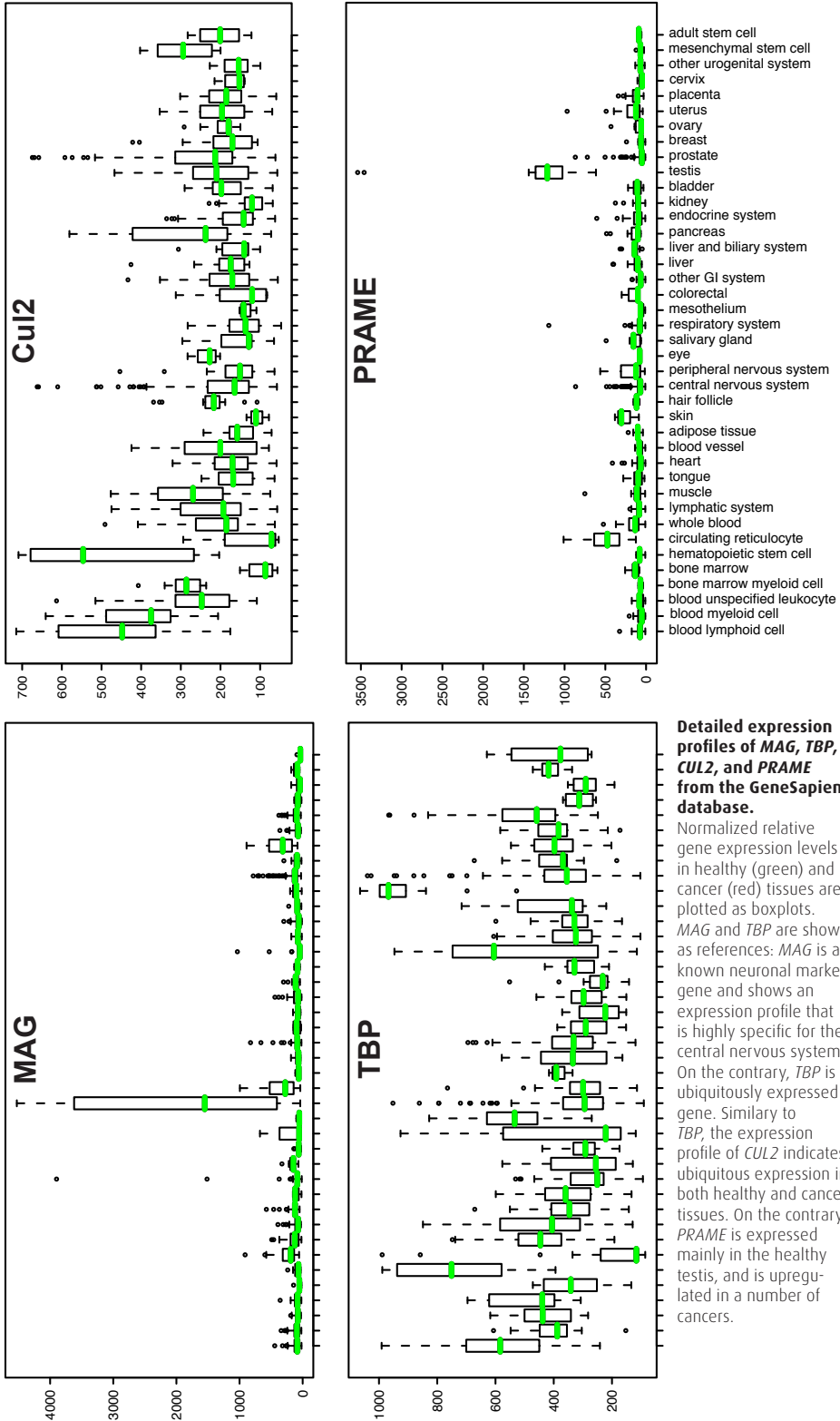Available online at www.plosone.com

Analysis script S1.
PERL script mapPep2Prot_v0.3.pl for analysis of mass spectrometry data.
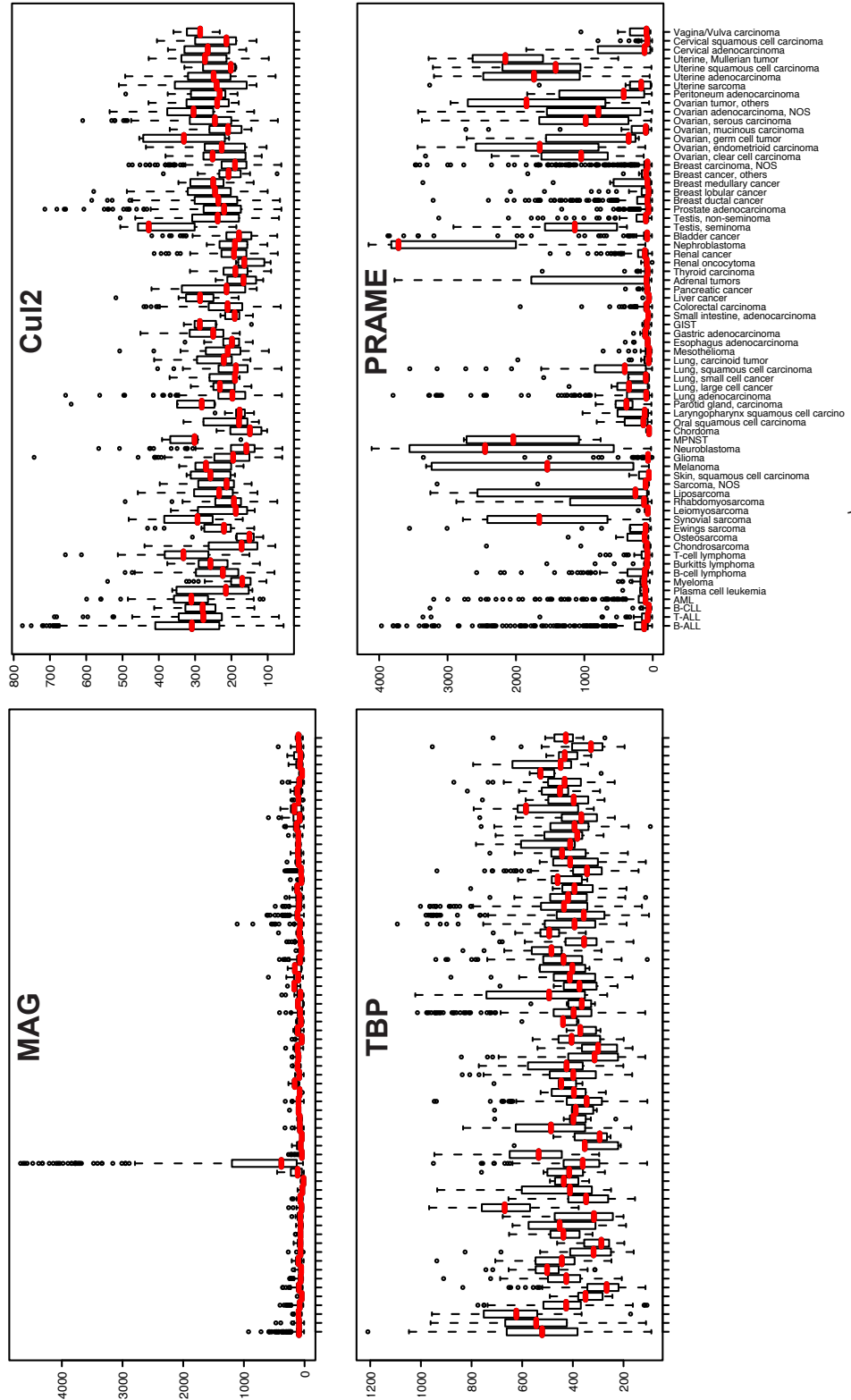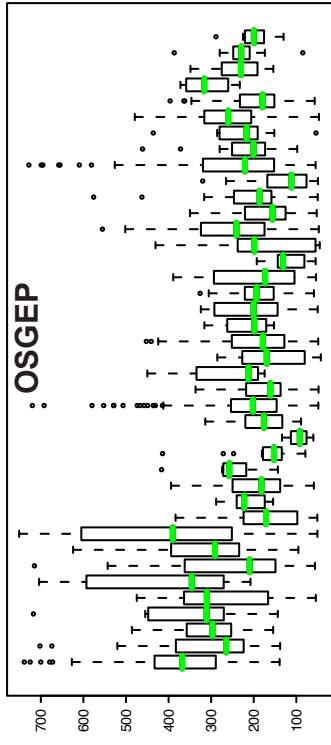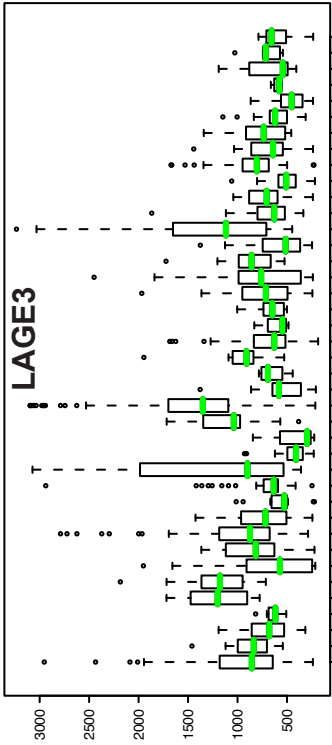Available online at www.plosone.com

Healthy tissues

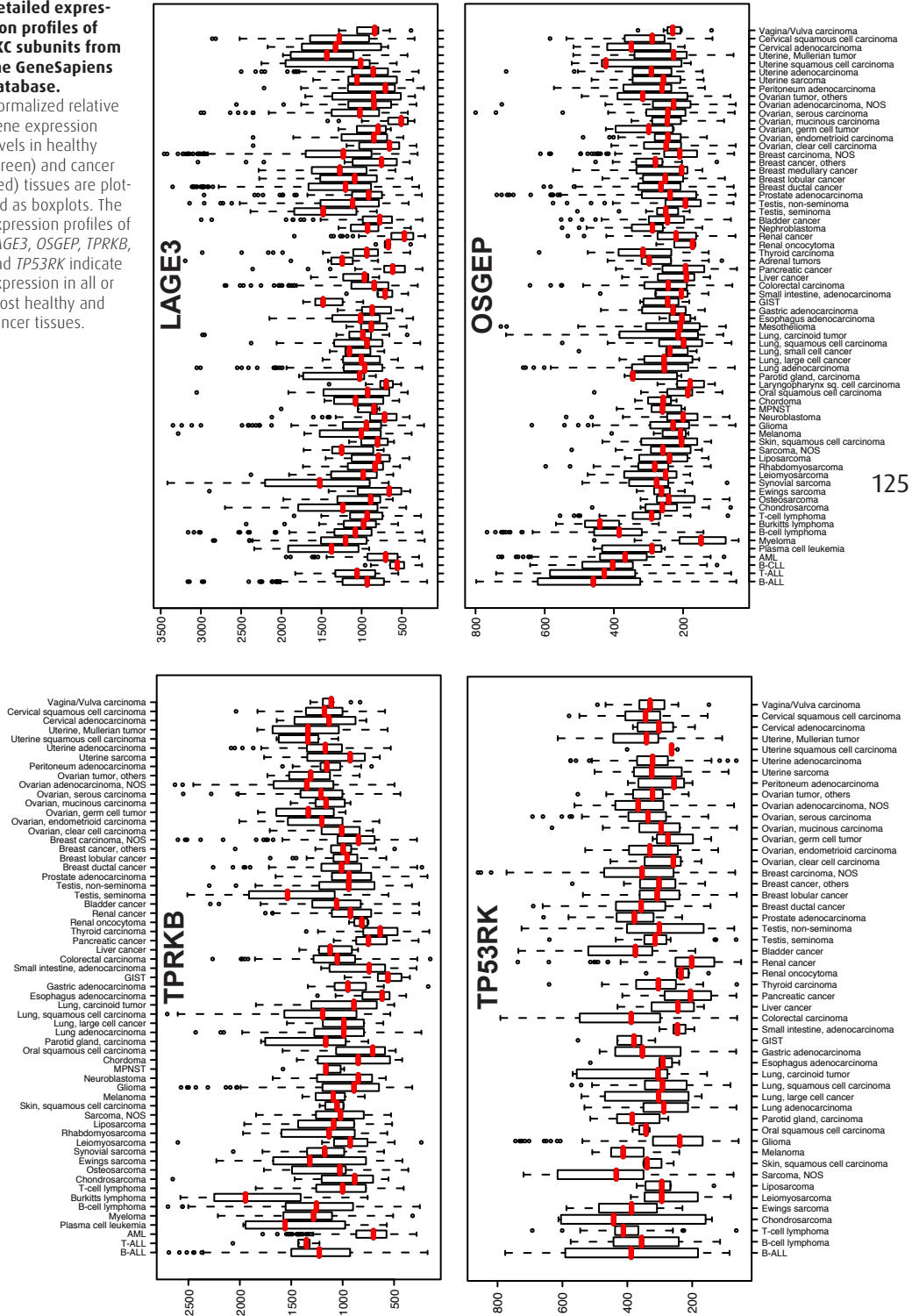| Tissue labels (for all four panels) |
|---|
| adult stem cell |
| mesenchymal stem cell |
| other urogenital system |
| cervix |
| placenta |
| uterus |
| ovary |
| breast |
| prostate |
| testis |
| bladder |
| kidney |
| endocrine system |
| pancreas |
| liver and biliary system |
| liver |
| other GI system |
| colorectal |
| mesothelium |
| respiratory system |
| salivary gland |
| eye |
| peripheral nervous system |
| central nervous system |
| hair follicle |
| skin |
| adipose tissue |
| blood vessel |
| heart |
| tongue |
| muscle |
| lymphatic system |
| whole blood |
| circulating reticulocyte |
| hematopoietic stem cell |
| bone marrow |
| bone marrow myeloid cell |
| blood unspecified leukocyte |
| blood myeloid cell |
| blood lymphoid cell |

Detailed expression profiles of *MAG*, *TBP*, *CUL2*, and *PRAME* from the GeneSapiens database. Normalized relative gene expression levels in healthy (green) and cancer (red) tissues are plotted as boxplots. *MAG* and *TBP* are shown as references: *MAG* is a known neuronal marker gene and shows an expression profile that is highly specific for the central nervous system. On the contrary, *TBP* is a ubiquitously expressed gene. Similarly to *TBP*, the expression profile of *CUL2* indicates ubiquitous expression in both healthy and cancer tissues. On the contrary, *PRAME* is expressed mainly in the healthy testis, and is upregulated in a number of cancers.

Cancer tissues

Cul2

MAG

PRAME

TBP

Vagina/Vulva carcinoma
Cervical squamous cell carcinoma
Cervical adenocarcinoma
Uterine, Mullerian tumor
Uterine squamous cell carcinoma
Uterine adenocarcinoma
Uterine sarcoma
Peritoneum adenocarcinoma
Ovarian tumor, others
Ovarian adenocarcinoma, NOS
Ovarian, serous carcinoma
Ovarian, mucinous carcinoma
Ovarian, germ cell tumor
Ovarian, endometrioid carcinoma
Ovarian, clear cell carcinoma
Breast carcinoma, NOS
Breast cancer, others
Breast medullary cancer
Breast lobular cancer
Breast ductal cancer
Prostate adenocarcinoma
Testis, non-seminoma
Testis, seminoma
Bladder cancer
Nephroblastoma
Renal cancer
Renal oncocytoma
Thyroid carcinoma
Adrenal tumors
Pancreatic cancer
Liver cancer
Colorectal carcinoma
Small intestine, adenocarcinoma
GIST
Gastric adenocarcinoma
Esophagus adenocarcinoma
Mesothelioma
Lung, carcinoid tumor
Lung, squamous cell carcinoma
Lung, small cell cancer
Lung, large cell cancer
Lung adenocarcinoma
Parotid gland, carcinoma
Laryngopharynx squamous cell carcino
Oral squamous cell carcinoma
MPNST
Chordoma
Neuroblastoma
Glioma
Melanoma
Skin, squamous cell carcinoma
Sarcoma, NOS
Liposarcoma
Rhabdomyosarcoma
Leiomyosarcoma
Synovial sarcoma
Ewings sarcoma
Osteosarcoma
Chondrosarcoma
T-cell lymphoma
Burkitts lymphoma
B-cell lymphoma
Myeloma
Plasma cell leukemia
AML
B-CLL
T-ALL
B-ALL

123

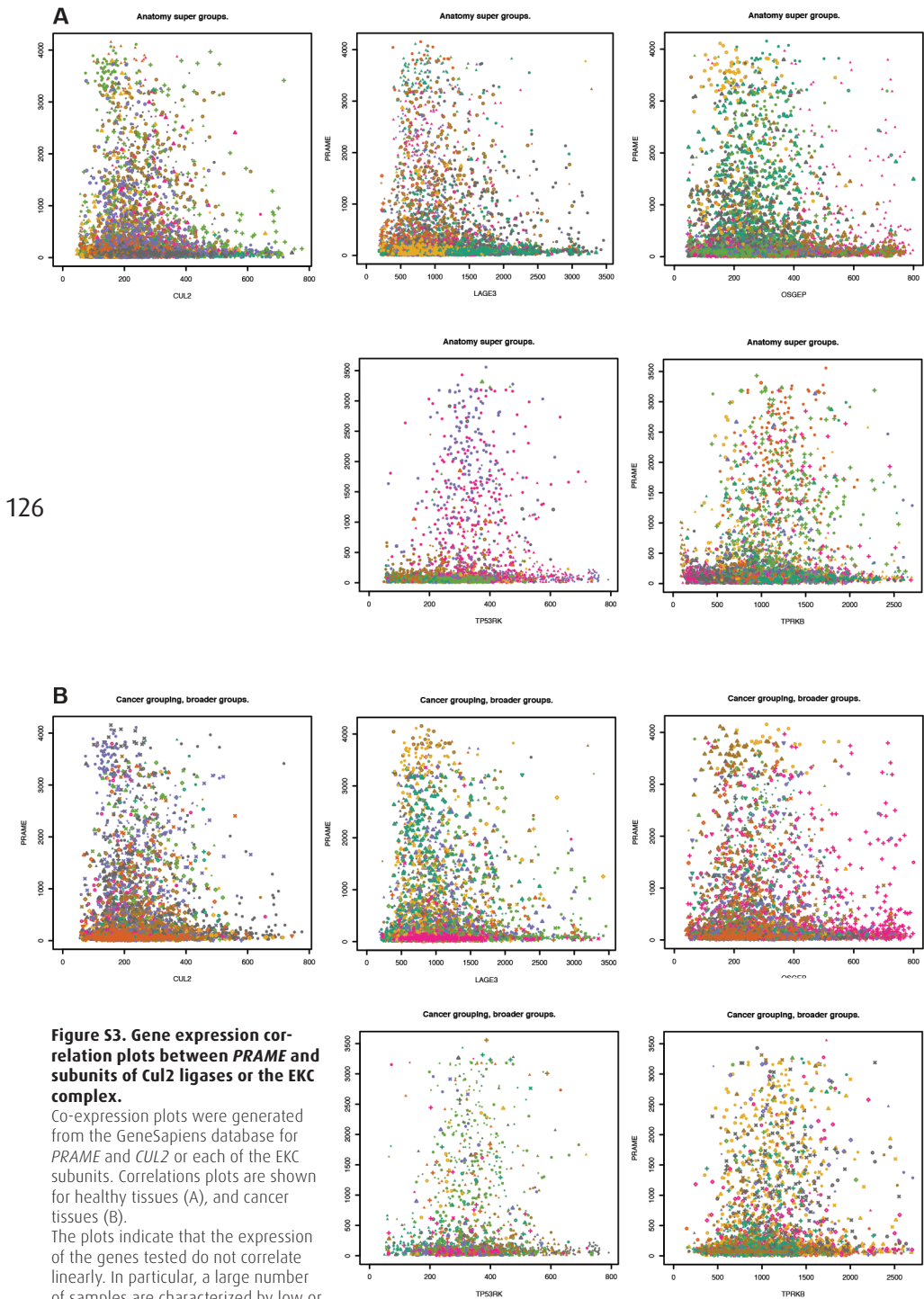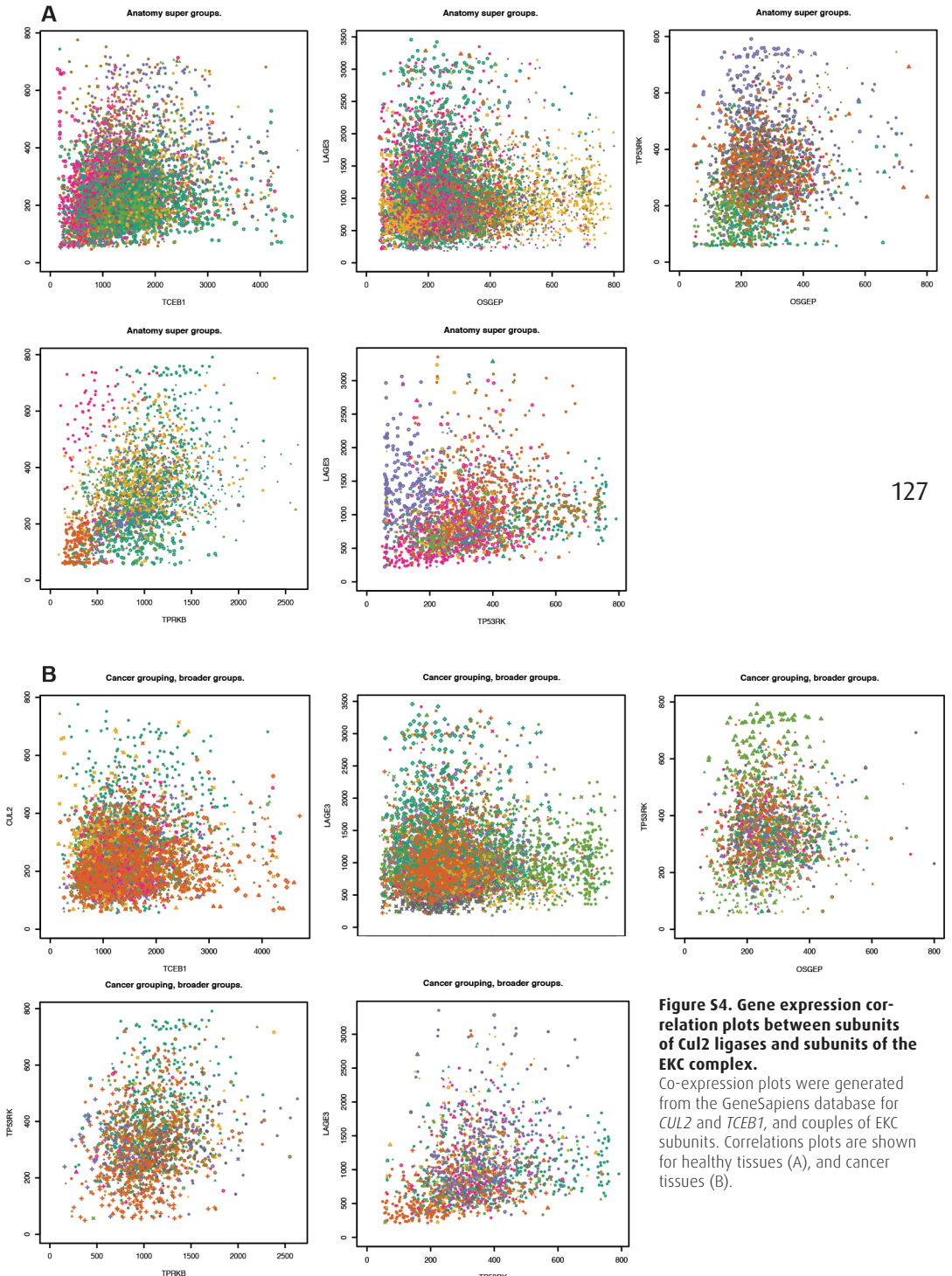# Cancer tissues

Detailed expression profiles of EKC subunits from the GeneSapiens database. Normalized relative gene expression levels in healthy (green) and cancer (red) tissues are plotted as boxplots. The expression profiles of *LAGE3, OSGEP, TPRKB,* and *TP53RK* indicate expression in all or most healthy and cancer tissues.

125

A

126

B

**Figure S3. Gene expression cor-
relation plots between *PRAME* and
subunits of Cul2 ligases or the EKC
complex.**
Co-expression plots were generated
from the GeneSapiens database for
PRAME and *CUL2* or each of the EKC
subunits. Correlations plots are shown
for healthy tissues (A), and cancer
tissues (B).
The plots indicate that the expression
of the genes tested do not correlate
linearly. In particular, a large number
of samples are characterized by low or
no *PRAME* expression, while *CUL2* and
EKC subunits are expressed at higher
levels.

Figure S4. Gene expression correlation plots between subunits of Cul2 ligases and subunits of the EKC complex.
Co-expression plots were generated from the GeneSapiens database for *CUL2* and *TCEB1*, and couples of EKC subunits. Correlations plots are shown for healthy tissues (A), and cancer tissues (B).

PRAME

DEVELOP–MENT OF REAGENTS AND TOOLS

for the study of the human oncoprotein PRAME

Adalberto Costessi, Rieka Stunnenberg, Esther Tijchon, Hendrik G. Stunnenberg

Department of Molecular Biology, Nijmegen Centre for Molecular Life Sciences, Radboud University, Nijmegen, The Netherlands

ABSTRACT

Expression of the human oncoprotein PRAME has been investigated in several tumours and numerous cohorts of cancer patients. However, studies of PRAME function at the molecular and cellular level are scarce and are also hindered by the lack of validated antibodies. In this chapter, we report the development and validation of PRAME antibodies raised against five different peptides. We successfully raised sera and affinity-purified antibodies for use in western blot, immunoprecipitation and ChIP-seq with high sensitivity and specificity. We established that PRAME is a chromatin-associated protein and we generated the first genome-wide ChIP-seq binding profiles of epitope-tagged as well as endogenous PRAME in K562 cells. Finally, we investigated the reported interplay between PRAME and retinoic acid signaling in the A375 melanoma cell line using two independent silencing constructs and complementation with a knockdown-resistant PRAME. In contrast to published reports, our data suggest that the growth phenotype identified in *PRAME* knockdown cells treated with retinoic acid is likely caused by "off-target" knockdown of a gene unrelated to PRAME by one of the silencing constructs used.

INTRODUCTION

The human *PRAME* gene was first identified in 1998 as the gene coding for the antigen responsible for the anti-tumour immune response in a melanoma patient (Ikeda *et al*, 1997). Since then, numerous studies have investigated the transcript expression levels of *PRAME* in patients affected by different types of cancers, establishing that this gene is often overexpressed in a wide range of malignancies. Investigation of the molecular and cellular functions of PRAME, however, has been hindered by the lack of validated reagents for use in experimental research. For instance, several assays rely on antibodies with high specificity for the protein of interest, and it takes significant efforts to develop and test antibodies when these are not available from commercial suppliers. Therefore, we invested on the development of a toolbox of reagents, including antibodies, expression plasmids, and suitable epitope tags.

In this chapter, we report the development of a number of rabbit polyclonal antibodies against human PRAME and their validation for western blot, immunoprecipitation, and ChIP assays. We further discuss strategies that can be undertaken to generate genome-wide ChIP-seq binding profiles of proteins of interest, and we present an epitope-tagged approach based on a double TY1 epitope that we successfully applied for ChIP-seq profiling of PRAME in K562 cells. A recent study reported that downregulation of PRAME sensitized A375 melanoma cells to retinoic acid-induced growth arrest, suggesting a repressive function for PRAME on retinoic acid signaling (Epping *et al*, 2005). Here, we report data generated with two independent knockdown constructs, which do not support the published model and rather suggest that the phenotype observed is likely caused by an "off-target" effect of the knockdown vector used in the literature.

RESULTS AND DISCUSSION

Development of anti-PRAME peptide antibodies

At the beginning of our study, antibodies for human PRAME were not commercially available. Therefore, we performed an immunization protocol in rabbits to raise anti-PRAME peptide antibodies (see Table 1 and Figure 1). Peptides 2, 3, 4, and 5 were designed in regions of PRAME that showed low sequence homology with other PRAME-like proteins, in order to reduce the chance of cross-reactivity; peptide 7 was previously described (Ikeda *et al*, 1997). Each peptide was administered to two rabbits in parallel, and crude sera were collected every three weeks for a maximum of ten harvests (S1 to S10).

Sera were first tested in dot blot assays to assess if the sera contained antibodies directed against the peptides administered. The results showed that all rabbits mounted an efficient immune response against the corresponding peptide (data not shown).

Next, we tested the sera in western blot assays using protein extracts from 293T cells transiently transfected with pcDNA5-TAG-PRAME (StrepII-TEV-MYC-3xHA-PRAME, 73 kDa) and control untransfected cells. Importantly, 293T cells do not express significant levels of endogenous PRAME. Figure 2 shows the results obtained with sera from bleed S5 (panel A) and bleed S8 (panel

FIGURE 1 **Peptides used for the generation of PRAME antibodies.** The locations of the peptides used in the immunization protocol are highlighted on the complete amino acid sequence of human PRAME.

```
  1   MERRRLWGSI QSRYISMSVW TSPRRLVELA GQSLLKDEAL AIAALELLPR ELFPPLFMAA

 61   FDGRHSQTLK AMVQAWPFTC LPLGVLMKGQ HLHLETFKAV LDGLDVLLAQ EVRPRRWKLQ

                         Peptide 2          Peptide 7                      Peptide 3
121   VLDLRKNSHQ DFWTVWSGNR ASLYSFPEPE AAQPMTKKRK VDGLSTEAEQ PFIPVEVLVD

181   LFLKEGACDE LFSYLIEKVK RKKNVLRLCC KKLKIFAMPM QDIKMILKMV QLDSIEDLEV

                                      Peptide 4
241   TCTWKLPTLA KFSPYLGQMI NLRRLLLSHI HASSYISPEK EEQYIAQFTS QFLSLQCLQA

301   LYVDSLFFLR GRLDQLLRHV MNPLETLSIT NCRLSEGDVM HLSQSPSVSQ LSVLSLSGVM

                        Peptide 5
361   LTDVSPEPLQ ALLERASATL QDLVFDECGI TDDQLLALLP SLSHCSQLTT LSFYGNSISI

421   SALQSLLQHL IGLSNLTHVL YPVPLESYED IHGTLHLERL AYLHARLREL LCELGRPSMV

481   WLSANPCPHC GDRTFYDPEP ILCPCFMPN
```

A

C PR | PR C | C PR | PR C | C PR | PR C | C PR | C PR | PR C

82 —
64 —
49 —
37 —

V19 V20 V22 V23 V24 V25 V26 V27 V28

B

C PR | PR C | C PR | PR C | C PR | PR C | C PR | PR C | PR C

V19 V20 V22 V23 V24 V25 V26 V27 V28

C

293T-TAG-PRAME / K562 wt | 293T-TAG-PRAME / K562 wt | 293T-TAG-PRAME / K562 wt

TAG-PRAME —
PRAME —

Preimm V19 | V19/S5 | V19/S5 + peptide

D

K562-TAG-PRAME / K562 wt / 293T-TAG-PRAME | K562-TAG-PRAME / K562 wt / 293T-TAG-PRAME | K562-TAG-PRAME / K562 wt / 293T-TAG-PRAME

TAG-PRAME
PRAME
*

Preimm V28 | V28/S5 | V28/S5 + peptide

133

Validation of crude sera in western blot assays.
(A-B) Western blot experiments with PRAME crude sera using extracts from HEK293T cells (indicated as C, control) and from HEK293T transiently transfected with pcDNA5-TAG-PRAME (indicated as PR). Results are shown for sera obtain with bleed 5 (A) and bleed 8 (B). About 20 micrograms of whole cell protein extracts were loaded for the control cell line, and about 40 micrograms for TAG-PRAME transfected cells. Sera were used at 1:1000 dilution in 5% skimmed milk and PBS/T, and secondary antibody conjugated to HRP. The two main bands recognized for TAG-PRAME are indicated with arrows. Sera recognizing TAG-PRAME are underlined.
(C) Peptide competition assay with serum V19. Whole cell extracts from the cell lines indicated were probed with 1:1000 dilutions of: preimmune serum from rabbit V19 (Preimm V19), V19/S5 serum, or V19/S5 serum (7 μl) after a 2 hours incubation at 37°C with 50 μg peptide 2 and 100 μl PBS.
(D) Peptide competition with serum V28, as in (C). Asterisk indicates unrelated bands.

FIGURE 2

B). Sera V19, V20, V22, V27 and V28 successfully detected TAG-PRAME, while the other sera did not work efficiently in these assays.

The V19 and V20 sera recognized also an additional band of about 50-55 kDa. Notably, this band was not detected by antibodies raised against other peptides, nor by the pre-immune serum from rabbit V19 (Fig. 2C), suggesting that reactivity against this band was mediated by antibodies raised by the immunization protocol against peptide 2. Consistent with this hypothesis, a

peptide competition experiment revealed that the band was not recognized when the V19 serum was pre-incubated with peptide 2 (Fig. 2C). Therefore, we conclude that the V19 and V20 sera cross-react with a 50-55 kDa protein unrelated to PRAME, but containing an epitope similar or identical to peptide 2.

Remarkably, V28 (but not V27) sera recognized two additional bands with an apparent molecular weight of 40-45 kDa (Fig. 2A-B). These bands were specifically recognized in 293T cells transiently transfected with pcDNA5-TAG-PRAME or K562 cells transduced with LZRS-TAG-PRAME, but not in the control untransfected cells (Fig. 2D). The preimmune serum from rabbit V28 detected the same bands (Fig. 2D), indicating that the antibodies against these proteins were already present in the serum of rabbit V28 before the immunization protocol, and were not raised by the immunization with peptide 7. This was confirmed by a peptide competition experiment in which V28 serum was pre-incubated with peptide 7 before western blot detection: as expected, detection of PRAME was inhibited, but the two lower "aspecific" bands were still efficiently detected (Fig. 2D).

TABLE 1.   **Peptides used to generate PRAME antibodies and corresponding sera.**

Peptide ID	Peptide sequence	Peptide position	Rabbits ID	Affinity-purified Ab
Peptide 2	NRASLYSFPEPEA	139-151	V19, V20	P1 (V19/S5)
Peptide 3	FIPVEVLVDLFL	172-183	V21, V22	P2 (V22/S5)
Peptide 4	LLLSHIHASSYIS	265-277	V23, V24	-
Peptide 5	RASATLQDLVFDE	375-387	V25, V26	-
Peptide 7	FPEPEAAQPMTKKRKVDG	146-163	V27, V28	P3 (V27/S6) P4 (V28/S5)

Purification of anti-PRAME antibodies

The best performing sera V19/S5, V22/S5, V27/S6 and V28/S5 were further subjected to two different purification protocols. The first protocol consisted of tandem precipitations with caprilic acid and ammonium sulphate to deplete abundant serum proteins and enrich in immunoglobulins. Caprilic acid in mildly acidic conditions precipitates most serum proteins, with the exception of IgG molecules. The soluble antibodies are then precipitated with ammonium sulphate and resuspended in PBS. These IgG-enriched fractions were referred to as "clean sera": V19cl, V20cl, V22cl, V27cl, V28cl.

The second protocol was a peptide affinity purification that specifically

selects the antibodies that have affinity for the peptide used in the rabbit immunization. The peptide is first crosslinked to an insoluble agarose matrix to generate an affinity column, through which the serum is flowed. Antibodies that recognize the peptide bind to the insoluble matrix, while unrelated antibodies and serum proteins flow-through. Eventually, the immobilized antibodies are eluted from the column with acidic glycine, which destabilizes the antibody-peptide interactions. The antibodies purified with this approach were labelled "purified antibodies": P1 to P4 (Table 1).

The crude sera, the flow-through fractions from the affinity column and the purified antibodies (P1-P4) were tested in ELISA to assess the efficiency of the protocol (data not shown). The results confirmed that P1 (derived from V19/S5 serum) and P2 (from V22/S5) were enriched in specific antibodies, while the flow-through fractions had little reactivity against the peptides, as expected for an efficient affinity purification. The purified antibodies P3 (from V27/S6) and P4 (from V28/S5) showed also a good reactivity against the peptides; however, also the flow-through fractions displayed reactivity against peptide 7, indicating that a fraction of PRAME antibodies did not bind efficiently the affinity column and were still present in the flow-through. Cleaned sera and affinity purified antibodies were tested in western blot assays in the same conditions used previously for the crude sera (Figure 3).

Taken together, our experiments established that antibodies raised against peptide 2 (V19, V19cl, V20, V20cl and P1) efficiently recognized TAG-PRAME but also an unrelated protein with an apparent molecular weight similar to PRAME (Figures 2 and 3). Therefore, these antibodies are not suitable for detection of the endogenous PRAME protein, but can be used for detection of epitope-tagged PRAME.

Similarly, sera V22 and V22cl reproducibly detected TAG-PRAME and an additional band of about 40-45 kDa. This band was less prominent in western blots with the purified antibody P2, where an aspecific band running just above TAG-PRAME was identified (Figure 3B).

Sera V23, V24, V25, V26 performed poorly in all experiments and did not react against TAG-PRAME in western blots (Figure 2).

All antibodies raised against peptide 7 (V27, V27cl, V28, V28cl, P3 and P4) displayed a very strong reactivity against PRAME (Figures 2 and 3). A band of 55 kDa was detected with lower intensity as compared to TAG-PRAME in transfected 293T cells but not in untransfected cells: this bands most likely represents untagged PRAME expressed from the transiently
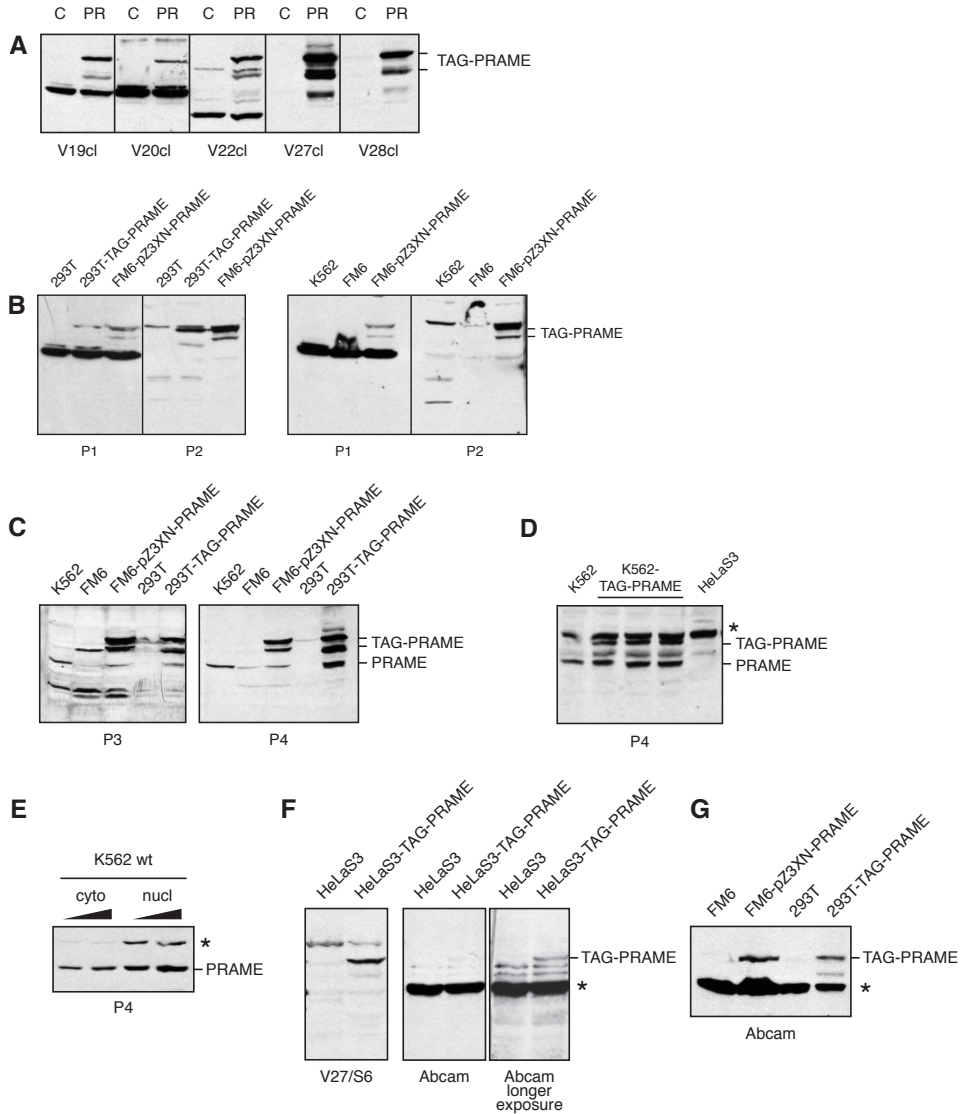
Validation of cleaned sera and affinity-purified antibodies in western blot assays.
(A) Western blot with cleaned sera on whole cell extracts from HEK293T cells (C) and HEK293T cells transiently transfected with pcDNA5-TAG-PRAME (PR).
(B) Western blot with affinity-purified antibodies P1 and P2 on whole cell protein extracts from the cell lines indicated.

(C) Western blot with affinity-purified antibodies P3 and P4 on whole cell protein extracts from the cell lines indicated. Epitope-tagged and endogenous PRAME are indicated.
(D) Western blot on nuclear protein extracts separated on a 4-12% Novex gradient gel. An unrelated protein detected by P4 antibody is indicated with an asterisk.

(E) Nuclear and cytoplasmic extracts from K562 cells show that the unrelated band detected by the P4 antibody is mainly present in the nuclear fraction.
(F-G) Whole cell protein extracts were probed with V27/S6 serum or Abcam PRAME antibody in parallel. The asterisk indicates an unrelated protein detected by Abcam antibody.

Figure 3

136

transfected plasmid, since the starting ATG codon of PRAME was still present in the pcDNA5-TAG-PRAME construct.

The P3 and P4 purified antibodies performed very well with whole cell extracts of several cell lines: a single band at the expected molecular weight of PRAME was identified in K562 and FM6 melanoma cells, which express endogenous gene, while no band was detected in the PRAME-negative 293T cells (Figure 3C). Subsequent experiments, however, revealed that the P4 antibody recognized also an unrelated protein with the same apparent molecular weight of TAG-PRAME in nuclear extracts of K562 cells. This unrelated protein was less evident in whole cell extracts (Fig. 3C), cytoplasmic extracts (Fig. 3E), and chromatin fraction (Fig. 7B).

During the course of our study, a commercial antibody became available against human PRAME (ab32185, Abcam). Western blot experiments revealed that this antibody could recognize epitope-tagged PRAME (Figure 3F-G). However, this antibody displayed a significant cross-reactivity with a protein of similar molecular weight as PRAME in all cell lines tested. Therefore, antibody ab32185 (Abcam) could not be used for detection of the endogenous PRAME protein, although this antibody has been used in recent publications (De Carvalho *et al*, 2011; Partheen *et al*, 2009; Quintarelli *et al*, 2008; Partheen *et al*, 2008; Passeron *et al*, 2009).

FIGURE 4

Validation of P4 antibody specificity with PRAME knockdown cells.
(A) The target sites of knockdown constructs pRS-P1 and pRS-P2 are indicated on the coding sequence of human PRAME.
(B) Western blot on the cell lines indicated. HeLaS3-TAG-PRAME cl.3 and K562-TAG-PRAME cl.12 lines were transduced with the pRS knockdown constructs indicated.
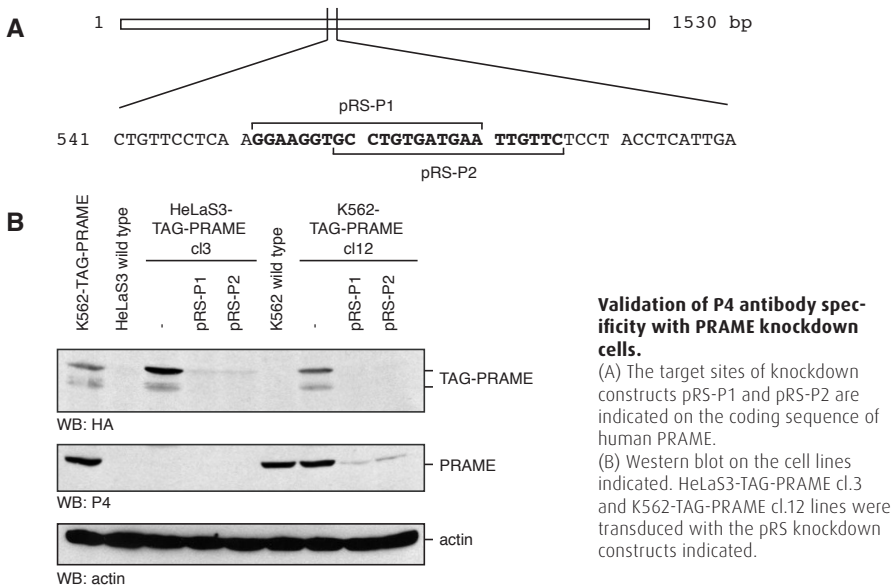
PRAME knockdown

The specificity of the affinity purified P4 antibody was further tested upon knockdown of PRAME. For this purpose, stable cell lines were made using two independent retroviral pRetroSuper constructs targeting PRAME: pRS-PRAME1 and pRS-PRAME2 (further indicated as pRS-P1 and pRS-P2; Figure 4). Double-stable cell lines were made by transduction of zeocin-resistant K562-TAG-PRAME cl.12 and HeLaS3-TAG-PRAME cl.3 with pRS-P1 or pRS-P2 retroviral particles and selection of polyclonal cell populations with puromycin. Western blot analysis with HA antibody established that both knockdown constructs efficiently silenced TAG-PRAME (Figure 4). As expected, the 55 kDa band recognized by P4 antibody was also downregulated, demonstrating that this band corresponds to the endogenous PRAME protein. The same experiment showed that PRAME is not detectable in HeLaS3 wild type cells in these conditions, which is consistent with cDNA gene expression analyses (Chapter 2, Figure 2D).

PRAME antibodies in immunoprecipitation assays

Finally, PRAME antibodies were tested in immunoprecipitation assays using nuclear extracts of K562-TAG-PRAME cells (Figure 5). Immunoprecipitation without antibodies, or with rabbit HA antibodies (Abcam) were performed as negative and positive controls, respectively. Supernatant and IP fractions were analyzed by western blotting with mouse HA.Y11 antibody. The results

FIGURE 5
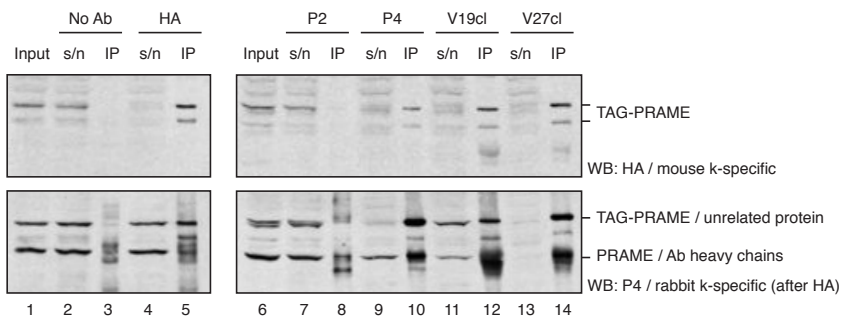
Validation of PRAME antibodies in immunoprecipitation assays. Immunoprecipitation experiments on nuclear extracts from K562-TAG-PRAME cl.12 cells. The following antibodies were used: HA Abcam (1 μl), P2 (8 μl), P4 (5 μl), V19cl (5 μl), V27cl (5 μl). Equivalent amounts (18%) of nuclear extract (input), supernatant after IP, and IP fractions (beads boiled after IP) were separated by SDS-PAGE and probed with HA.Y11 antibody and secondary antibody specific for mouse k chains. Next, the same membranes were probed with P4 antibody using a secondary antibody specific for rabbit k chains.

FIGURE 6

Experimental strategy to generate ChIP-seq binding profiles for a factor of interest.
The scheme presents a general strategy to obtain a genome-wide binding profile and the validation experiments required (modified from Brinkman & Stunnenberg, 2009).
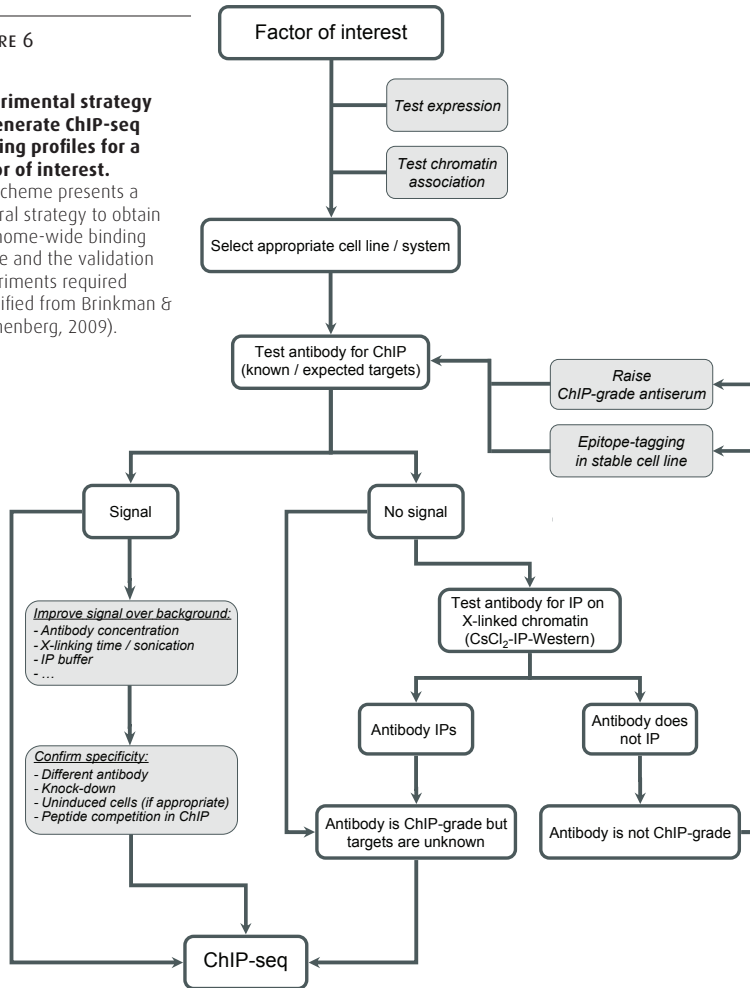
show that P4, V19cl and V27cl antibodies successfully immunoprecipitated TAG-PRAME, while P2 purified antibody did not work in this assay. Moreover, also the endogenous PRAME protein was efficiently immunoprecipitated by V27cl, since it was depleted from the supernatant (Figure 5, lane 13). Notably, the immunoprecipitated endogenous PRAME protein is difficult to visualize in the IP fractions, since it runs at the same height as the heavy chains of the antibodies that are highly abundant in such fractions. Taken together, our results established that PRAME antibodies raised against peptide 2 and 7 performed efficiently also in immunoprecipitation assays.

PRAME association to chromatin

PRAME was recently reported to associate with the retinoic acid receptor

alpha (RARA) and the polycomb protein EZH2, and to mediate repression of transcription of RARA target genes (Epping *et al*, 2005). However, PRAME binding to chromatin was reported only for the RARB promoter in the retinoic-acid responsive A375 melanoma cell line (Epping *et al*, 2005). Notably, the K562 cell line was reported to be unresponsive to retinoic acid treatment because of low endogenous expression of retinoic acid receptors (Robertson *et al*, 1991). Evidence for the association of PRAME with chromatin can be obtained using ChIP and ChIP-seq, though a major hurdle is the absence of validated ChIP-grade antibodies against PRAME.

A general approach to obtain a genome-wide binding profile for a factor of interest is illustrated by the decision tree of Figure 6 (Brinkman & Stunnenberg, 2009). First of all, an appropriate cell system needs to be selected, which involves testing expression of the factor of interest and establishing its association with chromatin. If antibodies and (putative) target sites are available, these can be tested in ChIP-qPCR assays. If a signal is obtained by ChIP-qPCR, this can be further improved by optimizing the antibody concentration, the buffers used, crosslinking and sonication conditions. Subsequently, the specificity of the signals can be validated by using antibodies raised against different epitopes of the protein, knock-out or knockdown systems, or epitope-tagged versions of the protein of interest. When these steps are completed, it is possible to proceed with a genome-wide profiling experiment.

If no signal is obtained in an exploratory ChIP-qPCR assay, either the antibody used is not ChIP-grade, or the presumed genomic sites investigated are not bound by the protein of interest. The ChIP-grade properties of the antibody can be tested performing an immunoprecipitation assay on crosslinked chromatin followed by western blot, or preferably on crosslinked chromatin purified by caesium chloride density centrifugation, which separates free soluble proteins and nucleic acids from the sample, and enriches in chromatin-associated proteins. If the target protein can be IP'ed, the antibody is likely ChIP-grade and a "blind" ChIP-seq experiment can be performed. If putative binding sites are discovered, these can be used in ChIP-qPCR experiments to optimize the assay conditions before performing a new profiling experiment.

If the antibody does not IP the protein of interest, the antibody is not ChIP-grade and alternative approaches can be undertaken, i.e. raising new (ChIP-grade) antibodies, or applying epitope tagging in stable cells lines with retroviral expression systems or BAC transgenics.

140

A

1 - Free protein
2 - Free protein
3 - Chromatin
4 - Chromatin
5 - Free nucleic acids

K562-TAG-PRAME cl. 12

1 2 3 4 5 M

K562 wt

1 2 3 4 5

4 kb
2 kb
1 kb
0,5 kb

B

K562-TAG-PRAME cl. 12

Input chromatin
Whole cell extract
Nuclear extract
1 2 3 4 5

TAG-PRAME

WB: HA

WB: Histone H3

K562 wt

Input chromatin
Whole cell extract
Nuclear extract
1 2 3 4 5

*
*
PRAME

WB: P4

WB: Histone H3

141
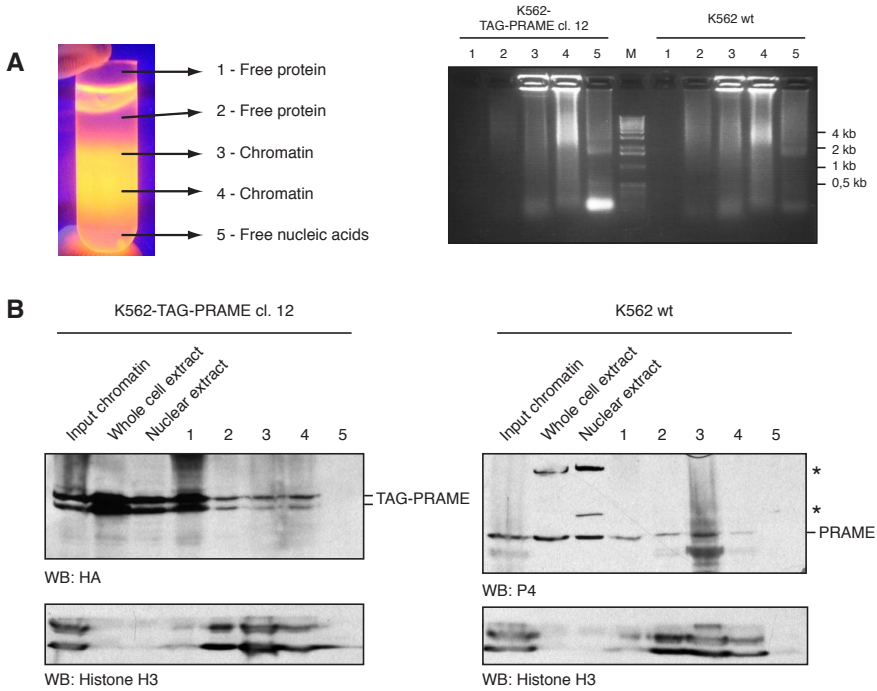
PRAME associates to chromatin in vivo.
(A) Left, photograph under UV light of crosslinked chromatin from K562 cells after CsCl$_2$ gradient centrifugation in the presence of EtBr. Right, the fractions collected (1-5) for K562 wild type and K562-TAG-PRAME cl.12 were analyzed by electrophoresis on 1% agarose gel to assess their nucleic acid content.
(B) Western blot analysis of the CsCl$_2$ fractions analyzed in (A).

FIGURE 7

In our study, we first investigated the association of PRAME to chromatin in the K562 cell line. We prepared formaldehyde-crosslinked chromatin from K562 wild type and K562-TAG-PRAME cells and fractionated it on a CsCl$_2$ density gradient by ultracentrifugation. This technique separates the free soluble proteins from the chromatin-associated proteins based on their different densities. Five different fractions were collected as indicated in Figure 7A and analyzed by agarose gel electrophoresis and western blot. Agarose gel electrophoresis confirmed that chromatin was enriched in fractions 3 and 4, as indicated by the intense staining present in the wells, which represents crosslinked chromatin that cannot enter the gel matrix (Fig. 7A). As expected, western blotting showed that the core histone H3 was also enriched in these fractions (Fig. 7B). TAG-PRAME and endogenous PRAME were present in the free-protein fraction (fraction 1), but they were also detected in the chromatin fractions, consistent with their association with chromatin *in vivo*

A

*Bam*HI

Kozak sequence

TY1
TY1
ERa

*Bgl*II
*Sal*I
*Eco*RI
*Xho*I

gag

SV40 early promoter

PuroR

pTTE
5213 bp

5' LTR

3' LTR

AmpR

origin

FIGURE 8
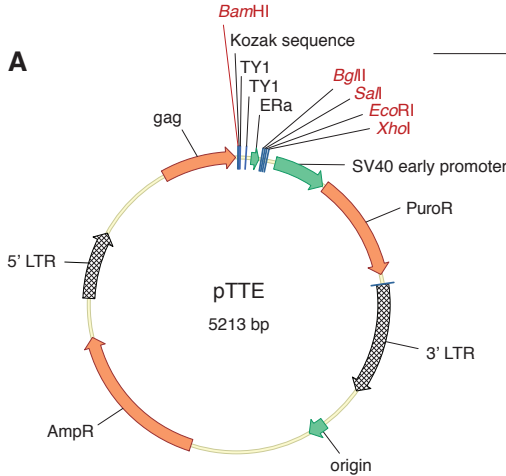
The TTE (TY1-TY1-ER) epitope tag.
(A) Map of the pTTE plasmid, and coding sequence of the TY1 and ER epitope tags in orange and blue, respectively.
(B) Western blot on K562 stable cell lines carrying the pTTE empty vector (negative control) or the pTTE-PRAME expression vector.

+1

TY1 epitope

TY1 epitope

BamHI | Met | Glu Val His | Thr Asn Gln | Asp Pro Leu | Asp | Ala | Glu | Val His Thr Asn

1 GGATCCGCCA TG GAGGTGCA CACCAACCAG GACCCACTGG ACGCC GAGGT GCACACCAAC

ERα epitope

Gln Asp Pro | Leu Asp | Ser Leu | Gln Lys Tyr | Tyr Ile Thr | Gly | Glu Ala | Glu Gly Phe Pro

1 CAGGACCCAC TGGAC TCCTT GCAGAAGTAT TACATCACGG GAGAGGCAGA GGGTTTCCCT

BglII   SalI   EcoRI   XhoI

Ala Thr Val | Ser Asp Leu Pro | Val Asp Ala | Glu Phe His | Met Leu Glu

1 GCCACGGTCT CAGATCTGCC AGTCGACGCC GAATTCCATA TGCTCGAG

B

TTE-PRAME
PRAME
(WB: P4)

(WB: B6)

ERα
TTE-PRAME
(WB: PRAME)

(Fig. 10B). Consistent with previous western blot results, a protein unrelated to PRAME with a higher apparent molecular size was recognized by the P4 antibody in the nuclear extracts. Importantly, this protein was not detected in the input or fractionated chromatin fractions, suggesting that the use of V27/P4 antibodies for ChIP assays should not result in cross-reactivity with proteins other than PRAME (Costessi *et al*, 2011).

In the absence of validated binding sites for PRAME in the K562 cell line, we setup an epitope-tagging approach to identify PRAME binding sites in an unbiased way. For this purpose, we established K562 stable cell lines expressing PRAME fused to a composite N-terminal TTE tag, including two TY1 epitopes and one ERalpha epitope, against which efficient ChIP-grade mono-
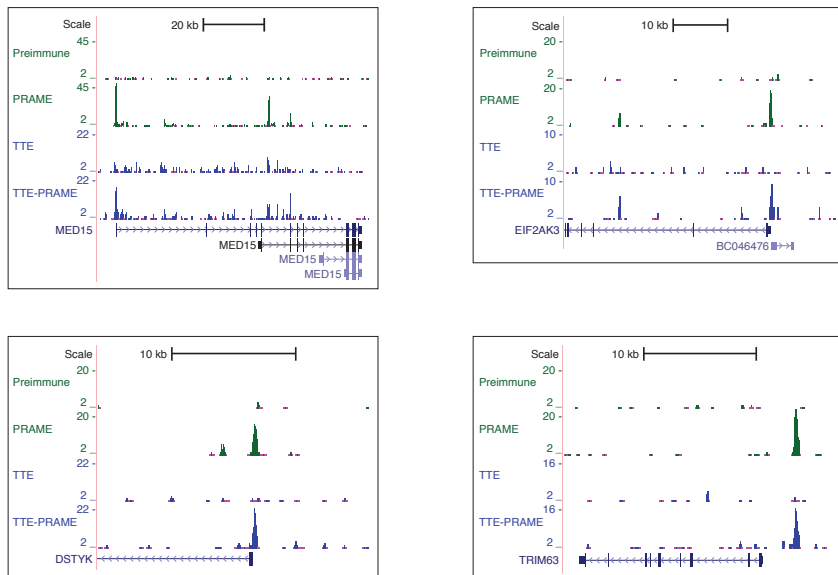
Genome-wide profiling of TTE-PRAME and endogenous PRAME association with chromatin.
UCSC genome browser views of ChIP-seq profiles generated with preimmune serum (negative control), PRAME V27 serum on K562 wild type cells, BB2 antibodies on K562-TTE control cells (TTE) and K562-TTE-PRAME cells (TTE-PRAME). Specific peaks identified at the promoters of the *MED15, EIF2AK3, DSTYK* (previously known as *RIPK5*), and *TRIM63* genes are shown.

143

clonal antibodies are available: BB2 and F3A6, respectively (Fig. 8A). Expression of TTE-PRAME was tested by western blot in the polyclonal cell line K562-TTE-PRAME and the control cell line K562-TTE carrying the empty vector (Fig. 8B). Next, we performed an exploratory ChIP-seq experiment using the BB2 antibody (against the TY1 tag) and chromatin from K562-TTE control and K562-TTE-PRAME cells. Sequencing was performed on a Solexa GAIIx with a single-read protocol for 36 cycles and generated 4 059 488 reads for the TTE control and 3 615 534 reads for TTE-PRAME after alignment to the human hg18 reference genome. Reads numbers were normalized to the lowest number and average profiles were visualized in the UCSC genome browser. Manual inspection of these profiles revealed a number of loci with specific enrichment of sequence tags for TTE-PRAME (Figure 9). ChIP-qPCR assays were performed with primer sets targeting these putative binding sites, and the assay conditions were optimized to improve the signal-over-background ratio. The results confirmed that the sites tested are *bona fide* binding sites for TTE-PRAME (see chapter 3).

We then performed ChIP-qPCR experiments using our PRAME antibodies and chromatin from K562 wild type cells. As shown in chapter 3 (Supplementary Figure 1), the V27cl serum gave good recoveries, indicating that it is ChIP-grade and establishing that also the endogenous PRAME protein binds to the genomic loci tested.

Based on these results, we performed a ChIP-seq profiling on K562 wild type cells using the PRAME antibody V27, or the preimmune serum from the same rabbit as negative control. We obtained 3 101 777 sequence reads for PRAME and 2 815 020 sequence reads for the preimmune serum. Screenshots in Figure 9 show genomic regions bound by both TTE-PRAME as well as endogenous PRAME.

Next, we used the MACS algorithm (Zhang *et al*, 2008) to identify binding sites for TTE-PRAME and PRAME using their corresponding control tracks as negative controls. We iterated the peak calling with different stringency settings: the "e-value" was set to 5, 6, or 7, where 5 corresponds to relaxed settings (that can result in a higher false-positives rate), while an e-value of 7 corresponds to more stringent settings. Table 2 shows that significantly more peaks were called for the endogenous PRAME than for TTE-PRAME.

TABLE 2.   **Number and overlap of binding sites identified for TTE-PRAME and endogenous PRAME.**

e-value	TTE-PRAME peaks	PRAME peaks	Overlap TTE-PRAME with PRAME peaks
e-5 (m16)	7517	24611	41% (3084)
e-6 (m16)	2467	14153	56% (1380)
e-7 (m16)	1142	9228	63% (579)

This is consistent with the lower recoveries obtained for TTE-PRAME in ChIP-qPCR experiments and a higher signal-to-noise ratio for the PRAME serum than the BB2 antibody. We next calculated the overlap between binding sites identified for TTE-PRAME and endogenous PRAME. Notably, the percentage of overlapping peaks was dependent on the stringency applied to the peak calling: the overlap was 41% for e5 peaks, 56% for e6 peaks, and 63% for e7 peaks. This suggests that low e-values (low stringency) result in larger pools of putative binding sites, likely containing a higher percentage of false positives, while higher e-values (high stringency) result in the identification of smaller numbers of high-confidence peaks, although binding sites might be missed.

Figure 10

A

B

C
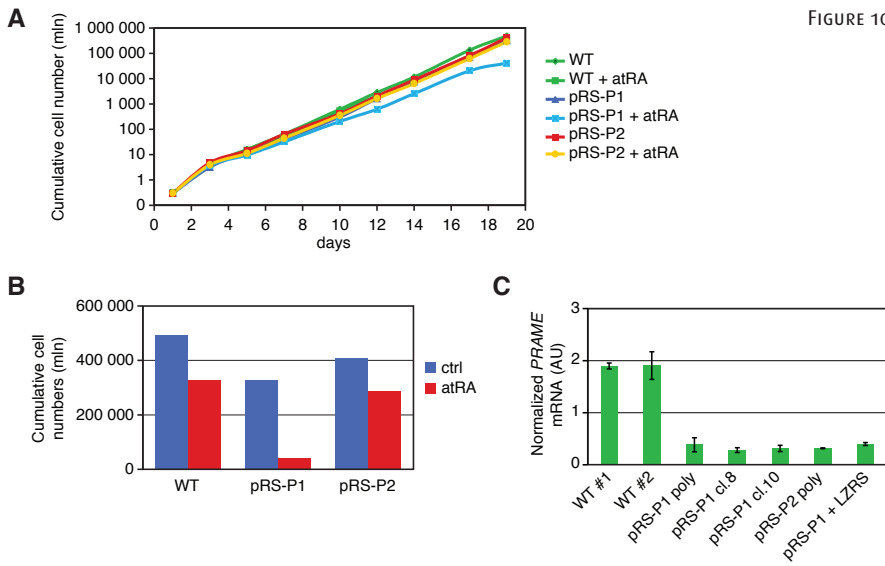
Cell proliferation assays upon PRAME knock-down.
(A) Growth curve of A375 cell lines upon treatment with 1µM atRA or vehicle (ethanol). A375-pRS-P1 and A375-pRS-P2 cells were kept under puromycin selection.
(B) Cumulative cell numbers at the end of the growth assay (day 19).
(C) Gene expression analysis of *PRAME* in A375 cell lines used for the growth assay (wild type, pRS-P1and pRS-P2), two single cell clones derived from the A375-pRS-P1 polyclonal population, and double-stable for pRS-P1 and LZRS control vector. Expression was measured in triplicate and normalized to a pool of 3 control genes (AU, arbitrary units).

Take together, our data demonstrate that the oncoprotein PRAME is a chromatin-associated factor, we validated an epitope-tagging approach as well as a PRAME antibody for ChIP-seq analyses and generated the first genome-wide profile of PRAME association to chromatin.

PRAME and retinoic acid signaling

Recent work indicated that overexpression of *PRAME* prevents the growth arrest and apoptosis induced by histone deacetylase inhibitors (HDACIs) (Epping *et al*, 2007). The authors proposed a model in which HDACI-induced growth arrest requires an active retinoic acid signaling pathway, which PRAME would repress via direct protein-protein interactions with the retinoic acid receptor alpha (RARalpha), recruitment of the EZH2 polycomb protein and eventual transcriptional repression of RARalpha target genes (Epping *et al*, 2005).

In support of their model, the authors reported that knockdown of PRAME in A375 melanoma cells caused downregulation of the retinoic acid receptor alpha (RARa) at the protein level (Epping *et al*, 2005, Figure 6F),

A

B

C
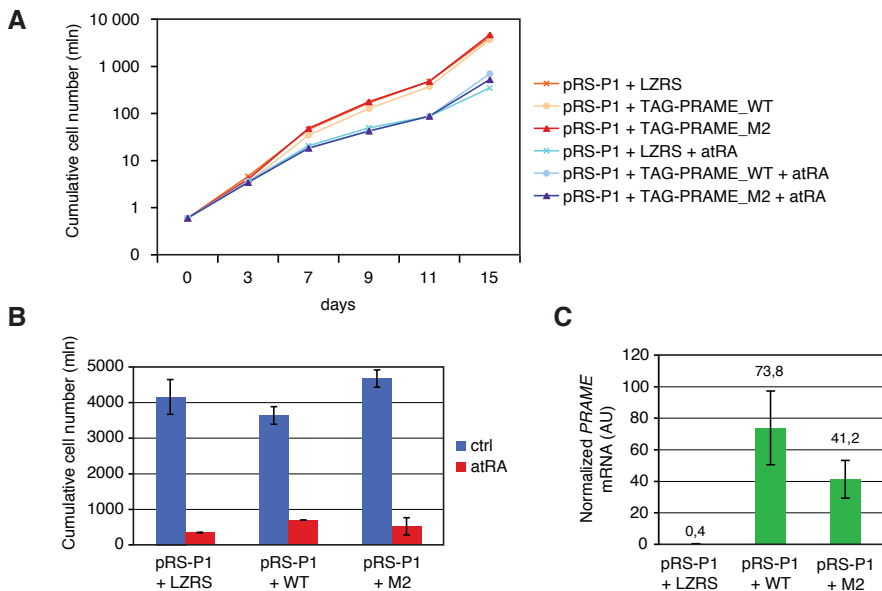
Cell proliferation assays and knockdown-resistant PRAME.
(A) Growth curve of A375-pRS-P1 polyclonal cells further transduced with LZRS control vector, LZRS-TAG-PRAME_WT_KDR, or LZRS-TAG-PRAME_M2_KDR. Cells were treated with 1µM atRA or vehicle (ethanol), and data are averages of two independent plates.
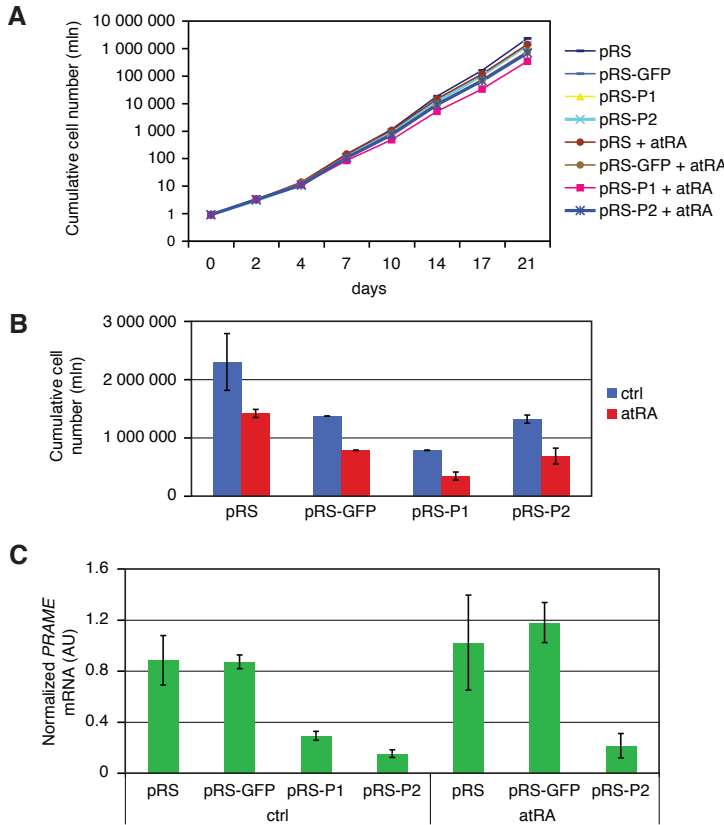(B) Cumulative cell numbers at the end of the growth assay (day 15).

(C) Gene expression analysis of *PRAME* in A375 double-stable cell lines. Expression was measured in triplicate and normalized to a pool of 3 control genes (AU, arbitrary units).

Figure 11

which was interpreted as a sign of derepression of retinoic acid signaling. Remarkably, the same experiment presented in a follow-up publication showed no difference in RARalpha levels upon PRAME knockdown (Epping *et al*, 2007, Figure 4E). Although the authors did not comment on this difference, this discrepancy raises doubts about the reproducibility of the reported RARalpha-PRAME interplay.

Further, the authors reported that knockdown of *PRAME* sensitized A375 cells to the growth inhibiting effects of retinoic acid treatment (Epping *et al*, 2005). Importantly, complementation of the cells with a knockdown-resistant version of PRAME was shown to re-establish a normal growth curve, indicating that expression of PRAME confers resistance to retinoic acid-induced growth arrest.

We have recently characterized PRAME as a component of Cullin2-based E3 ubiquitin ligases (see chapter 3 and Costessi *et al*, 2011). In order to confirm whether PRAME can act as a repressor of the retinoic acid signaling as reported, we performed growth assays upon downregulation of *PRAME* in

A (A) Growth curve of A375 stable cell lines upon treatment with 1µM atRA or vehicle (ethanol).

Cell proliferation assays upon PRAME knockdown and control vectors.
(A) Growth curve of A375 stable cell lines upon treatment with 1µM atRA or vehicle (ethanol).
(B) Cumulative cell numbers at the end of the growth assay (day 21).
(C) Gene expression analysis of *PRAME* in A375 cell lines used for the growth assay. Expression was measured in triplicate and normalized to a pool of 3 control genes (AU, arbitrary units).

Figure 12

the presence or absence of retinoic acid.

First, we generated A375 stable knockdown cell lines by retroviral transduction using two different silencing constructs, pRS-P1 and pRS-P2 and compared them to wild type untransduced cells. Construct pRS-P1 was the one used in the publication of Epping *et al.* (Epping *et al*, 2005), while pRS-P2 is an independent construct targeting a region of the PRAME transcript partially overlapping with the region targeted by pRS-P1 (Figure 4). Cells were treated with a final concentration of 1 µM all-trans retinoic acid (atRA) or the equal volume of vehicle (ethanol), and knockdown cells were kept under puromycin selection during the course of the experiment. Cumulative cell numbers reported in Figure 10A-B reflect a growth inhibition on A375-pRS-P1 cells upon retinoic acid treatment. On the contrary, A375-pRS-P2 cells were only slightly affected as compared to the wild type control cells, even though *PRAME* was downregulated to similar levels by both silencing constructs (Figure 10C).

Next, we generated wild type and BC-box mutant versions of PRAME that are resistant to knockdown (KDR) in order to test if ectopic re-expression of PRAME could rescue the growth phenotype, as previously reported (Epping *et al*, 2005, Supplementary Figure 3). For this purpose, we used the LZRS retroviral system to express either wild type or the M2 BC-box mutant TAG-PRAME in A375-pRS-P1, and the empty LZRS vector was used as negative control. Unexpectedly, a growth assay on these cell lines revealed that complementation of A375-pRS-P1 cells with wild type PRAME did not rescue their growth phenotype, in disagreement with the published data (Figure 11A-B). Expression of TAG-PRAME_KDR was confirmed by cDNA expression analysis (Figure 11C).

Given this discrepancy, we setup an independent knockdown experiment in which two control lines were also included using the empty pRS vector and the pRS-GFP vector expressing a silencing hairpin against GFP. New PRAME

knockdown cell lines were generated at the same time with pRS-P1 and pRS-P2 vectors. Growth curves of these lines are reported in Figure 12A-B.

First, cells carrying the empty vector grew better than the cells expressing the GFP control hairpin. This result is consistent with reports that overexpression of silencing hairpins can overload the silencing machinery and therefore lead to growth defects and cell mortality *in vitro* and *in vivo*, independent from the downregulation of specific target genes (Grimm *et al*, 2006). This result further underscores the critical importance of proper controls for knockdown experiments.

Second, no significant difference was found between A375-pRS-GFP control and A375-pRS-P2 knockdown cells, while A375-pRS-P1 cells were clearly affected: almost no A375-pRS-P1 cells survived after three weeks of treatment with retinoic acid, consistent with the first experiment. Untreated A375-pRS-P1 cells displayed a slower growth rate than other cell lines and this might be caused by an increased sensitivity to the low levels of retinoic acid present in the medium used.

In order to investigate the lack of phenotype in A375-pRS-P2 lines, we measured *PRAME* expression levels during the course of the experiment (day 17). The results displayed in Figure 12C established that both pRS-P1 and pRS-P2 vectors lead to a significant and similar downregulation of *PRAME*. Therefore, the lack of phenotype for A375-pRS-P2 cells could not be explained by inefficient knockdown of *PRAME* by this vector as compared to

pRS-P1; atRA-treated A375-pRS-P1 cells could not be tested because too few cells were present in the plates.

Taken together, independent experiments with polyclonal cell populations consistently showed that vector pRS-P1 caused a severe growth phenotype in A375 melanoma cells exposed to 1 μM retinoic acid. However, the pRS-P2 vector did not affect cell growth compared to the pRS-GFP control, even though it silenced *PRAME* expression with equal efficiency as pRS-P1 (~80%).

Moreover, complementation of A375-pRS-P1 cells with knockdown-resistant TAG-PRAME did not alleviate the growth phenotype. Formally, we cannot exclude that overexpression of TAG-PRAME_KDR and the epitope tag used might have prevented an efficient rescue of the growth phenotype. However, this epitope tag has been successfully used for the purification of native PRAME-containing protein complexes, showing that this tag does not alter the protein structure of PRAME or its ability to establish protein-protein interactions.

In contrast to the report of Epping *et al.*, our data suggest that the growth phenotype induced by the pRS-P1 silencing vector is highly unlikely caused by down-regulation of *PRAME*, but rather by down-regulation of an unrelated gene which is specifically targeted by construct pRS-P1, but not by pRS-P2.

CONCLUSIONS

In this chapter, we described the development of antibodies against the human oncoprotein PRAME, we tested the reported interplay between PRAME and the retinoic acid signaling using knockdown systems, we established the in vivo chromatin association of PRAME and reported the first genome-wide ChIP profiling of PRAME.

We immunized a total of ten rabbits against five different peptides designed in regions of human PRAME with low sequence similarity to other annotated or predicted PRAME-like proteins (Figure 1 and Table 1). Sera and affinity-purified antibodies were extensively tested in dot blot, western blot, immunoprecipitation and ChIP assays and a number of them were characterized by high sensitivity and specificity (Figures 2-5). Taken together, our data indicate that antibodies raised against peptide 2 (V19, V20, P1) and peptide 3 (V22, P2) can be efficiently used to detect epitope-tagged versions of PRAME by western blot, but are not suited for endogenous PRAME because of cross-reactivity with other proteins of similar size. On the contrary, antibodies raised against peptide 7 (V27, V28, P3, P4) are suitable for detection of both epitope-tagged as well as endogenous PRAME. Furthermore, V19, V27, P4 (Figure 5) and V28 (not shown) antibodies performed successfully in immunoprecipitations assays. Notably, our antibodies performed better than a commercially available PRAME antibody, which had been used in a number of recent publications. Our results underscore the importance of extensive testing of both specificity and sensitivity of antibodies.

We further discussed a general experimental strategy to generate ChIP-seq binding profiles of a factor of interest (Figure 6), which we applied to our study of the PRAME oncoprotein. We established that PRAME is present in the crosslinked-chromatin fraction in caesium chloride gradient experiments (Figure 7) and we successfully applied an epitope-tagging approach with a double TY1 epitope to discover genomic binding sites for TTE-PRAME with a "blind" ChIP-seq profiling (Figure 8 and 12). Moreover, we validated the ChIP-profiling properties of one of our PRAME antibodies, and we generated the first profiles for the endogenous PRAME protein in K562 cells (Figure 9).

We investigated the reported function of PRAME as dominant repressor of retinoic acid signaling. In particular, we aimed to investigate if the interaction of PRAME with Cullin2-based ubiquitin ligases plays in role in the

reported repression. We performed cell proliferation assays in the presence of all trans retinoic acid (atRA) on A375 melanoma cells, similarly to recent studies from the Bernards lab (Epping *et al*, 2005). While published work included only one silencing vector, we used two different knockdown plasmids in our experiments (pRS-P1 and pRS-P2). Notably, the regions of the PRAME transcript which are targeted by the two silencing plasmids showed a significant overlap (Figure 4) and both constructs resulted in ~80% downregulation of *PRAME* transcript (Figures 4B, 6-8).

Unexpectedly, several independent experiments reproduced the reported growth phenotype only in A375-pRS-P1 cells, but not in A375-pRS-P2 cells (Figure 10 and 12). Moreover, complementation with a knockdown-resistant PRAME did not rescue the growth phenotype of A375-pRS-P1 cells in our assays (Figure 11), although a rescue was previously reported (Epping *et al*, 2005). Since wild type PRAME did not rescue the growth phenotype, we could not investigate the potential role of a BC-box mutant PRAME in this assay. Taken together, our data suggest that the growth phenotype induced by the pRS-P1 silencing vector might not be caused by down-regulation of *PRAME*, but rather by down-regulation of an unrelated transcript which is specifically targeted by construct pRS-P1, but not by pRS-P2. A genome-wide mRNA-seq profiling experiment on these cell lines could potentially identify this gene and shed light on the reported growth phenotype.

MATERIALS AND METHODS

Antigen preparation and rabbit immunization

PRAME peptides with a leading ATA group were synthesized at the Dept. Organic Chemistry (Radboud University) and checked by mass spectrometry. The peptides were first conjugated to KLH (Imject Mariculture Keyhole Limpet Hemocyanin, Pierce, product #77600) according to standard protocols.

A total of ten rabbits (two rabbits for each peptide) were subjected to a first immunization by lymphnode injection of 200 µg of KLH-peptide and Freund adjuvant. Every three weeks, an immunization booster was performed by subcutaneous injection of 400 µg of KLH-peptide and Freund adjuvant. Blood was harvested two weeks every each immunization booster. Rabbit V21 died prematurely during the experiment before the fifth serum harvest. The rabbits were subjected to 8 cycles of booster immunization and blood collection, except rabbits V19, V22, V27, and V28 which were subjected to a total of 10 cycles.

Antibody purification with caprilic acid/ammonium sulphate precipitation

4 ml of rabbit sera were mixed with 2 volumes (8 ml) of 60 mM sodium acetate buffer pH 4.0 (for 250 ml of buffer: 1,23 g/l sodium acetate trihydrate, 3,06 g/l acetic acid). The pH was adjusted to 4,8, if necessary, and 300 µl caprilic acid were added drop-wise. The solution was incubated for 30 minutes on a rotating wheel at room temperature, and centrifuged for 20 minutes at 5000 g, 10 °C. This step precipitates abundant serum proteins, while IgG molecules remain soluble. The top lipid layer and pellet were discarded, the supernatant carefully decanted (~11-12 ml) and 150-200 µl Tris 2M (pH 11) was added to neutralize the acidic pH. A half volume (5,5-6 ml) of saturate ammonium sulphate pH 7,0-7,2 (132 g in 250 ml water) was added slowly to the supernatant, and the solution was incubated for 4 h in rotation at 4 °C. This pre-precipitation step removes large protein aggregates and any proteins that may precipitate with low concentrations of ammonium sulphate. After 30 minutes centrifugation at 3000 g at 4 °C, the supernatant was carefully transferred to a new tube, and enough ammonium sulphate (5,5-6 ml) was added slowly to obtain a 50% ammonium sulphate saturation. This step precipitates IgG molecules. The solution was incubated in rotation overnight at 4 °C, and centrifuged for 30 minutes, 3000 g, 4 °C. The pellet containing IgG was resuspended with 0,8-1 ml PBS and dialyzed versus three changes of PBS overnight. On average, the final antibody solution doubles its volume after dialysis. Dialysis tubing was prepared by boiling twice for 5 min in 5 mM EDTA, 200 mM sodium bicarbonate, and eventually rinsed in sterile water.

Peptide affinity purification of antibodies

An affinity column was prepared by covalent binding of the peptides used in the immunization protocol to the agarose matrix SulfoLink (#2040, Pierce). Crude sera were then passed through the columns, so that peptide-binding IgG bind to the matrix. After several washes, the IgG were eluted with 100 mM glycine pH 2,7. Next, a gel filtration step (G25, Hightrap, Acta, GE Healthcare) was performed to remove glycine from the eluates and exchange the buffer for PBS (NaCl 0,14 M, KCl 0,0027 M, KH_2PO_4 0,001 M, Na_2HPO_4 0,01 M) pH 7,4 containing NaN_3 0,05% w/v and ProClin300 0,05% w/v. The crude sera, flow-through from the affinity column (unbound antibodies) and the purified antibodies were tested by ELISA (see Supplementary Materials).

Western blot

Proteins were separated by conventional SDS-PAGE or NuPAGE 4-12% BisTris gels (Invitrogen) run with MES buffer. After blotting, protein transfer was checked by staining of the membranes with Ponceau S. The PRAME sera were diluted 1:1000 in standard PBS/T with 5% skimmed milk, and incubated for 1 hour at room temperature. Western blots were visualized by ECL (GE Healthcare).

Cell culture, stable cell lines and growth assays

Stable cell lines were generated as described (Costessi *et al*, 2011). The pRetroSuper plasmid carries a Puromycin resistance gene, while LZRS carries zeocin resistance. For cell proliferation assays, between

300 000 and 900 000 A375 cells were seeded in 9 cm tissue culture dishes. Cells were passages every 2-3 days, when at least one plate reached ~70% confluence and cell densities measured (Beckman Coulter). Cells were detached with 1 ml 0,5% Trypsin-EDTA (GIBCO), incubated 5 min at 37°C and 3 ml medium were added (10% FCS, 1% PenStrep cocktail, DMEM 1640 with GlutaMAX-I, GIBCO). All plates were diluted with the same factor, so that the fastest growing cells would be confluent after 2-3 days. Either 1 µM atRA or the same volume of ethanol 96% (vehicle) was added to the medium when cells were split. Stable cell lines carrying pRS vectors were kept on puromycin selection (2 µg puromycin/ml medium).

Chromatin isolation CsCl$_2$ gradient

Chromatin was isolated from 150 million K562 and K562-TAG-PRAME cl.12 cells essentially as described (Costessi *et al*, 2011) with the following modifications: cells were crosslinked for 30 min in 40 ml volume and sonicated in 2 ml buffer D without SDS (1mM EDTA, 0,5 mM EGTA, 20 mM HEPES and protease inhibitors) for 15 min, because SDS is not compatible with caesium chloride gradient sedimentation. Next, 2 ml of chromatin were mixed with 1,25 g CsCl$_2$ (final concentration 568 mg CsCl$_2$/ml), 70 µl sarcosyl 15% and 1-2 µl EtBr, and centrifuged at 260 000 g for 48 h or overweekend. Samples from different layers were collected and dialysed o/n against buffer D with 5% glycerol.

ACKNOWLEDGMENTS

REFERENCES

Brinkman A & Stunnenberg H (2009) Strategies for epigenome analysis. In *Epigenomics*, Martienssen RA Greally JM & Ferguson-Smith EAC (eds) Springer Available at: http://www.springer.com/biomed/book/978-1-4020-9186-5

Costessi A, Mahrour N, Tijchon E, Stunnenberg R, Stoel MA, Jansen PW, Sela D, Martin-Brown S, Washburn MP, Florens L, Conaway JW, Conaway RC & Stunnenberg HG (2011) The tumour antigen PRAME is a subunit of a Cul2 ubiquitin ligase and associates with active NFY promoters. *EMBO J* **30**: 3786–3798

De Carvalho DD, Binato R, Pereira WO, Leroy JMG, Colassanti MD, Proto-Siqueira R, Bueno-Da-Silva AEB, Zago MA, Zanichelli MA, Abdelhay E, Castro FA, Jacysyn JF & Amarante-Mendes GP (2011) BCR-ABL-mediated upregulation of PRAME is responsible for knocking down TRAIL in CML patients. *Oncogene* **30**: 223–233

Epping MT, Wang L, Edel MJ, Carlée L, Hernandez M & Bernards R (2005) The human tumor antigen PRAME is a dominant repressor of retinoic acid receptor signaling. *Cell* **122**: 835–847

Epping MT, Wang L, Plumb JA, Lieb M, Gronemeyer H, Brown R & Bernards R (2007) A functional genetic screen identifies retinoic acid signaling as a target of histone deacetylase inhibitors. *Proc Natl Acad Sci USA* **104**: 17777–17782

Grimm D, Streetz KL, Jopling CL, Storm TA, Pandey K, Davis CR, Marion P, Salazar F & Kay MA (2006) Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* **441**: 537–541

Ikeda H, Lethé B, Lehmann F, van Baren N, Baurain JF, de Smet C, Chambost H, Vitale M, Moretta A, Boon T & Coulie PG (1997) Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory receptor. *Immunity* **6**: 199–208

Partheen K, Levan K, Osterberg L, Claesson I, Fallenius G, Sundfeldt K & Horvath G (2008) Four potential biomarkers as prognostic factors in stage III serous ovarian adenocarcinomas. *Int J Cancer* **123**: 2130–2137

Partheen K, Levan K, Osterberg L, Claesson I, Sundfeldt K & Horvath G (2009) External validation suggests Integrin beta 3 as prognostic biomarker in serous ovarian adenocarcinomas. *BMC Cancer* **9**: 336

Passeron T, Valencia JC, Namiki T, Vieira WD, Passeron H, Miyamura Y & Hearing VJ (2009) Upregulation of SOX9 inhibits the growth of human and mouse melanomas and restores their sensitivity to retinoic acid. *J. Clin. Invest.* **119**: 954–963

Quintarelli C, Dotti G, De Angelis B, Hoyos V, Mims M, Luciano L, Heslop HE, Rooney CM, Pane F & Savoldo B (2008) Cytotoxic T lymphocytes directed to the preferentially expressed antigen of melanoma (PRAME) target chronic myeloid leukemia. *Blood* **112**: 1876–1885

Robertson KA, Mueller L & Collins SJ (1991) Retinoic acid receptors in myeloid leukemia: characterization of receptors in retinoic acid-resistant K-562 cells. *Blood* **77**: 340–347

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W & Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137

DISCUS–
SION

157

In this thesis, we explored the molecular function(s) of the human oncoprotein PRAME using an innovative combination of biochemical protein-complex purification techniques and next-generation sequencing. We dissected the protein-protein interaction network of PRAME applying epitope-tagging and retroviral transduction in cell lines (chapter 2). Mass spectrometry analysis provided the first unbiased list of PRAME interactors, including all subunits of Cullin2-based ubiquitin ligases (chapter 2). These interactions were further characterised in chapter 3, where we identified PRAME as a BC-box protein and substrate receptor component of Cullin2-based ubiquitin ligases. This data associated for the first time a defined biochemical activity to the PRAME oncoprotein. Furthermore, we used specific antibodies and ChIP-sequencing techniques to generate the first genome-wide profiling of PRAME interactions with chromatin. We discovered that PRAME is specifically recruited to transcriptionally and epigenetically active promoter regions bound by NFY and nearby enhancers, implicating PRAME in the NFY signaling pathway. Optimisation of the protein affinity purification protocols and a complex walking approach allowed us to identify and characterise the human EKC/KEOPS complex and its interactions with Cul2-PRAME ligases (chapter 4). In chapter 5, we described the development and validation of PRAME antibodies that have been a crucial tool for our study, and we discussed in detail the strategy we undertook to generate the genome-wide ChIP-seq profiles presented in chapter 3. Finally, we reported the results of cell proliferation assays that do not support the previously reported function of PRAME as dominant repressor of retinoic acid signaling (chapter 5).

158

The results presented in this thesis are discussed in more detail in this chapter together with an outlook on the future perspectives of this field of research.

Epitope-tagging technology and dissection of protein complexes

During the past two decades, epitope tagging has become an indispensable technology in many fields of research, including biochemistry, cellular and molecular biology, and biotechnology. Affinity fusion tags have been used as research tools for detection, purification, and immobilisation of recombinant proteins expressed in bacteria, yeast, insect cells, plants, and mammalian cells (Nilsson *et al*, 1997). We applied this technology to purify

PRAME-containing protein complexes and identify its protein interactors by mass spectrometry analysis. In chapter 2, we described our approach based on the LZRS retroviral expression system and the StrepII-TEV-MYC-3xHA composite epitope tag that we used throughout our study. We setup a single-step immunoprecipitation protocol to purify PRAME-containing protein complexes using crosslinked MYC antibodies and peptide elution (chapters 2 and 3). The yield of the protein-complex purifications were further increased by the use of crosslinked HA antibodies and acidic glycine elution (chapter 4). Notably, this improvement of the protocol allowed us to significantly reduce the amount of cells and protein extracts needed for one purification experiment, and we could scale down from about 10 liters of cell culture (~10^10 cells) to about 3 liters (~3x10^9 cells) for each purification experiment.

Taken together, our mass spectrometry data combined with biochemical experiments characterized PRAME as a BC-box protein and component of Cullin2-based E3 ubiquitin ligases. The interactions between PRAME and the Cullin complex are mediated by conserved amino acid motifs (the so-called "BC-box" and "Cul2-box") at the N-terminus of PRAME. In vitro experiments with purified recombinant proteins further established that a reconstituted Cullin2–PRAME complex is a catalytically active ubiquitin ligase. Moreover, we discovered that PRAME bridges Cul2 ligases to EKC via direct interactions between PRAME and the OSGEP and LAGE3 subunits, supporting a functional link between these complexes (chapter 4).

Cullin-based ubiquitin ligases

Cullin complexes (also referred to as cullin-RING ligases, CRL) are multi-subunit enzymes and constitute the largest known class of ubiquitin ligases (Petroski & Deshaies, 2005; Lipkowitz & Weissman, 2011; Okumura *et al*, 2012). Owing to the different cullin proteins expressed in human cells and the great diversity of their substrate-receptor subunits, hundreds of distinct cullin-based E3 ubiquitin ligases are thought to exist, which establishes these enzymes as key mediators of post-translational protein regulation (Petroski & Deshaies, 2005). CRLs have been implicated in a wide range of cellular processes, including cell cycle regulation, DNA replication, gene transcription, and development. Moreover, substrate receptor subunits of these ligases are targeted by viral and bacterial pathogens to subvert normal cellular processes. Although most target proteins that bind to CRL substrate receptors are ubiquitinated as a prelude to degradation, proteolysis is not the only

outcome, and ubiquitination can also lead to post-translational regulation without alteration of protein levels (Petroski & Deshaies, 2005).

All CRLs share a similar modular architecture with a central cullin protein of elongated shape. The cullin N-terminus binds the substrate-selector module, while the C-terminus interacts with a RING box protein (Petroski & Deshaies, 2005). The N-terminal domain of Cullin2 interacts with the heterodimer ElonginB-ElonginC to bind a BC-box protein, which in turn recruits the ubiquitination substrate to the ligase (Okumura *et al*, 2012). The Von Hippel-Lindau tumour suppressor (VHL) is one of the most well-studied BC-box proteins and mutations in its coding sequence lead to renal carcinomas because of deregulation of the hypoxia-inducible factor-1α (HIF1α) transcription factor (Okumura *et al*, 2012). In standard normoxic conditions HIF1α is hydroxylated, which creates a binding site for VHL and subsequent ubiquitination by a Cul2-VHL ubiquitin ligase and degradation by the proteasome.

In hypoxic conditions, however, HIF1α is not hydroxylated and cannot be recognised by the Cul2-VHL ligase, resulting in increased protein levels and upregulation of genes mediating adaptation to low oxygen tension, including factors stimulating the formation of new blood vessels. Loss-of-function mutations of VHL result in constitutive higher levels of HIF1α and aberrant activation of its target genes, contributing to angiogenesis and tumour development (Okumura *et al*, 2012).

Identification of the substrate(s) of Cul2-PRAME ubiquitin ligases constitutes a crucial step in elucidating the detailed molecular mechanisms and pathways of this so-far elusive oncoprotein and should be the priority of follow-up studies. Our mass spectrometry data indicated that the EKC complex is the most abundant and stable interactor of Cul2-PRAME ligases in several cell lines and EKC subunits are therefore potential substrates of PRAME-mediated ubiquitination. As expected for BC-box proteins, PRAME was itself ubiquitinated in our experiments, but we could not detect a PRAME-mediated ubiquitination of EKC components. However, we cannot exclude that our negative results were caused by technical limitations of the assays used. Therefore, we believe that more and probably different assays should be developed to assess whether EKC subunits can be substrates of PRAME-mediated ubiquitination.

Importantly, our present knowledge of CRLs indicates that interactions

between ligases and substrates are often dependent on post-translation modifications of the substrates like proline hydroxylation, phosphorylation, or glycosylation (Verma *et al*, 1997; Yoshida *et al*, 2002; Ivan *et al*, 2001). It is therefore likely that PRAME substrates are not (or at very low level) amongst the stable interactors identified in our mass spectrometry data, which were generated from cell lines cultured in standard conditions. Specific signaling events might be necessary to trigger the interaction between PRAME and its substrate(s). Moreover, it appears likely that substrates of PRAME are expressed in a tissue-specific manner, and this would potentially account for the reported tissue-specific effects of PRAME upregulation. The apparent discrepancies of PRAME being associated to good as well as to bad prognosis cancers might be explained by differential expression of PRAME targets in different tumour types.

Intriguingly, we found a complete conservation of the BC-box and Cul2-box at the N-terminus of all the annotated PRAME-like proteins. This suggests that all members of the large PRAME protein family are likely to function as substrate recognition subunits of cullin-based ubiquitin ligases, possibly displaying different degrees of specificity and/or redundancy in substrate selection. It is tempting to speculate that the reported roles of PRAME-like genes in early development and stemness might be mediated by functional interactions with CRLs and ubiquitination of proteins involved in these pathways.

PRAME and the retinoic acid signaling

A function of PRAME as dominant repressor of the retinoic acid signaling was reported by the Bernards lab (Epping *et al*, 2005). Using a genetic screen, the authors found that ectopic expression of the retinoic acid receptor alpha (RARA) or PRAME prevented the growth arrest and apoptosis induced by histone deacetylase inhibitors (HDACIs) (Epping *et al*, 2007). They provided evidence supporting a model in which HDACI-induced growth arrest requires an active retinoic acid signaling pathway, which PRAME would repress via direct protein-protein interactions with the retinoic acid receptor alpha and the EZH2 polycomb protein, resulting in transcriptional repression of retinoic acid target genes like *RARB* and *CDKN1A* (also known as p21) (Epping *et al*, 2005; 2007). However, the suggested mechanisms do not seem to play a general role, since higher *PRAME* levels were found to correlate with

lower *RARA* levels in breast cancer patients (Epping *et al*, 2008). Moreover, the retinoic acid target genes *RARB* and *CDKN1A* were not repressed in acute myeloid leukemia patients with high *PRAME* levels (Steinbach *et al*, 2007).

Epping *et al*. reported that the C-terminus of PRAME (aa. 416-509) interacts directly with the ligand-binding domain (LBD) of RARalpha in GST pull downs and a mammalian two-hybrid system (Epping *et al*, 2005). The authors further showed that the C-terminal amino acid motif LRELL in PRAME was required for the interaction with the LBD of RARa; reciprocal mutations of the LBD of RARa were not included in the study. Using the full-length PRAME protein, however, Wadelin *et al* could not detect any interaction with the LBD of RARa using similar GST pull down or yeast two-hybrid assays (Wadelin *et al*, 2010). In agreement with the findings of Wadelin *et al*, we did not succeed in reproducing the interaction between RARa and PRAME in coimmunoprecipitations assays upon transient transfections in 293T cells, or upon immunoprecipitation of TAG-PRAME from stable cell lines in the presence or absence of ligand.

Epping *et al* further described that knockdown of PRAME in A375 human melanoma cells resulted in a higher sensitivity to the administration of retinoic acid (Epping *et al*, 2005) and HDACI (Epping *et al*, 2007). Notably, knockdown of PRAME in these cells was shown to cause downregulation of the retinoic acid receptor alpha (RARA) at the protein level (Epping *et al*, 2005, Figure 6F) and this was interpreted as a derepression of the retinoic acid signaling pathway. However, the same experiment presented in a follow-up publication showed no difference in RARA levels upon PRAME knockdown (Epping *et al*, 2007, Figure 4E), casting doubts about the reported effects. As described in chapter 5, we successfully reproduced the retinoic acid phenotype upon PRAME knockdown in A375 cells. However, this effect was limited to one of the two knockdown constructs used, despite a similar down-regulation of *PRAME* expression, and we could not rescue the phenotype by ectopic expression of knockdown-resistant version of PRAME. Our data are consistent with an "off-target" effect of the pRS-P1 knockdown construct. In other words, the phenotype observed is likely caused by down-regulation of a transcript unrelated to PRAME, which is specifically targeted by the published construct pRS-P1, but not by pRS-P2. A genome-wide mRNA-seq profiling experiment on these cell lines could potentially identify this gene and shed light on the cause of the growth phenotype observed.

Our protein-complex purifications clearly established that Cullin2 ligases and EKC subunits are the most abundant interactors of PRAME in a number of human cell lines kept in standard culture conditions, while we did not detect RARa or EZH2. However, we cannot exclude that PRAME might establish interactions with RARa and EZH2 upon activation of particular signaling pathways. Notably, Epping *et al* performed their genetic screen in p53-deficient mouse embryonic fibroblasts stably expressing an oncogenic RasV12 gene (RasV12 MEFs) (Epping *et al*, 2007). It is therefore possible that the activation of Ras pathways plays an unforeseen role in PRAME pathways.

In order to disentangle the conflicting result, purifications of PRAME-containing protein complexes from RasV12 MEFs should be performed to identify the most abundant PRAME interactors in this cell system in the presence and absence of retinoic acid. Moreover, the effect of PRAME BC-box mutations on the sensitivity to HDACI should also be tested in the same cells.

The enigmatic EKC/KEOPS complex

Our biochemical experiments demonstrate that the EKC complex is conserved and expressed also in human cells. Several crystal structures of yeast and archaeal EKC subunits were reported and supported a model in which the complex dimerises around the Pcc1/LAGE3 subunit assuming a V shape. Our experimental data together with preliminary structural modelling of the human subunits are consistent with a similar dimeric model for the human EKC. Furthermore, we detected specific interactions between the conserved EKC orthologues and an uncharacterised protein of small size, C14orf142, suggesting that this might be a novel EKC subunit or interactor.

The EKC or KEOPS complex has been first identified in yeast by two independent laboratories that described a role for the complex in either transcription regulation (Kisseleva-Romanova *et al*, 2006) or telomere maintenance (Downey *et al*, 2006). More recently, the complex has been involved in the biosynthesis of a chemical modification of tRNAs that is critical for efficient protein synthesis (Daugeron *et al*, 2011; Yacoubi *et al*, 2011; Srinivasan *et al*, 2011).

Human EKC comprises four evolutionarily conserved subunits (yeast and human names are given): Pcc1p/LAGE3, Kae1p/OSGEP, Bud32p/TP53RK, Cgi121p/TPRKB. Additionally, the yeast complex comprises a fifth subunit called Gon7p or Pcc2p, which seems to be fungi-specific. While *PCC1/LAGE3*, *BUD32/TP53RK* and *CGI121/TPRKB* are present in archaea and eukarya, but

not in bacteria, orthologues of *KAE1/OSGEP* are present in all three domains of life (although often annotated with different names). In fact, a study from 2004 found that *KAE1/OSGEP* was one of the about 60 genes present in all genomes sequenced until then, with the only exceptions of the reduced genomes of the endosymbiotic bacteria *Carsonella ruddii* and *Sulcia muelleri* (Galperin & Koonin, 2004; Galperin, 2008). Moreover, *KAE1/OSGEP* is represented in bacteria by two different homologues that were shown to establish protein-protein interactions: YeaZ (specific to bacteria) and YgjD (orthologue of Kae1/OSGEP) (Handford *et al*, 2009).

Phylogenetic analysis of 767 Kae1 homologues grouped these proteins in four distinct clusters (Hecker *et al*, 2007):

1. archaeal Kae1;
2. eukaryotic orthologues of yeast Kae1p and human OSGEP;
3. bacterial orthologues of E. coli YgjD and eukaryotic orthologues of Qri7p/OSGEPL1;
4. bacterial orthologues of E. coli YeaZ.

Interestingly, the eukaryotic orthologues of Kae1p/OSGEP and the archaeal Kae1 proteins appeared significantly more similar to one another than to the two bacteria-specific subfamilies (Hecker *et al*, 2007). Yeast Qri7p and human OSGEPL1 are nuclear-encoded paralogues that localize to mitochondria, itself a bacterial symbiont (Oberto *et al*, 2009). The clustering of Qri7p/OSGEPL1 within the bacterial orthologues of YgjD supports the bacterial origin of these genes. Unlike the YgjD/*KAE1/OSGEP* gene family, *PRAME* genes seem to have appeared significantly later in evolution and to have undergone several rounds of gene duplication events in the rodent and hominin lineages (Birtle *et al*, 2005; Rhesus Macaque Genome Sequencing and Analysis Consortium *et al*, 2007). The peculiar evolutionary history of these gene families suggests that *PRAME* might function as a modulator of EKC.

The function of YgjD/*KAE1/OSGEP* proteins has been uncertain for many years and is still unclear, and functions have been ascribed to these proteins that later appeared to be incorrect. First, the Gcp protein was cloned from *Pasteurella haemolytica* and suggested to be an O-sialoglycoprotein endopeptidase, hence the names OSGEP and KAE1 (kinase-associated endopeptidase) (Abdullah *et al*, 1991; Abdullah *et al*, 1992). However, a peptidase activity was never confirmed by independent studies (Hecker *et al*, 2007). Second, a study of the archaeal orthologue of Kae1 from *Pyrococcus abyssii* detected apurinic-endonuclease activity *in vitro* (Hecker *et al*, 2007).

Although Srinivasan *et al* detected this enzymatic activity for yeast Kae1p expressed and purified from E. coli, they reported that the endonuclease activity was caused by a low level contamination from the highly active E. coli enzyme Fpg that copurified with Kae1p (Srinivasan *et al*, 2011).

Two independent laboratories identified the yeast EKC/KEOPS complex starting from genetic clues. The laboratory of Domenico Libri isolated Pcc1, a protein of small size with no predicted functional domains in a search for suppressors of a splicing defect and demonstrated that its function is not related to splicing (Kisseleva-Romanova *et al*, 2006). Their genetic and biochemical experiments led to the description of the five-subunit EKC complex, the integrity of which was required for normal induction of pheromone- and galactose-responsive genes.

A parallel genome-wide screen in yeast identified the *cgi121* null mutant as a suppressor of the telomere-capping defect of *cdc13-1* cells (Downey *et al*, 2006). Deletion of *CGI121* or other EKC/KEOPS subunits resulted in short telomeres in an otherwise wild-type context.

Since kae1, pcc1, and pcc2/gon7 temperature-sensitive strains stop dividing at the non-permissive temperature already after a few divisions, Kisseleva *et al*. suggested that telomere maintenance is unlikely to be the primary function of the complex. The authors did not exclude that the telomere phenotype might result from an altered transcriptional regulation, or that the complex might be involved in some general chromatin remodelling that would eventually affect both gene loci and their transcription as well as the telomeres (Kisseleva-Romanova *et al*, 2006).

Most recently, the group of Crécy-Lagarde used *in silico* comparative genomics approaches to identify a critical function for Kae1 in the universal t^6A biosynthesis pathway (Yacoubi *et al*, 2011). This pathway mediates the chemical modification of position 37 of all tRNAs that pair with ANN codons (thus, including the first ATG of protein-coding genes), with the only exception of prokaryotic fMet tRNA. The modification consists of a carbonyl group and a threonine attached to the amino group of adenine at nucleotide 37 immediately 3' of the anticodon, and is necessary to prevent mispairing between the first base of the codon and the third base of the anticodon (Yacoubi *et al*, 2011).

Work from three independent laboratories established that the t^6A modification activity was dependent on Pcc1p, the ATPase activity of Kae1p,

and the kinase activity of Bud32p, but not on Cgi121p (Srinivasan *et al*, 2011; Yacoubi *et al*, 2011; Daugeron *et al*, 2011). Moreover, Qri7p was found to mediate modification of mitochondrial tRNAs (Yacoubi *et al*, 2011).

Although the most recent data show that the EKC/KEOPS complex is required for t^6A synthesis, this might not be its only role and its function in telomere maintenance might be unrelated to t^6A modification of tRNAs. More specifically, it was hypothesised that the predicted chaperone activity of this complex might contribute to several processes, t^6A formation being just one of them (Srinivasan *et al*, 2011).

Outlook on the interplay between PRAME, EKC and NFY

Studying the relationships between PRAME, Cul2 ligases and EKC, we discovered that PRAME recruited Cullin2-based ligases to EKC via direct interactions with OSGEP and LAGE3. In fact, we consistently found that PRAME established stable interactions with OSGEP and LAGE3, but not with TP53RK or TPRKB. These results might represent a PRAME-mediated disruption of an otherwise complete EKC complex, or association of PRAME to an already existing EKC subassembly. Characterisation of the amino acid residues involved in these interactions should be prioritised in follow-up studies in order to shed light on the precise stoichiometry of the complexes. In particular, it is not known whether PRAME influences the conformation of OSGEP and LAGE3 and whether PRAME might cause steric hindrance and thus inhibit the formation of a complete dimeric EKC. Ideally, resolution of crystal structures containing PRAME, OSGEP and LAGE3 could provide ultimate clues on the three-dimensional relationships of these proteins.

No data is available yet on EKC function(s) in human cells and the absence of efficient knockdown or knock-out models have severely hindered functional analyses of EKC in higher eukaryotes so far. Considering the extreme conservation of EKC subunits, however, it is reasonable to expect that at least some of the functions and pathways identified for this complex in yeast could be conserved in human cells. However, the complex seems to have acquired also additional functions during evolution, like phosphorylation of the proto-oncogene p53 by TP53RK. Future studies will have to address the connections between apparently different functions ascribed to this complex, like phosphorylation of the p53 oncogene, chemical modification of tRNAs, roles in telomere maintenance and physical association with promoter regions.

Kae1p and Pcc1p were found to associate with active chromatin and to regulate expression of some of the yeast promoters to which they are re-cruited (Kisseleva-Romanova *et al*, 2006). In a yeast strain deficient for Pcc1p the activator Gal4p was still found on chromatin, but recruitment of Gal1p, the SAGA histone acetyltransferase and Mediator were inhibited, suggesting a role for EKC in gene transcription. We found that epitope-tagged EKC sub-units in human cells can specifically associate with chromatin at active pro-moter regions bound by PRAME and NFY. Unfortunately, stable knockdown strategies did not detect a clear function of EKC or PRAME on transcriptional regulation of those loci. As we elaborated in chapter 4, likely different cell systems and assays are needed to investigate the functional role(s) of PRAME and EKC on chromatin. As soon as an assay will become available to evidence EKC functions in human cells, the role of PRAME on such functions should be tested. For example, EKC function could be measured in a PRAME-negative cell line before and after ectopic expression of PRAME.

Our genome-wide ChIP-seq analyses revealed that PRAME is strongly enriched at promoters bound by the transcription factor NFY, and binding of PRAME predicts a permissive epigenetic landscape and transcriptional activity at these regions. Moreover, we found that PRAME is also significantly enriched at adjacent enhancer regions, suggesting that it could participate in promoter-enhancer interactions. Although our genome-wide ChIP-seq pro-files are consistent with a role for PRAME in active transcription, more efforts are required to provide a definitive answer to the burning question if PRAME can function as a transcriptional modulator.

Since neither PRAME nor EKC proteins possess known DNA-binding do-mains, it is not clear whether they can bind DNA directly and if the transcrip-tion factor NFY might play a direct or indirect role in the targeting of a Cul2-PRAME-EKC complex to chromatin. However, we did not find NFY subunits in our mass spectrometry interaction data of PRAME or EKC. Future studies should investigate the molecular determinants of the association of PRAME and EKC to chromatin and the role of NFY, and whether there is a hierarchy in their binding.

Analyses of gene expression levels in human samples indicate that most normal adult tissues express NFY, EKC and Cul2, but not PRAME; therefore, it is expected that all (or a vast majority) of PRAME-expressing cells also express NFY, EKC and Cul2. Taking into consideration these expression pat-terns and the ancient nature of EKC as compared to the later appearance of

PRAME, we speculate that EKC could fulfil housekeeping (PRAME-independent) functions also in human cells and that it could interact with chromatin in a PRAME-independent manner. PRAME expression would then lead to the recruitment of a Cul2-based ubiquitin ligase to EKC genomic loci by virtue of direct protein-protein interactions, resulting in ubiquitination of proteins present at these sites, potentially affecting EKC pathways and/or NFY signaling. This hypothesis could be tested by comparing the chromatin-binding profiles of EKC subunits in PRAME-negative and PRAME-expressing cells, and development of ChIP-grade antibodies against EKC proteins will be critical tools for such studies.

Intriguingly, several studies suggested important roles for NFY in stem cell biology (Grskovic *et al*, 2007) and mouse PRAME-like genes were found to be involved in embryonic development and embryonic stem cells (Bortvin *et al*, 2003). It should therefore be studied if PRAME-like proteins can also

play roles in EKC and NFY pathways in a redundant manner with PRAME, or if they rather display specificities of their own. We foresee that the dissection of the protein-protein interaction networks of PRAME-like proteins will definitely provide stimulating and exciting new leads for the field.

Overall, the experimental work outlined in this thesis establishes a biochemical and molecular function for the PRAME gene family and provides a new twist to both PRAME and EKC biology. Our results provide an important contribution to understanding the molecular mechanisms in which these proteins are active, both in healthy and tumour cells.

REFERENCES

Abdullah KM, Lo R & Mellors A (1991) Cloning, nucleotide sequence, and expression of the Pasteurella haemolytica A1 glycoprotease gene. J Bacteriol 173: 5597–5603

Abdullah KM, Udoh EA, Shewen PE & Mellors A (1992) A neutral glycoprotease of Pasteurella haemolytica A1 specifically cleaves O-sialoglycoproteins. *Infect Immun* **60**: 56–62

Birtle Z, Goodstadt L & Ponting C (2005) Duplication and positive selection among hominin-specific PRAME genes. *BMC Genomics* **6**: 120

Bortvin A, Eggan K, Skaletsky H, Akutsu H, Berry DL, Yanagimachi R, Page DC & Jaenisch R (2003) Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. *Development* **130**: 1673–1680

Daugeron M-C, Lenstra TL, Frizzarin M, Yacoubi El B, Liu X, Baudin-Baillieu A, Lijnzaad P, Decourty L, Saveanu C, Jacquier A, Holstege FCP, de Crécy-Lagard V, van Tilbeurgh H & Libri D (2011) Gcn4 misregulation reveals a direct role for the evolutionary conserved EKC/KEOPS in the t6A modification of tRNAs. *Nucleic Acids Res* **39**: 6148–6160

Downey M, Houlsworth R, Maringele L, Rollie A, Brehme M, Galicia S, Guillard S, Partington M, Zubko MK, Krogan NJ, Emili A, Greenblatt JF, Harrington L, Lydall D & Durocher D (2006) A genome-wide screen identifies the evolutionarily conserved KEOPS complex as a telomere regulator. *Cell* **124**: 1155–1168

Epping MT, Hart AAM, Glas AM, Krijgsman O & Bernards R (2008) PRAME expression and clinical outcome of breast cancer. *Br J Cancer* **99**: 398–403

Epping MT, Wang L, Edel MJ, Carlée L, Hernandez M & Bernards R (2005) The human tumor antigen PRAME is a dominant repressor of retinoic acid receptor signaling. *Cell* **122**: 835–847

Epping MT, Wang L, Plumb JA, Lieb M, Gronemeyer H, Brown R & Bernards R (2007) A functional genetic screen identifies retinoic acid signaling as a target of histone deacetylase inhibitors. *Proc Natl Acad Sci USA* **104**: 17777–17782

Galperin MY (2008) Social bacteria and asocial eukaryotes. *Environ Microbiol* **10**: 281–288

Galperin MY & Koonin EV (2004) "Conserved hypothetical" proteins: prioritization of targets for experimental study. *Nucleic Acids Res* **32**: 5452–5463

Grskovic M, Chaivorapol C, Gaspar-Maia A, Li H & Ramalho-Santos M (2007) Systematic identification of cis-regulatory sequences active in mouse and human embryonic stem cells. *PLoS Genet* **3**: e145

Handford JI, Ize B, Buchanan G, Butland GP, Greenblatt J, Emili A & Palmer T (2009) Conserved network of proteins essential for bacterial viability. *J Bacteriol* **191**: 4732–4749

Hecker A, Leulliot N, Gadelle D, Graille M, Justome A, Dorlet P, Brochier C, Quevillon-Cheruel S, Le Cam E & Forterre P (2007) An archaeal orthologue of the universal protein Kae1 is an iron metalloprotein which exhibits atypical DNA-binding properties and apurinic-endonuclease activity in vitro. *Nucleic Acids Res* **35**: 6042–6051

Ivan M, Kondo K, Yang H, Kim W, Valiando J, Ohh M, Salic A, Asara JM, Lane WS & Kaelin WG (2001) HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. *Science* **292**: 464–468

Kisseleva-Romanova E, Lopreiato R, Baudin-Baillieu A, Rousselle J-C, Ilan L, Hofmann K, Namane A, Mann C & Libri D (2006) Yeast homolog of a cancer-testis antigen defines a new transcription complex. *EMBO J* **25**: 3576–3585

Lipkowitz S & Weissman AM (2011) RINGs of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis. *Nat Rev Cancer* **11**: 629–643

Nilsson J, Ståhl S, Lundeberg J, Uhlén M & Nygren PA (1997) Affinity fusion strategies for

detection, purification, and immobilization of recombinant proteins. *Protein Expr. Purif.* **11**: 1–16

Oberto J, Breuil N, Hecker A, Farina F, Brochier-Armanet C, Culetto E & Forterre P (2009) Qri7/OSGEPL, the mitochondrial version of the universal Kae1/YgjD protein, is essential for mitochondrial genome maintenance. *Nucleic Acids Res* **37**: 5343–5352

Okumura F, Matsuzaki M, Nakatsukasa K & Kamura T (2012) The Role of Elongin BC-Containing Ubiquitin Ligases. *Front Oncol* **2**: 10

Petroski MD & Deshaies RJ (2005) Function and regulation of cullin-RING ubiquitin ligases. *Nat Rev Mol Cell Biol* **6**: 9–20

Rhesus Macaque Genome Sequencing and Analysis Consortium, Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER, Remington KA, Strausberg RL, Venter JC, Wilson RK, Batzer MA, Bustamante CD, Eichler EE, Hahn MW, Hardison RC, Makova KD, Miller W, Milosavljevic A, Palermo RE, *et al* (2007) Evolutionary and Biomedical Insights from the Rhesus Macaque Genome. *Science* **316**: 222–234

Srinivasan M, Mehta P, Yu Y, Prugar E, Koonin EV, Karzai AW & Sternglanz R (2011) The highly conserved KEOPS/EKC complex is essential for a universal tRNA modification, t6A. *EMBO J* **30**: 873–881

Steinbach D, Pfaffendorf N, Wittig S & Gruhn B (2007) PRAME expression is not associated with down-regulation of retinoic acid signaling in primary acute myeloid leukemia. *Cancer Genet Cytogenet* **177**: 51–54

Verma R, Annan RS, Huddleston MJ, Carr SA, Reynard G & Deshaies RJ (1997) Phosphorylation of Sic1p by G1 Cdk required for its degradation and entry into S phase. *Science* **278**: 455–460

Wadelin F, Fulton J, McEwan PA, Spriggs KA, Emsley J & Heery DM (2010) Leucine-rich repeat protein PRAME: expression, potential functions and clinical implications for leukaemia. *Mol. Cancer* **9**: 226

Yacoubi El B, Hatin I, Deutsch C, Kahveci T, Rousset J-P, Iwata-Reuyl D, Murzin AG & de Crécy-Lagard V (2011) A role for the universal Kae1/Qri7/YgjD (COG0533) family in tRNA modification. *EMBO J* **30**: 882–893

Yoshida Y, Chiba T, Tokunaga F, Kawasaki H, Iwai K, Suzuki T, Ito Y, Matsuoka K, Yoshida M, Tanaka K & Tai T (2002) E3 ubiquitin ligase that recognizes sugar chains. *Nature* **418**: 438–442

PRA

PROJECT TIMELINE

173

Scientific research projects usually span several years during complementary research lines and technical approaches are combined in order to dissect and understand the molecular mechanisms, proteins or genes of interest. A typical scientific project can be described as a collection of experimental tasks performed at different points in time and often in parallel, whose results are eventually bundled together in a scientific report (i.e. an article in scientific journals, a presentation at a conference, a PhD thesis). Scientific reports are generally structured in a logical and/or chronological order: an initial finding is extended with follow-up experiments, and then a final conclusion or model is presented. Notably, articles in peer-reviewed journals focus on the results obtained, while the time and resources required to obtain them are traditionally not discussed. The efforts required to raise an antibody, to setup a novel technique or negative results are rather discussed at meetings and conferences, or at informal gatherings among researchers. On the other hand, the timeline and the financial and human resources needed for a research project have become standard requirements of increasing importance when submitting project proposals to agencies responsible for research funds.

The aim of this appendix is to provide an overview of our PRAME research project in a chronological context and describe how the different lines of research developed over time, and often in parallel. The main activities and milestones are presented in schemes called Gantt chart, a project management tool that represents the start, duration and end of each activity.

Overall, the research presented in this thesis can be divided in four main lines or tasks (Figure 1):

  i. the development of PRAME antibodies,
  ii. the identification and characterisation of the protein interactors of PRAME,
  iii. the identification and characterisation of PRAME interactions with the EKC complex,
  iv. the investigation of PRAME association to chromatin.

As depicted in Figure 2, the project started in the fall of 2005 with a focus on the first two tasks: PRAME antibody development and protein complex purification. The peptides needed to raise the antibodies were designed and synthesised, the immunisation scheme started in March 2006 and lasted about eight months until the last serum was harvested. The antibody charac-
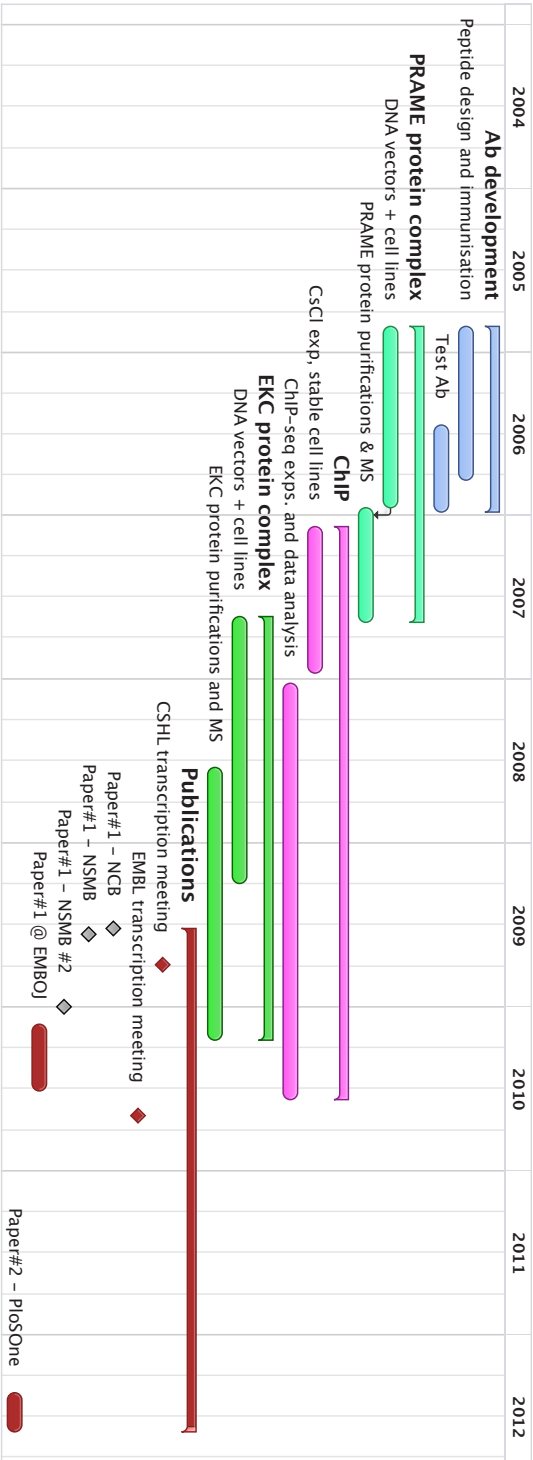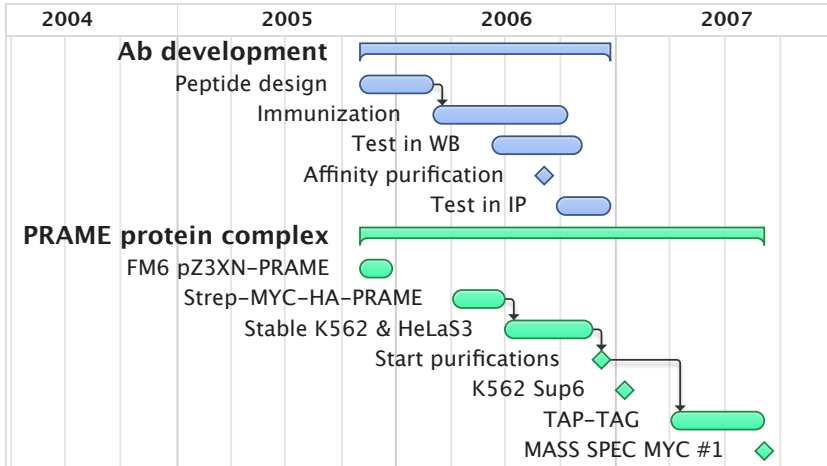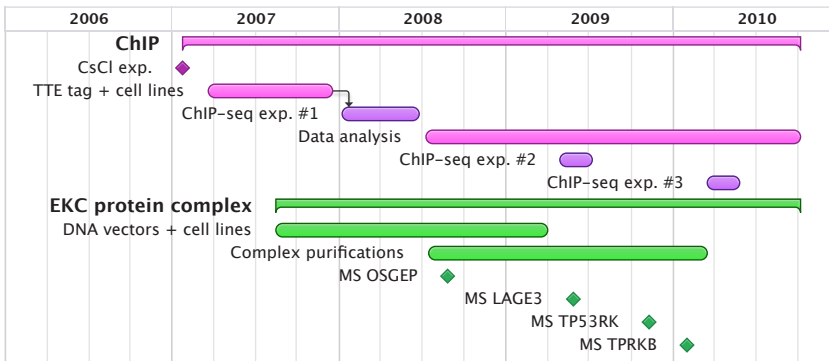
Figure 1

Figure 1

FIGURE 2

Figure 2

FIGURE 3

Figure 3

terisation and testing started as soon as the first sera became available (May 2006) and continued through the second half of 2006, when conditions were setup for use of sera and affinity-purified antibodies in western blot and immunoprecipitation assays (chapter 5).

In parallel, we attempted PRAME protein-complex purifications using FM6 stable cell lines that were already available in the lab. As described in chapter 2, this strategy was not successful because of a too low efficiency and we therefore moved to a different epitope tag (StrepII-Myc-3xHA)

combined with the use of suspension cell lines. In the first half of 2006 we generated the retroviral expression plasmids required and we setup the stable transfection protocol. In the second half of 2006 we generated single-cell clones stably expressing epitope-tagged PRAME and established that PRAME is part of high molecular weight complexes by size-exclusion chromatography (K562 Sup6 in Fig. 2). In parallel, we setup the nuclear/cytoplasmic protein extraction protocol and the immunoprecipitation strategy for protein complex purifications (TAP-TAG). These resulted in the first mass spectrometry data of PRAME interactors in July 2007 using a single MYC IP strategy (chapter 3). TAG-PRAME purification experiments were repeated several times later in the project, also in combination with optimisation of single-step HA IP strategy (chapter 4). In total, almost 2 years of experimental work were required to identify the first protein-protein interactors of PRAME. Based on these results, we started in September 2007 a close collaboration with the Conaway lab (Stowers Institute, USA).

Early in 2007 we started to study the association of PRAME to chromatin (Figure 3). As reported in chapter 5, we performed first caesium chloride gradient experiments and then setup the TTE epitope tag approach for ChIP assays. During the course of 2007, we developed the retroviral pTTE vector, cell lines stably expressing TTE-PRAME and towards the end of the year we performed the first genome-wide ChIP-sequencing experiments for TTE-PRAME (ChIP-seq #1). Analysis of the data identified the first putative binding sites which have been instrumental to setup and optimise ChIP assays for the endogenous PRAME protein in wild type K562 cells (ChIP-seq #2). Bioinformatical analysis revealed a link to the NFY transcription factor (chapter 3), whose binding sites were mapped in ChIP-seq experiments in early 2010 (ChIP-seq #3).

In the second half of the project we also characterised the discovered interactions of PRAME with Cullin2-based E3 ubiquitin ligases and with the human EKC complex (Figure 3). These experiments started with the generation of expression plasmids for all four human EKC subunits (Q3 2007), establishment of stable cell lines, ChIP assays and protein complex purifications and mass spectrometry (Figure 3). Between the end of 2008 and the beginning of 2010, mass spectrometry data for all four subunits were collected and analyzed, while biochemical experiments were performed in parallel.

Our results were first presented at the CSHL Eukaryotic transcription meeting in August 2009 and at the EMBL transcription meeting in August 2010 (Figure 1). After submissions of our work to Nature Cell Biology (NSB) and Nature Structural and Molecular Biology (NSMB), we published chapter 3 of this thesis in EMBO Journal (Q2 of 2010), and chapter 4 in PlosOne (Q3 of 2012).

SUMMARY

The genome present in the nucleus of each human somatic cell is made of about 2 meters of DNA organized in 23 chromosome pairs with a total of 6 billion basepairs. About 5% of the human genome contains protein-coding genes that can be activated to produce proteins under strict regulation. The proteins present in a cell (also called the "proteome") are not lonely islands, but they rather interact with each other in dynamic networks and multiprotein complexes. Importantly, deregulation of gene expression programs and protein interaction networks can lead to disease and cancer.

The focus of this PhD thesis is to unravel the molecular function of the human oncoprotein PRAME (preferentially expressed antigen of melanoma) using cell line model systems and a powerful combination of proteomic and genomic techniques. At the beginning of this study, PRAME was known to be frequently overexpressed in tumours and high expression levels were found to correlate with poor clinical outcome in melanoma, neuroblastoma, and breast cancer patients. Moreover, evidence that expression of PRAME is restricted mainly to tumour cells has designated it as a promising therapeutic target. However, efforts to develop strategies to interfere with the oncogenic activity of PRAME are hindered by a lack of information about its molecular function(s). Notably, analysis of the human genome revealed that a large family of more than twenty PRAME-like genes exists, whose functions are also unclear.

In order to identify the molecular functions of PRAME, we developed innovative high-throughput biochemical approaches to dissect the protein-protein interaction network of PRAME (chapter 2). By applying epitope-tagging in human cell lines and mass spectrometry analysis of purified protein complexes, we discovered that PRAME is a stable component of Cullin2-based E3 ubiquitin ligases (chapter 3). Ubiquitin ligases mediate the addition of the small protein ubiquitin to target proteins, altering their function and often leading to their degradation. Our results established a first strong functional link between PRAME and a defined biochemical pathway. Using the same epitope-tagging strategy we discovered interactions between Cullin2-PRAME ligases and subunits of the conserved EKC/KEOPS complex that we characterised in chapter 4. In particular, knock-down experiments revealed that PRAME functions as a bridge between Cullin2 ligases and the EKC complex.

In parallel to the proteomic techniques, we set up chromatin immuno-precipitation assays coupled to next-generation sequencing technologies (ChIP-seq). For this purpose, we raised specific PRAME antibodies in rabbits to investigate the binding profile of PRAME to the human genome (chapter 3 and 5). The data analysis showed that PRAME (chapter 2) as well as EKC sub-units (chapter 4) are specifically enriched at transcriptionally and epigeneti-cally active promoter regions that are also bound by the transcription factor NF-Y, and at enhancer elements. These results suggest that PRAME might play a role in the NF-Y signaling pathway.

Overall, the experimental work outlined in this thesis establishes a biochemical and molecular function for the PRAME gene family and provides a new twist to both PRAME and EKC biology. Our results provide an important contribution to understand the molecular mechanisms in which these pro-teins are active, both in healthy and tumour cells, and thus lay a foundation for novel therapeutic strategies against cancer.

SAMENVATTING

Het genoom in de kern van elke menselijke somatische cel bestaat uit ongeveer 2 meters DNA en is ingedeeld in 23 chromosoom paren die in totaal 6 miljard basenparen bevatten. Ongeveer 5% van het menselijke genoom bevat coderende genen die eiwitten kunnen produceren, wat gebeurt onder strikte regulatie. De eiwitten die aanwezig zijn in een cel (ook wel "proteoom" genoemd) zijn niet geïsoleerde eilanden, maar zijn betrokken bij dynamische netwerken en maken deel uit van multi-eiwit complexen. Ontregeling van genexpressie en eiwit interacties kan leiden tot verschillende ziektes, waaronder kanker.

Het onderwerp van dit proefschrift betreft een uitgebreide studie naar de moleculaire functie van het kanker-gerelateerde eiwit PRAME (preferentially expressed antigen of melanoma) met behulp van cellijn modelsystemen en een krachtige combinatie van proteomische en genomische technieken. Aan het begin van deze studie was bekend dat PRAME vaak actief aanwezig is in kankercellen en dat een hoge expressie van PRAME geassocieerd is met een slechte klinische prognose voor melanoom-, neuroblastoom- en borstkankerpatiënten.

PRAME vervult de rol van een veelbelovend therapeutisch doelwit aangezien expressie van dit eiwit voornamelijk beperkt is tot kankercellen. Echter, het ontwikkelen van strategieën om de activiteit van PRAME te beïnvloeden wordt gehinderd door een gebrek aan informatie over de moleculaire functie(s) van PRAME. Verder heeft de analyse van het menselijk genoom een grote familie van meer dan twintig PRAME-achtige genen aan het licht gebracht, waarvan de functies nog onduidelijk zijn.

Om de moleculaire functies van PRAME te onderzoeken, hebben wij innovatieve high-throughput biochemische technieken ontwikkeld om het eiwit-eiwit interactienetwerk van PRAME te ontrafelen (hoofdstuk 2). Door toepassing van epitoop-tagging in menselijke cellijnen en massaspectrometrie van gezuiverde eiwitcomplexen, hebben wij ontdekt dat PRAME een stabiele component is van Cullin2 E3 ubiquitine ligases (hoofdstuk 3). Ubiquitine ligases koppelen het kleine eiwit ubiquitine aan doeleiwitten en veranderen

zodoende hun functie of markeren ze om afgebroken te worden. Onze resultaten hebben voor het eerst een sterke relatie gelegd tussen PRAME en een specifieke biochemische activiteit. Met dezelfde epitoop-tagging strategie hebben wij laten zien dat Cullin2-PRAME ligases een interactie aangaan met het EKC/KEOPS complex (hoofdstuk 4). In het bijzonder hebben knock-down experimenten aangetoond dat PRAME fungeert als een brug tussen Cullin2 ligases en het EKC complex.

Naast de proteomische technieken hebben wij een techniek ontwikkeld waarbij chromatine-immunoprecipitaties gevolgd worden door next-generation sequencing (ChIP-seq). Wij hebben specifieke PRAME antilichamen in konijnen opgewekt om het bindingsprofiel van PRAME op het menselijke genoom te onderzoeken (hoofdstuk 3 en 5). De data analyse liet zien dat PRAME (hoofdstuk 2) en het EKC complex (hoofdstuk 4) specifiek verrijkt zijn

op transcriptionele en epigenetisch actieve promotors van genen die ook gebonden zijn door de transcriptiefactor NF-Y, als ook op enhancer elementen. Deze resultaten suggereren dat PRAME een rol zou kunnen spelen bij de NF-Y signaleringsroute.

Het experimentele werk beschreven in dit proefschrift heeft een nieuwe biochemische en moleculaire functie voor de PRAME gen-familie vastgesteld en geeft een vernieuwd inzicht aan zowel PRAME en EKC biologie. Onze resultaten leveren een belangrijke bijdrage aan het begrijpen van de moleculaire mechanismen waarin deze eiwitten actief zijn, zowel in gezonde als in kankercellen. Hiermee wordt een basis gelegd voor het ontwikkelen van nieuwe therapeutische strategieën tegen kanker.

RIASSUNTO

Il genoma presente nel nucleo di ogni cellula somatica umana è costituito da circa 2 metri di DNA organizzato in 23 coppie di cromosomi che contengono un totale di 6 miliardi di paia di basi. Circa il 5% del genoma umano costituisce geni codificanti proteine che possono essere attivati sotto stretta regolazione. Le proteine presenti in una cellula (nel loro insieme chiamate anche "proteoma") non sono isole separate ed indipendenti, ma piuttosto interagiscono tra di loro in reti dinamiche e fanno parte di complessi multiproteici. In particolare, la disregolazione dell'espressione genica e delle interazione tra proteine può portare a malattie e cancro.

L'obiettivo di questa tesi di dottorato è di rivelare la funzione molecolare della oncoproteina umana PRAME (preferentially expressed antigen melanoma, antigene preferenzialmente espresso in melanoma) utilizzando linee cellulari come sistema modello ed una potente combinazione di tecniche di proteomica e genomica. All'inizio di questo studio, PRAME era noto per essere frequentemente attivo in tumori e livelli elevati di espressione erano risultati correlati con prognosi negative in pazienti di melanoma, neuroblastoma e cancro al seno. Inoltre, il fatto che PRAME è espresso principalmente in cellule tumorali e non in cellule sane ha indicato PRAME come un obiettivo terapeutico promettente. Tuttavia, gli sforzi per sviluppare strategie terapeutiche per interferire con l'attività oncogenica di PRAME sono ostacolati dalla mancanza di informazione sulle sue funzioni molecolari. Inoltre, analisi del genoma umano ha rivelato l'esistenza di una grande famiglia di più di venti geni simili a PRAME (PRAME-like genes), le cui funzioni non sono ancora chiare.

Al fine di identificare le funzioni di PRAME a livello molecolare, abbiamo sviluppato approcci biochimici innovativi e high-throughput per identificare la rete di interazioni proteina-proteina di PRAME (capitolo 2). Applicando un approccio di "epitope-tagging" in linee cellulari umane e analisi di spettrometria di massa dei complessi proteici purificati, abbiamo scoperto che PRAME è un componente stabile di E3 ubiquitin ligasi basati sulla proteina Cullin2 (capitolo 3). Le ubiquitin ligasi mediano l'aggiunta della piccola proteina ubiquitina a proteine bersaglio, alterando la loro funzione o spesso causandone la degradazione. I nostri risultati hanno quindi stabilito per la prima volta un

forte relazione funzionale tra PRAME ed un'attività biochimica ben definita. Utilizzando le stesse tecniche abbiamo scoperto interazioni tra Cullin2-PRAME ligasi e subunità del complesso EKC/KEOPS, che abbiamo caratterizzato nel capitolo 4. In particolare, esperimenti di knock-down hanno rivelato che PRA-ME funziona come un ponte tra le Cullin2 ubiquitin ligasi ed il complesso EKC.

In parallelo alle tecniche di proteomica, abbiamo sviluppato saggi di im-munoprecipitazione della cromatina seguiti da sequenziamento high-throu-ghput di nuova generazione (ChIP-seq). A questo scopo, abbiamo generato anticorpi specifici contro PRAME in conigli allo scopo di indagare il profilo di legame di PRAME al genoma umano (capitolo 3 e 5). L'analisi dei dati ha mostrato chiaramente che PRAME (capitolo 2) e subunità di EKC (capitolo 4) legano specificamente elementi enhancer e promotori di geni attivi che sono anche legati dal fattore di trascrizione NF-Y. Questi risultati suggeriscono che

PRAME potrebbe giocare un ruolo nella via di segnale di NF-Y.

Nel complesso, il lavoro sperimentale descritto in questa tesi stabilisce una funzione biochimica e molecolare per PRAME e la vasta famiglia di geni PRAME-simili e fornisce una nuova svolta nella biologia sia di PRAME che del complesso EKC. I nostri risultati forniscono un importante contributo alla comprensione dei meccanismi molecolari in cui queste proteine sono attive, sia in cellule sane che tumorali, e offrono così possibilità di sviluppare nuove strategie terapeutiche.

CURRICULUM VITAE

Adalberto Costessi was born on April 23rd in Trieste, Italy. He developed a keen interest for scientific subjects and philosophy at the secondary school Liceo Scientifico Galileo Galilei in Trieste, where he graduated cum laude.

In 1998 he enrolled at the Medical Biotechnology curriculum at the University of Trieste in order to study the molecular mechanisms of biological life and disease. Besides the regular studies, he fulfilled the role of vice-president (2002) and president (2003) of the CNSB, the Italian association of biotechnology students. In these roles, Adalberto promoted a national project to evaluate the quality of the university education in biotechnology, coordinated several projects of public communication of biotechnology, and actively participated in the discussions about GMO and patents regulations. In those years, a peculiar interest for space-related research and weightlessness brought Adalberto in close contact with the European Space Agency (ESA). In 2000 his project proposal on the study of fruit flies in weightlessness was selected for participation to the 3rd Student Parabolic Flight Campaign, awarding him the opportunity to perform the experiment during the weightlessness conditions of a parabolic flight.

In 2002 Adalberto joined the Molecular Biology lab of Gianluca Tell at the University of Trieste where he conducted experimental research on the molecular responses of human osteoblasts to extracellular nucleotides and mechanical stress, leading to two peer-reviewed publications and a graduation cum laude in 2003.

In 2002 Adalberto's research proposal "BOP - Bone proteomics" to study the molecular responses of human osteoblasts in microgravity was awarded the first prize at the International "Success 2002 Student contest" by the European Space Agency. This brought him to The Netherlands in 2004, at the ESA research centre in Noordwijk. There he contributed to the development of an innovative cell-culture system and trained the astronauts that carried out the BOP experiment during a space mission to the International Space Station in April 2005.

In October of the same year, Adalberto decided to start a PhD research project with the feet on the ground under the supervision of prof. Henk Stunnenberg at the Molecular Biology Department of the Radboud University Nijmegen, The Netherlands. During five years of experimental research on the human oncoprotein PRAME he applied and developed biochemical

techniques to unravel the protein-protein interaction network of PRAME as well as high-throughput sequencing approaches to identify the binding sites of PRAME on the human genome.

Since January 2011, Adalberto works as Product Specialist Genome Analysis at BaseClear (Leiden, The Netherlands), an independent accredited laboratory specialized in DNA analysis services. He leads the Next-Generation Sequencing laboratory and is responsible for the application of the newest related technologies.

PUBLICATIONS

Maczuga P, Lubelski J, van Logtenstein R, Borel F, Blits B, Fakkert E, **Costessi A**, Butler D, van Deventer S, Petry H, Koornneef A, Konstantinova P.
Embedding siRNA sequences targeting Apolipoprotein B100 in shRNA and miRNA scaffolds results in differential processing and in vivo efficacy.
Molecular Therapy. 2012 Oct 23.

Costessi A, Mahrour N, Sharma V, Stunnenberg R, Stoel MA, Tijchon E, Conaway JW, Conaway RC, Stunnenberg HG.
The human EKC/KEOPS complex is recruited to Cullin2 ubiquitin ligases by the human tumour antigen PRAME.
PLoS One, 2012;7(8):e42822.

Costessi A, Mahrour N, Tijchon E, Stunnenberg R, Stoel MA, Jansen PW, Sela D, Martin-Brown S, Washburn MP, Florens L, Conaway JW, Conaway RC, Stunnenberg HG.
The tumour antigen PRAME is a component of Cullin2 ubiquitin ligases and is recruited to active NFY promoters.
EMBO Journal. 2011 Aug 5;30(18):3786-98.

187

Schonenborg R, **Costessi A**, Schiller P.
Hardware development for the bone proteomics experiment.
Microgravity Science and Technology, 19 (5-6), pp. 219-224 (2007)

Costessi A, Pines A, D'Andrea P, Romanello M, Damante G, Cesaratto L, Quadrifoglio F, Moro L, Tell G.
Extracellular nucleotides activate Runx2 in the osteoblast-like HOBIT cell line: A possible molecular link between mechanical stress and osteoblasts' response.
Bone. 2005 Mar;36(3):418-32.

Risso A, Tell G, Vascotto C, **Costessi A**, Arena S, Scaloni A, Cosulich ME.
Activation of human T lymphocytes under conditions similar to those that occur during exposure to microgravity: A proteomics study.
Proteomics. 2005 May;5(7):1827-37.

ACKNOWLEDGEMENTS

Here we go, the last page of the book that officially closes a 5-year-long intense adventure. An important chapter of my life both personally as well as professionally from many points of view: the wonderful people I met, the places I lived in, the challenges and opportunities, the disappointing experiments as well as the exciting results and, of course, this beautiful book. Starting up pretty much from scratch and successfully rounding up such a challenging PhD project in a top research lab is one of those achievements that prompt you to stop for a second, look back, and think at how this all happened... Personally, the more I think about it, the more I appreciate that this experience would not have been possible without a great blend of wise and helpful people around me that really deserve a big "Thank you!". Truly.

My first thoughts and gratitude go to the wonderful and dedicated teachers and professors that supported me, pushed me, challenged me, and inspired me throughout my education. This thesis is dedicated to all of you! A special thought goes to the late Gabriella Greblo, my maths and physics teacher in high school that left us far too soon. You didn't miss a chance to feed and encourage my scientific curiosity and, moreover, you have been an important example of integrity and dedication that I will never forget. I know you would have been proud of me and this book.

Among my university teachers I would like to especially mention Gianluca Tell for the great time I spent in your lab in Trieste and later in Udine during my third year internship and my master thesis. In your lab I made my first moves in the world of experimental molecular biology and you provided me with lots of opportunities and support to test my ideas. Moreover, your door was always open for scientific discussions, even with my space-related projects for the European Space Agency, when my feet where literally not so much on the ground!

And then a big hug to Alex and Michela, of course! The friendship with Alex started in Gianluca's lab, when you were a PhD student and my direct supervisor (the good old times!) And a couple of years later you and Michela decided to find a job in The Netherlands and follow me to the north. Are you guys ever gonna stop blaming me for the Dutch weather? Ok, ok, one more grappa and we all forget the rain...

And then from Trieste, via Noordwijk, I got to Nijmegen. Henk, I vividly remember our first meeting at an open day of the NCMLS when your enthu-

188

siasm and interest in the basic molecular mechanisms made me put aside the PhD positions I already had in my pocket. In the end, the PRAME project turned out to be just a little bit more challenging than the initial plan "tag the protein, push it in the cells, pull out the interactors, function is solved, publish in Nature, done"… However, this project gave me the opportunity to learn and setup so many different techniques that I really enjoyed jumping from the 40-liters scale cell cultures to refined biochemistry, from DNA cloning to next-gen sequencing and bioinformatical analyses. Henk, on day one you gave me straight away such a large degree of freedom that I could explore many directions, even though at times it felt like swimming in the ocean without the shore in sight. That freedom meant for me also a big responsibility for the project, from which I learnt for sure a lot. You gave me great opportunities to work with your latest high-tech "toys for big boys" and I dare to state that I eventually reached a pretty solid ground and I am proud of this work. Your focus for the big picture is something I have always appreciated and have taken with me. Thanks also for the inputs at some critical points in the project, like that Blast search you did one night, opening the EKC side of the story, and your enthusiasm to hear my latest results at any time of day or night, whatever the time zone. All in all, thank you for the opportunities you provided and your contribution!

I would also like to thank the members of the Conaway lab at Stowers and in particular Nawel, Joan and Ron for the fruitful collaboration in the last years of my PhD project. It was nice that we could join forces and I enjoyed very much meeting you in person as well. I hope I will have the chance to visit you at Stowers one day! I wish you all the best with your future research!

In a kind of chronological order I should actually start by acknowledging Willem Welboren. Most people don't actually know that Willem played a key role in my choice for the PRAME project. At the time of my PhD application to Henk, I had a preference for the estrogen receptor project. However, Willem had also chosen the same project and since he started earlier than me in the lab, he could get that one, and so I eventually started working on PRAME. Willem, I am so glad you took on that project, as I could not have handled five years of ChIPing and ChIPing and ChIPing as well as you did! Thank you also a lot for being my Bioinformatics Master and introducing me to the bioinformatic sides of my MacBook!

Xavier and Michiel: you made the first steps in the PRAME project before I joined the lab, though at some point you dumped PRAME in favour of other proteins that happened to be easier to study, and left this interesting protein for me... Well, given the amount of work that was needed to get a grasp on PRAME you made a very wise decision at the time! But thanks for the advices at the start of the project and the first constructs and cell lines you made. I would like to acknowledge also prof. Rene Bernards (NKI) and Mirjam Epping, with whom the PRAME project was setup.

Esther, Marieke en Rieka: the PRAME girls, my three undergraduate students. All of you gave a very significant contribution to the work presented in this book and it was a real pleasure to work together with you. I hope you are still enjoying the cooking books! I wish you all the best for your careers!

Then my U-buddies, Arjen and Hendrik. Arjen, it has been a pleasure to share the U-bench with you in my first years of PhD, and the laughs with your student Jorrit. I appreciated your eye for detail and your strive for writing beautiful and efficient analysis scripts. Thanks for always taking the time to answer my questions, engage in technical discussions and help with my own analysis code. Hendrik, thank you too for the exchange and help with analyses, code and shell scripts and for trying to understand my complicated biochemistry. I miss your humour!

Thanks to the members of the lunch group in its varying composition over the years: Pascal, Evelien, Blanca, Rike, Ozren, Melysia, Vikram, Sarita... It was so nice to have a daily relaxing moment together in between experiments, when we could share expectations, hopes, frustrations, exciting results as well as our life outside the lab. I miss our fights with the personnel at the canteen, and I am sure they miss us too!

Blanca, you are one of the most knowledgeable people I have met, you always had an answer to my questions and I learnt so much from you. It has been a real pleasure to work with you, thank you so much for the countless advices at important steps in my project. I am glad we still manage to meet up regularly together with the other girls. Melysia, you have been a great neighbour in the lab, and how many laughs did we have together. Shall we open that KFC one day? Evelien, we worked on very similar project and it was always nice to discuss results and help each other through the technical challenges. Rike, thanks for our long chats especially in my last months in the lab, best wishes for your busy career! Hvala lijepa Ozren, my Adriatic buddy, for the nice time in the lab! Adrian, Pascal, Vikram, thank you so much for

your support with the mass spectrometry experiments along my whole PhD.

Adri, que tal? Muchas gracias for the late evening chats and our attempts to keep up our French together and then regularly ending up talking Spanish or Italian. Thanks also for your important help with raising the antibodies. I hope to visit you again soon.

Maria, Josephine, thanks for making the lab a nice and efficient work place and for your kindness when practicing my Dutch with you. Eva, thanks for organizing so well the Illumina sequencing, in my present work I got to realize how important that is. Anita, thank you for all the cloning work you performed for my project, your contribution definitely helped me a lot and it was a pleasure to work with you!

A big big thank goes to everybody in the Mol Bio lab, and I apologize if my memory is too short to acknowledge everyone. All of you contributed to my work either with deep technical discussion, with your experience and tips, or with the nice and collegial atmosphere. Waseem with the sharp biochemistry discussions, Joost with important suggestions on my first paper and tips on how to best present sequencing data, Blaise for your PERL scripting abilities, Mena for the help in the lab and for introducing me to the real pastiera napoletana, Dik and Jo, Richard, Wieteke, Femke, Stefanie, Robert, Max, Simon, Kees-Jan, Joke, Fiona, Colin, Gert-Jan, Marion.... thanks for everything. I would like to acknowledge prof. Nakatani for challenging my biochemical purification abilities (and losing the bet). Silvia and Federica, thanks for the nice times together up in the tower, I wish you all the best and hope to see you soon again.

When I first arrived in Nijmegen, I had the chance to meet two wonderful people: Cathy and Martin. I had a wonderful time living in your house and I cannot wait to hear the stories and experiences of your next travel and have some good laughs together.

The sport and dance lessons at Indoor Action in Arnhem proved to be one of the best anti-stress ever during my PhD years. Marije and Nanda, thank you for all the fun! I didn't imagine I could get so addicted to your lessons, and I miss you so much.

I would like to thank my colleagues at BaseClear for the nice working environment and for the support in the last phases of my PhD. Especially the Next-Gen lab team: Erwin, Diana, Marten, Walter, Derek, but also Danny, Evelien, Schelto, Marthje, Bas and Erna and all the other colleagues.

René and André: we always have so much fun together and you don't

miss a chance to teach me the weirdest Dutch words and expressions. I love you guys! Wij gaan snel weer de bloemetjes buiten zetten. Maarten, thanks for beautifying my research with this book.

Finally, I want to thank my family. Sono estremamente grato ai miei genitori e a mio fratello per il supporto e amore incondizionato in tutti questi anni. Questo percorso e questo libro non sarebbero stati possibili senza i vostri sacrifici e senza l'appoggio e la motivazione che mi avete dimostrato in ogni occasione. Dirk, my love, thank you for your unconditional love and support. I am so happy to have you at my side and to share the best moments in life with you.

Adalberto

The human protein PRAME (preferentially expressed antigen of melanoma) is frequently overexpressed in tumours. Importantly, high expression levels are associated to poor clinical outcome in melanoma, neuroblastoma, and breast cancer. However, its function and molecular mechanisms in normal and cancer cells are not yet well understood and the cellular pathways affected by PRAME are still elusive. Although PRAME has been indicated as a promising potential target for cancer therapy and anti-tumour immunization protocols, efforts to develop strategies to interfere with the oncogenic activity of PRAME are hindered by a lack of information about its function.

The aim of this thesis was to investigate the molecular function of human PRAME using cell line model systems and a powerful combination of proteomic and genomic techniques. We developed innovative biochemical approaches to dissect the protein-protein interaction network of PRAME, and we setup chromatin immunoprecipitation assays coupled to next-generation sequencing technologies to reveal binding profile of PRAME to the human genome in cancer cells.