Biochemical and Biophysical Research Communications 430 (2013) 547–553

Contents lists available at SciVerse ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Tudor domains of the PRC2 components PHF1 and PHF19 selectively bind to histone H3K36me3

Su Qin [a], Yahong Guo [a], Chao Xu [a], Chuanbing Bian [a], Minfei Fu [a], Sarah Gong [a], Jinrong Min [a,b,*]

[a] Structural Genomics Consortium, University of Toronto, 101 College St., Toronto, Ontario M5G 1L7, Canada
[b] Department of Physiology, University of Toronto, Toronto, Ontario M5S 1A8, Canada

ARTICLE INFO

Article history:
Received 26 November 2012
Available online 7 December 2012

Keywords:
Tudor
H3K36me3
PHF1
MTF2
PHF19
PRC2

ABSTRACT

PRC2 is the major H3K27 methyltransferase and is responsible for maintaining repressed gene expression patterns throughout development. It contains four core components: EZH2, EED, SUZ12 and RbAp46/48 and some cell-type specific components. In this study, we focused on characterizing the histone binding domains of PHF1 and PHF19, and found that the Tudor domains of PHF1 and PHF19 selectively bind to histone H3K36me3. Structural analysis of these Tudor domains also shed light on how these Tudor domains selectively bind to histone H3K36me3. The histone H3K36me3 binding by the Tudor domains of PHF1, PHF19 and likely MTF2 provide another recruitment and regulatory mechanism for the PRC2 complex. In addition, we found that the first PHD domains of PHF1 and PHF19 do not exhibit histone H3K4 binding ability, nor do they affect the Tudor domain binding to histones.

1. Introduction

The polycomb repressive complex 2 (PRC2) is the major H3K27 methyltransferase and is responsible for maintaining repressed gene expression patterns throughout development. The PRC2 complex consists of four core components: EZH2, EED, SUZ12 and RbAp46/48 [1–3]. The function of the *Drosophila EZH2* homologue E(z) has been long implicated in the repression of segment identity genes during development [4–10], and X-inactivation, germline development, stem cell pluripotency and differentiation, and cancer metastasis later on [11]. Sequence analysis reveals that EZH2 contain an evolutionarily conserved sequence motif of 130 amino acids, SET domain (named after Suv39, E(z) and Trithorax), which was later shown to be the catalytic domain of histone methylation [12] and attracts enormous attention since then [13]. EZH2 has virtually no histone methylation activity on its own, but exhibits robust methylation activity as a complex [1,2]. EED is a WD40 repeat domain protein, which was shown to directly interact with EZH2 both *in vitro* and *in vivo* in 1998 [14,15]. The N-terminal domain of EED is also essential for histone H3K27 trimethylation of H3K27 because an N-terminally truncated EED can still form a complex with EZH2, but this complex cannot carry out trimethylation of histone H3 any more [16]. In addition to binding to EZH2, we and others have also displayed that EED is able to bind to histone H3K27me3, suggesting a mechanism for PRC2 to propagate and spread the H3K27me3 mark to daughter strands during cell division [17–19].

SUZ12 is another essential component of the PRC2 complex, which is required for active methyltransferase activity, Hox gene silencing, and embryonic stem cell differentiation [20,21]. SUZ12 directly binds to the promoters of the PRC2 target genes probably through its Zinc-finger domains, and Suz12 expression is increased in human colon tumors [22]. In the mammalian PRC2 complex, there is a pair of highly homologous WD40 proteins RbAP46/48, which is shown to be essential for cell survival and patterning in *Drosophila* development, and knockdown of the Drosophila RbAP46/48 P55 causes severe reduction in histone H3K27me3 methylation level [23]. Although AEBP2 is not essential for PRC2 activity, it is required for optimal enzymatic activity [21]. AEBP2 may regulate the development of the neural crest cells through the PRC2-mediated epigenetic mechanism [24]. Very recently, AEBP2 was suggested to play an allosteric role in regulating PRC2 activity and gene silencing (eLife online).

In the human genome there are three Polycomblike (PCL) genes: PHF1 (PCL1), MTF2 (PCL2) and PHF19 (PCL3). Polycomblike (PCL) gene was first identified over 30 years ago and it was show to be required for the maintenance of normal identities in many of the body segments during *Drosophila* development [25], and it was named as Polycomblike due to its similar phenotypes to another developmental control gene Polycomb in the mutant clonal analysis experiments [25]. Polycomb is a component of PRC1 and selectively recognizes the histone H3K27me3 [26,27]. Domain analysis of PHF1, MTF2 and PHF19 unveil that they also contain some potential histone binding domains, such as Tudor and PHD domains

* Corresponding author at: Structural Genomics Consortium, University of Toronto, 101 College St., Toronto, Ontario M5G 1L7, Canada. Fax: +1 416 9460588.
E-mail address: jr.min@utoronto.ca (J. Min).

0006-291X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bbrc.2012.11.116

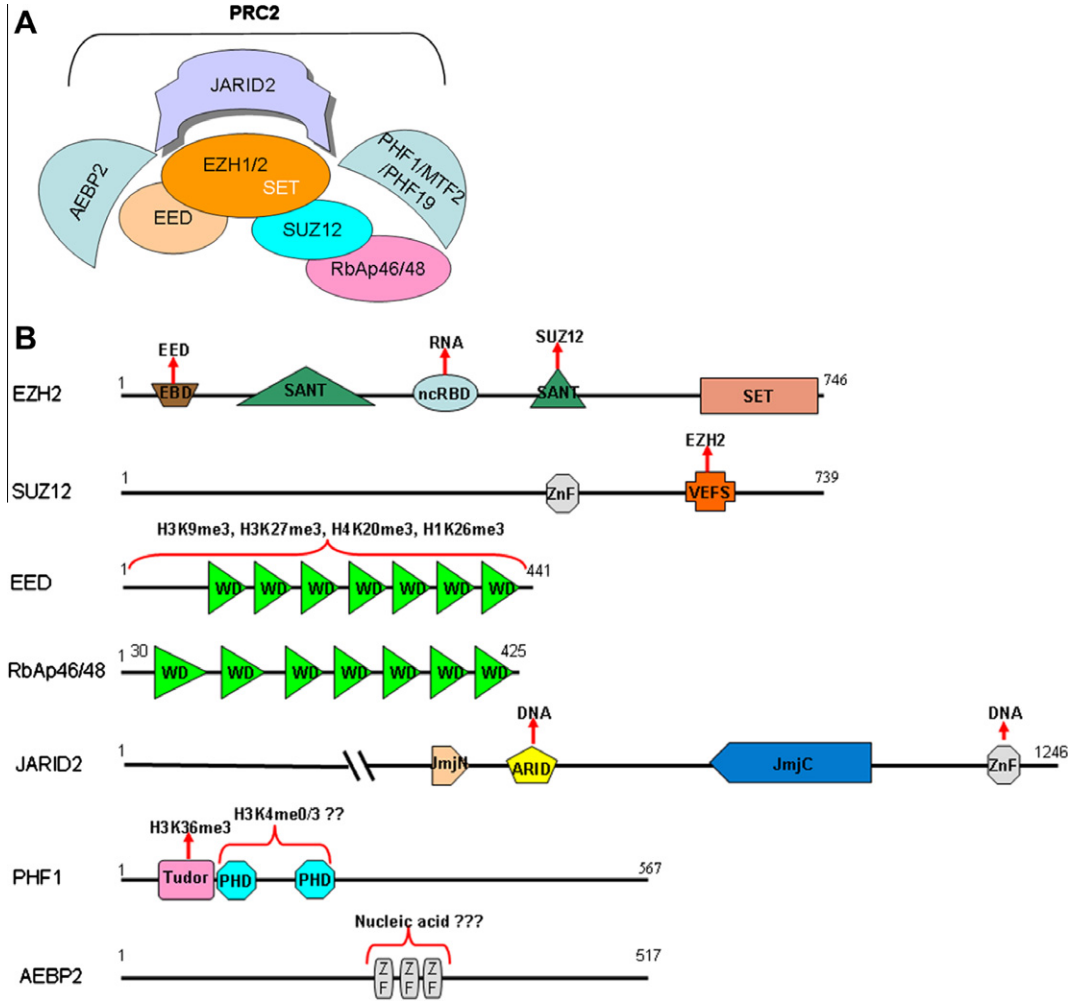

Fig. 1. Composition of Human PRC2 complex. (A) Architecture of the PRC2 complex. (B) Characterized domains with potential functions are indicated for each PRC2 component. EBD, EED binding domain; ncRBD, non-coding-RNA-binding domain; SANT, SWI3, ADA2, N-CoR and TFIIIB DNAbinding domain; SET, Su(var)3–9, enhancer of zeste, trithorax domain; VEFS, conserved among VRN2–EMF2–FIS2–SU(Z)12; WD, short ~40 amino acid motifs; ZF, zinc finger; ARID, AT-rich interacting domain; Tudor, Tudor domain; PHD, Plant Homeo Domain.

(Fig. 1) [28]. PCL was identified as a PRC2 component in a 1 Megadalton complex in *Drosophila* nuclear extracts [29], and both *Drosophila* PCL and mammalian PHF1 were shown to be required for efficient H3-K27 trimethylation and Hox gene silencing [30–32]. Interestingly, PHF1 is also implicated in the genome maintenance processes because it was recruited to DSBs (**D**ouble **S**trand **B**reaks) immediately after the irradiation [33]. MTF2 is present in the PRC2 complex in embryonic stem cells, and regulates the transcriptional networks during mouse embryonic stem cell self-renewal and differentiation [34]. MTF2 recruits the PRC2 complex to the inactive X chromosome and target loci in embryonic stem cells [35]. Surprisingly, MTF2 could also activate the Cdkn2a gene and promote cellular senescence, implicating that MTF2 could also suppress the catalytic activity of PRC2 locally [36]. PHF19 is also shown to interact with PRC2 and recruits it to CpG islands and contributes to embryonic stem cell self-renewal [37]. Taken together, the common function of PHF1, MTF2 and PHF19 is recruitment of the PRC2 complex to its target genes for histone methylation, Hox gene silencing and ESC (**E**xtra **S**ex **C**ombs) self-renewal. In addition to PHF1, MTF2 and PHF19, Jarid2 was recently identified as the embryonic stem cell specific PRC2 regulatory subunit, which is required for embryonic stem cell differentiation (Fig. 1) [38].

Although it is implicated that PHF1, MTF2 and PHF19 play a role in the PRC2 recruitment, the molecular mechanism of targeting is still unclear. Considering that each of these 3 proteins contains a Tudor domain and two PHD domains, we hypothesize that these potential histone binding domains may recruit PRC2 to its target genes by binding the histones. In the present study, we are going to focus on the Tudor domain and the first PHD domain by characterizing their histone binding ability and structural features.

2. Results and discussion

2.1. The Tudor domains of PHF1, MTF2 and PHF19 selectively bind to histone H3K36me3

In the human genome, there are at least 36 Tudor domain containing proteins. This number may still grow as more Tudor domains are identified and confirmed by means of structural and sequence analysis, such as SGF29 [39]. The difficulty in identifying Tudor domains is due to their low sequence conservation, although their structure folds are highly conserved. Based on the binding ligands, the Tudor domains can be classified as two groups: the methyl-arginine binding Tudor and the methyl-lysine binding Tudor domains. Both methyl-arginine binding Tudor and methyl-lysine binding Tudor domains may function as a single Tudor or Tandem Tudors. For instance, TDRD3 is a single Tudor protein which preferentially recognizes the asymmetrical dimethylated