Biochimica et Biophysica Acta 1829 (2013) 1126–1135

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Review

Peptidylarginine deiminases in citrullination, gene regulation, health and pathogenesis

Shu Wang [a,b,c], Yanming Wang [a,b,c,*]

[a] Graduate Program in Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA, 16802 USA
[b] Center for Eukaryotic Gene Regulation, Pennsylvania State University, University Park, PA, 16802 USA
[c] Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA, 16802 USA

ARTICLE INFO

Article history:
Received 17 May 2013
Received in revised form 3 July 2013
Accepted 8 July 2013
Available online 13 July 2013

Keywords:
Peptidylarginine deiminase
Histone
Gene regulation
Cancer
Autoimmunity

ABSTRACT

Peptidylarginine deiminases are a family of enzymes that mediate post-translational modifications of protein arginine residues by deimination or demethylimination to produce citrulline. *In vitro*, the activity of PADs is dependent on calcium and reductive reagents carrying a free sulfhydryl group. The discovery that PAD4 can target both arginine and methyl-arginine for citrullination about 10 years ago renewed our interest in studying this family of enzymes in gene regulation and their physiological functions. The deregulation of PADs is involved in the etiology of multiple human diseases, including cancers and autoimmune disorders. There is a growing effort to develop isoform specific PAD inhibitors for disease treatment. However, the regulation of the activity of PADs *in vivo* remains largely elusive, and we expect that much will be learned about the role of these enzymes in a normal life cycle and under pathology conditions.

1. Introduction

1.1. The early discovery of PADs as protein citrullination enzymes

Citrulline residues were detected in early 1960s from polypeptide hydrolysates of hair inner root sheath cells and medullary cells [1]. Since there is no citrulline tRNA *in vivo*, citrulline can only be produced by enzymatic modifications after protein synthesis. Subsequently, the peptidylarginine deiminases (PADs, also called PADIs) were identified, which convert protein arginine residues to citrulline in a calcium dependent manner ([2], reviewed in [3]). PADs are involved normal functions in the immune and reproduction systems as well as in the skin (see recent review in these particular areas in [4–6]). In each of the mammalian vertebrate genomes, five highly conserved PADs exist, including PAD1, 2, 3, 4 and 6. (Fig. 1A). In the last decade or so, more research on the gene features, tissue-specific distribution and preferred substrates of PADs has been performed (Table 1). Human PAD4 can bind five calcium ions (Fig. 1B). Several flexible parts of PAD4 form stable secondary structures after the binding of calcium and substrate, indicating that calcium stabilizes the conformation of PAD4 and may facilitate the formation of the active site cleft (Fig. 1C and D). The calcium binding sites in PAD4 are conserved in PAD1, -2, and -3 except for PAD6. Thus, calcium is an important regulator of the active PAD enzymes [7].

1.2. The five PADs

1.2.1. PAD1

PAD1 is mainly expressed in epidermis and uterus [8,9]. PAD1 deiminates keratin K1 and is involved in the cornification of epidermal tissues [10].

1.2.2. PAD2

PAD2 is widely expressed in multiple tissues, including secretory glands, brain, uterus, spleen, pancreas, skeletal muscle [11–16]. The expression of PAD2 can be regulated at both mRNA splicing and protein translation levels [15,16]. Myelin basic protein of the central nervous system and vimentin in skeletal muscle and macrophages are long known substrates of PAD2 [16,17]. Recently, β and γ-actins were identified as PAD2 substrates in human neutrophils [18]. PAD2 is mainly a cytoplasmic protein, but a fraction of PAD2 may become nuclear in canine and human mammary epithelial cells [19,20]. Nuclear PAD2 may citrullinate histones H3 and H4 [19,21–23], suggesting a role of this protein in gene regulation.

1.2.3. PAD3

PAD3 is localized to epidermis and hair follicles [24–26]. PAD3 is colocalized with trichohyalin, a structural protein in the inner root sheath and medulla of hair follicles [25,26]. In addition, PAD3 colocalizes with profilaggrin and filaggrin in the epidermis [26]. PAD3 targets filaggrin, which interacts with keratin intermediate filament to regulate epidermal homeostasis in the granular layer and lower stratum corneum of human epidermis [2,25,26]. Deimination of filaggrin and

* Corresponding author at: Graduate Program in Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA, 16802 USA. Tel.: +1 814 865 3775.
E-mail address: yuw12@psu.edu (Y. Wang).

1874-9399/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbagrm.2013.07.003

trichohyalin *in vitro* by recombinant PAD3 further supports that PAD3 is involved in regulating epidermis functions [25,26].

1.2.4. PAD4

PAD4 (also called PADV and PADI4) is detected mainly in white blood cells including granulocytes and monocytes under normal physiological conditions [16,27,28]. However, in a wide range of tumors of various tissue origins, the overexpression of PAD4 was detected, suggesting that PAD4 plays a role in tumorigenesis [29,30]. PAD4 is localized primarily in the nucleus and contains a nuclear localization signal sequence at its N-terminus. PAD4 citrullinates a range of nuclear proteins, such as histones H2A, H3 and H4, ING4, p300/CBP, nucleophosmin and nuclear lamin C, thereby playing an important role in nuclear functions [31–36].

1.2.5. PAD6

PAD6 was originally identified from mouse eggs and embryos and was named ePAD (egg PAD) [37]. PAD6 regulates oocyte cytoskeletal sheet formation and female fertility [38]. Recently, it is found that PAD6 localizes to the cytoplasmic lattices and regulates the function of microtubules during early embryo development [39]. In human tissues, PAD6 is mainly restricted to the ovary, testis and peripheral blood leucocytes [40]. Interestingly, unlike the other PADs, PAD6 has lost some of the conserved Ca^{2+} binding residues and the active center cysteine residue is also different from other PADs [41], suggesting that PAD6 is likely not an active deiminase.

Recent studies have revealed the role of PADs in physiological and pathological conditions. In the following sessions, we will focus on the role of PADs in gene regulation, innate immunity, cancers and autoimmune diseases. We will further discuss the potential of PADs as druggable targets for disease treatment.

2. PAD4 and PAD2 in gene regulation

2.1. PAD4 in gene regulation

In eukaryotic cells, nuclear DNA is organized with two of each histones H3, H4, H2B, and H2A to form a nucleosome core particle, which is further organized with the linker DNA and histone H1 to form the 10 nm chromatin fiber and folded to form the higher order chromatin structures. The nuclear structure and 3D organization is under continuous remodeling to adapt to the physiological and environmental changes that the cells are exposed to. Because of this structural organization, histone modifications, including methylation, acetylation, phosphorylation and citrullination, work as a signaling network to provide the on the off signal for gene expression and/or a landing platform for effector protein binding [42–44].

Histone modifying enzymes with opposite activities counteract each other's effect, such as histone acetyltransferases (HATs) and histone deacetylases (HDACs), kinases and phosphatases [45,46]. Until the Lys demethylases LSD1 and JmjC domain-containing dioxygenase, as well as the Arg deiminase PAD4 were identified [47,48], histone methylation on Arg and Lys residues was considered as static rather than dynamic because of the low turnover rate of the methyl groups [49]. PAD4 antagonizes CARM1 (also called PRMT4) and PRMT1 mediated histone H3 and H4 Arg methylation through a reaction dubbed as demethylimination in reflecting the removal of the methyl-imine group from monomethyl-arginine residues [33]. The activity of PAD4 on asymmetrical dimethyl-arginine residues is very low. We have observed that PAD4 prefers methyl-Arg in histone proteins over methyl-Arg in short peptides as substrates, suggesting the substrate has an allosteric effect on PAD4 [33]. CARM1 and PRMT1 function as transcription coactivators by catalyzing histone Arg monomethylation and asymmetrical dimethylation [42,50–53]. By antagonizing Arg methylation, PAD4 functions as a transcription corepressor. In the case of ER target genes in the breast cancer MCF-7 cells, PAD4 regulates histone Arg methylation via its citrullination activity on the gene promoters [33,54]. Interestingly, the modification of histone Arg residues on the ER (estrogen receptor) target promoters fluctuates over time after estradiol treatment, whereby the increase in histone citrullination correlates with the decrease in histone Arg methylation, indicating that opposite enzymes are alternatively working on the ER target gene promoters [55].

In addition to the ER target genes, our group has found that PAD4 interacts with the tumor suppressor and transcription factor p53 and functions as a corepressor to regulate the expression of multiple p53 target genes [56,57]. Before DNA damage, a high level of histone citrullination and PAD4 was detected on the promoter of p53 target genes, such as p21/CIP1/WAF1, GADD45 and PUMA [56,57]. After DNA damage, PAD4 association and histone citrullination decreases on these gene promoters with a concomitant increase in histone Arg methylation, suggesting that citrullination and arginine methylation counteract each other's function to regulate gene expression.

The role of histone acetylation in gene activation has been long established. A landmark 2004 paper with rich *in vitro* biochemical analyses showed that protein Arg methyltransferases and histone acetyltransferase p300/CBP cooperatively activate the p53-mediated transcription [42]. Reversely, we found that PAD4 and HDAC2 associate with p53 in a dynamic fashion [56]. Before DNA damage, both PAD4 and HDAC2 are associated with p53 target gene promoters, while they dissociate from the gene promoters after DNA damage allowing the activation of the p53 target genes, such as p21, GADD45 and PUMA [56]. Taken together, we propose that histone Arg modification in concert with histone Lys acetylation forms a molecular switch on the p53 target gene promoters for gene regulation (Fig. 2).

Although the corepressor function of PAD4 has been well established, it may also play a coactivator role in a promoter context dependent manner. A genome wide ChIP-chip study of PAD4 promoter association in MCF-7 cells found that PAD4 is enriched on the promoter regions of actively transcribed genes [58]. Motif analyses found that many of the PAD4 bound genes contain potential binding sites for Elk-1, a member of the ETS oncogene family [58]. It was proposed that PAD4 interacts with and citrullinates Elk-1 thereby facilitating Elk-1 phosphorylation to activate transcription [58]. Additionally, PAD4 can target histone H3 Arg8 for citrullination and subsequently affects the binding of HP1 (heterochromatin protein 1) to the nearby H3 Lys9 methylation site [59,60]. The dissociation of HP1 from its binding cognate sites after citrullination activates transcription in multiple sclerosis patients [60]. In addition, the dissociation of HP1 likely regulates chromatin decondensation during the formation of neutrophil extracellular traps (NETs) [59].

The hypothesis of the “histone code” predicts that histone modifications may function in a combinatorial manner to regulate chromatin biology [61]. Two adjacent modification sites were proposed to form a binary code to antagonize or synergize the function of each other [62]. The effect of histone H3 Ser10 phosphorylation on the function of H3 Lys9 methylation and HP1 in cell cycle and gene regulation was well studied [63,64]. The new finding of the effect of H3 Arg8 citrullination on HP1 binding to H3 Lys9 methylation greatly enriched the binary code concept and highlights the dense cell signaling information that a cluster of histone residues carries (Fig. 3).

2.2. PAD2 in gene regulation

That PAD2 can citrullinate histone H4 Arg3 *in vitro* raises a role of PAD2-mediated histone citrullination in transcriptional regulation [21]. The expression of PAD2 is regulated by estrogen in vertebrate uterus and pituitary gland [65–67]. Recent data from the Coonrod's group support that PAD2 responses to cellular signals to regulate transcription via histone citrullination. Only during the diestrus phase of the reproduction cycle, PAD2 in mammary gland epithelial cells was found to citrullinate the histone H3 N-terminus [19].