FEBS Letters xxx (2015) xxx–xxx

journal homepage: www.FEBSLetters.org

Review

Architectural proteins, transcription, and the three-dimensional organization of the genome

Caelin Cubeñas-Potts, Victor G. Corces *

Department of Biology, Emory University, 1510 Clifton Rd NE, Atlanta, GA 30322, USA

ARTICLE INFO

Article history:
Received 2 April 2015
Revised 7 May 2015
Accepted 9 May 2015
Available online xxxx

Edited by Wilhelm Just

Keywords:
Chromatin
Epigenetics
Topologically associating domain
Insulator
CCCTC-binding factor

ABSTRACT

Architectural proteins mediate interactions between distant sequences in the genome. Two well-characterized functions of architectural protein interactions include the tethering of enhancers to promoters and bringing together Polycomb-containing sites to facilitate silencing. The nature of which sequences interact genome-wide appears to be determined by the orientation of the architectural protein binding sites as well as the number and identity of architectural proteins present. Ultimately, long range chromatin interactions result in the formation of loops within the chromatin fiber. In this review, we discuss data suggesting that architectural proteins mediate long range chromatin interactions that both facilitate and hinder neighboring interactions, compartmentalizing the genome into regions of highly interacting chromatin domains.

1. Introduction

Multiple lines of evidence indicate that chromosome organization is a contributor to gene expression regulation [1,2]. The use of 3C-derived approaches to detect intra- and inter-chromosome interactions has led to the observation that individual chromosomes are highly organized structures. Chromatin interactions decrease with increasing linear genomic distance and occur non-randomly across the chromosome length [3–6]. Based on the frequency of these interactions, chromosomes can be divided into distinct regions of highly interacting chromatin, named topologically associating domains (TADs), which engage in few long-range interactions with loci in other TADs [7]. Architectural proteins, also known as insulator proteins, appear to play a critical role in the three-dimensional organization of the genome. Here we discuss known architectural proteins in *Drosophila* and mammals, and describe evidence suggesting that architectural proteins regulate long range chromatin contacts and ultimately, gene expression. Current results suggest that architectural proteins have two inter-related functions, genome compartmentalization and the facilitation of interactions between regulatory elements. Finally, we end with a discussion of the molecular mechanisms regulating interactions between distant architectural protein binding sites.

2. Architectural proteins

The roles of architectural proteins in genome organization and function can be explained by their ability to facilitate the formation of long-range contacts between DNA sequences. In *Drosophila*, 11 different DNA binding architectural proteins have been identified, each recognizing a unique DNA motif: CCCTC-binding factor (CTCF), Suppressor of Hairy-wing (Su(Hw)), Boundary Element Associated Factor 32 (BEAF-32), DNA Replication Related Element binding Factor (DREF), Transcription Factor IIIC (TFIIIC), Z4 (also called Putzig), Early Boundary Activity DNA-binding Factor (Elba), Pita (also called Spotted dick), Zinc Finger Interacting with CP190 (ZIPIC), Insulator binding factor 1 (Ibf1), and Insulator binding factor 2 (Ibf2) [8–15]. ChIP-seq experiments demonstrating co-occupancy in the genome as well as coimmunoprecipitation studies have demonstrated that DNA binding architectural proteins interact with accessory proteins, which do not recognize specific DNA motifs [9,10,16]. The accessory proteins identified in *Drosophila* include Centrosomal Protein 190 (CP190), Modifier of mdg4 (Mod(mdg4)), Rad21 (a component of the cohesin complex), Cap-H2 (a component of the condensin II complex), the long

Abbreviations: ChIP-seq, chromatin immunoprecipitation-sequencing; 3C, chromosome conformation capture; ChIA-PET, chromatin interaction analysis by paired-end tag sequencing; STARR-seq, self-transcribing active regulatory region sequencing

* Corresponding author. Fax: +1 404 727 2880.
E-mail address: vcorces@emory.edu (V.G. Corces).

http://dx.doi.org/10.1016/j.febslet.2015.05.025
0014-5793/

isoform of Female sterile homeotic on chromosome 1 (Fs(1)h-L), Lethal (3) malignant brain tumor (L3mbt), and Chromator (also called Chriz) [10]. Notably, depletion of either CTCF or CP190 reduced the chromatin interactions in the *Abd-B* locus by 3C analysis, suggesting that both DNA-binding and accessory architectural proteins can functionally contribute to chromatin looping interactions in cells [17]. Of particular interest, a recent *in vitro* analysis has provided a model for how DNA-binding and accessory architectural proteins function together to mediate long range chromatin interactions. Purified BEAF-32 protein was capable of binding its DNA motif, but only formed intermolecular interactions between two BEAF-32 motifs in the presence of the accessory proteins Chromator or CP190 [18]. However, how DNA-binding and accessory architectural proteins interact to mediate intra-chromosomal interactions within a cell remains to be determined.

The number of architectural proteins characterized in mammals is not as extensive as in *Drosophila*. Multiple lines of evidence indicate that CTCF and cohesin are mediators of chromatin interactions and thus, architectural proteins. Functional studies evaluating the function of CTCF and cohesin in chromatin looping within individual genomic loci were the first studies to indicate these proteins are regulators of chromatin interactions [19–22]. More recently, CTCF and cohesin depletion studies have shown that loss of these architectural proteins reduces genome-wide chromatin interactions [23–25]. Furthermore, ChIA-PET analysis, a technique that maps the chromatin interactions occurring between loci occupied by a specific protein, characterized a subset of the chromatin interactions that occur between cohesin and CTCF occupied sites, indicative that cohesin and CTCF may play a role mediating chromatin interactions [26,27]. In addition to CTCF and cohesin, the cohesin interacting proteins Nipbl (the protein responsible for loading cohesin rings onto DNA), and Mediator have also been implicated as mammalian architectural proteins due to their interactions with cohesin and enrichment at enhancer–promoter contact sites [28,29]. Similar to *Drosophila*, a series of proteins that co-localize or directly interact with CTCF have been identified in mammals, including Yin Yang 1 (YY1), Kaiso, Chromodomain Helicase-DNA-binding protein 8 (CHD8), Poly ADP-Ribose Polymerase 1 (PARP1), MYC-associated zing-finger protein (MAZ), jun-D proto-oncogene (JUND), ZNF143, nucleophosmin, the PR domain zinc-finger protein 5 (PRDM5), and TFII-I [30,31]. Through their association with CTCF, it is possible that CTCF interacting proteins also function as architectural proteins but additional experiments are required to address this hypothesis. Notably, there is a growing body of evidence that the CTCF-interacting protein ZNF143 is a mammalian architectural protein. The genomic occupancy of ZNF143 was shown to highly correlate with CTCF sites forming chromatin loops and ZNF143 depletion studies demonstrated a functional role for this protein in mediating long range chromatin interactions [32,33]. Beyond the cohesin and CTCF interacting proteins, a number of other potential architectural proteins have been characterized as proteins required for chromatin loop formation within specific genomic loci like CHD6 in the CFTR locus or Ldb1 in the β-globin locus [34,35]. However, the significance of CHD6 and Ldb1 as genome-wide regulators of chromatin interactions is not currently known. In summary, these data suggest that, similar to *Drosophila*, mammals express a wide array of potential architectural proteins important for regulating long range interactions that should be subject to additional characterization.

In mammalian cells, the architectural protein cohesin is an important contributor to the regulation of nuclear size and organization. For example, depletion of cohesin in astrocytes caused an approximately 25% increase in the volume of the entire nucleus, possibly due to a loss of chromatin interactions and interphase chromatin organization [23]. In a reciprocal analysis, deletion of the cohesin regulator Wapl caused excessive cohesin and CTCF occupancy on chromatin and a striking hypercondensation of interphase DNA was observed [36]. It is interesting to speculate that the accumulation of cohesin and CTCF on chromatin resulted in excessive chromatin interactions responsible for the condensation phenotype. However, additional characterization is required to determine if these changes in nuclear morphology were direct effects of altered chromatin interactions.

3. General mechanisms of architectural protein function

Many functions of architectural proteins can be explained by their ability to mediate interactions between distant loci and form chromatin loops. Architectural proteins directly binding enhancers and promoters increase the contact frequency between regulatory elements by forming stable protein–protein interactions between them, consistent with the classical model of enhancer–promoter interactions (Fig. 1A) [32,37–40]. In addition, polymer simulations have suggested that an architectural protein interaction can facilitate neighboring interactions by two additional mechanisms [41,42]. First, regulatory elements that are looped out by two interacting architectural protein-bound loci have higher contact frequencies, indicative that a chromatin loop highly interacts within itself (Fig. 1B) [41,42]. Secondly, the genomic elements flanking a chromatin loop are brought into closer proximity by architectural proteins, reducing the linear genomic distance between them and increasing their contact frequency (Fig. 1C) [41]. However, additional evidence is required to determine if the simulation studies are representative of mechanisms of architectural protein-mediated facilitation in cells. In addition to facilitating some interactions, the establishment of a chromatin loop by architectural proteins also precludes other interactions. The original function ascribed to architectural proteins was their ability to insulate promoters from the effect of regulatory sequences such as enhancers, consistent with a function in hindering chromatin interactions [19,43–51]. Furthermore, a reduction of contact frequency for interactions between sequences located within a chromatin loop and sequences present outside the loop was observed in polymer simulation studies, supporting the notion that an architectural protein-mediated interaction will hinder a subset of

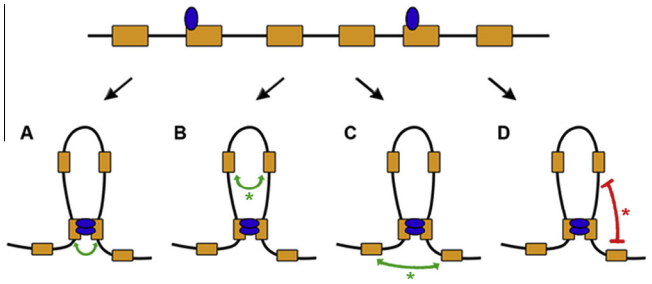

Fig. 1. The consequences of a single architectural protein interaction. Architectural proteins organize regulatory elements within the genome. The facilitating and inhibitory effects of a single architectural protein interaction between two genomic loci are shown. Regulatory elements are shown as gold boxes and architectural proteins are in blue. Facilitating interactions are depicted as green arrows, while insulating interactions are in red. (A) Architectural proteins bound to regulatory elements promote interactions between the regulatory sequences. In addition, polymer simulation studies have suggested that a single architectural protein interaction affects neighboring interactions as well (denoted by *). (B) Regulatory elements within chromatin loops formed by architectural protein interactions may interact more frequently. (C) By reducing the linear genomic distance between loci flanking a chromatin loop, architectural proteins may promote their interactions. (D) Regulatory elements within chromatin loops are insulated from interactions with elements outside the chromatin loops.

Please cite this article in press as: Cubeñas-Potts, C. and Corces, V.G. Architectural proteins, transcription, and the three-dimensional organization of the genome. FEBS Lett. (2015), http://dx.doi.org/10.1016/j.febslet.2015.05.025