BBA - General Subjects xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

BBA - General Subjects

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

Deciphering the role of dimer interface in intrinsic dynamics and allosteric pathways underlying the functional transformation of DNMT3A

Zhongjie Liang[a], Junchi Hu[b,c], Wenying Yan[a], Hualiang Jiang[b], Guang Hu[a,*], Cheng Luo[b,*]

[a] Center for Systems Biology, Soochow University, Suzhou 215006, China
[b] Drug Discovery and Design Center, CAS Key Laboratory of Receptor Research, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zuchongzhi Road, Shanghai 201203, China
[c] University of Chinese Academy of Sciences, 19 Yuquan Road, Beijing 100049, China

ARTICLE INFO

Keywords:
Conformational dynamics
Allosteric communication
Elastic network models, network theory, coevolution analysis

ABSTRACT

Background: DNMT3A, as de novo DNA methyltransferase, is essential for regulating gene expression through cellular development and differentiation. The functions of DNMT3A rely on its oligomeric states and allosteric regulations between its catalytic domain and binding partners. Despite recent resolution of autoinhibitory and active DNMT3A/3L crystal structures, the mechanism of their functional motions and interdomain allostery in regulating the activity remains to be established.

Methods: The hybrid approach, comprising Elastic Network Models coupled with information theory, Protein Structure Network, and sequence evolution analysis was employed to investigate intrinsic dynamics and allosteric properties of DNMT3A resolved in autoinhibitory and active states.

Results: The conformational transition between two states is characterized by global motions, and the homodimer displays the similar dynamic properties as tetramer, acting as the basic functional unit. The hinge residues with restricted fluctuations are clustered at the dimer interface, which are predicted to enjoy remarkably efficient signal transduction properties. The allosteric pathways through the dimer interface are achieved by a cascade of interactions predominantly involving conserved and co-evolved residues.

Conclusions: Our results suggest that structural topology coupled with global motions indicates the structural origin of the functional transformation of DNMT3A. The comprehensive analysis further highlights the pivotal role of the dimer interface of DNMT3A both in defining the quaternary structure dynamics and establishing interdomain communications.

General significance: Understanding the global motions of DNMT3As not only provides mechanical insights into the functions of such molecular machines, but also reveals the mediators that determine their allosteric regulations.

1. Introduction

DNA methylation, occurring at the 5-position of cytosine residues within CG dinucleotides, is a critical epigenetic modification controlling chromatin stability, genomic imprinting, cellular differentiation, transcriptional regulation, and memory formation in mammals [1,2]. Aberrant methylation patterning is a driving force in the onset and progression of several diseases, in particular cancer [3]. The methyl group of the coenzyme S-adenosyl-*L*-methionine (SAM) is transferred to cytosine residues of DNA by DNA methyltransferases (MTases). Generally, MTases could be simply classified into de novo and maintenance methyltransferases. As de novo MTases, DNMT3 family consists of two active DNMTs, DNMT3A and DNMT3B, and one regulatory factor, DNMT3-Like protein (DNMT3L), which does not have catalytic activity and has been identified as a stimulator for the catalytic activity of DNMT3 enzymes [4,5]. Whereas DNMT1 preferentially methylates hemimethylated CpG sites, acting as maintenance methyltransferase [6,7].

All mammalian DNMTs have similar architectural arrangements, including a large multi-domain N-terminal part of variable sizes, with regulatory and targeting functions, and a smaller C-terminal part, comprising the catalytic center. The N-terminal parts, comprising several domains with regulatory functions, are quite different between DNMT1 and DNMT3 enzymes. In DNMT3A and 3B, the N-terminal part contains two defined domains, including a PWWP domain binding with Histone 3 (H3) trimethylated at K36, and a cysteine-rich

* Corresponding authors.
E-mail addresses: huguang@suda.edu.cn (G. Hu), cluo@simm.ac.cn (C. Luo).

https://doi.org/10.1016/j.bbagen.2018.04.015
Received 1 November 2017; Received in revised form 3 April 2018; Accepted 13 April 2018

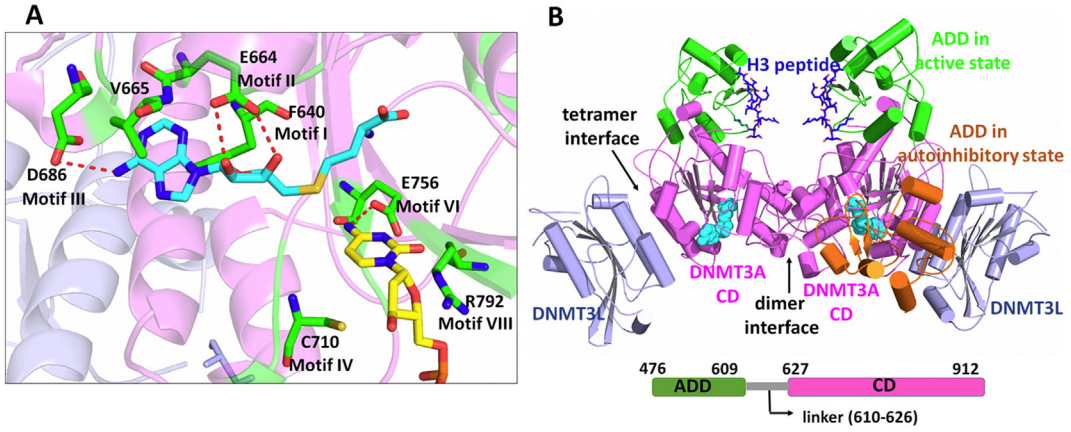

Fig. 1. The detailed catalytic site of CD in DNMT3A and structural representation of complex DNMT3A/3L heterotetramer. (A) The detailed interactions in SAM and cytosine binding site, with conserved motifs labeled. (B) Structure of the heterotetramer of DNMT3A/3L (PDB ID: 4U7P for autoinhibitory state, 4U7T for active state). Herein, the interface involved in DNMT3L and CD of DNMT3A is referred as the tetramer interface, and the interface involved in both CDs is referred as the dimer interface. The ADD domain is shown both in autoinhibitory state (orange) and catalytically active state (green), and only one ADD domain in autoinhibiory is shown for clarity. H3 peptide (blue) binds with ADD domain in active state. SAM is represented as the cyan spheres. Corresponding color-coded domain organization of DNMT3A is schematically shown at the bottom.

ATRX–DNMT3–DNMT3L (ADD) domain binding with H3 tails unmodified at K4, which are involved in interactions with other proteins and in targeting to heterochromatin [8–11]. The connection between DNA methylation and histone modification has thus long been established. In recent structural and biochemical research, the arrangement of these domains is involved in the allosteric regulation of the enzyme activity, specificity and localization, playing a central role in the regulation of the biological functions of DNMTs [12–14]. The C-terminal domains of DNMTs contain ten amino acid motifs conserved among prokaryotic and eukaryotic DNA-(cytosine-C5)-MTases, adopting a fold characteristic for SAM binding, DNA recognition and binding, target base flipping and catalysis [15]. In the catalytic domain (CD) of DNMT3A (Fig. 1A), the adenine ring and ribose moiety of SAM make flanking van der Waals contacts and H-bonds with the residues in motif I (F640-D641-G642), motif II (E664-V665) and motif III (D686-V687). Concerning the cytosine binding site, motif IV (P709-C710-N711), motif VI (E756-N757-V758H) and VIII (R790-X-R792-X-F) contains the crucial residues C710, E756 and R792 responsible for the methylation catalysis [16,17]. The variable region between motifs VIII and IX, also called target recognition domain (TRD), is small in DNMT3A, but defines the DNA sequence specificity as well as the base to be methylated within the target sequence [18].

In 2007, the first structure of the complex C-terminal domains of DNMT3A/3L was resolved in human [19]. Recently, the novel and more complete structures of DNMT3A in complex with the C-terminal domain of DNMT3L (C^{DNMT3L}), in two alternative conformations (PDB ID: 4U7P and 4U7T), have been determined [20]. 4U7P represents an autoinhibitory state without H3 peptide, whereas 4U7T represents an active state binding with H3. The complexes form a linear heterotetramer, consisting of two DNMT3A subunits in the middle and two DNMT3L subunits on both sides (Fig. 1B). The CD of DNMT3A contains two kinds of protein-protein interface: the dimer interface involved in both CDs and the tetramer interface involved in DNMT3A/3L interactions. The dimer interface, supposed to compensate for the small TRD region, increases the size of DNA binding interface and indicates the important roles for DNA recognition and binding. Concerning the tetramer interface, DNMT3L was originally discovered as a stimulator of DNMT3A [21], and could change the sub-nuclear localization of the enzyme in the recent research, influencing multiple aspects of DNMT3A function [22]. DNMT3A consists of ADD and CD domains connected by a long flexible “linker” (A610-V626), which allows the protein to assume a wide range of conformations. As shown in Fig. 1B, the ADD domain of DNMT3A interacts with and inhibits the enzymatic activity of CD through blocking the binding of DNA. With H3 binding and a large movement of ADD domain, DNMT3A transforms from autoinhibitory state to active state. Both states represent a breakthrough in the research of structure-function of DNMT3A, which firstly provides the molecular mechanism underlying the positive correlation between DNA methylation and H3 without modifications.

In DNMT3A, the wealth information from the outstanding crystal study indicates that the catalytic activity of CD is under allosteric control by the rearrangements of ADD domains [20]. In addition, the regulation of DNMT3A function by the oligomerization states has also attracted much attention [20,22–24]. However, the question of whether the large conformational change of ADD domain is the intrinsic dynamics of DNMT3A/3L tetramer or induced by H3 binding remains to be answered. The role of the interfacial interactions and overall contact topology in the inherent dynamical properties of different oligomerizations remains challenging. For example, is the switch of ADD domain in the two binding sites of CD an intrinsic dynamical ability of monomer that is conserved and exploited in the dimer/tetramer complex? Or, is it an acquired mode of motion originating from the topology of the dimer/tetramer complex and not accessible to monomer? To answer these questions, the intrinsic dynamics in different oligomeric states and allosteric pathways underlying functional transformations are the focus in our research.

Nowadays, computational methods are being increasingly applied to protein structures to investigate their complex biological functions. Among them, Elastic Network Models (ENMs) and Protein Structure Networks (PSNs) are two kinds of coarse-grained approaches, which can fast probe protein structures and dynamics efficiently, especially for large protein molecular machines [25]. To date, the Gaussian Network Model (GNM) [26] and the Anisotropic Network Model (ANM) [27] proposed by Bahar et al. [28], are two most widely used ENM methods. Both the GNM and the ANM consider a protein as an elastic network connected by springs with uniform force constants. Thus, their major advantage against molecular dynamics (MD) simulation is using few global modes to describe the intrinsic dynamics of large molecular systems. The further study proved that intrinsic dynamics or global modes would reveal significant features of the mechanistic basis for allostery [29,30]. In addition, integrating information-theoretic approaches into GNM could elucidate the most efficient pathways of signal transmission favored by the overall architecture [31,32]. The PSN considers a protein as a graph, in which nodes correspond to residues and edges represent different interactions between two nodes. Various measures of node centrality (betweenness and closeness) have