Structural Bioinformatics Project 4

Anabel Yong, University of Edinburgh

DNMT1 methylates hemi-methylated CpGs but not unmethylated CpGs

Schematic diagram summarizing what the mouse DNMT1-DNA structures reveal about molecular mechanism of DNMT1

DNMT1 uses the residues C1229 and E1269 in coordination with S-adenosyl L-methionine(SAM) to methylate CpG Cytosine in hemi-methylated DNA (Figure A) via nucleophilic attack. C1229 carries out a nucleophilic attack on C6 of Cytosine and forms a sulfur bond to stabilize the base in its flipped position. At the same time nitrogen (N3) picks up a proton from E1269, generating the 1st transition state. After this, N3 gets deprotonated by E1269, leading to a nucleophilic attack of SAM by the cytosine ring to transfer the SAM methyl group to Cytosine (Figure B), resulting in 2nd transition state where methylated cytosine is bound to Cys1229 via sulfur bond, and SAH (Figure C). The hydrophobic residues Cys1501, Leu1502, Leu1514 and Met1535 form a shallow concave surface interacting with 5-methyl group of mC6 through van der Waals forces (Figure D). The aromatic ring of Trp1512 is slid into the DNA major groove and is partially stacking with the mC6 upon the complex formation. The side chain of Lys1537 occupies the empty space on the parental strand via the insertion through DNA major groove. The side chain of Met1235 occupies the empty space vacated by the flipped-out fC71 on the target strand (Figure G). Catalytic helix gets displaced in the DNMT1 bound to hemimethylated DNA compared to the DNMT1 bound to unmethylated DNA (Figure E and F). Unmethylated DNA is excluded from the active site of DNMT1 by the binding of CXXC domain, whereas the presence of the acidic CXXC-BAH1 linker (D703-D711) positioned directly between the DNA and active site prevents entrance of DNA into catalytic pocket. Furthermore, BAH2-TRD loop anchors TRD (a region of the methyltransferase domain) in retracted position and prevents it from binding in the unmethylated DNA major groove.

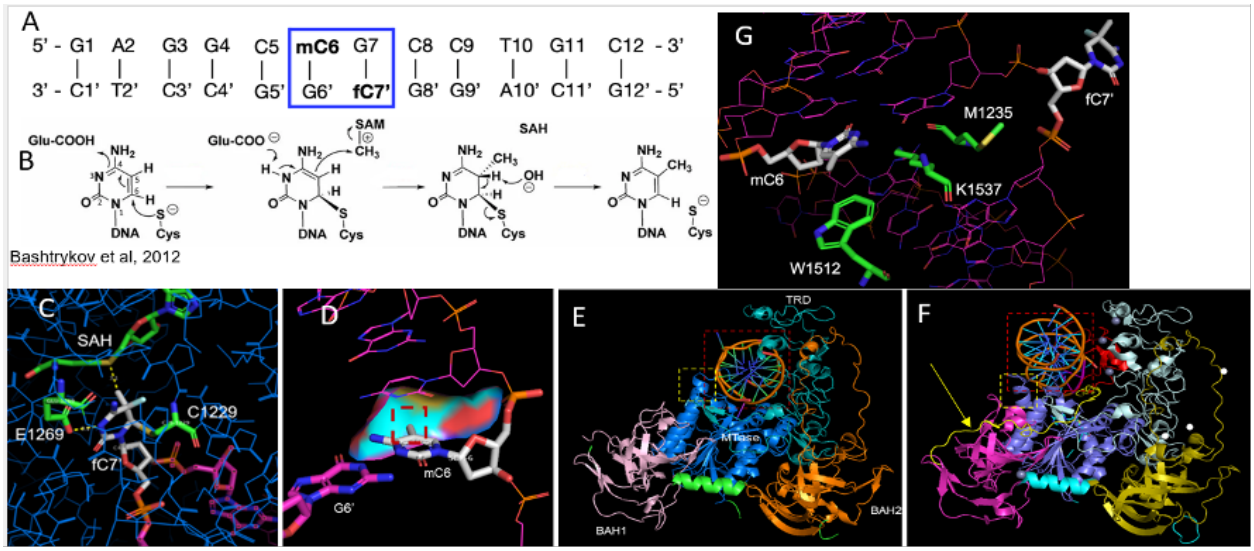

Figure 1: (A) Sequencing system of the 12-bp hemi-methylated DNA duplex at CpG site, the blue bracket presents the hemi-CpG site in which fC7' corresponds to the flipped out fluorocytosine. (B) Proposed catalytic mechanism of DNA methylation by DNMT1. (C) Pymol representation of the DMNT1 catalytic side, residues E1269 and C1229, fC7' DNA base and SAM shown in sticks (4DA4). (D) Pymol representation of the hydrophobic concave pocket of DNMT1, red bracket indicates the methyl group of mC6 (4DA4). (E, F) Pymol structures comparison of DNMT1 bound to hemimethylated DNA (4DA4) and unmethylated DNA (3PT6), yellow brackets show the catalytic helix and red brackets show the bound DNA duplex. (G) Pymol representation of the insertion of Met1235 and Lys1537 side chains into the empty space and the partial stacking of Trp1512 indole group and mC6 in DMNT1 complex, CXXC-BAH1 linker indicated with a yellow arrow (4DA4).

Molecular Mechanisms of DNMT1 structure that remain unanswered: how do we experimentally solve this?

Function of BAH1/2 domains remain unknown. There are indirect evidence indicating how they seem to function as a protein-protein interaction module for gene silencing but there is no direct involvement of the BAH domains in the methylation process from papers which have inferred functionality of BAH domains. BAH1/2 domains provide accessible surface area for binding other proteins but specific binding proteins to these domains are not known. The mechanism of DNMT1 recruitment to replication foci via PCNA interaction remains unknown. There have been a few models proposing an interaction between H3K23, DNMT1 and UHRF1 (adaptor protein) or direct interactions between DNMT1 and SUV39H1, SUV39H2 or G9a functioning in targeting histone methyltransferases to replication foci during replication.

Experimental Protocol- with biochemical and molecular experiments to solve this issue.

UHRF1 is a proposed adaptor protein for DNMT1 replication fork recruitment. Visualisation of UHRF1 and DNMT1 in S phase would be conducted using the mouse ESCs model with either GFP-tags or fluorescently labelled specific antibodies. Images would be taken in early, mid and late S phase, and comparison of localisation of the UHRF1 and DNMT1 signals. Additionally, a DNMT1 mutant with an abolished UHRF1 binding site would be produced to compare colocalization of WT and mutant DNMT1 with UHRF1. A pull-down assay would also be conducted to detect the proteins involved in the DNMT1 recruitment to the replication fork. DNMT1 would be labelled with a high affinity tag, such as biotin, and DNMT1 would be the bait protein. The cells would be lysed during S phase and the protein complex immobilised on the column. Non-interacting proteins would be washed away and target protein complexes eluted. SDS-PAGE and MS would be performed to analyse the interacting partners of DNMT1. Noteworthy, fluorescent labelling may infer the functionality of both proteins, giving unreliable results. The pull-down assay might also produce unreliable results from non-specific binding. To investigate BAH domain function, homology search of BAH1/2 and look for homologues with annotated function. Multiple sequence alignment would help to determine conserved regions and visualization via CONSURF analysis that could be important for protein-protein interactions. We can infer the functionality of the BAH domains through these protein-protein interactions. To identify the interactions of BAH1 or BAH2 domains, a single protein of one of the domains would be constructed, and a pulldown with antibody for C-terminal myc tag. This Co-IP experiment would pulldown additional proteins interacting with BAH1/2 domain, which would then be identified via MS. From protein databases, we can identify these proteins which specifically bind to BAH1/2 domains. Mutation of conserved residues of BAH 1/2 is done. If these mutations affect other proteins binding we can see the absence of these proteins in Co-IP experiment. From here, we can compare the proteins which aren't bound to mutated BAH domains, and determine which residues involved in binding proteins. BAH domain functionality is inferred from interacting partners.