Reading and writing DNA methylation

Albert Jeltsch

By recruiting the Dnmt1 DNA methyltransferase to hemimethylated DNA, the ubiquitin-like with PHD and ring finger domains 1 (UHRF1) protein plays an important part in DNA methylation. The structures of the SRA domain of UHRF1 in complex with hemimethylated DNA show that the methylated cytosine is flipped out of the DNA helix, as observed previously with DNA methyltransferases.

In mammalian DNA, about 70% of the cytosine residues within CG sequences are methylated at the cytosine-C5 position in a cell type–specific pattern[1,2]. Methylation of CG-rich sequences in the promoter regions of genes, so-called CG islands, represses gene expression and represents an important epigenetic signal that is essential for development. Three groups report structures of the SET and RING-associated (SRA) domain of the UHRF1 protein (also known as NP95 and ICBP90)[3–5], which has an important role in DNA methylation[6,7]. Unexpectedly, the structures of the SRA domain in specific complex with hemimethylated DNA (**Fig. 1**) show that the protein flips the 5-methylcytosine out of the DNA helix, which recapitulates a conformational change of the DNA that was discovered with DNA methyltransferases 15 years ago[8]. Thus, both readers and writers of DNA methylation make use of a similar structural approach when interacting with DNA.

The molecular basis of the inheritance of DNA methylation signals lies in the fact that the CG site is palindromic and either methylated on both strands or completely unmethylated. Therefore, the pattern of DNA methylation is not destroyed by DNA replication, but converted into a pattern of unmethylated and hemimethylated CG sites. An enzyme that specifically methylates hemimethylated CG sites can act as an epigenetic copy machine and reconstitute the original DNA methylation pattern[9,10]. Dnmt1, the first mammalian DNA methyltransferase isolated and cloned[11,12], shows properties expected for such a maintenance enzyme: it has specificity for methylation of hemimethylated DNA (2-fold to 200-fold, depending on the substrate DNA sequence and experimental assay system) and methylates DNA in a highly processive reaction[13,14].

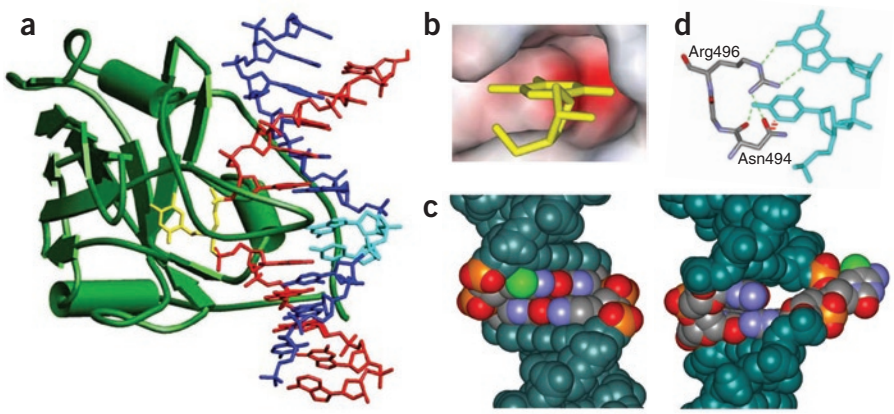

Figure 1 Structures of the SRA domain with hemimethylated DNA. (**a**) Model of the SRA domain bound to hemimethylated DNA[3–5]. The protein is shown schematically and colored green. The methylated DNA strand is colored red, with the flipped 5-methylcytosine highlighted in yellow. The unmethylated strand of the DNA is colored blue, and the cytosine and guanine bases of the hemimethylated CG site are highlighted in cyan. (**b**) Picture of the 5-methylcytosine (yellow), which is swung out of the DNA helix and inserted into a deep binding pocket of the SRA domain (shown in surface representation colored according to electrostatic potential). (**c**) Structural models of a hemimethylated CG site in B-DNA (left) and after base flipping in the SRA domain (right). The CG base pair is colored according to atom type, except for the methyl group, which is colored green. The flanking DNA is colored dark green-blue. The solvent-accessible surface area, calculated by WebLab Viewer Pro 4.0 using a probe radius of 1.4 Å, is 477 Å^2 for the hemimethylated CG site in B-DNA and 789 Å^2 after flipping of the methylcytosine base. (**d**) Recognition of the guanine and cytosine in the unmethylated DNA strand (both in cyan). The guanine is contacted by Arg496, which is positioned by Asn494. Hydrogen bonds are shown as dashed green lines, and the close proximity between the cytosine C5 atom and Asn496 is indicated in red. The picture was generated with WebLab Viewer Pro 4.0 using PDB 2ZKF.

Despite its intrinsic preference for hemimethylated DNA, the question remained of how Dnmt1 would find its target sites on the DNA. Cellular localization studies showed that the enzyme is mainly concentrated at sites of active DNA replication, so-called DNA replication foci[15]. Later it was observed that Dnmt1 interacts with PCNA[16], a major component of the DNA replication fork. The PCNA interaction turned out to be supportive but not necessary for DNA methylation: after removal of the PCNA interaction domain in Dnmt1, DNA re-methylation after replication was delayed but still possible[17,18]. The interaction of Dnmt1 with PCNA is transient[19], suggesting that Dnmt1 is not a stable component of the replication fork. In this model, PCNA recruits Dnmt1 to sites of DNA replication and loads the enzyme onto the DNA, where it can search for target sites by linear diffusion. UHRF1 was identified only last year as an additional factor involved in the recruitment of Dnmt1 to newly replicated DNA[6,7]. It specifically binds to hemimethylated DNA via its SRA domain and also interacts with Dnmt1. Deletion of UHRF1 leads to a striking loss of DNA methylation, indicating that UHRF1 is an essential component of the maintenance DNA methylation

Albert Jeltsch is at the Biochemistry Laboratory, School of Engineering and Science, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany.
e-mail a.jeltsch@jacobs-university.de

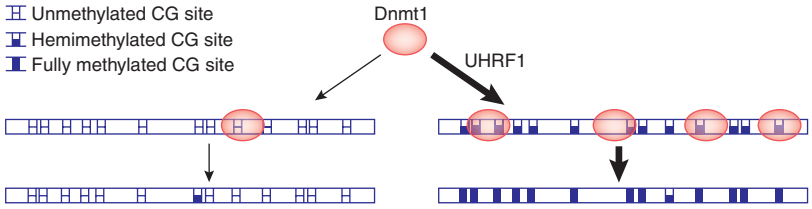

Figure 2 Double-sieve mechanism to ensure accurate maintenance methylation based on Dnmt1 recruitment and specificity. UHRF1 preferentially recruits Dnmt1 to hemimethylated DNA regions, which had been methylated before DNA replication. Dnmt1 shows higher catalytic activity on hemimethylated DNA as compared to unmethylated DNA.

machinery. In addition to its role in recruiting Dnmt1 to hemimethylated DNA, UHRF1 also recruits histone deacetylases[20] and interacts with histone H3 lysine 9 (H3K9)-methylated histone tails[21]. Furthermore, UHRF1 harbors a ubiquitin ligase domain that is important for its biological function, but the physiological target(s) of ubiquitination have not been characterized so far[22]. These results suggest that UHRF1 acts as an epigenetic integration platform involved in the cross-talk between histone deacetylation, H3K9 methylation and DNA methylation, all silencing signals that have been shown to function synergistically.

It is an interesting detail of the structural analyses that the SRA domain of UHRF1 flips the methylated cytosine out of the DNA helix and buries it in a pocket in the protein[3–5] (**Fig. 1b**). Base flipping was initially discovered in DNA methyltransferases, which flip their target base to get sufficient access to catalyze the chemical reaction[23]. The SRA domain uses base flipping as the solution to a challenging task in molecular recognition: despite its short recognition sequence, consisting of only 2 base pairs, it has to form a stable complex with hemimethylated DNA that allows discrimination of the three different methylation states of the DNA—unmethylated, methylated in one DNA strand (hemimethylated) and methylated in both DNA strands (fully methylated), which all differ by single methyl groups. Base flipping helps DNA recognition by the SRA domain by almost doubling the solvent-accessible surface area of the hemimethylated CG site[5] (**Fig. 1c**), thus allowing for the formation of more specific contacts between the protein and its target sequence. After flipping, the methylated cytosine is recognized by several hydrogen bonds and van der Waals contacts formed to the methyl group. The orphaned guanosine is contacted closely by Arg496 with two hydrogen bonds (**Fig. 1d**). Interestingly, there are only a few contacts formed to the second base pair of the recognition sequence. This does not cause specificity problems, because in mammals DNA methylation is observed only in a CG context. However, contacts to the second base pair are necessary for the discrimination between hemimethylated and fully methylated CG sites, which requires recognition of the unmethylated cytosine in the second strand. This is achieved by the positioning of Asn494, which buttresses the conformation of Arg496, in close proximity to the C5 atom of the cytosine. If this C5 atom were methylated, Asn494 would be moved away and the interaction of Arg496 with the DNA would be disrupted[3,4] (**Fig. 1c**).

Because the crystal structure analyses show that binding of UHRF1 to a hemimethylated site would block the target strand for methylation, a direct interaction of UHRF1 and Dnmt1 with one hemimethylated CG site is unlikely. Therefore, UHRF1 might recruit Dnmt1 to a hemimethylated DNA domain where the enzyme could search for hemimethylated target sites by linear diffusion on the DNA. After dissociation of UHRF1, the site initially blocked by UHRF1 would become accessible to Dnmt1. The specific recruitment of Dnmt1 to hemimethylated DNA by UHRF1 could improve the overall accuracy of maintenance methylation, because genome-wide DNA methylation analyses have indicated that methylation levels of DNA show a bimodal distribution, consisting of either unmethylated or densely methylated regions[24–26]. Thus, DNA mainly contains two types of sequences after replication: stretches with many hemimethylated CG sites, which are targets of maintenance methylation; and stretches of unmethylated DNA that should not become subject to DNA methylation. The accuracy of maintenance methylation could be ensured in a double-sieve mechanism based on the methylation-specific recruitment of Dnmt1 to hemimethylated DNA and the intrinsic preference of the enzyme to methylate hemimethylated CG sites (**Fig. 2**).

In summary, the structures of the SRA domain bound to hemimethylated DNA provide a fascinating snapshot of the molecular details of the recognition of hemimethylated DNA by an epigenetic reading domain. Future structural and biochemical work will address the questions of how the UHRF1 protein interacts with Dnmt1, how its other functions are linked to its binding to hemimethylated DNA and what its contribution to the accuracy of the maintenance of DNA methylation is in quantitative terms.

1. Klose, R.J. & Bird, A.P. *Trends Biochem. Sci.* **31**, 89–97 (2006).
2. Reik, W. *Nature* **447**, 425–432 (2007).
3. Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y. & Shirakawa, M. *Nature* advance online publication, doi:10.1038/nature07249 (03 September 2008).
4. Avvakumov, G.V. *et al. Nature* advance online publication, doi: 10.1038/nature07273 (03 September 2008).
5. Hashimoto, H. *et al. Nature* advance online publication, doi:10.1038/nature07280 (03 September 2008).
6. Bostick, M. *et al. Science* **317**, 1760–1764 (2007).
7. Sharif, J. *et al. Nature* **450**, 908–912 (2007).
8. Klimasauskas, S., Kumar, S., Roberts, R.J. & Cheng, X. *Cell* **76**, 357–369 (1994).
9. Riggs, A.D. *Cytogenet. Cell Genet.* **14**, 9–25 (1975).
10. Holliday, R. & Pugh, J.E. *Science* **187**, 226–232 (1975).
11. Gruenbaum, Y., Cedar, H. & Razin, A. *Nature* **295**, 620–622 (1982).
12. Bestor, T., Laudano, A., Mattaliano, R. & Ingram, V. *J. Mol. Biol.* **203**, 971–983 (1988).
13. Jeltsch, A. *Curr. Top. Microbiol. Immunol.* **301**, 203–225 (2006).
14. Goyal, R., Reinhardt, R. & Jeltsch, A. *Nucleic Acids Res.* **34**, 1182–1188 (2006).
15. Leonhardt, H., Page, A.W., Weier, H.U. & Bestor, T.H. *Cell* **71**, 865–873 (1992).
16. Chuang, L.S. *et al. Science* **277**, 1996–2000 (1997).
17. Egger, G. *et al. Proc. Natl. Acad. Sci. USA* **103**, 14080–14085 (2006).
18. Spada, F. *et al. J. Cell Biol.* **176**, 565–571 (2007).
19. Schermelleh, L. *et al. Nucleic Acids Res.* **35**, 4301–4312 (2007).
20. Papait, R. *et al. Mol. Biol. Cell* **18**, 1098–1106 (2007).
21. Karagianni, P., Amazit, L., Qin, J. & Wong, J. *Mol. Cell. Biol.* **28**, 705–717 (2008).
22. Jenkins, Y. *et al. Mol. Biol. Cell* **16**, 5621–5629 (2005).
23. Cheng, X. *Annu. Rev. Biophys. Biomol. Struct.* **24**, 293–318 (1995).
24. Eckhardt, F. *et al. Nat. Genet.* **38**, 1378–1385 (2006).
25. Weber, M. *et al. Nat. Genet.* **39**, 457–466 (2007).
26. Meissner, A. *et al. Nature* **454**, 766–770 (2008).