Basis for DNA Methyltransferase Recognition and Methylation Specificity

**Background**
DNA methylation has been recognized to play a key role in the regulation of transcription and gene expression through silencing. Therefore, it is not surprising that DNA methylation has been suggested to play a role in cellular disease and dysfunction (such as cancer). In mammals, methylation occurs at the cytosine C5 position by enzymes called DNA-methyltransferases (DNMTs). There have been 2 types of mammalian DNMTs identified. The first type of DNMT is responsible for de novo methylation and consists of the DNMT3a and DNMT3b proteins. The second type of DNMT functions in the maintenance of methylation patterns and the methylase protein responsible is DNMT1. The structure of an auto-inhibited DNMT1 protein in complex with dsDNA has previously been determined (3pt6) and more recently the structure of an active DNMT1-dsDNA complex was published showing a more intimate complex with hemimethylated dsDNA (4da4). In addition, as Carolina presented in journal club, the structure of DNMT3a in complex with the regulatory protein DNMT3L has also been solved (2qrv) (no DNMT3a-DNA complex has been determined... yet).

**Question**
How is the binding/methylation activity of DNMTs defined and what happens in case of disease/dysfunction associated with aberrant DNA methylation?.... and can such a question be answered by a simple structural explanation???

**Hypothesis**
Sequence-dependent DNA shape plays a key role in DNA-methyltransferase recognition and in defining enzymatic specificity (needs work).

Or in other words... the DNA shape readout determines methyltransferase binding and therefore methylation.
Linear map of DNMT1 and DNMT3a-b proteins. Regions involved in DNA recognition of DNMT1 include the CXXC (zinc-finger-like) domain and the target recognition domain (TRD) within the methyltransferase domain.

Structure of DNMTs. Structures of DNMTs colored by domains shown in linear map.

Linear protein map of solved structures. Structural domains used in crystallization.
Minor groove interactions of the DNMT1 CXXC domain (3pt6).

Summary
DNA methylation plays an important role in many different biological pathways. Understanding the basis of DNA methylation in terms of what drives DNA binding and enzyme specificity poses an interesting biological question. From structural data, or at least from what structural data is present, it could be possible that DNA plays a role in determining DNMT binding and perhaps methylation specificity. If such an explanation holds true, it could be also be a good hypothesis for explaining the basis of disease in cases of abnormal silencing patterns resulting from DNA mutations/polymorphisms. However, all-in-all,
this is just a very simplistic exploration of DNMTs and DNA structure and there are many questions and considerations still left unaddressed.

Weekly Papers + Summary and Applications (purp = previous week’s papers, red = next week’s papers)

Degner et al. “DNaseI sensitivity QTLs are a major determinant of human expression variation” Nature 16 February 2012

Chromatin organization is a major determinant in regulating gene expression. This study used DNaseI sensitivity to generate a genome-wide map of chromatin accessibility. DNaseI-seq read-depth was significantly correlated with genotype and the presence of SNPs. It also was shown that DNaseI sensitive loci were strongly enriched at transcription factor binding sites. A couple of interesting questions could be:

- How are SNPs causing a change in chromatin structure (as measured by DNaseI sensitivity) and what are the biological outcomes of such change (especially in diseases such as cancer)?
- I thought it was interesting that DNaseI sensitive loci overlapped with TF binding sites… given this information and Carol’s data on DNA shape and DNaseI activity, is there anything that can be inferred about the structure (in terms of DNA shape) of TF binding sites compared to all other DNA as a whole?

Mullen et al., “Master transcription factors determine cell-type-specific responses to TGF-β signaling” Cell, October 28, 2011

Tissue specific master transcription factors Oct4, Myod1, and PU.1 associate with and direct Smad3 binding in response to TGF-β signaling thus orchestrating the cell-type-specific response. It’s interesting to think about how DNA structure may (or may not) give rise to tissue specific responses. The base readout for Smad3 differs quite a bit when associated (not direct) with tissue specific MTFs. Perhaps this association guides Smad3 to recognize a specific sequence-dependent (and tissue specific???) shape readout.

Smad3 ChIP-seq results (note that this represents the MTF/Smad3 complex)

Xie et al. “Base resolution analysis of sequence and parent-of-origin dependent DNA methylation in the mouse genome” Cell, 17 February 2012