Structural and biochemical insight into the mechanism of dual CpG site binding and methylation by the DNMT3A DNA methyltransferase

Max Emperle[#], Disha M. Bangalore[#], Sabrina Adam, Stefan Kunert, Hannah S. Heil, Katrin G. Heinze, Pavel Bashtrykov, Ingrid Tessmer*, & Albert Jeltsch*

Supplemental information

Supplemental text

Supplemental text S1: Automated SFM image analysis.

Supplemental figures

Supplemental figure S1: Enzyme preparations and methylation of substrate libraries by DNMT3A and DNMT3A/3L.
Supplemental figure S2: Schematic figure of the SFM DNA substrates.
Supplemental figure S3: Additional SFM data obtained with free DNA.
Supplemental figure S4: Analysis of MM co-methylation.
Supplemental figure S5: Comparison of flanking sequence preferences of DNMT3A extracted from the MW co-methylation data with previously published flanking specificity data of DNMT3A.
Supplemental figure S6: Total DNA length distributions.
Supplemental figure S7: Height and length distributions of single heterotetramers bound to the D6, D9 and D12 substrates.
Supplemental figure S8: Detail of the side-by-side DNMT3A/3L model.
Supplemental figure S9: Statistical analysis of the distribution of co-methylation events by DNMT3A.
Supplemental figure S10: Statistical analysis of the distribution of co-methylation events by DNMT3A/3L.
Supplemental figure S11: Side-by-side DNMT3A/3L model explaining two tetramer complexes binding to the D12 substrate.

Supplemental tables

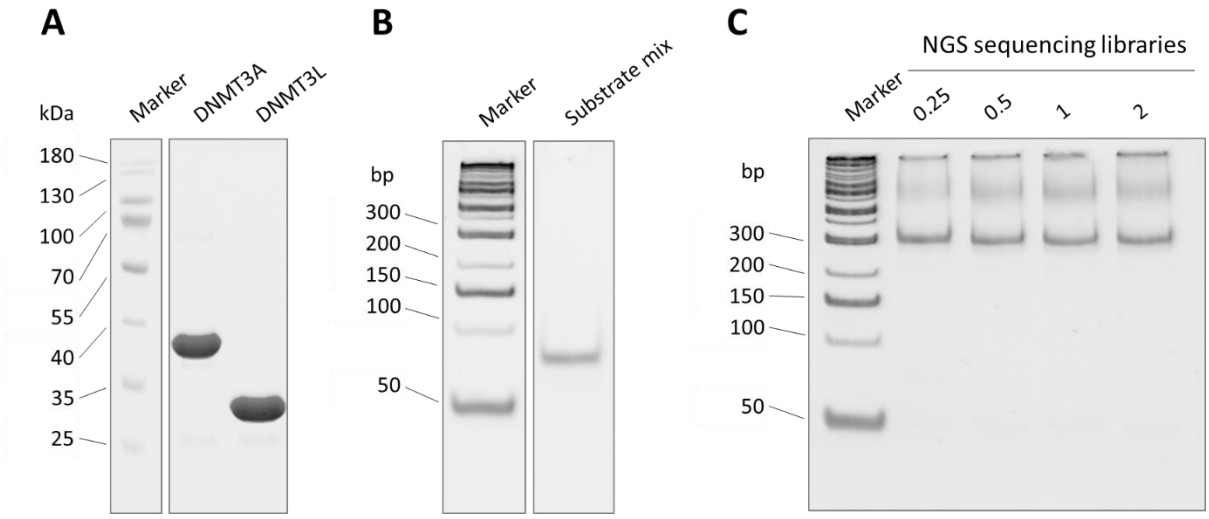
Supplemental table S1. Sequences of the oligonucleotide substrates used in this study.
Supplemental table S2: Total number of NGS reads in the different experiments conducted at different concentrations of DNMT3A and DNMT3A/3L and in the no-enzyme control reaction.
Supplemental table S3: Total number of co-methylation events in the different experiments conducted with DNMT3A and DNMT3A/3L.

Supplemental text

Supplemental text S1: Automated SFM image analysis. For automated analyses of protein positions on the DNA and DNA bend angles in the protein-DNA complexes, we developed a MatLab tool that represents an extension to our previous automated DNA bend angle analysis routine (1). The MatLab script is available at Open Science Foundation (OSF) (https://osf.io/76e9s/). SFM images of protein-DNA samples were processed in the SFM (MFP-3D) software (plane fitted and flattened to 3rd order) and exported in Tiff format for further analysis. The images were cropped to remove the frame in the original Tiff exports. The resulting images were loaded into ImageJ (2) for selection of DNA substrates by a height cut-off filter. Selected DNAs were then skeletonized in FIESTA software (3) by 2 nm segment lines along the center of the DNA strands obtained by a Gaussian fit through sections along the DNA contours. Segmented skeletons were then smoothed by interpolation in our MatLab script. In parallel, protein peaks in the images were located in ImageJ by a height cut-off filter and the application of the object counter function. The obtained protein peak coordinates were then fed into MatLab together with the DNA skeleton lines.

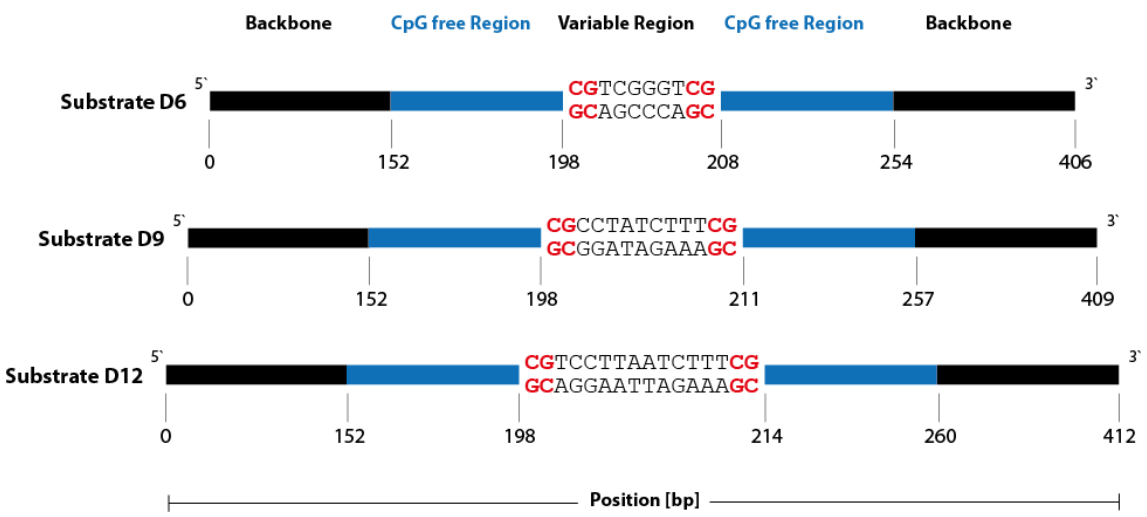
The MatLab script determines DNA-bound proteins by using a selection criterion of maximum allowed distance between peak centers (peak coordinates) and DNA skeleton lines. This cut-off distance is determined as the sum of the peak radii for unbound proteins in the images (measured manually in ImageJ, 10 nm for DNMT3A/3L molecules in our images) and the DNA full maximum half width (FMHW) from the Gaussian section fits in the FIESTA skeletonization procedure (see above). Complete lengths of skeleton lines (DNA contour lengths) were also simultaneously measured and only DNAs with the correct length were included in the further analyses to exclude broken DNAs in which the position of the CpGs would deviate from the 50% DNA length position. Theoretical DNA length was calculated based on the number of base pairs (0.34 nm/bp) and all lengths within two standard deviations from the center of the distribution of measured lengths around this theoretical value were allowed (>122 nm for the ~400 bp fragments in these studies, see Supplemental figure S4 for DNA length distributions). Since the two ends of our DNA are indistinguishable (because they are not labeled in our substrates), the MatLab script measures the distance from the center of protein peaks to the closer DNA end. Protein positions were thus plotted to 50% of DNA length (folded over at the center of the DNA substrates) using Origin software (OriginLab Corporation, Northampton, USA). DNA bend angles are measured by the script selectively at each of the protein positions on DNA. Protein peak positions are output correlated with their respective DNA bend angles. Protein peaks are also numbered in an output image to allow further correlated volume measurements for the individual protein complexes on the DNA.

Supplemental figures and figure legends



Supplemental figure S1: Enzyme preparations and methylation of substrate libraries by DNMT3A and DNMT3A/3L. A) Examples of the purified DNMT3A and DNMT3L proteins. The figure shows a 12% SDS polyacrylamide gel stained with Coomassie BB. All lanes were taken from the same gel. **B)** Examples of the mixtures of all substrates used for the methylation reactions. The image shows a 10% polyacrylamide gel run in TPE and stained with GelRed. Both lanes were taken from the same gel. **C)** Examples of the final sequencing libraries after bisulfite conversion, hairpin ligation, and the two PCR steps adding barcodes and indices used for Illumina NGS sequencing. In this picture, the samples related to the DNMT3A methylation reactions are shown. The image shows a 10% polyacrylamide gel run in TPE and stained with GelRed.

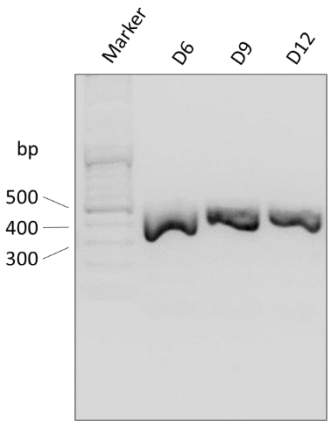
A



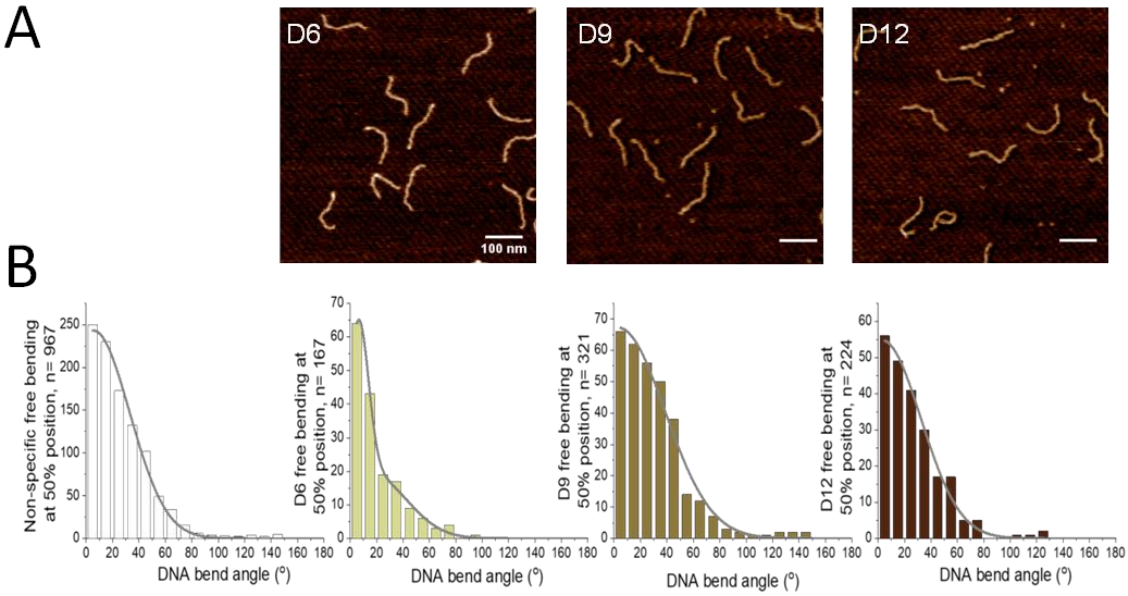
B **Sequence of substrate D12**

```
TTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGAT      50
AACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGGT     100
ACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCC     150
TTAAAAGGAGGCCCATTAGAGTCCTGTCTCTGTTTGATGGAATTGCAACG     200
TCCTTAATCTTTCGGGGTACTTGGTGCTCAAGGAGTTGGGTATTAAAGTG     250
GAAAAGTACTAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGC     300
TCGAGCATGCATCTAGAGGGCCCAATTCGCCCTATAGTGAGTCGTATTAC     350
AATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGT     400
TACCCAACTTAA
```

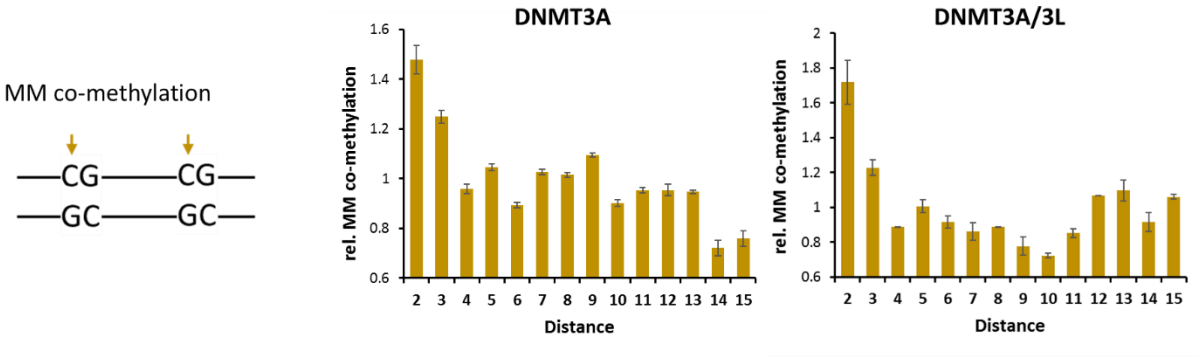
C



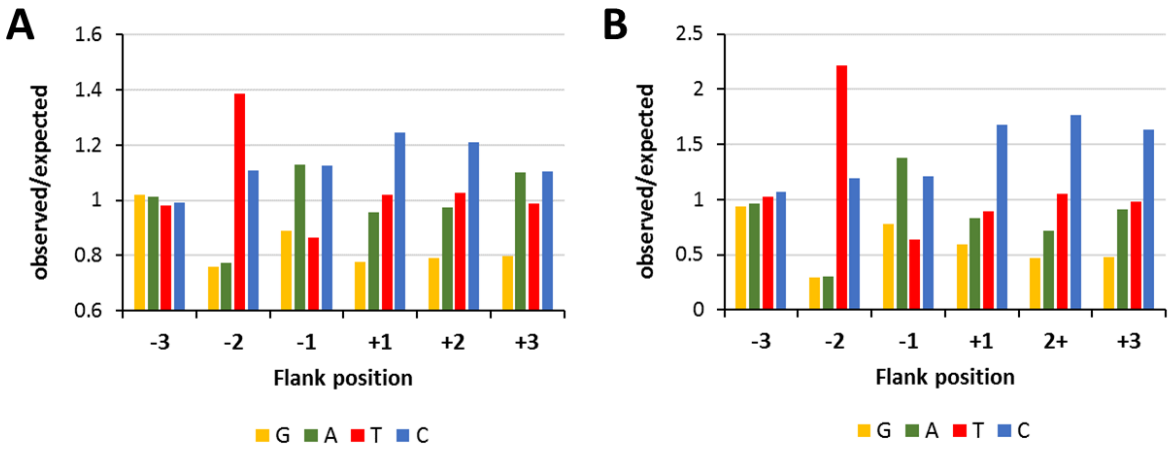
Supplemental figure S2: Schematic figure of the SFM DNA substrates. A) General design and sequences of the central pair of CpG sites (colored in red) in the D6, D9 and D12 substrates. **B)** Sequence of the D12 substrate as an example. The central pair of CpG sites is shaded in grey and the CpG free region is colored in blue. CpG sites are colored in red. **C)** Gel images of the purified SFM substrates. The image shows a 10% polyacrylamide gel run in TPE and stained with GelRed.



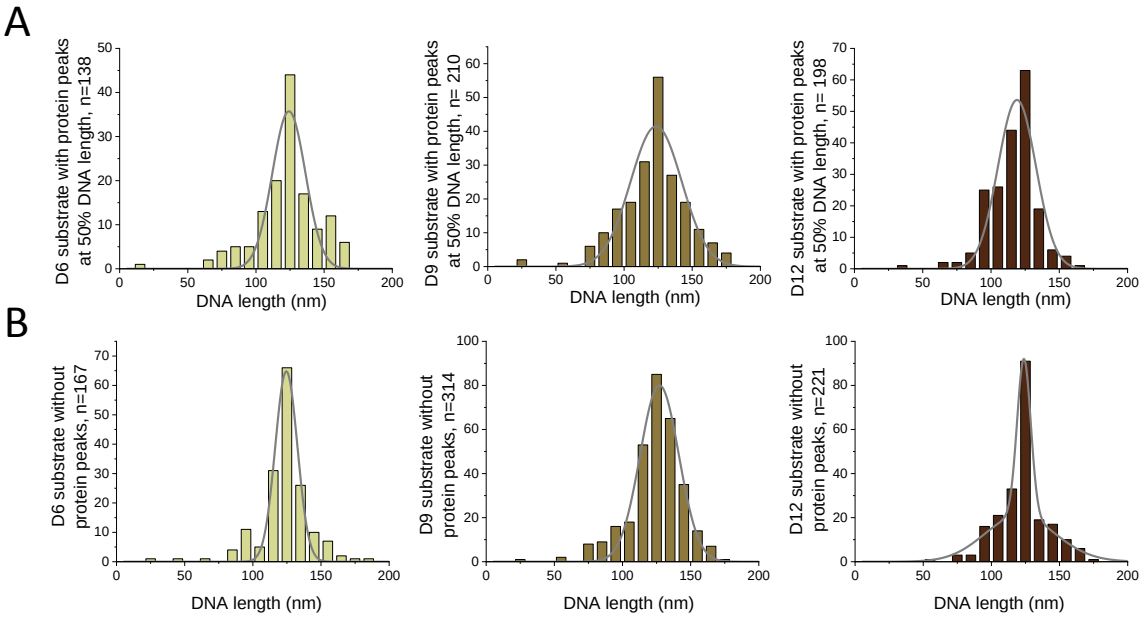
Supplemental figure S3: Additional SFM data obtained with free DNA. A) SFM images of D6, D9 and D12 DNA substrates captured in the absence of protein. **B)** DNA bending angles in free DNA. The three colored histograms show bending of the D6, D9, and D12 substrates at 50% of the DNA length in the absence of protein. The white bars show bending of DNA at non-specific sites (left panel, taken from (1)). All distributions are similar to standard B-form DNA indicating that no intrinsic DNA bending is induced by the CpG pairs.



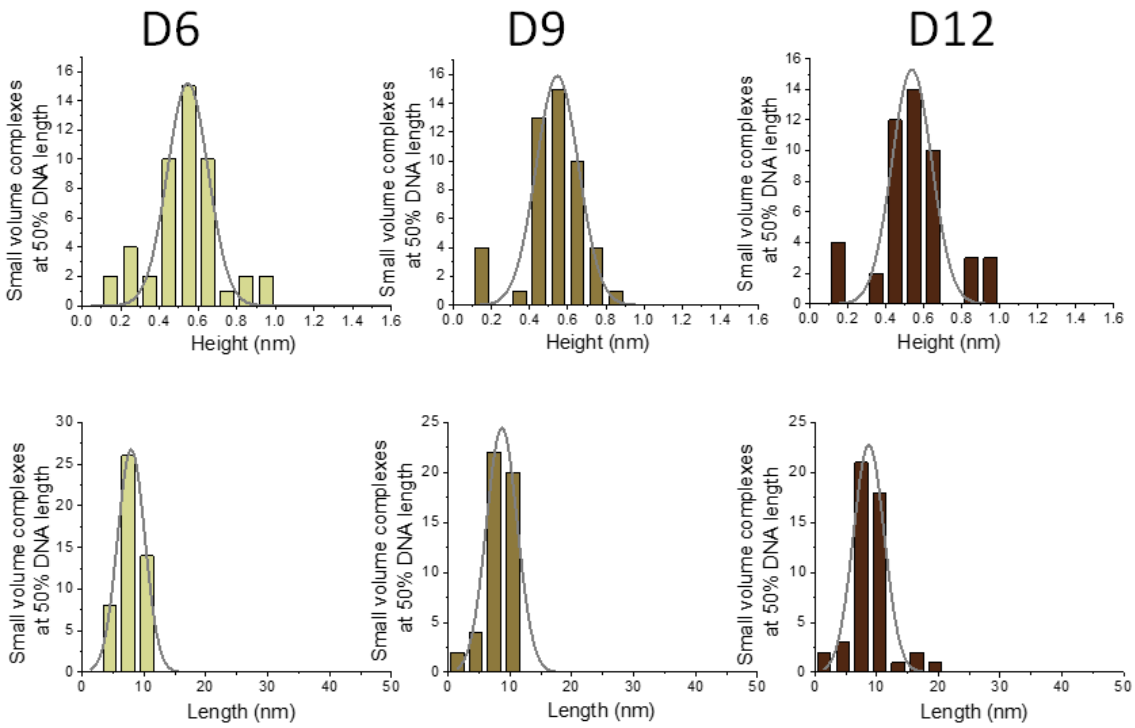
Supplemental figure S4: Analysis of MM co-methylation. Frequencies of MM co-methylation of CpG sites in variable distances. The DNMT3A data show averages of 4 normalized experiments at different enzyme concentrations. The DNMT3A/3L data show averages of 2 normalized experiments at different enzyme concentrations. In each case, the error bars show the SEM.



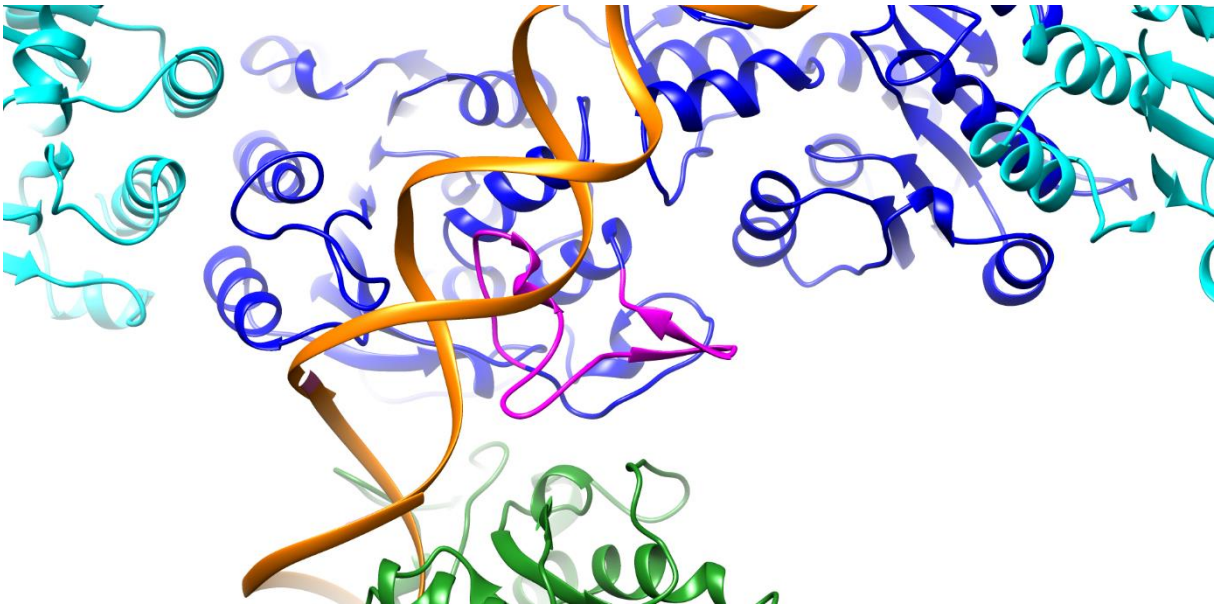
Supplemental figure S5: Comparison of flanking sequence preferences of DNMT3A extracted from the MW co-methylation data with previously published flanking specificity data of DNMT3A (4). A) Averaged observed/expected frequencies of all nucleotides flanking methylated sites in substrates with MW co-methylation in distances of 9-15 base pairs. **B)** Observed/expected frequencies of all bases in the 25% most preferred NNNCGNNN flanks observed by Gao et al. (2020). The graph was recalculated based on the published data.



Supplemental figure S6: Total DNA length distributions. A) DNMT3A/3L complexes bound at the 50% position and **B)** unbound DNA for the three DNA substrates D6, D9, and D12. All values were determined by the MatLab routine.



Supplemental figure S7: Height and length distributions of single DNMT3A/3L heterotetramers bound to the D6, D9 and D12 substrates. In each case, complexes with volumes of ~100 nm^3 corresponding to single heterotetramers were selected and analyzed separately. The average lengths and heights of the different DNA bound complexes are 8.4 ± 0.4 nm and 0.54 ± 0.01 nm, respectively, with standard deviations given for comparison between the different substrates.



Supplemental figure S8: Detail of the side-by-side DNMT3A/3L model. The figure shows an enlargement of the complex shown in Figure 8C with a loop (R831-K855) identified to be involved in the multimerization of DNMT3A on DNA in a previous study (5) highlighted in magenta in one of the blue DNMT3A subunits of complex one, showing that it approaches one of the DNMT3A subunits in complex 2 (green).

MM co-methylation

Distance	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	1													
3	1.5E-02	1												
4	1.1E-03	1.6E-04	1											
5	1.8E-03	1.1E-03	5.7E-03	1										
6	8.0E-04	8.7E-05	7.1E-03	4.2E-04	1									
7	1.6E-03	4.4E-04	4.4E-02	7.0E-02	4.6E-04	1								
8	2.0E-03	7.2E-04	6.9E-02	3.5E-02	3.7E-04	5.3E-01	1							
9	2.1E-03	7.8E-04	9.0E-03	2.6E-01	1.5E-04	2.6E-01	7.9E-02	1						
10	3.6E-04	1.6E-05	1.4E-03	1.0E-04	4.8E-02	2.5E-04	5.4E-04	1.4E-04	1					
11	6.3E-04	4.7E-05	4.5E-03	2.6E-04	5.3E-01	4.0E-04	5.3E-04	1.5E-04	1.2E-01	1				
12	3.1E-04	4.9E-05	8.1E-02	2.0E-03	5.9E-01	1.1E-02	1.8E-02	5.3E-03	6.1E-02	3.6E-01	1			
13	1.1E-03	4.2E-04	5.0E-03	1.3E-03	1.8E-01	1.0E-03	2.2E-04	4.8E-04	1.5E-01	6.1E-01	2.3E-01	1		
14	2.5E-05	9.6E-06	6.5E-04	1.1E-04	2.4E-03	4.1E-04	8.1E-04	3.6E-04	4.3E-03	2.4E-03	8.6E-04	5.7E-03	1	
15	4.2E-05	1.7E-05	1.7E-03	2.4E-04	7.9E-03	8.8E-04	1.6E-03	7.3E-04	2.6E-02	9.4E-03	4.0E-03	1.7E-02	1.5E-01	1

p-value
>0.05
<0.05
<1E-02
<1E-03

Global probability of observing 74 of the 91 pairwise comparisons with $p<0.05$: $p<10^{-99}$

MW co-methylation

Distance	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	1													
3	2.4E-03	1												
4	3.4E-05	1.0E-02	1											
5	2.2E-05	3.3E-02	2.1E-01	1										
6	1.4E-05	7.8E-03	8.1E-01	1.2E-01	1									
7	6.5E-03	6.9E-02	1.7E-03	1.0E-03	9.2E-04	1								
8	8.1E-02	2.7E-02	5.1E-04	1.4E-03	4.8E-04	2.7E-01	1							
9	1.3E-01	5.9E-03	1.8E-03	3.4E-03	2.0E-03	2.6E-02	3.4E-02	1						
10	5.8E-04	4.5E-01	1.3E-03	1.8E-03	5.6E-04	5.1E-02	4.3E-02	1.1E-02	1					
11	6.4E-03	4.1E-02	6.2E-04	2.1E-04	2.3E-04	2.2E-01	4.8E-01	3.0E-02	2.1E-02	1				
12	1.5E-02	3.0E-04	5.9E-05	1.9E-04	7.8E-05	4.2E-03	2.7E-03	6.3E-01	7.4E-04	3.8E-03	1			
13	3.0E-02	8.1E-04	1.6E-05	4.0E-06	5.1E-06	7.2E-04	9.2E-03	4.8E-01	5.9E-05	2.5E-04	9.8E-02	1		
14	4.9E-01	1.5E-03	5.6E-05	1.4E-04	5.4E-05	1.3E-02	4.6E-02	2.1E-01	1.6E-03	1.4E-02	3.1E-02	2.1E-01	1	
15	4.0E-02	1.5E-02	1.2E-04	1.3E-04	5.9E-05	1.0E-01	8.8E-01	4.0E-02	9.6E-03	2.4E-01	3.1E-03	1.7E-03	3.2E-02	1

p-value
>0.05
<0.05
<1E-02
<1E-03

Global probability of observing 71 of the 91 pairwise comparisons with $p<0.05$: $p<10^{-99}$

WM co-methylation

Distance	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	1													
3	8.2E-01	1												
4	1.7E-02	2.0E-02	1											
5	1.7E-01	2.3E-01	5.8E-01	1										
6	2.8E-02	3.6E-02	3.1E-01	6.0E-02	1									
7	1.7E-03	1.8E-03	9.4E-04	1.9E-03	9.1E-03	1								
8	1.3E-02	1.4E-02	3.3E-04	3.2E-04	1.6E-01	3.2E-02	1							
9	2.1E-03	2.1E-03	3.6E-04	4.9E-06	2.2E-03	1.7E-01	2.1E-04	1						
10	2.1E-04	2.0E-04	1.3E-03	3.2E-04	2.6E-04	7.6E-03	1.1E-03	2.6E-02	1					
11	2.2E-04	2.1E-04	9.4E-04	3.6E-05	7.7E-05	1.5E-03	9.7E-05	2.1E-03	2.6E-01	1				
12	3.7E-05	3.4E-05	9.4E-04	8.7E-05	2.1E-05	1.7E-04	1.4E-04	7.7E-04	4.9E-03	1.3E-02	1			
13	1.0E-04	9.3E-05	2.2E-04	3.0E-07	1.1E-05	5.2E-05	4.3E-07	2.7E-06	5.4E-04	2.6E-04	4.0E-02	1		
14	2.7E-05	2.3E-05	9.4E-04	1.1E-06	1.7E-06	6.6E-06	8.3E-07	2.9E-06	4.7E-05	1.8E-05	9.2E-04	1.8E-03	1	
15	6.6E-05	5.8E-05	1.8E-06	3.0E-06	7.0E-06	4.1E-05	3.5E-06	2.0E-05	6.2E-04	4.9E-04	9.5E-02	4.1E-01	1.2E-03	1

p-value
>0.05
<0.05
<1E-02
<1E-03

Global probability of observing 80 of the 91 pairwise comparisons with $p<0.05$: $p<10^{-99}$

Supplemental figure S9: Analysis of the distribution of co-methylation events by DNMT3A. A model assuming co-methylation occurred by independent events would predict equal probabilities of co-methylation in all distances. This model was evaluated by comparing the co-methylation frequencies by DNMT3A in the different distances by pairwise T-Tests using MS Excel using two-tailed and unequal variance (heteroscedastic) settings. P-values < 0.05 were taken as evidence for a significant difference that argues against an equal probability of co-methylation at all distances. For a global analysis, the probability of finding N or more pairwise comparisons with p-values below 0.05 was determined using a binomial distribution. Overall, these data demonstrate that the distribution of co-methylation events is not in agreement with a model assuming co-methylation by independent events.

MM co-methylation

Distance	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	1													
3	1.4E-01	1												
4	8.7E-02	1.5E-02	1											
5	6.9E-02	1.5E-01	2.2E-01	1										
6	6.8E-02	9.5E-02	3.1E-01	4.5E-01	1									
7	3.9E-02	2.3E-01	9.5E-01	4.5E-01	6.2E-01	1								
8	8.6E-02	2.5E-03	3.8E-01	2.3E-01	3.7E-01	8.7E-01	1							
9	3.6E-02	7.9E-02	2.7E-01	7.9E-02	1.2E-01	5.1E-01	2.4E-01	1						
10	6.4E-02	1.9E-02	1.1E-01	6.2E-02	6.4E-02	4.8E-01	6.2E-02	9.0E-01	1					
11	4.9E-02	8.5E-02	6.7E-01	1.4E-01	2.6E-01	9.1E-01	5.4E-01	3.2E-01	2.3E-01	1				
12	8.4E-02	2.7E-01	1.3E-01	3.2E-01	1.4E-01	3.0E-01	1.2E-01	4.9E-02	3.7E-02	6.9E-02	1			
13	6.5E-02	6.1E-01	2.8E-01	5.1E-01	3.6E-01	2.8E-01	2.9E-01	1.3E-01	1.7E-01	2.1E-01	8.9E-01	1		
14	8.9E-02	8.2E-03	5.6E-02	2.6E-01	4.3E-01	8.3E-01	5.3E-01	2.3E-01	8.8E-02	4.8E-01	1.4E-01	3.0E-01	1	
15	1.0E-01	1.4E-01	7.7E-02	3.0E-01	1.2E-01	3.1E-01	5.3E-02	7.0E-02	1.0E-02	8.1E-02	9.5E-01	8.7E-01	8.4E-02	1

p-value
>0.05
<0.05
<1E-02
<1E-03

Global probability of observing 10 of the 91 pairwise comparisons with p<0.05: p=0.0156

MW co-methylation

Distance	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	1													
3	5.9E-03	1												
4	8.3E-03	3.5E-02	1											
5	4.6E-02	2.1E-01	3.1E-01	1										
6	1.7E-02	8.5E-02	3.4E-01	6.8E-01	1									
7	7.2E-03	2.2E-01	4.2E-02	1.7E-01	8.1E-02	1								
8	8.1E-02	5.3E-02	1.2E-02	4.7E-02	1.8E-02	1.0E-01	1							
9	7.8E-02	4.9E-02	1.1E-02	4.7E-02	1.8E-02	9.6E-02	9.7E-01	1						
10	6.7E-03	1.4E-01	3.1E-02	1.4E-01	6.5E-02	4.8E-01	1.1E-01	1.0E-01	1					
11	7.2E-02	1.1E-01	1.9E-02	5.8E-02	2.7E-02	2.0E-01	4.2E-01	4.0E-01	2.4E-01	1				
12	5.8E-02	3.0E-02	1.4E-02	7.3E-03	1.2E-02	4.0E-02	2.4E-02	2.5E-02	3.8E-02	1.8E-02	1			
13	3.2E-02	1.2E-02	1.8E-02	4.4E-02	2.6E-02	2.9E-03	4.9E-02	4.8E-02	6.1E-03	5.2E-02	1.1E-01	1		
14	2.3E-01	1.1E-01	3.1E-02	3.8E-02	3.3E-02	1.6E-01	6.4E-01	6.5E-01	1.7E-01	3.2E-01	1.7E-02	9.6E-02	1	
15	9.6E-02	6.7E-02	1.4E-02	4.5E-02	2.0E-02	1.2E-01	9.7E-01	9.4E-01	1.3E-01	4.5E-01	2.1E-02	5.6E-02	6.3E-01	1

p-value
>0.05
<0.05
<1E-02
<1E-03

Global probability of observing 44 of the 91 pairwise comparisons with p<0.05: $p<10^{-99}$

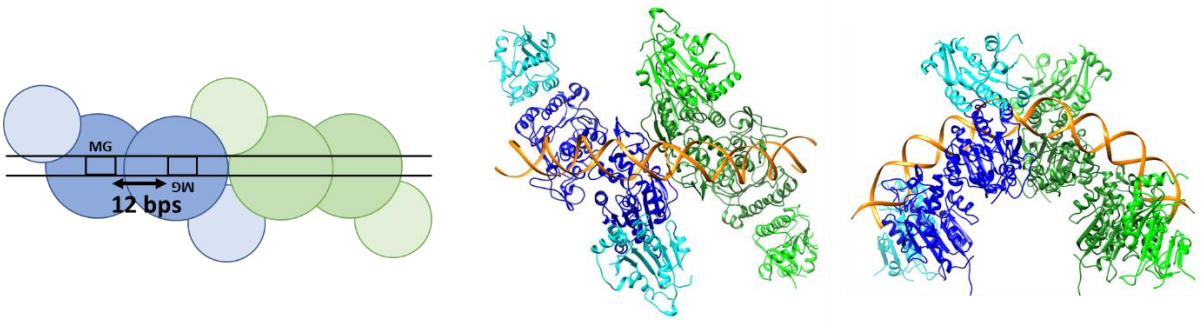
WM co-methylation

Distance	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	1													
3	2.1E-01	1												
4	1.2E-01	1.6E-02	1											
5	2.1E-01	5.6E-01	6.2E-02	1										
6	2.1E-01	7.3E-01	1.7E-02	4.7E-01	1									
7	3.2E-02	1.2E-01	2.0E-01	8.6E-02	1.2E-01	1								
8	4.5E-02	1.6E-02	2.9E-02	7.0E-03	2.4E-02	4.2E-01	1							
9	6.8E-02	1.6E-02	2.3E-02	3.5E-02	1.1E-02	3.6E-01	5.4E-01	1						
10	3.0E-02	6.2E-02	1.0E-01	3.2E-02	7.0E-02	7.5E-01	4.5E-01	3.4E-01	1					
11	3.5E-02	6.1E-02	1.1E-01	3.0E-02	7.0E-02	9.8E-01	2.5E-01	2.4E-01	6.5E-01	1				
12	3.5E-02	2.3E-02	3.7E-02	8.3E-03	3.1E-02	3.0E-01	4.8E-01	7.7E-01	2.7E-01	1.6E-01	1			
13	2.8E-02	1.1E-01	1.6E-01	8.3E-02	1.1E-01	5.7E-01	8.7E-01	6.9E-01	7.2E-01	5.2E-01	6.2E-01	1		
14	3.0E-02	5.4E-02	9.0E-02	2.6E-02	6.2E-02	6.4E-01	5.4E-01	3.8E-01	8.5E-01	5.2E-01	3.1E-01	8.2E-01	1	
15	4.9E-02	1.5E-01	4.1E-01	1.0E-01	1.6E-01	2.8E-01	1.2E-01	1.5E-01	1.6E-01	2.2E-01	8.5E-02	1.8E-01	1.4E-01	1

p-value
>0.05
<0.05
<1E-02
<1E-03

Global probability of observing 25 of the 91 pairwise comparisons with p<0.05: $p=1.9\times10^{-12}$

Supplemental figure S10: Analysis of the distribution of co-methylation events by DNMT3A/3L. A model assuming co-methylation occurred by independent events would predict equal probabilities of co-methylation in all distances. This model was evaluated by comparing the co-methylation frequencies by DNMT3A/3L in the different distances by pairwise T-Tests using MS Excel using two-tailed and unequal variance (heteroscedastic) settings. P-values < 0.05 were taken as evidence for a significant difference that argues against an equal probability of co-methylation at all distances. For a global analysis, the probability of finding N or more pairwise comparisons with p-values below 0.05 was determined using a binomial distribution. Overall, these data demonstrate that the distribution of co-methylation events is not in agreement with a model assuming co-methylation by independent events.



Supplemental figure S11: Side-by-side DNMT3A/3L model explaining two tetramer complexes binding to the D12 substrate. On the D12 substrate one DNMT3A/3L tetramer can bind to both CpG sites and a second tetramer can bind next to it on the DNA. The resulting model is identical to the side-by-side binding model presented in Figure 8C, with the exception that the binding pattern to the CpG sites is different.

Supplemental tables

Supplemental table S1. Sequences of the oligonucleotide substrates used in this study.

Distance	Sequence
2	GAGTGTGACTAGGCTCTCACTGCC*NNNNNNNNNN*CG*NN*CG*NNNNNNNNNNN*GAGAGGAGACCTAGTGAGAAG
3	GAGTGTGACTAGGCTCTCACTGCC*NNNNNNNNNNN*CG*NNN*CG*NNNNNNNNNN*GAGAGGAGACCTAGTGAGAAG
4	GAGTGTGACTAGGCTCTCACTGCC*NNNNNNNNN*CG*NNNN*CG*NNNNNNNNNN*GAGAGGAGACCTAGTGAGAAG
5	GAGTGTGACTAGGCTCTCACTGCC*NNNNNNNNN*CG*NNNNN*CG*NNNNNNNNN*GAGAGGAGACCTAGTGAGAAG
6	GAGTGTGACTAGGCTCTCACTGCC*NNNNNNNN*CG*NNNNNN*CG*NNNNNNNNN*GAGAGGAGACCTAGTGAGAAG
7	GAGTGTGACTAGGCTCTCACTGCC*NNNNNNNN*CG*NNNNNNN*CG*NNNNNNNN*GAGAGGAGACCTAGTGAGAAG
8	GAGTGTGACTAGGCTCTCACTGCC*NNNNNNN*CG*NNNNNNNN*CG*NNNNNNNN*GAGAGGAGACCTAGTGAGAAG
9	GAGTGTGACTAGGCTCTCACTGCC*NNNNNNN*CG*NNNNNNNNN*CG*NNNNNNN*GAGAGGAGACCTAGTGAGAAG
10	GAGTGTGACTAGGCTCTCACTGCC*NNNNNN*CG*NNNNNNNNNN*CG*NNNNNNN*GAGAGGAGACCTAGTGAGAAG
11	GAGTGTGACTAGGCTCTCACTGCC*NNNNNN*CG*NNNNNNNNNNN*CG*NNNNNN*GAGAGGAGACCTAGTGAGAAG
12	GAGTGTGACTAGGCTCTCACTGCC*NNNNN*CG*NNNNNNNNNNNN*CG*NNNNNN*GAGAGGAGACCTAGTGAGAAG
13	GAGTGTGACTAGGCTCTCACTGCC*NNNNN*CG*NNNNNNNNNNNNN*CG*NNNNN*GAGAGGAGACCTAGTGAGAAG
14	GAGTGTGACTAGGCTCTCACTGCC*NNNN*CG*NNNNNNNNNNNNNN*CG*NNNNN*GAGAGGAGACCTAGTGAGAAG
15	GAGTGTGACTAGGCTCTCACTGCC*NNNN*CG*NNNNNNNNNNNNNNN*CG*NNNN*GAGAGGAGACCTAGTGAGAAG

Supplemental table S2: Total number of NGS reads in the different experiments conducted with DNMT3A and DNMT3A/3L and in the no-enzyme control reaction.

Distance	#reads							
	no enzyme	DNMT3A				DNMT3A/3L		sum
		0.25 µM	0.5 µM	1 µM	2 µM	0.125 µM	0.25 µM	
2	16184	19162	16548	14160	16042	4262	3700	90058
3	18174	22172	19514	16776	18722	4818	4218	104394
4	19106	23128	20378	17300	19588	4820	4398	108718
5	28102	34650	30066	26402	29310	7346	6418	162294
6	11340	13682	11938	9844	11667	2674	2670	63815
7	15716	18998	16768	14382	16590	4070	3546	90070
8	13550	16614	14678	12640	14645	3350	3114	78591
9	13122	15884	14108	12070	13432	3334	3084	75034
10	10878	12700	11008	9626	10938	2690	2508	60348
11	14272	16684	14754	13180	14438	3398	3300	80026
12	12892	16340	14158	12294	13972	3406	2942	76004
13	12802	15780	13884	11860	13373	3280	3092	74071
14	17262	20558	17832	15102	16813	4404	4076	96047
15	14192	16784	14762	12742	14349	3636	3426	79891
sum	217592	263136	230396	198378	223879	55488	50492	1239361

Supplemental table S3: Total number of co-methylation events in the different experiments conducted with DNMT3A and DNMT3A/3L.

DNMT3A

Distance	2 µM				1 µM				0.5 µM				0.25 µM			
	#Seq	#MM	#MW	#WM	#Seq	#MM	#MW	#WM	#Seq	#MM	#MW	#WM	#Seq	#MM	#MW	#WM
2	16042	3814	1466	2001	14160	2424	1038	1185	16548	1219	462	546	19162	345	139	146
3	18722	4159	1374	2395	16776	2410	823	1315	19514	1236	410	646	22172	336	89	168
4	19506	3393	960	2244	17300	2123	607	1265	20378	1010	286	557	23128	263	61	146
5	29132	5467	1632	3661	26402	3440	1016	2114	30066	1590	444	932	34650	461	118	232
6	11608	1920	593	1385	9844	1083	322	696	11938	528	143	339	13682	153	40	92
7	16518	2965	1291	1776	14382	1818	804	960	16768	854	409	432	18998	237	110	96
8	14570	2672	1179	1632	12640	1529	706	909	14678	744	359	412	16614	205	122	99
9	13350	2572	1342	1377	12070	1565	896	748	14108	719	467	326	15884	195	165	82
10	10886	1746	780	1039	9626	979	504	544	11008	458	247	216	12700	135	60	57
11	14372	2438	1200	1279	13180	1429	748	684	14754	651	370	300	16684	176	97	68
12	13904	2385	1472	1061	12294	1454	1003	566	14158	636	518	223	16340	171	161	55
13	13290	2148	1451	844	11860	1295	896	425	13884	620	460	210	15780	166	118	47
14	16720	2302	1627	833	15102	1281	969	426	17832	614	541	202	20558	152	170	43
15	14276	2045	1233	881	12742	1202	794	456	14762	584	404	223	16784	135	97	60

DNMT3A/3L

Distance	0.25 µM				0.125 µM			
	#Seq	#MM	#MW	#WM	#Seq	#MM	#MW	#WM
2	3700	387	182	193	4262	194	90	100
3	4218	339	123	184	4818	151	58	84
4	4398	273	76	154	4820	115	30	68
5	6418	467	151	304	7346	190	58	125
6	2670	184	45	116	2674	66	25	46
7	3546	191	106	67	4070	108	57	43
8	3114	194	128	61	3350	82	56	21
9	3084	178	127	51	3334	65	56	21
10	2508	138	77	46	2690	54	39	25
11	3300	209	127	68	3398	75	52	34
12	2942	213	222	52	3406	102	126	18
13	3092	256	183	73	3280	88	88	18
14	4076	258	156	94	4404	108	87	30
15	3426	251	127	119	3636	107	67	39

Supplemental references

1. Bangalore, D.M., Heil, H.S., Mehringer, C.F., Hirsch, L., Hemmen, K., Heinze, K.G. and Tessmer, I. (2020) Automated AFM analysis of DNA bending reveals initial lesion sensing strategies of DNA glycosylases. *Sci Rep*, **10**, 15484.
2. Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*, **9**, 671-675.
3. Ruhnow, F., Zwicker, D. and Diez, S. (2011) Tracking single particles and elongated filaments with nanometer precision. *Biophys J*, **100**, 2820-2828.
4. Gao, L., Emperle, M., Guo, Y., Grimm, S.A., Ren, W., Adam, S., Uryu, H., Zhang, Z.M., Chen, D., Yin, J. *et al.* (2020) Comprehensive structure-function characterization of DNMT3B and DNMT3A reveals distinctive de novo DNA methylation mechanisms. *Nature communications*, **11**, 3355.
5. Rajavelu, A., Jurkowska, R.Z., Fritz, J. and Jeltsch, A. (2012) Function and disruption of DNA methyltransferase 3a cooperative DNA binding and nucleoprotein filament formation. *Nucleic Acids Res*, **40**, 569-580.