DNA methyltransferase DNMT3A forms interaction networks with the CpG site and flanking sequence elements for efficient methylation

Michael Dukatz, Marianna Dittrich, Elias Stahl, Sabrina Adam, Alex de Mendoza, Pavel Bashtrykov, & Albert Jeltsch*

Supplemental information

Supplemental Figures
Supplemental Figure 1. Exemplary images of the purified DNMT3A proteins.
Supplemental Figure 2. +2/+3 flank preferences for different bases at the +1 flank position.
Supplemental Figure 3. Detailed outer flank sequence preference profiles of DNMT3A mutants.
Supplemental Figure 4. Exemplary data of the fits of CpG and non-CpG methylation of WT and selected mutants.

Supplemental Tables
Supplemental Table 1. Compilation of conditions, read counts and methylation levels of the Deep Enzymology experiments conducted with the CpG substrate.
Supplemental Table 2. Correlation of flanking preference profiles of CG methylation.
Supplemental Table 3. Compilation of conditions, read counts and methylation levels of the Deep Enzymology experiments conducted with the outer flank substrates.
Supplemental Table 4. Correlation of outer flank flanking preference profiles in CpG methylation.
Supplemental Table 5. Compilation of conditions, read counts and methylation levels of the Deep Enzymology experiments conducted with the CpN substrate.
Supplemental Table 6. Correlation of flanking preference profiles of CpN methylation.
Supplemental Figure 1. Exemplary images of the purified DNMT3A proteins. Shown are Coomassie BB stained SDS-polyacrylamide gels.
Supplemental Figure 2. +2/+3 flank preferences for different bases at the +1 flank position. Preferences for all 16 dinucleotides at the +2/+3 flank positions were determined for each possible base at the +1 flank site.
Supplemental Figure 3. Detailed outer flank sequence preference profiles of DNMT3A WT and mutants. Shown are -8 to -4 and +4 to +8 profiles of the obs/exp values of bases in the methylated sequences.

**WT**

| Substrate          | Profile |
|--------------------|---------|
| GTACGTAC (PB955)   | ![Graph](#) |
| CTACGGCA (MD221)   | ![Graph](#) |
| GTCCGCGA (MD222)   | ![Graph](#) |
| ATTCGATG (PB954)   | ![Graph](#) |
| TGCCGTTG (PB956)   | ![Graph](#) |
### R836A

| Substrate                          | Profile |
|------------------------------------|---------|
| GTACGTAC (PB955)                   | ![Graph](image1) |
| CTACGGCA (MD221)                   | ![Graph](image2) |
| GTCCGGCA (MD222)                   | ![Graph](image3) |
| ATTCGATG (PB954)                   | ![Graph](image4) |
| TGCCGTTG (PB956)                   | ![Graph](image5) |
### S837A

| Substrate                      | Profile |
|--------------------------------|---------|
| GTACGTAC (PB955)               | ![Graph](image1) |
| CTACGGCA (MD221)               | ![Graph](image2) |
| GTCCGCGA (MD222)               | ![Graph](image3) |
| ATTCGATG (PB954)               | ![Graph](image4) |
| TGCCGTTG (PB956)               | ![Graph](image5) |
### Supplemental Figure 3 (cont.)

#### S881A

| Substrate                  | Profile |
|----------------------------|---------|
| GTACGTAC (PB955)           | ![Graph](image1) |
| CTACGGCA (MD221)           | ![Graph](image2) |
| GTCCGCGA (MD222)           | ![Graph](image3) |
| ATTCGATG (PB954)           | ![Graph](image4) |
| TGCCGTTG (PB956)           | ![Graph](image5) |
### Supplemental Figure 3 (cont.)

#### R882A

| Substrate                | Profile |
|--------------------------|---------|
| GTACGTAC (PB955)         | ![Bar Chart](image1.png) |
| CTACGGCA (MD221)         | ![Bar Chart](image2.png) |
| GTCCGCGA (MD222)         | ![Bar Chart](image3.png) |
| ATTCGATG (PB954)         | ![Bar Chart](image4.png) |
| TGCCGTTG (PB956)         | ![Bar Chart](image5.png) |
Supplemental Figure 3 (cont.)

**R882H**

| Substrate                        | Profile |
|----------------------------------|---------|
| GTACGTAC (PB955)                 | ![Graph](image) |
| CTACGGCA (MD221)                 | ![Graph](image) |
| GTCCGCGA (MD222)                 | ![Graph](image) |
| ATTCGATG (PB954)                 | ![Graph](image) |
| TGCCGTGG (PB956)                 | ![Graph](image) |
Supplemental Figure 3 (cont.)

L883A

| Substrate                  | Profile |
|----------------------------|---------|
| GTACGTAC (PB955)           | ![Graph](image1)
| CTACGGCA (MD221)           | ![Graph](image2)
| GTCCGCGA (MD222)           | ![Graph](image3)
| ATTGATG (PB954)            | ![Graph](image4)
| TGCGTTG (PB956)            | ![Graph](image5)
Supplemental Figure 3 (cont.)

R887A

| Substrate                  | Profile |
|----------------------------|---------|
| GTACGTAC (PB955)           | ![Graph](image1) |
| CTACGGCA (MD221)           | ![Graph](image2) |
| GTCCGCGA (MD222)           | ![Graph](image3) |
| ATTCGATG (PB954)           | ![Graph](image4) |
| TGCCGTTG (PB956)           | ![Graph](image5) |
Supplemental Figure 4. Exemplary data of the fits of CpG and non-CpG methylation of WT and selected mutants.
**Supplemental Table 1.** Compilation of conditions, read counts and methylation levels of the Deep Enzymology experiments conducted with the CpG substrate. All reactions were incubated for 60 min.

| Enzyme       | Repeat | c(E) [µM] | #Met | #Unmet | #Total | Fraction methylated |
|--------------|--------|-----------|------|--------|--------|---------------------|
| DNMT3A WT    | R1     | 2         | 19273| 19836  | 39109  | 0.493               |
|              | R2     | 4         | 20569| 18826  | 39395  | 0.522               |
|              | R2     | 6         | 21699| 15664  | 37363  | 0.581               |
| R836A        | R1     | 1         | 47153| 200732 | 247885 | 0.190               |
|              | R2     | 2         | 54092| 181774 | 235866 | 0.229               |
|              | R1     | 2         | 25913| 40713  | 66626  | 0.389               |
|              | R2     | 4         | 28367| 32248  | 60615  | 0.468               |
| N838A        | R1     | 3         | 39517| 18157  | 57674  | 0.685               |
|              | R2     | 6         | 36620| 10981  | 47601  | 0.769               |
| S837A        | R1     | 2         | 25913| 40713  | 66626  | 0.389               |
|              | R2     | 4         | 28367| 32248  | 60615  | 0.468               |
| S881A        | R1     | 3         | 39517| 18157  | 57674  | 0.685               |
|              | R2     | 6         | 36620| 10981  | 47601  | 0.769               |
| R882A        | R1     | 10        | 1414 | 63856  | 65270  | 0.022               |
|              | R2 1)  | 4         | 16219| 50199  | 66418  | 0.244               |
| R882H        | R1     | 3         | 4815 | 51921  | 56736  | 0.085               |
|              | R2     | 6         | 6833 | 66271  | 73104  | 0.093               |
| L883A        | R1     | 0.5       | 39999| 22472  | 62471  | 0.640               |
|              | R2     | 1         | 33733| 20818  | 54551  | 0.618               |
| R887A        | R1     | 2         | 20876| 28581  | 49457  | 0.422               |
|              | R2     | 4         | 15433| 27659  | 43092  | 0.358               |
| no enzyme    | R1     | 0         | 68   | 34969  | 35037  | 0.002               |

1) MBP-tag cleaved off.
**Supplemental Table 2.** Pearson correlation factors of the -5 to +5 regions of the flanking sequence profiles determined in the independent experimental repeats.

| Enzyme     | R1 vs. R2 |
|------------|-----------|
| WT         | 0.985     |
| R836A      | 0.996     |
| S837A      | 0.987     |
| N838A      | 0.998     |
| S881A      | 0.991     |
| R882A      | 0.920     |
| R882H      | 0.995     |
| L883A      | 0.993     |
| R887A      | 0.993     |

| Enzyme     | Average vs. Gao et al., 2020 |
|------------|-------------------------------|
| WT         | 0.936                         |
### Supplemental Table 3. Compilation of conditions, read counts and methylation levels of the Deep Enzymology experiments conducted with the outer flank substrates.

| Enzyme | Substrate | Sequence context | Repeat | C(E) [µM] | t (min) | #Met | #Unmet | #Total | Methylated (%) |
|--------|-----------|------------------|--------|-----------|---------|------|--------|--------|----------------|
| **DNMT3A** | | | | | | | | | |
| PB954 | ATT CG ATG | R836A | 1 | 0.5 | 60 | 23719 | 183904 | 207623 | 11.42% |
| PB955 | GTA CG TCA | R2 | 1 | 0.5 | 60 | 55777 | 136128 | 187405 | 29.76% |
| PB956 | TGC CG TGT | R836A | 1 | 0.5 | 60 | 7487 | 177460 | 189497 | 4.05% |
| MD221 | CTA CG GCA | R1 | 1 | 0.5 | 60 | 4806 | 18826 | 23632 | 20.14% |
| MD222 | GTC CG GCA | R1 | 1 | 0.5 | 60 | 9697 | 31722 | 42419 | 23.41% |
| PB954 | ATT CG ATG | R836A | 1 | 0.5 | 60 | 7487 | 14572 | 20645 | 70.58% |
| PB955 | GTA CG TCA | R1 | 1 | 0.5 | 60 | 6775 | 44427 | 51202 | 13.23% |
| PB956 | TGC CG TGT | R1 | 3 | 0 | 30 | 15846 | 62891 | 78737 | 20.13% |
| MD221 | CTA CG GCA | R1 | 3 | 0 | 30 | 61744 | 35780 | 97524 | 63.31% |
| MD222 | GTC CG GCA | R1 | 3 | 0 | 30 | 26805 | 41183 | 67888 | 39.48% |
| **R836A** | | | | | | | | | |
| PB954 | ATT CG ATG | R836A | 1 | 0.5 | 60 | 28853 | 16848 | 35701 | 24.7% |
| PB955 | GTA CG TCA | R1 | 2 | 0 | 30 | 15206 | 6368 | 85674 | 89.15% |
| PB956 | TGC CG TGT | R1 | 2 | 0 | 30 | 17031 | 67182 | 88885 | 24.7% |
| MD221 | CTA CG GCA | R1 | 2 | 0 | 30 | 16632 | 13635 | 30284 | 54.92% |
| MD222 | GTC CG GCA | R1 | 2 | 0 | 30 | 60958 | 12274 | 104459 | 68.39% |
| **S837A** | | | | | | | | | |
| PB954 | ATT CG ATG | R836A | 1 | 0.5 | 60 | 40210 | 33259 | 75269 | 55.81% |
| PB955 | GTA CG TCA | R2 | 2 | 0 | 30 | 5206 | 6368 | 85674 | 89.15% |
| PB956 | TGC CG TGT | R2 | 2 | 0 | 30 | 53939 | 8663 | 10202 | 80.07% |
| MD221 | CTA CG GCA | R2 | 2 | 0 | 30 | 61882 | 4453 | 66315 | 93.23% |
| MD222 | GTC CG GCA | R2 | 2 | 0 | 30 | 37184 | 33187 | 40371 | 17.79% |
| **R882A** | | | | | | | | | |
| PB954 | ATT CG ATG | R836A | 1 | 0.5 | 60 | 17031 | 67182 | 88885 | 24.7% |
| PB955 | GTA CG TCA | R1 | 2 | 0 | 30 | 15206 | 6368 | 85674 | 89.15% |
| PB956 | TGC CG TGT | R1 | 2 | 0 | 30 | 17031 | 67182 | 88885 | 24.7% |
| MD221 | CTA CG GCA | R1 | 2 | 0 | 30 | 16632 | 13635 | 30284 | 54.92% |
| MD222 | GTC CG GCA | R1 | 2 | 0 | 30 | 60958 | 12274 | 104459 | 68.39% |
| **S882H** | | | | | | | | | |
| PB954 | ATT CG ATG | R836A | 1 | 0.5 | 60 | 40210 | 33259 | 75269 | 55.81% |
| PB955 | GTA CG TCA | R2 | 2 | 0 | 30 | 5206 | 6368 | 85674 | 89.15% |
| PB956 | TGC CG TGT | R2 | 2 | 0 | 30 | 53939 | 8663 | 10202 | 80.07% |
| MD221 | CTA CG GCA | R2 | 2 | 0 | 30 | 61882 | 4453 | 66315 | 93.23% |
| MD222 | GTC CG GCA | R2 | 2 | 0 | 30 | 37184 | 33187 | 40371 | 17.79% |
| **LB83A** | | | | | | | | | |
| PB954 | ATT CG ATG | R836A | 1 | 0.5 | 60 | 23719 | 183904 | 207623 | 11.42% |
| PB955 | GTA CG TCA | R1 | 2 | 0 | 30 | 15206 | 6368 | 85674 | 89.15% |
| PB956 | TGC CG TGT | R1 | 2 | 0 | 30 | 17031 | 67182 | 88885 | 24.7% |
| MD221 | CTA CG GCA | R1 | 2 | 0 | 30 | 16632 | 13635 | 30284 | 54.92% |
| MD222 | GTC CG GCA | R1 | 2 | 0 | 30 | 60958 | 12274 | 104459 | 68.39% |
| **R887A** | | | | | | | | | |
| PB954 | ATT CG ATG | R836A | 1 | 0.5 | 60 | 23719 | 183904 | 207623 | 11.42% |
| PB955 | GTA CG TCA | R1 | 2 | 0 | 30 | 15206 | 6368 | 85674 | 89.15% |
| PB956 | TGC CG TGT | R1 | 2 | 0 | 30 | 17031 | 67182 | 88885 | 24.7% |
| MD221 | CTA CG GCA | R1 | 2 | 0 | 30 | 16632 | 13635 | 30284 | 54.92% |
| MD222 | GTC CG GCA | R1 | 2 | 0 | 30 | 60958 | 12274 | 104459 | 68.39% |
| **no enzyme** | | | | | | | | | |
| PB954 | ATT CG ATG | R836A | 1 | 0.5 | 60 | 23719 | 183904 | 207623 | 11.42% |
| PB955 | GTA CG TCA | R1 | 2 | 0 | 30 | 15206 | 6368 | 85674 | 89.15% |
| PB956 | TGC CG TGT | R1 | 2 | 0 | 30 | 17031 | 67182 | 88885 | 24.7% |
**Supplemental Table 4.** Correlation of outer flank flanking preference profiles in CpG methylation. Pearson correlation factors of the -8 to +8 regions of the flanking sequence profiles determined in the independent experimental repeats.

| Enzyme  | R1 vs. R2 |
|---------|-----------|
|         | CTACGGCA  | GTCCGCGA | ATTGATG | GTACGTCA | TGGCGTTG |
| WT      | 0.951     | 0.923    | 0.976   | 0.821    | 0.955    |
| R836A   | 0.879     | 0.966    | 0.915   | 0.913    | 0.950    |
| S837A   | 0.909     | 0.971    | 0.931   | 0.943    | 0.945    |
| S881A   | 0.781     | 0.749    | 0.978   | 0.799    | 0.974    |
| R882A   | 0.966     | 0.875    | 0.914   | 0.945    | 0.718    |
| R882H   | 0.919     | 0.891    | 0.976   | 0.853    | 0.910    |
| L883A   | 0.947     | 0.945    | 0.981   | 0.950    | 0.944    |
| R887A   | 0.666     | 0.950    | 0.954   | 0.932    | 0.965    |
# Supplemental Table 5. Compilation of conditions, read counts and methylation levels of the Deep Enzymology experiments conducted with the CpN substrate.

| Enzyme | Context | Repeat | c(E) [μM] | t [min] | #Met | #Unmet | #Total | Methylated (%) |
|--------|---------|--------|-----------|---------|------|--------|--------|----------------|
|        |         |        |           |         |      |        |        |                |
| DNMT3A | WT      |        |           |         |      |        |        |                |
|        | R836A   |        |           |         |      |        |        |                |
|        | N838A   |        |           |         |      |        |        |                |
|        | S837A   |        |           |         |      |        |        |                |
|        | S881A   |        |           |         |      |        |        |                |
|        | R882A   |        |           |         |      |        |        |                |
|        | R882H   |        |           |         |      |        |        |                |
|        | L883A   |        |           |         |      |        |        |                |
|        | R887A   | no enzyme |        |         |      |        |        |                |

Supplemental Table 5. Compilation of conditions, read counts and methylation levels of the Deep Enzymology experiments conducted with the CpN substrate.
**Supplemental Table 6.** Correlation of flanking preference profiles of CpN methylation. Pearson correlation factors of the -4 to +4 regions of the flanking sequence profiles determined in the independent experimental repeats.

| Enzyme | R1 vs. R2 |  |  |  | CpG (CN) vs. CpG (CG) |
|--------|-----------|---|---|---|------------------------|
|        | CpG | CpA | CpT | CpC |                        |
| WT     | 0.952 | 0.985 | 0.962 | 0.700 | 0.951                  |
| R836A  | 0.882 | 0.985 | 0.954 | 0.957 | 0.948                  |
| S837A  | 0.969 | 0.987 | 0.987 | 0.793 | 0.956                  |
| N838A  | 0.993 | 0.819 | 0.670 | low  | 0.998                  |
| S881A  | 0.971 | 0.992 | 0.992 | 0.899 | 0.958                  |
| R882A  | 0.973 | 0.802 | low  | low  | 0.966                  |
| R882H  | 0.984 | 0.941 | 0.648 | low  | 0.949                  |
| L883A  | 0.986 | 0.942 | 0.903 | 0.450 | 0.942                  |
| R887A  | 0.990 | 0.967 | 0.834 | low  | 0.874                  |