3130–3139 *Nucleic Acids Research, 2018, Vol. 46, No. 6* *Published online 6 March 2018*
doi: 10.1093/nar/gky168

The DNMT3A R882H mutant displays altered flanking sequence preferences

Max Emperle[1], Arumugam Rajavelu[1], Stefan Kunert[1], Paola B. Arimondo[2], Richard Reinhardt[3], Renata Z. Jurkowska[1] and Albert Jeltsch[1,*]

[1]Department of Biochemistry, Institute of Biochemistry and Technical Biochemistry, Stuttgart University, Allmandring 31, 70569 Stuttgart, Germany, [2]CNRS ETaC FRE3600, Bât. IBCG. 118, Route de Narbonne, 31062 Toulouse cedex 9, France and [3]Max-Planck-Genomzentrum Köln, Carl-von-Linné-Weg 10, 50829 Köln, Germany

Received December 20, 2017; Revised February 20, 2018; Editorial Decision February 23, 2018; Accepted February 27, 2018

ABSTRACT

The DNMT3A R882H mutation is frequently observed in acute myeloid leukemia (AML). It is located in the subunit and DNA binding interface of DNMT3A and has been reported to cause a reduction in activity and dominant negative effects. We investigated the mechanistic consequences of the R882H mutation on DNMT3A showing a roughly 40% reduction in overall DNA methylation activity. Biochemical assays demonstrated that R882H does not change DNA binding affinity, protein stability or subnuclear distribution of DNMT3A. Strikingly, DNA methylation experiments revealed pronounced changes in the flanking sequence preference of the DNMT3A-R882H mutant. Based on these results, different DNA substrates with selected flanking sequences were designed to be favored or disfavored by R882H. Kinetic analyses showed that the R882H favored substrate was methylated by R882H with 45% increased rate when compared with wildtype DNMT3A, while methylation of the disfavored substrate was reduced 7-fold. Our data expand the model of the potential carcinogenic effect of the R882H mutation by showing CpG site specific activity changes. This result suggests that R882 is involved in the indirect readout of flanking sequence preferences of DNMT3A and it may explain the particular enrichment of the R882H mutation in cancer patients by revealing mutation specific effects.

INTRODUCTION

Cancers acquire somatic DNA mutations as well as epigenetic changes, which ultimately confer a growth advantage to the affected cells (1–5). These two mechanisms of carcinogenesis are directly connected in cases when mutations occur in epigenetic enzymes like DNA and histone methyltransferases and demethylases. Such mutations have recently been observed in many human cancers (6–9). In 2010, first studies have reported particularly high frequencies of missense mutations in the DNMT3A DNA methyltransferase in acute myeloid leukemia (AML) (10,11). They were confirmed by several other follow up reports (reviews: (9,12)) showing that up to 25% of all AML patients contain DNMT3A mutations, and even 37% of all AML patients with a normal karyotype (10). Recurrent DNMT3A mutations have also been identified at lower frequencies in other cancers, including myelodysplastic syndrome (MDS) and T-cell lymphomas (reviews: (9,12)). DNMT3A mutations in AML usually occur heterozygously in only one allele that is accompanied by an intact wildtype allele and they show a strong enrichment of missense mutations (73% of all mutations, Supplementary Figure S1A). In particular arginine 882 in the catalytic C-terminal domain of DNMT3A is often mutated (65% of all missense mutations in DNMT3A, Supplementary Figure S1B), it is most often exchanged by histidine (66% of all R882 mutations). DNMT3A mutations are a negative prognostic marker correlated with a reduced time to treatment failure, shorter duration of complete remission and decreased disease-free survival in cytogenetically normal AML (12,13). Recent studies demonstrated that DNMT3A mutations occur early in pre-leukemic cells and they act as driver mutations of the carcinogenic process (12–15).

The structure of the C-terminal domain of DNMT3A (DNMT3AC) in complex with the C-terminal domain of its stimulator DNMT3L forms a linear heterotetramer consisting of two DNMT3L subunits (at the edges of the tetramer) and two DNMT3A molecules (in the center) (16) (Supplementary Figure S1C). The DNMT3A/3L interface is called FF interface and also used for multimerization

*To whom correspondence should be addressed. Tel: +49 711 685 64390; Fax: +49 711 685 64392; Email: albert.jeltsch@ibc.uni-stuttgart.de
Present addresses:
Arumugam Rajavelu, Rajiv Gandhi Center for Biotechnology (RGCB), Trivandrum 695014, Kerala, India.
Renata Z. Jurkowska, BioMed X Innovation Center, Im Neuenheimer Feld 583, D-69120 Heidelberg, Germany.

of DNMT3A (17,18). The inner DNMT3A/3A interface, called RD interface, also forms the DNA binding site of the DNMT3A complex (16,17). This is a polar interface mediating a reversible oligomerization of DNMT3A FF-interface dimers into a mixture of dimers, tetramers and higher aggregates depending on the buffer conditions and enzyme concentrations (17–19). The R882H mutation is located in the RD interface implying potential effects on multimerization of DNMT3A and its DNA interaction (Supplementary Figure S1C).

Despite intensive work, the molecular and biochemical mechanism of leukemia initiation through the DNMT3A-R882H mutation has not yet been fully clarified (13). The high prevalence of this particular missense mutation, together with the low frequency of non-sense and frameshift mutations indicates that it must have a specific mechanistic effect at the molecular level. Different studies have determined the catalytic activity of the purified R882H mutant and often observed 50–70% residual activity of the mutant enzyme (11,20,21). Later, studies reported that the R882H mutation exhibits a dominant negative effect in cells and *in vitro*. Kim *et al.* showed that the exogenously expressed murine R878H mutant (corresponding to human R882H) interacts with the wildtype enzyme, but it was less efficiently methylating major satellite repeats in mouse ES cells (22). Russler-Germain *et al.* described that mixed enzyme preparations obtained after co-expression of wildtype and R882H in mammalian cells only showed a weak residual activity, suggesting that the complex formation of wildtype and R882H subunits inactivates the wildtype subunits in a dominant negative manner (23). However, it remains unclear, why the heterozygous R882H mutation is so strongly enriched and why it is not accompanied by a larger number of cases showing other forms of loss-of-function mutations like frameshifts or non-sense mutations in both alleles of the *Dnmt3a* gene.

MATERIALS AND METHODS

Site-directed mutagenesis, protein expression and purification

Mutagenesis was performed using the megaprimer site-directed mutagenesis method (24) and confirmed by restriction marker analysis and DNA sequencing. The C-terminal domain of human DNMT3A (Q9Y6K1.4) (amino acids 612–912) and its R882H mutant were cloned into pET28+ (Novagen) creating an N-terminal His_6-tag fusion. His-tagged DNMT3AC and R882H were overexpressed in BL21 (DE3) Codon+ RIL *Escherichia coli* cells (Stratagene) and purified as described (25). Briefly, cells were grown in TB medium until an A_{600nm} of 0.6 was reached and protein expression was induced for 12 h at 20°C by addition of 0.5 mM isopropyl-1-thio-β-D-galactopyranoside. The proteins were purified at high micromolar concentrations using Ni–NTA agarose and stored in 20 mM HEPES pH 7.5, 200 mM KCl, 0.2 mM DTT, 1 mM EDTA and 10% glycerol at –80°C. The concentrations of the proteins were determined by UV spectrophotometry and confirmed by densitometric analysis of Coomassie stained SDS-polyacrylamide gels.

Methyltransferase activity assay

The methyltransferase activity of DNMT3AC and the R882H mutant was measured using 100 nM of a biotinylated 509mer containing 58 CpG sites using the following primers (Bt-AGA TTA GGG AAG GGG GTG TG and AAG ATC CTT TCA AGG CCT CAG) as described (25) (Supplementary Figure S2). To study the flanking sequence effect on the activity of DNMT3AC and R882H, DNA methylation kinetics were measured using 1 μM biotinylated double stranded 30-mer oligonucleotides containing a single CpG site, which were designed to be preferred or disfavored by R882H. In both substrates, the lower DNA strand was biotinylated. The correct annealing of the individual strands was confirmed on polyacrylamide gels stained with gel red. To study the specific methylation of one CpG site in one DNA strand, the oligonucleotide substrates were used in hemimethylated form and with the central CpG site in fully methylated form (26). The fully methylated substrate could only be methylated at non-CpG sites in the flanking regions. The rate of methylation of the upper DNA strand at the target site was calculated by subtracting the activity observed with the fully methylated substrate from the activity observed with the hemimethylated substrate. Consideration of the fully methylated control was necessary to determine methylation rates at single CpG sites, because DNMT3A shows a high activity at some non-CpG sites (27), which otherwise would obscure the results.

With both substrates, DNA methylation was measured by the incorporation of tritiated methyl groups from radioactively labeled AdoMet (Perkin Elmer) into the biotinylated substrate using an avidin–biotin methylation plate assay (28). The methylation reactions were carried out in methylation buffer (20 mM HEPES pH 7.5, 1 mM EDTA, 50 mM KCl, 0.05 mg/ml bovine serum albumin) at 37°C using 2 μM DNMT3AC wildtype or mutant enzyme. The reactions were started by adding 0.76 μM AdoMet. The initial slope of the enzymatic reaction was determined by linear regression. To determine the thermal stability of DNMT3AC and R882H, one sample mix was prepared for each enzyme without DNA and split into two reaction tubes. In one tube the DNA methylation reaction was immediately started by adding DNA and AdoMet. The second tube was kept at 22°C or 37°C for 30 min before starting the methylation reaction.

In vitro bisulfite DNA methylation studies were carried out using the 509mer DNA fragment and primers specific for the converted DNA (GTG GGA TTT GGT TTT GTT TTG TAT TTT and GAA TAT TAC TAC TAC CCT CCT CCT TCT CAA TTT AAC) covering 56 CpG sites as described (25,29) and analyzed with BISMA (30,31).

DNA binding analysis of DNMT3A

DNA binding was investigated by fluorescence depolarization using a Cy5-labeled 29mer oligonucleotide substrate (5′-ACT TGC AAC GGT CCT AAC CGT CAC CTC TT-3′) as described (25,32). The binding reactions were carried out in binding buffer containing 20 mM HEPES pH 7.5, 100 mM KCl, 1 mM EDTA, 2 nM of Cy5-labeled DNA and increasing concentration of protein after incubation at room temperature for 5 min. Each data point was measured

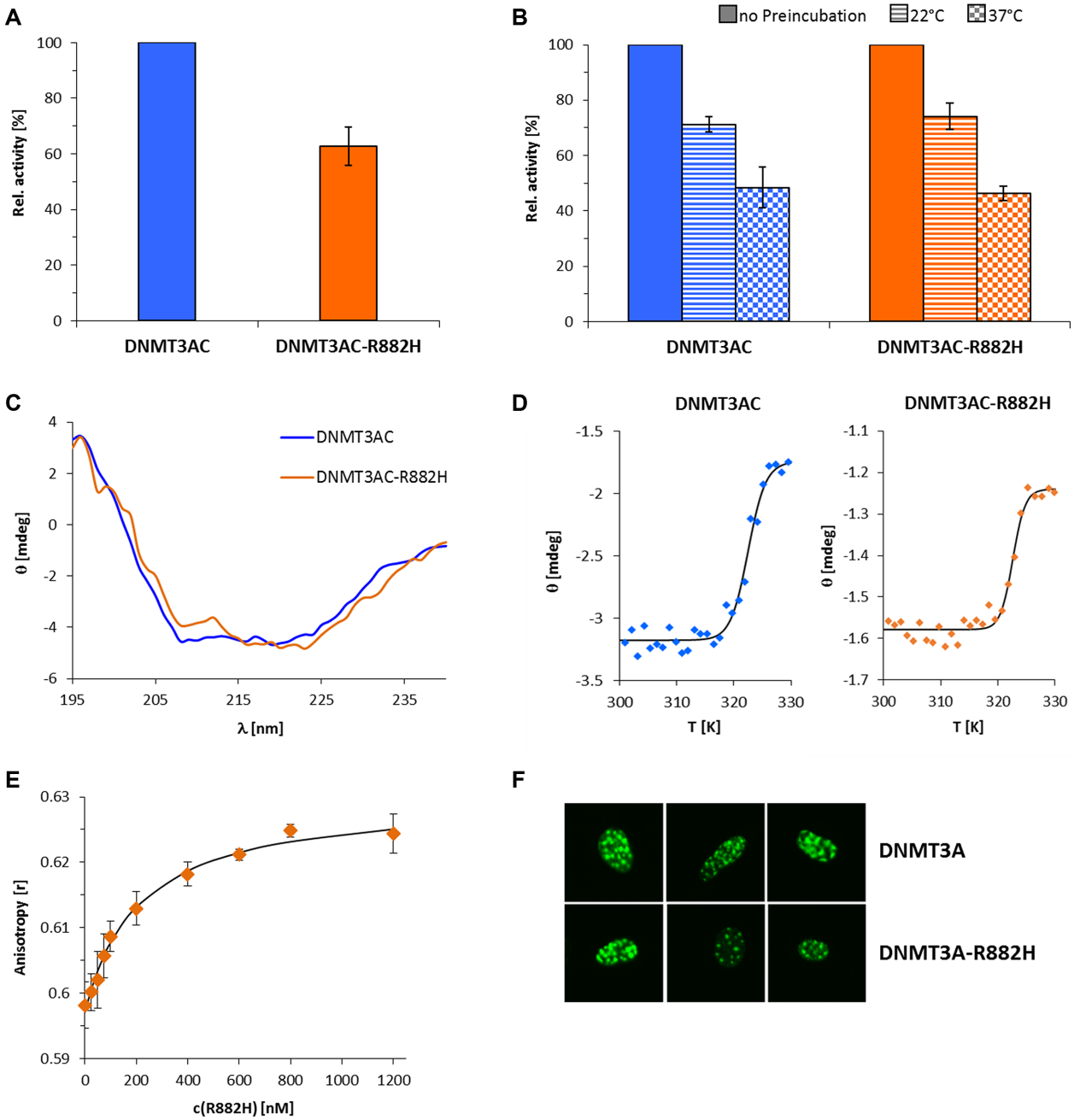


Figure 1. Catalytic activity and characterization of the R882H mutant. (**A**) DNA methylation activity of DNMT3A catalytic domain and the R882H mutant using a 509mer DNA substrate. The bars represent the average of 5 experiments, error bars indicate the standard error of the mean. See also Supplementary Figure S3. (**B**) Protein stability of wildtype DNMT3AC and R882H analyzed by the loss of enzymatic activity during a pre-incubation of the protein for 30 min at different temperatures. Error bars indicate the SEM based on 2 independent experiments. (**C**) CD spectra of wildtype DNMT3AC and R882H. The similar shape of the spectra indicates comparable folding of both proteins. (**D**) Thermal stability of DNMT3AC and R882H analyzed by CD melting at protein concentrations of 22 μM for DNMT3AC and 8 μM for R882H. Quantitative analysis revealed identical melting temperatures of 49.3 ± 0.4°C and 49.5 ± 0.6°C for DNMT3AC and R882H. (**E**) DNA binding titration of the R882H mutant to a 29mer DNA substrate. The data points show the average of two experiments and the standard error. The line shows a fit of the data with a K_D of 209 nM. (**F**) Heterochromatic localization of EYFP fused murine DNMT3A containing the R878H mutation (corresponding to R882H in human DNMT3A).

in triplicate and the average anisotropy values were taken for the analysis. Data were least squares fitted using an equilibrium binding model to determine the dissociation constant (K_D) of DNA binding. Error analysis was based on the separate analysis of experimental repeats. DNA binding of wildtype DNMT3AC and R882H was measured side by side, the results obtained with wildtype DNMT3AC have already been published (25).

Subnuclear localization of DNMT3A

The subnuclear localization of DNMT3A was analyzed by fluorescence microscopy as described using EYFP fusions of murine DNMT3A (18,32). To study the localization of R882H, the R878H mutation was introduced into the murine DNMT3A (which corresponds to R882H in the human protein).

Circular dichroism (CD) spectroscopy

Folding of purified proteins was analyzed by CD spectroscopy. CD measurements were performed using a J-815 circular dichroism spectrophotometer (JASCO Corporation, Tokyo, Japan) in 200 mM KCl, 20 mM HEPES pH 7.5, 10% glycerol using 10 μM purified His-tagged DNMT3AC or R882H. The spectra were collected at 21°C using a 0.1 mm cuvette in a wavelength range between 195 and 240 nm. For each sample, at least 30 scans were collected and averaged. CD melting experiments were performed in dialysis buffer using enzyme concentrations as indicated. The CD signal was measured at a wavelength of 222 nm in a 0.1 mm cuvette in the temperature range from 20 to 70°C applying a temperature increase of 2°C/min. The melting temperature was determined as described (33).

RESULTS

The catalytic domain of DNMT3A (DNMT3AC) is enzymatically active (34) indicating that it is a suitable model system to study the effects of DNMT3A mutations on enzymatic properties and the RD domain interface (16). To search for the enzymatic effects of DNMT3A R882H mutations, we have prepared the R882H mutant in the context of DNMT3AC and purified the proteins after recombinant expression in *E. coli* (Supplementary Figure S1D).

Characterization of the properties of DNMT3AC-R882H

In vitro methylation experiments using identical concentrations of DNMT3AC wildtype and R882H showed that the R882H mutant has a residual activity of 62% with a 509mer DNA substrate containing 58 CpG sites, which was amplified from the CpG island upstream of the human SUHW1 gene (Figure 1A, Supplementary Figures S2 and S3). This result is comparable to residual activities of R882H observed in other studies (11,20,21). Since it represents a relatively mild reduction in activity, we were curious to study if other enzymatic properties of DNMT3A were altered by the mutation. To determine the effect of the mutation on protein stability, we pre-incubated the wildtype and mutant enzymes at 22 or 37°C before measuring the enzymatic activities. This caused a reduction of catalytic activity, which reflects the amount of enzyme denaturation during the pre-incubation interval. However, we did not observe any difference in thermal stability between wildtype and R882H in these experiments (Figure 1B), indicating that the mutation does not destabilize the enzyme folding. To directly investigate a potential effect of the R882H mutation on protein folding, far UV circular dichroism spectra of the purified DNMT3AC and R882H were determined. The CD spectra were superimposable within the level of noise indicating a similar folding of wildtype and mutant proteins (Figure 1C). To further examine the potential influence of the mutation on general protein stability, CD-melting experiments were conducted demonstrating an identical melting temperature of 49.3 ± 0.4°C and 49.5 ± 0.6°C for DNMT3AC and R882H (Figure 1D) which confirms the kinetic data shown in Figure 1B. As the R882H mutation is located in the DNA binding interface of DNMT3A, we also measured the DNA binding constant of R882H by fluorescence depolarization using a 29mer DNA containing two CpG sites. Our data show that the R882H mutant binds the substrate with a K_D of 209±19 nM (Figure 1E), which is only slightly weaker than the wildtype enzyme (K_D = 183±19 nM) measured under identical conditions with protein preparation that were purified in parallel (25). Previously it had been shown, that disruption of the RD interface by an R881A mutation in murine DNMT3A (the corresponding residue in human DNMT3A is R885) resulted in a loss of the normal localization of EYFP fused DNMT3A at DAPI dense heterochromatin in NIH3T3 cells (17,18). To investigate if the R882H mutation leads to a similar deterioration of the RD interface, we have prepared a murine DNMT3A R878H mutant (corresponding to the human R882H) fused to EYFP, expressed the protein in NIH3T3 cells and studied its subnuclear localization by fluorescence microscopy. However, there was no detectable difference in the sub-nuclear localization between the wildtype DNMT3A and R882H EYFP fusion proteins (Figure 1F).

Overall, the lack of changes in DNA binding, enzyme localization and stability indicates that the DNMT3A/3A RD interface is not massively deranged by the R882H mutation. This conclusion is in agreement with the moderate reduction in catalytic activity of the R882H mutant, because disruption of the RD interface by a R881A mutation in murine DNMT3AC (corresponding to human R885) completely inactivated the enzyme (17).

Altered flanking sequence preferences of DNMT3A R882H

Based on the position of the R882H mutation in the DNA binding interface of DNMT3A, we next investigated if this mutation influences the specificity of the DNA interaction of DNMT3A. It has been well documented that DNMT3A shows pronounced differences in the methylation activity of CpG sites depending on their flanking sequence (26,35–37), and CpG sites have been identified which could not be methylated by DNMT3A at all (26). Moreover, the R882A mutation had already been shown to alter the relative preferences of the methylation of three DNA substrates with different flanking sequences (38). We used the purified R882H mutant protein for our experiments, because it is not possible to prepare DNMT3AC with a mixed wildtype/R882H

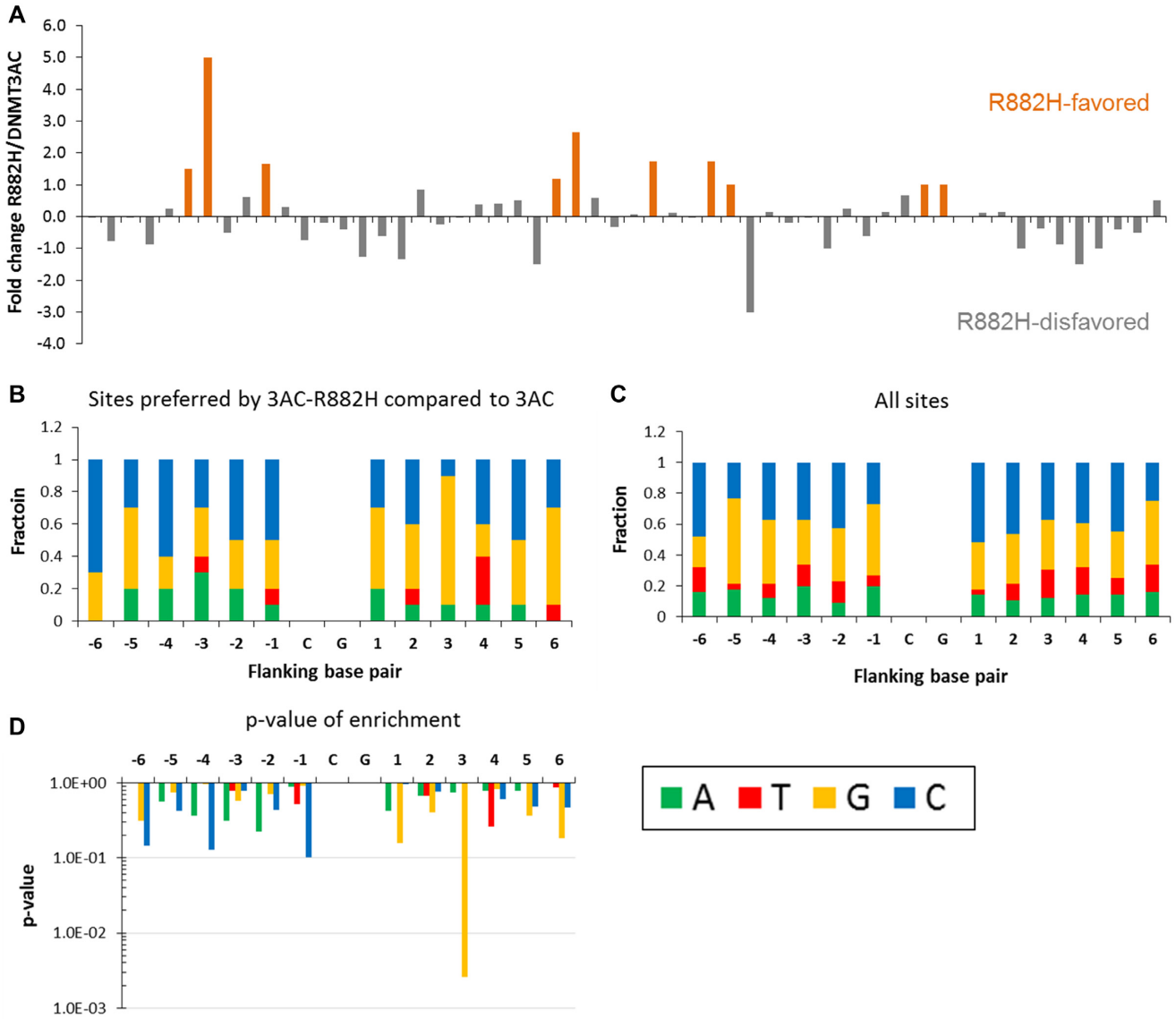


Figure 2. Flanking sequence analysis of R882H. (**A**) Bisulfite analysis of the methylation of 56 CpG sites of the 509mer PCR fragment. The figure shows for each site (1–56), the ratio of the relative methylation levels observed in the wildtype (wt) and R882H samples. In case of zero methylation by the wildtype, ratios were set to 5.0 (for details see Supplementary Table S1). (**B**) Distribution of bases at the flanking sequence positions in the 10 sites preferred by R882H. (**C**) Distribution of bases at the flanking sequence positions of all sites on the DNA substrate. (**D**) *P*-values of enrichment of bases at all flanking sites in the set of preferred sites (for details see Supplementary Table S2).

RD interface. We conducted methylation kinetics with a fragment of the 509mer substrate used in the radioactive methylation assays as DNA substrate, which allows to study the methylation of 56 CpG sites in different flanking contexts, and measured DNA methylation at all CpG sites by wildtype DNMT3AC and R882H by bisulfite conversion and sequencing. Wildtype DNMT3A showed strong differences in the methylation of different sites in agreement with previous studies (25,26). In agreement with the general reduction of R882H activity, most CpG sites exhibited lower methylation levels by R882H. However, a comparison of the relative methylation levels of different CpG sites between wildtype and the R882H revealed a strong change of the flanking sequence preferences of the mutant (Figure 2A). Because of this, some sites showed a much more pronounced reduction of methylation activity than the 38% average reduction in activity reported above. Even more strikingly, some sites were methylated up to five times faster by R882H than by the wildtype enzyme. This was mainly observed at sites which were weakly methylated by wildtype DNMT3AC, but still methylated with good yields by R882H. We extracted the –6 to +6 flanks of the 10 sites most preferred by R882H when compared to wildtype and observed a strong overrepresentation of G at the +3 position (Figure 2B and C). Less pronounced enrichments at other positions were detected after a statistical analysis comparing the distribution of bases at each flanking position in the preferred sites with the overall distribution in all sites (Figure 2D). These results clearly indicate that the flank-

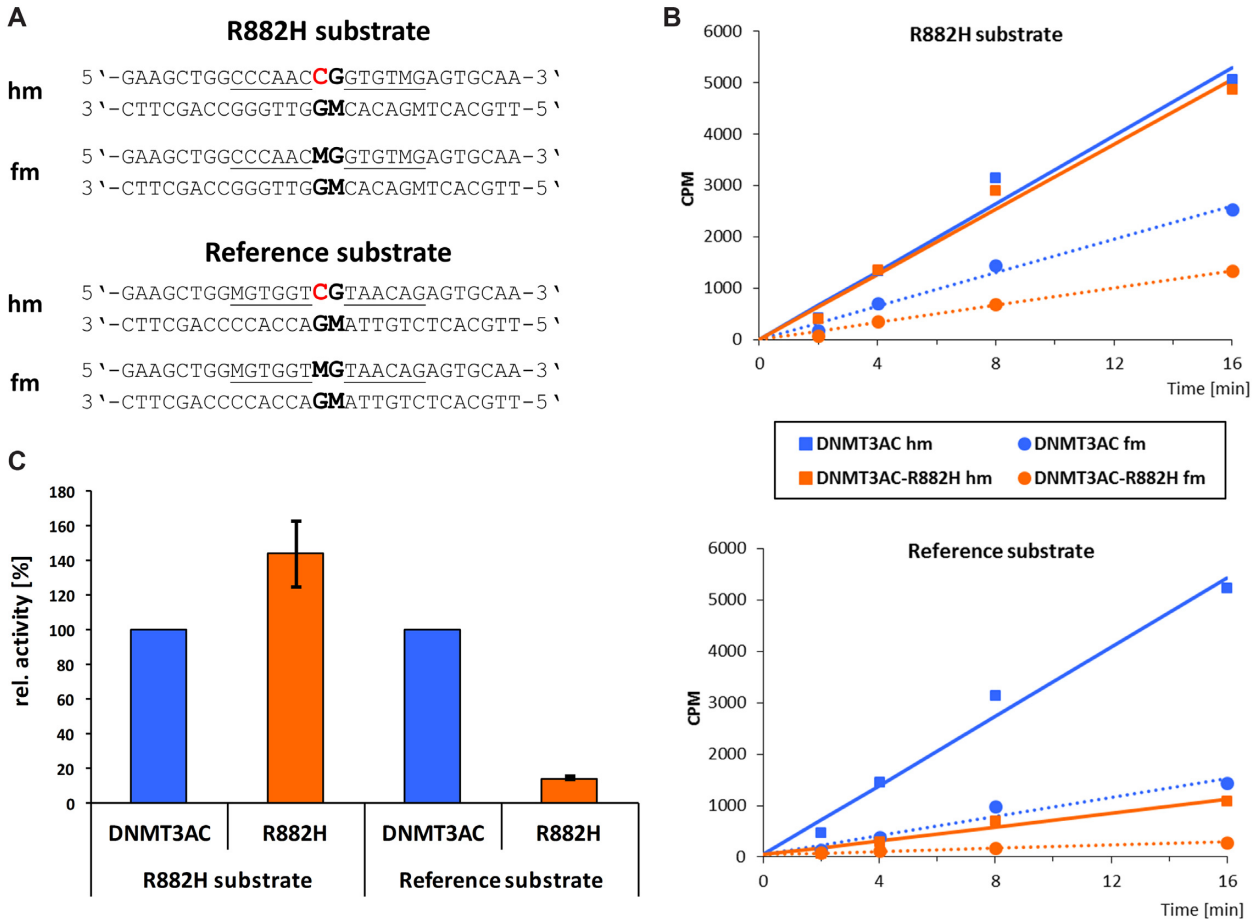


Figure 3. Design of a preferred R882H substrate. (**A**) Based on the sequences of the 10 most preferred methylation sites by R882H, a substrate with R882H optimized flanks was designed together with a reference substrate not containing R882H preferred residues. For each substrate, a hemimethylated version (hm) and a fully methylated (fm) version was used to allow measurement of the specific methylation activity in the upper DNA strand of the central CpG site. 'M' denotes 5-methyldeoxycytosine, the target C is highlighted in red. (**B**) Example kinetic data of the methylation of the hm (squares, continuous lines) and fm (circles, dotted lines) substrates by DNMT3AC (blue) and R882H (orange). (**C**) Average target site methylation rates of the R882H substrate and reference substrate by wildtype DNMT3AC and DNMT3AC-R882H. To determine the specific methylation of the central CpG site in the upper DNA strand, the methylation of the fm substrate is subtracted from the methylation of the hm substrate. Error bars indicate the SEM based on four independent experiments.

ing sequences of R882H differ substantially from wildtype DNMT3AC.

Methylation kinetics with designed oligonucleotide substrates

Based on the analysis described in the last section, we designed model substrates containing one central CpG site with flanks optimized for methylation of the upper DNA strand by R882H ('R882H substrate') and flanks not containing features preferred by R882H ('reference substrate') (Figure 3A). To study only the specific methylation of the central CpG in the upper strand, we used the designed substrates in hemimethylated form. Moreover, it has been repeatedly observed, that DNMT3A also methylates other cytosine residues in oligonucleotide substrates, which are not situated in a CpG context (26,27,39,40). This additional activity could obscure the comparison of the methylation rates of the two substrates. Therefore, both designed substrates were used in an additional control form, where the central CpG site is methylated in both strands (fully methylated). Hence, these control substrates could only be methylated at other cytosine residues. To determine the methylation activity only for the central CpG in the upper strand, the activity observed with the fully methylated substrates was subtracted from the activity observed with the corresponding hemimethylated substrates (Figure 3B).

Methylation experiments of the two substrates in their hemi- and fully methylated forms with wildtype DNMT3AC and R882H revealed that methylation of the R882H substrate by the R882H mutant was 45% faster than by wildtype DNMT3AC (Figure 3C). In contrast, the residual activity of R882H on the reference substrate was <15% when compared to the wildtype enzyme, indicating that the lack of sequence features preferred by R882H has led to a very low activity. Strikingly, the overall difference of the catalytic activities of wildtype and R882H on these two substrates was ~10-fold. These results confirm the conclusions of the flanking sequence analysis of methylated sites by demonstrating that it is possible to design an R882H pre-

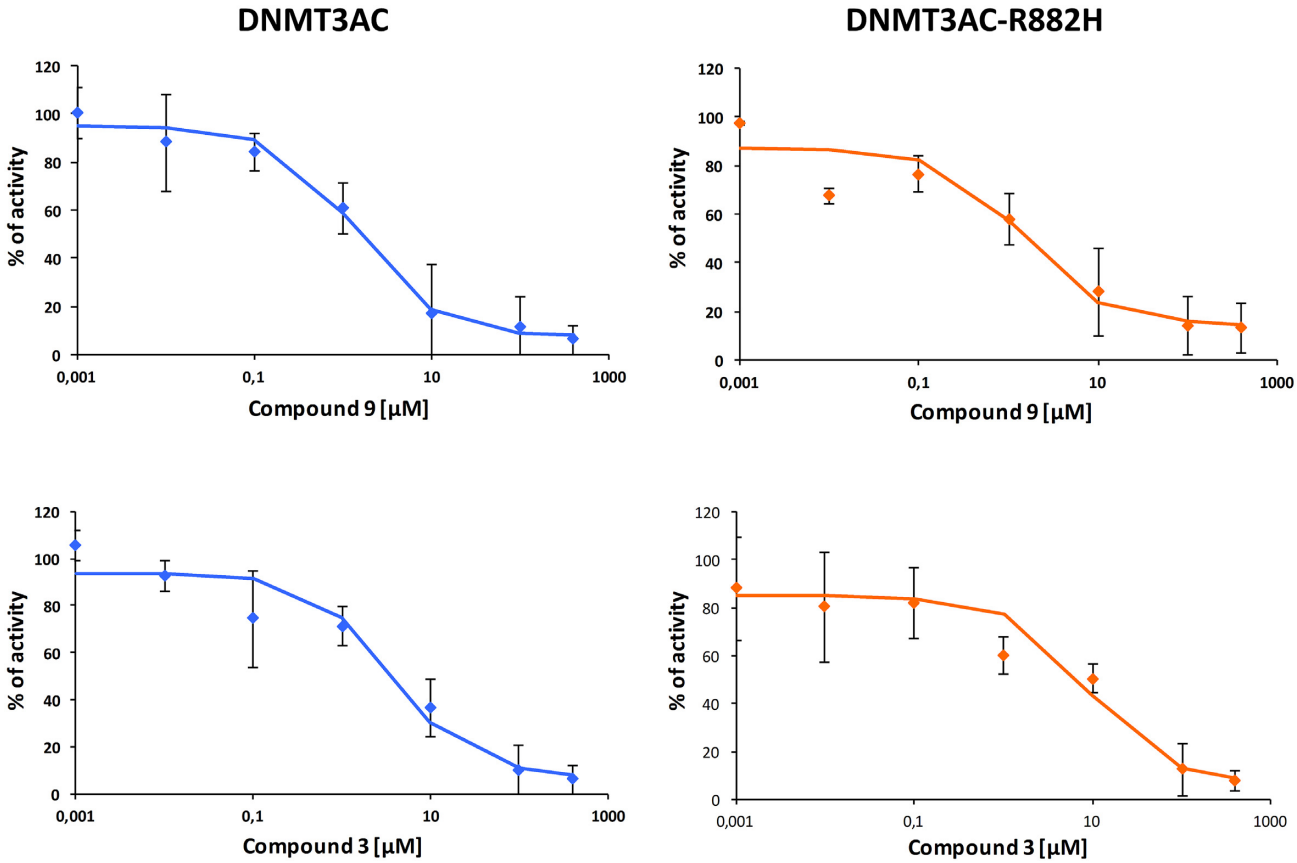
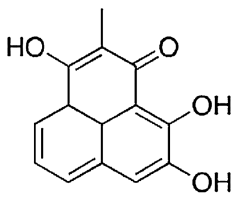
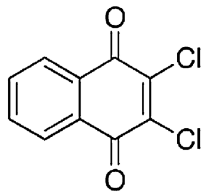


Figure 4. Inhibition of wildtype DNMT3AC and R882H by compound 3 and compound 9. Averages and standard errors from two independent experiments are shown. IC_{50}-values are listed in Table 1.

Table 1. Inhibition of wildtype and R882H mutant by DNMT3A inhibitors (43). IC_{50} values were derived from 2 independent experiments, in which inhibitor concentrations were titrated between 1 nM and 400 µM. Error margins denote the SD of the fit of the IC_{50} values. The titration curves are shown in Figure 4.

Compound structure	Name(s)	IC_{50} wildtype DNMT3A	IC_{50} R882H
	Compound 3	2.9 ± 1.2 µM	8.8 ± 1.0 µM
	Compound 9 Commercial name: DICHLONE	1.3 ± 0.6 µM	1.4 ± 1.2 µM

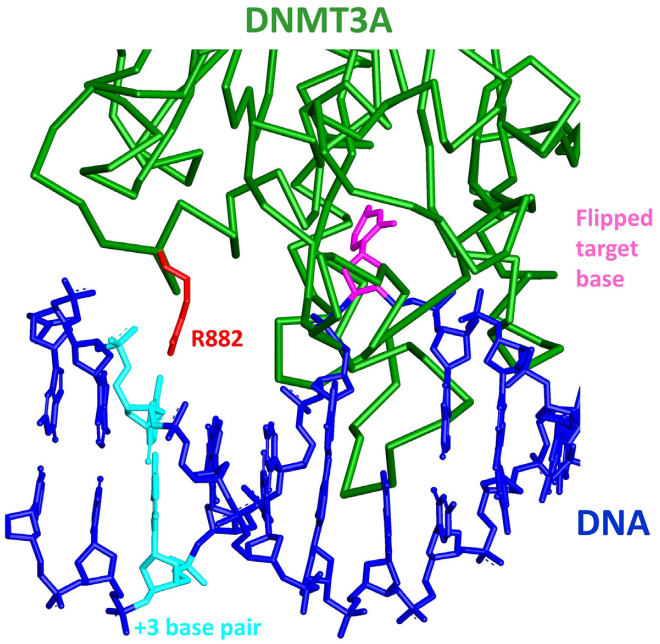


Figure 5. Details of a model of the DNMT3A-DNA complex structure. The Cα-backbone of DNMT3A (pdb: 2QRV) is shown in green, R882 is in red. DNA (shown in blue) was modelled into the DNMT3A structure using the DNA molecule from the M.HhaI-DNA complex structure (pdb: 1MHT) by superposition of the conserved catalytic methyltransferase residues of M.HhaI and DNMT3A. The flipped target cytosine is shown in pink and the +3 base pair in cyan.

ferred and a disfavored substrate on the basis of the flanking sequence preferences.

Inhibition of DNMT3A R882H by small compounds

Our data indicate that the R882H mutation causes a defined hyperactivity of DNMT3A at a subset of CpG sites. Hence, an inhibition of the mutated DNMT3A might be a therapeutic option. Initial studies indeed indicated that DNMT3A mutant cancer patients respond better to treatment with nucleosidic hypomethylating agents (41,42). To investigate the possibility of a direct inhibition of DNMT3A R882H, we tested two previously validated DNMT3A inhibitors (43) for their effects on the activity of the R882H mutant. As shown in Table 1 and Figure 4, one of them (compound 9) showed equal inhibition of both wildtype and mutant DNMT3AC, while the other (compound 3) showed a threefold reduced inhibition of the R882H mutant. Thus, these inhibitors, in particular compound 9, are still active against the mutant DNMT3A.

DISCUSSION

Despite the fact that the R882H mutation in the DNA binding interface of DNMT3A is highly prevalent in AML, its pathogenic mechanism has not been finally clarified at the molecular level (13). Previous data showed a moderately reduced activity of the mutant and a dominant negative effect of the R882H mutation (23). Here, we show that the R882H mutation leads to strong changes in the flanking sequence preferences of DNMT3A most prominently at the +3 position but also at other sites. DNMT3A has been shown previously to exhibit pronounced flanking sequence preferences in the methylation of different CpG sites (26,35–37), which could be due to flanking sequence effects on the DNA binding or conformational changes of the DNA during complex formation (like DNA bending or unwinding), but also effects of the flanking sequence on dynamic steps during catalysis, like target base flipping. Our findings suggests that the R882H exchange directly affects the DNA interaction of DNMT3A. To further explore this point, we prepared a model of the DNMT3A-DNA complex structure by using the DNA from the M.HhaI-DNA complex (44) after superposition of the conserved catalytic methyltransferase residues of DNMT3A and M.HhaI (Figure 5). Strikingly, in this model the R882 side chain is positioned directly adjacent to the +3 base pair in an ideal position to contact the DNA backbone at this site. Thereby, R882 could probe sequence dependent conformational preferences of the DNA substrate in a process reminiscent of the classic indirect readout of DNA sequence by proteins (45,46). This leads to a complex readout of the DNA sequence dependent bending and twisting properties of the DNA, mediated by contacts of the protein to the phosphodiester backbone of the DNA.

The change in flanking sequence preferences of R882H reported here uncovers an additional dimension in the effects of the R882H mutation on the methylation activity of DNMT3A by showing that the influence of the mutation is CpG site specific. In addition to the moderate global reduction of activity, our data show that specific sites are affected much more strongly and strong hypomethylation as well as hypermethylation can occur depending on the target site. Unfortunately, these changes in flanking sequence preferences observed in vitro cannot be directly compared with patient data, because in cells DNMT3A R882H can methylate the upper or lower DNA strand creating a hemimethylated CpG site. Afterwards, DNMT1 will methylate the second strand due to its high activity on hemimethylated CpG sites. Hence, in cellular methylation data sets, one cannot determine which strand was initially methylated by DNMT3A. As the flanking preferences are not palindromic, in cellular data all flanking signatures are mixed resulting in very poor statistics.

A gain-of-function effect of the R882H mutation potentially related to hyperactivity at particular sites is in agreement with its positive correlation with other mutations blocking DNA demethylation, which would enhance these effects. This refers to the co-occurrence of R882H with TET2 inactivating mutations in peripheral T-cell lymphomas (12,47), which were experimentally shown to cooperate in the induction of lymphoid malignancies in mice (47). Along the same line, IDH mutations (which indirectly reduce the activity of TET enzymes) are also positively correlated with DNMT3A R882H in AML and myeloproliferative neoplasm (3,12,48). Finally, some studies indicated that DNMT3A mutant patients responded better to treatment with hypomethlyating agents (41,42), which could be explained by the site specific gain-of-function model of R882H. Future work in hematopoietic stem cells and AML precursor cells, will be needed to identify the rele-

vant methylation sites which are hyper- and hypomethylated by R882H, connect them with corresponding genes, and find out, if inhibition of hypermethylation of CpG sites by DNMT3A-R882H in AML precursor cells may delay the onset of the disease.

DATA AVAILABILITY

All data are included in the manuscript and its supplemental material.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Thanks are due to Drs Christine Maulay-Bailly and Séverine Amand for the initial synthesis of compounds 3 and 9.
Author contributions: A.J., M.E. and A.R. devised the project. M.E. and A.R. conducted the biochemical assays with contributions from S.K. R.R. conducted DNA sequencing for the bisulfite methylation analysis. P.B.A. provided the DNMT3A inhibitors. R.Z.J. and A.J. supervised research. A.J. wrote the manuscript draft. All authors contributed to data interpretation and discussion, read and approved the final manuscript.

FUNDING

Deutsche Forschungsgemeinschaft Priority Program SPP1463 [JE 252/20-1]. The open access publication charge for this paper has been waived by Oxford University Press – *NAR* Editorial Board members are entitled to one free paper per year in recognition of their work on behalf of the journal.
Conflict of interest statement. None declared.

REFERENCES

1. Baylin,S.B. and Jones,P.A. (2011) A decade of exploring the cancer epigenome—biological and translational implications. *Nat. Rev. Cancer*, **11**, 726–734.
2. You,J.S. and Jones,P.A. (2012) Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell*, **22**, 9–20.
3. Alexandrov,L.B., Nik-Zainal,S., Wedge,D.C., Aparicio,S.A., Behjati,S., Biankin,A.V., Bignell,G.R., Bolli,N., Borg,A., Borresen-Dale,A.L. *et al.* (2013) Signatures of mutational processes in human cancer. *Nature*, **500**, 415–421.
4. Kandoth,C., McLellan,M.D., Vandin,F., Ye,K., Niu,B., Lu,C., Xie,M., Zhang,Q., McMichael,J.F., Wyczalkowski,M.A. *et al.* (2013) Mutational landscape and significance across 12 major cancer types. *Nature*, **502**, 333–339.
5. Suva,M.L., Riggi,N. and Bernstein,B.E. (2013) Epigenetic reprogramming in cancer. *Science*, **339**, 1567–1570.
6. Shih,A.H., Abdel-Wahab,O., Patel,J.P. and Levine,R.L. (2012) The role of mutations in epigenetic regulators in myeloid malignancies. *Nat. Rev. Cancer*, **12**, 599–612.
7. Aumann,S. and Abdel-Wahab,O. (2014) Somatic alterations and dysregulation of epigenetic modifiers in cancers. *Biochem. Biophys. Res. Commun.*, **455**, 24–34.
8. Kudithipudi,S. and Jeltsch,A. (2014) Role of somatic cancer mutations in human protein lysine methyltransferases. *Biochim. Biophys. Acta*, **1846**, 366–379.
9. Hamidi,T., Singh,A.K. and Chen,T. (2015) Genetic alterations of DNA methylation machinery in human diseases. *Epigenomics*, **7**, 247–265.
10. Ley,T.J., Ding,L., Walter,M.J., McLellan,M.D., Lamprecht,T., Larson,D.E., Kandoth,C., Payton,J.E., Baty,J., Welch,J. *et al.* (2010) DNMT3A mutations in acute myeloid leukemia. *N. Engl. J. Med.*, **363**, 2424–2433.
11. Yamashita,Y., Yuan,J., Suetake,I., Suzuki,H., Ishikawa,Y., Choi,Y.L., Ueno,T., Soda,M., Hamada,T., Haruta,H. *et al.* (2010) Array-based genomic resequencing of human leukemia. *Oncogene*, **29**, 3723–3731.
12. Yang,L., Rau,R. and Goodell,M.A. (2015) DNMT3A in haematological malignancies. *Nat. Rev. Cancer*, **15**, 152–165.
13. Roy,D.M., Walsh,L.A. and Chan,T.A. (2014) Driver mutations of cancer epigenomes. *Protein Cell*, **5**, 265–296.
14. Xu,J., Wang,Y.Y., Dai,Y.J., Zhang,W., Zhang,W.N., Xiong,S.M., Gu,Z.H., Wang,K.K., Zeng,R., Chen,Z. *et al.* (2014) DNMT3A Arg882 mutation drives chronic myelomonocytic leukemia through disturbing gene expression/DNA methylation in hematopoietic cells. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 2620–2625.
15. Shlush,L.I., Zandi,S., Mitchell,A., Chen,W.C., Brandwein,J.M., Gupta,V., Kennedy,J.A., Schimmer,A.D., Schuh,A.C., Yee,K.W. *et al.* (2014) Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature*, **506**, 328–333.
16. Jia,D., Jurkowska,R.Z., Zhang,X., Jeltsch,A. and Cheng,X. (2007) Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature*, **449**, 248–251.
17. Jurkowska,R.Z., Anspach,N., Urbanke,C., Jia,D., Reinhardt,R., Nellen,W., Cheng,X. and Jeltsch,A. (2008) Formation of nucleoprotein filaments by mammalian DNA methyltransferase Dnmt3a in complex with regulator Dnmt3L. *Nucleic Acids Res.*, **36**, 6656–6663.
18. Jurkowska,R.Z., Rajavelu,A., Anspach,N., Urbanke,C., Jankevicius,G., Ragozin,S., Nellen,W. and Jeltsch,A. (2011) Oligomerization and binding of the Dnmt3a DNA methyltransferase to parallel DNA molecules: heterochromatic localization and role of Dnmt3L. *J. Biol. Chem.*, **286**, 24200–24207.
19. Kareta,M.S., Botello,Z.M., Ennis,J.J., Chou,C. and Chedin,F. (2006) Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. *J. Biol. Chem.*, **281**, 25893–25902.
20. Yan,X.J., Xu,J., Gu,Z.H., Pan,C.M., Lu,G., Shen,Y., Shi,J.Y., Zhu,Y.M., Tang,L., Zhang,X.W. *et al.* (2011) Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat. Genet.*, **43**, 309–315.
21. Holz-Schietinger,C., Matje,D.M. and Reich,N.O. (2012) Mutations in DNA methyltransferase (DNMT3A) observed in acute myeloid leukemia patients disrupt processive methylation. *J. Biol. Chem.*, **287**, 30941–30951.
22. Kim,S.J., Zhao,H., Hardikar,S., Singh,A.K., Goodell,M.A. and Chen,T. (2013) A DNMT3A mutation common in AML exhibits dominant-negative effects in murine ES cells. *Blood*, **122**, 4086–4089.
23. Russler-Germain,D.A., Spencer,D.H., Young,M.A., Lamprecht,T.L., Miller,C.A., Fulton,R., Meyer,M.R., Erdmann-Gilmore,P., Townsend,R.R., Wilson,R.K. *et al.* (2014) The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. *Cancer Cell*, **25**, 442–454.
24. Jeltsch,A. and Lanio,T. (2002) Site-directed mutagenesis by polymerase chain reaction. *Methods Mol. Biol.*, **182**, 85–94.
25. Emperle,M., Rajavelu,A., Reinhardt,R., Jurkowska,R.Z. and Jeltsch,A. (2014) Cooperative DNA binding and protein/DNA fiber formation increases the activity of the Dnmt3a DNA methyltransferase. *J. Biol. Chem.*, **289**, 29602–29613.
26. Jurkowska,R.Z., Siddique,A.N., Jurkowski,T.P. and Jeltsch,A. (2011) Approaches to enzyme and substrate design of the murine Dnmt3a DNA methyltransferase. *Chembiochem*, **12**, 1589–1594.
27. Gowher,H. and Jeltsch,A. (2001) Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. *J. Mol. Biol.*, **309**, 1201–1208.
28. Roth,M. and Jeltsch,A. (2000) Biotin-avidin microplate assay for the quantitative analysis of enzymatic methylation of DNA by DNA methyltransferases. *Biol. Chem.*, **381**, 269–272.

29. Zhang,Y., Rohde,C., Tierling,S., Stamerjohanns,H., Reinhardt,R., Walter,J. and Jeltsch,A. (2009) DNA methylation analysis by bisulfite conversion, cloning, and sequencing of individual clones. *DNA Methylation: Methods Protoc.*, 177–187.
30. Rohde,C., Zhang,Y., Reinhardt,R. and Jeltsch,A. (2010) BISMA - fast and accurate bisulfite sequencing data analysis of individual clones from unique and repetitive sequences. *BMC Bioinformatics*, **11**, 230.
31. Rohde,C., Zhang,Y., Jurkowski,T.P., Stamerjohanns,H., Reinhardt,R. and Jeltsch,A. (2008) Bisulfite sequencing Data Presentation and Compilation (BDPC) web server—a useful tool for DNA methylation analysis. *Nucleic Acids Res.*, **36**, e34.
32. 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.
33. Greenfield,N.J. (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc.*, **1**, 2527–2535.
34. Gowher,H. and Jeltsch,A. (2002) Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases. *J. Biol. Chem.*, **277**, 20409–20414.
35. Lin,I.G., Han,L., Taghva,A., O'Brien,L.E. and Hsieh,C.L. (2002) Murine de novo methyltransferase Dnmt3a demonstrates strand asymmetry and site preference in the methylation of DNA in vitro. *Mol. Cell. Biol.*, **22**, 704–723.
36. Handa,V. and Jeltsch,A. (2005) Profound flanking sequence preference of Dnmt3a and Dnmt3b mammalian DNA methyltransferases shape the human epigenome. *J. Mol. Biol.*, **348**, 1103–1112.
37. Wienholz,B.L., Kareta,M.S., Moarefi,A.H., Gordon,C.A., Ginno,P.A. and Chedin,F. (2010) DNMT3L modulates significant and distinct flanking sequence preference for DNA methylation by DNMT3A and DNMT3B in vivo. *PLoS Genet.*, **6**, e1001106.
38. Gowher,H., Loutchanwoot,P., Vorobjeva,O., Handa,V., Jurkowska,R.Z., Jurkowski,T.P. and Jeltsch,A. (2006) Mutational analysis of the catalytic domain of the murine Dnmt3a DNA-(cytosine C5)-methyltransferase. *J. Mol. Biol.*, **357**, 928–941.
39. Ramsahoye,B.H., Biniszkiewicz,D., Lyko,F., Clark,V., Bird,A.P. and Jaenisch,R. (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 5237–5242.
40. Arand,J., Spieler,D., Karius,T., Branco,M.R., Meilinger,D., Meissner,A., Jenuwein,T., Xu,G., Leonhardt,H., Wolf,V. *et al.* (2012) In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases. *PLoS Genet.*, **8**, e1002750.
41. Traina,F., Visconte,V., Elson,P., Tabarroki,A., Jankowska,A.M., Hasrouni,E., Sugimoto,Y., Szpurka,H., Makishima,H., O'Keefe,C.L. *et al.* (2014) Impact of molecular mutations on treatment response to DNMT inhibitors in myelodysplasia and related neoplasms. *Leukemia*, **28**, 78–87.
42. Metzeler,K.H., Walker,A., Geyer,S., Garzon,R., Klisovic,R.B., Bloomfield,C.D., Blum,W. and Marcucci,G. (2012) DNMT3A mutations and response to the hypomethylating agent decitabine in acute myeloid leukemia. *Leukemia*, **26**, 1106–1107.
43. Ceccaldi,A., Rajavelu,A., Ragozin,S., Senamaud-Beaufort,C., Bashtrykov,P., Testa,N., Dali-Ali,H., Maulay-Bailly,C., Amand,S., Guianvarc'h,D. *et al.* (2013) Identification of novel inhibitors of DNA methylation by screening of a chemical library. *ACS Chem. Biol.*, **8**, 543–548.
44. Klimasauskas,S., Kumar,S., Roberts,R.J. and Cheng,X. (1994) HhaI methyltransferase flips its target base out of the DNA helix. *Cell*, **76**, 357–369.
45. Garvie,C.W. and Wolberger,C. (2001) Recognition of specific DNA sequences. *Mol. Cell*, **8**, 937–946.
46. Rohs,R., Jin,X., West,S.M., Joshi,R., Honig,B. and Mann,R.S. (2010) Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.*, **79**, 233–269.
47. Scourzic,L., Couronne,L., Pedersen,M.T., Della Valle,V., Diop,M., Mylonas,E., Calvo,J., Mouly,E., Lopez,C.K., Martin,N. *et al.* (2016) DNMT3A(R882H) mutant and Tet2 inactivation cooperate in the deregulation of DNA methylation control to induce lymphoid malignancies in mice. *Leukemia*, **30**, 1388–1398.
48. Shen,Y., Zhu,Y.M., Fan,X., Shi,J.Y., Wang,Q.R., Yan,X.J., Gu,Z.H., Wang,Y.Y., Chen,B., Jiang,C.L. *et al.* (2011) Gene mutation patterns and their prognostic impact in a cohort of 1185 patients with acute myeloid leukemia. *Blood*, **118**, 5593–5603.