Oligomerization of DNMT3A controls the mechanism of de novo DNA methylation

Celeste Holz-Schietinger¹, Douglas M. Matje², Madeleine Flexer Harrison² and Norbert O. Reich¹,²

¹Interdepartmental Program in Biomolecular Science & Engineering, and ²Department of Chemistry & Biochemistry University of California, Santa Barbara, California 93106-9510

Running title: Oligomerization of DNMT3A controls processivity

Correspondence should be addressed to: Norbert O. Reich, Email: reich@chem.ucsb.edu. Phone: (805) 893-8368. Fax (805) 893-4120

Background: de novo DNA methyl transferase 3A is essential for gene regulation, however its regulation is poorly understood.

Results: DNMT3A dimers are fully functional but eliminate methylation on clustered CpG sites (processive catalysis) that is observed with the wild type homotetramers.

Conclusion: Protein binding by DNMT3A or DNMT3L at the tetramer interface controls the catalytic properties of DNMT3A.

Significance: Provides a structural mechanism for the regulation of methylation patterning.

Background: de novo DNA methyltransferase 3A is essential for gene regulation, however its regulation is poorly understood.

Results: DNMT3A dimers are fully functional but eliminate methylation on clustered CpG sites (processive catalysis) that is observed with the wild type homotetramers.

Conclusion: Protein binding by DNMT3A or DNMT3L at the tetramer interface controls the catalytic properties of DNMT3A.

Significance: Provides a structural mechanism for the regulation of methylation patterning.

DNMT3A is one of two human de novo DNA methyltransferases essential for regulating gene expression through cellular development and differentiation. Here we describe the consequences of single amino acid mutations, including those implicated in the development of acute myeloid leukemia (AML) and myelodysplastic syndromes (MDS), at the DNMT3A:DNMT3A homotetramer and DNMT3A:DNMT3L heterotetramer interfaces. A model for the DNMT3A homotetramer was developed via computational interface scanning and tested using light scattering and electrophoretic mobility shift assays. Distinct oligomeric states were functionally characterized using fluorescence anisotropy and steady state kinetics. Replacement of residues that result in DNMT3A dimers, including those identified in AML patients, show minor changes in methylation activity but lose the capacity for processive catalysis on multisite DNA substrates, unlike the highly processive wild type enzyme. Our results are consistent with the bimodal distribution of DNA methylation in vivo and the loss of clustered methylation in AML patients. Tetramerization with the known interacting partner DNMT3L rescues processive catalysis, demonstrating that protein binding at the DNMT3A tetramer interface can modulate methylation patterning. Our results provide a structural mechanism for the regulation of DNMT3A activity and epigenetic imprinting.

DNA methylation is a critical epigenetic modification controlling chromatin stability, genomic imprinting, cellular differentiation, transcriptional regulation and memory formation in mammals (1-3). Initial (de novo) 5-methylcytosine (5mC) patterning is established by DNMT3A and the closely related DNMT3B (4, 5) and aberrant 5mC patterning is highly correlated with oncogenesis (6). The mechanism by which these enzymes are directed to methylate specific CpG sites in the genome remains unclear, but methylation patterning is likely mediated by the formation of higher order complexes. One known example is the DNMT-like protein, DNMT3L, a processivity factor (7, 8), that forms heterotetramers with DNMT3A to modulate the catalytic mechanism of DNMT3A (9-11). During early development, de novo methylation patterning is dependent on the presence of DNMT3L (12). Further characterization of the oligomeric states of DNMT3A is necessary for understanding how DNA methylation is established in mammalian genomes.

The recent crystal structure of the catalytic domain of DNMT3A with the catalytically inactive DNMT3L shows a DNMT3L:DNMT3A:DNMT3A:DNMT3L heterotetramer complex (9). The complex shows specific contacts at the DNMT3A homodimer interface (dimer interface) and dimerization brings
two enzyme active sites separated by approximately one helical turn, in B-form DNA (9). Still unclear is the physical basis of the DNMT3A homotetramer that is observed in gel filtration experiments in the absence of DNMT3L. Several lines of evidence suggest that the DNMT3A homotetramer is formed by DNMT3A monomers assuming the position occupied by DNMT3L in the heterotetramer crystal structure (tetramer interface) (13). While most characterized C5 methyltransferases function as monomers (14), the disruption of tetramerization and loss of activity with the Phe728Ala DNMT3A mutant led to the proposal that tetramerization is essential for methylation (9). However, homotetramerization does not provide additional DNA binding surfaces, nor does it bring more active sites into contact with the substrate, and it is not clear why tetramerization should be essential for catalysis.

Recent studies have demonstrated that more aggressive forms of acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS) are correlated with mutations in DNMT3A (15-17), including three at the tetramer interface (9). Patients with DNMT3A mutations have altered methylation patterns without a decrease in global DNA methylation levels (16). We hypothesized that DNMT3A is functional in the homodimeric form and mutations at the tetramer interface disrupt the ability of regulatory proteins to bind to the DNMT3A homodimer, resulting in aberrant methylation patterning. We sought to evaluate the functional significance of DNMT3A oligomerization at the tetramer interface and understand how the regulatory protein DNMT3L modulates DNMT3A function.

A homotetramer DNMT3A model was developed for the design of side chain mutations that disrupt tetramerization while minimizing changes in catalysis. Mutations at five individual residues are shown by light scattering and electrophoretic mobility assays to disrupt homotetramerization. Four mutants are dimeric on DNA and have activity comparable to the wild type DNMT3A and M.HhaI were purified from BioRex and nickel affinity columns, DNMT3L was purified from nickel affinity and chitin columns, all purified to > 95%. EcoRV was prepared as described in Hiller et al. (20). Three different wild type DNMT3A preparations were compared and showed less than 5% difference in $k_{cat}$. The catalytic domain of DNMT3A has similar kinetic parameters as the full-length enzyme, including $k_{cat}$, $K_m$DNA, $K_m$AdoMet, processivity and DNMT3L activation (8, 13), and was used for the DNMT3A/DNMT3L cocrystal structure (9). Both the homo- and DNMT3L hetero- oligomerization interfaces are located on the catalytic domain.

**EXPERIMENTAL PROCEDURES**

**Protein cloning and purification**— The catalytic domain of DNMT3A and full length DNMT3L were purified as stated in Holz-Schietinger et al. 2010 (8), M.HhaI was purified as stated in Matje et al. 2011 (18). Plasmids used for protein expression and site-directed mutagenesis include codon-optimized pET28a-hDNMT3A_CD (Δ1-611) (19), pET28a-M.HhaI and pTYB1-3L for hDNMT3L (13). Briefly, protein expression occurred in E. coli Rosetta2 (DE3) for hDNMT3L, M.HhaI and wild type DNMT3A catalytic domain and their mutants were expressed in T7Iq cells, induction occurred at 0.6 OD600 at room temperature for 5 hours with 1 mM IPTG. DNMT3A and M.HhaI were purified from BioRex and nickel affinity columns, DNMT3L was purified from nickel affinity and chitin columns, all purified to > 95%. EcoRV was prepared as described in Hiller et al. (20). Three different wild type DNMT3A preparations were compared and showed less than 5% difference in $k_{cat}$. The catalytic domain of DNMT3A has similar kinetic parameters as the full-length enzyme, including $k_{cat}$, $K_m$DNA, $K_m$AdoMet, processivity and DNMT3L activation (8, 13), and was used for the DNMT3A/DNMT3L cocrystal structure (9). Both the homo- and DNMT3L hetero- oligomerization interfaces are located on the catalytic domain.

**Mutagenesis design and computational modeling**— DNMT3A chain A of the PDB file 2QRV (9) was kept rigid while the backbone of chain D was aligned to the backbone of DNMT3L chain B using PyMol (Delano Scientific LLC). Chain D was then dragged into a random orientation ~10Å away from the tetramer interface of chain A. The complex was submitted to the RosettaDock server (21) to generate the initial tetramer interface model. After visual inspection, the best scoring models were re-submitted until all iterations repeatedly converged on a common model. This model was submitted to the Rosetta computational alanine scanning server (22, 23) to evaluate the contribution of individual residues to the tetramer interface. The side chains of residues
R771, E733, and R720 were manually optimized by choosing backbone dependent rotamers conformations that gave the highest scores using the computational alanine scanning protocol.

**DNA sequences**—The DNA used as substrates include duplex poly-dIdC (~1000 bp) (Sigma-Aldrich), plasmid pCpG 

TGGATATCTAGGGCGCTATGATATC

T-3’) duplex purchased from Integrated DNA Technologies HPLC purified. The recognition site for DNMT3A is underlined.

**Size-exclusion chromatography coupled to multi-angle light scattering**— Protein samples were passed through BioRad P6 resin pre-equilibrated with light scattering buffer (50 mM KH2PO4/K2HPO4 pH 7.8, with 200 mM NaCl, 1 mM DTT, 1 mM EDTA and 0.2% azide), then 100 µl of this sample (~30 µM of monomeric enzyme) was injected into an analytical size exclusion silica gel chromatography column 300 Å pore size (Wyatt Technologies) with a flow rate of 0.5 ml/min. The sample was run through a Waters 996 diode array measuring A260 and A280, optilab T-rEx differential refractometer (Wyatt), and triple-angle light-scattering detector miniDaWN TREOS (Wyatt). Molecular weight was calculated using ASTRA 5.3.4 software using a dn/dc value of 0.185. Standard error reported from at least two separate samples.

**Native gel mobility DNA shift assays**—DNMT3A wild type, mutants (150 nM or varied as indicated), or size control standards (M.HhaI and EcoRV, 150 nM) were incubated at room temperature for 15 minutes with 200 nM duplex 5’ 6-FAM GCbox30 that contains a single-site central CG site (DNMT3A) along with two sites for EcoRV and one site for M.HhaI binding. For DNMT3L super-shifting, DNMT3L was preincubated for 30 minutes with DNMT3A before the addition of DNA. Binding occurred in reaction buffer with 5 µM S-Adenosyl-L-homo cysteine (AdoHcy). Samples were run on native 4.5% polyacrylamide gels in 0.5 x TBE pH 7.8 at 250V for 35 minutes, visualized for fluorescein using a Typhoon scanner and data analyzed using ImageJ (25).

**Methylation assays**—DNMT3A methylation assays measured the amount of tritiated methyl groups transferred from cofactor S-Adenosyl- methionine (AdoMet) to the DNA by the enzyme. Reactions were carried out at 37°C in reaction buffer (50 mM KH2PO4/K2HPO4, 1 mM EDTA, 1 mM DTT, 0.2 mg/mL BSA and 20 mM NaCl at pH 7.8). Enzyme was at 150 nM total, which is 27 nM active enzyme (previously determined in Purdy et al. 2010) (19). DNA was the multiple site substrate poly-dIdC, used at saturation for kcat, KmAdoMet and DNMT3L activation. AdoMet was used at saturation for kcat, KmDNA, DNMT3L activation and processivity. Processivity assays and DNMT3L activation assays were carried out as described in Holz-Schietinger et al. 2010 (8). Briefly, processivity assays involve three separate reactions; positive control: 20 µM bp substrate, experiment: substrate followed at 20 minutes by chase DNA (pCpG) at 20-fold excess over substrate concentration, and negative control: chase and substrate at start of reaction. Varying DNMT3L 0.25- to 4-fold the concentration of DNMT3A was used to test DNMT3L activation. DNMT3A and DNMT3L were pre-incubated in reaction buffer with AdoMet for one hour at 37°C before starting the reaction with DNA. A 1:1 ratio with a one hour pre-incubation (with AdoMet) was used for all DNMT3L assays. All reactions were quenched by addition of 0.1 % SDS. Samples were spotted onto Whatman DE81 filters then washed, dried, and counted as described previously. For Km values data was fit to Michaelis-Menten equation, kcat and processivity data was fit to either linear regression or a fit to a nonlinear regression using Prism v5 (Graphpad). The fold stimulation was calculated by dividing the amount of product formed by DNMT3A with DNMT3L by product formed by DNMT3A without DNMT3L. Error bars are standard error from three reactions. Bar charts of kinetic values compared mutants to wild type using one-way ANOVA to determine p-value using Prism. M.HhaI kchem and Kd assays carried out as described in Matje et al. 2010 (18), Km and kcat values determined as described in Lindstrom et al. (26).

**Fluorescence anisotropy**— koff values were determined with enzyme (250 nM) and 5’ 6-FAM labeled GCbox30 duplex DNA at 20 nM, reaching maximum anisotropy, followed by adding in chase DNA (unlabeled GCbox30) at 100x concentration of labeled DNA. The decrease in anisotropy was measured with time on a PerkinElmer LS 55 fluorescence spectrometer using Fl Winlab.
software. Data was fit to a one-phase exponential decay \( Y = span e^{-kt} + \text{plateau} \) using GraphPad, from two independent experiments.

**RESULTS**

Generation of a DNMT3A homotetramer model— A DNMT3A homotetramer model (Fig. 1A-C) was generated for experimental evaluation by aligning the C\(\alpha\) backbone of a DNMT3A monomer to the DNMT3L position in the 2QRV (9) crystal structure (RMSD = 1.4 Å, supplemental Fig. 1A). This rough model was refined over several 1000-model iterations with the RosettaDock server (18) to optimize side chain interactions at the tetramer interface. After numerous simulations converged upon the same model (Fig. 1A), the refined complex was submitted to the Rosetta computational alanine interface scanning server (22, 23) to obtain approximations for the energetic contributions of individual residues to the stabilization of this interface (\(\Delta \Delta G\)). A summary of the homotetramer model and the contributions to tetramer stability are shown in Figure 1, supplemental Figure 3 and supplemental Table 4.

Computational alanine screening details an aromatic core at the center of the interface providing the highest energetic contribution to tetramerization, with several salt bridges surrounding the core also having significant contributions. One residue with a high \(\Delta \Delta G\) score, F732, has previously been shown to stabilize the DNMT3A homotetramer (9). Mutation to alanine results in a large reduction in methylation activity, which was interpreted to demonstrate that tetramerization is required for catalysis by DNMT3A. However, F732 resides at a structurally well-conserved region of the catalytic domain. While the DNMT3A catalytic domain (\(\Delta 1-611\)) has just 13% sequence identity to the well-characterized homolog M.HhaI, the tertiary structures show a remarkably small backbone RMSD of 2.3 Å across 220 residues of the catalytic domains (supplemental Fig. 1B). Sequence alignments of DNMT3A with monomeric bacterial C5 DNA methyltransferases reveal four conserved residues along the DNMT3A tetramer interface (Fig. 1D, supplemental Table 1) amongst otherwise highly divergent sequences. Of special interest is conservation of the LFF motif found at the hinge of the catalytic loop in both enzymes that includes F732 (Fig. 1D).

To determine whether this conserved residue might be directly involved in catalysis in a monomeric DNA methyltransferase, we removed the side chains of the homologous residues in M.HhaI. F101 and F102 in the M.HhaI cognate-tertiary complex (27) are nearly isomorphous with F731 and F732 in the DNMT3A complex (supplemental Fig. 2A). The purified F101A and F102A M.HhaI mutants showed respective decreases in \(k_{\text{cat}}\) of 8.3 and 13.4-fold from the wild type enzyme (supplemental Table 2). Preparations of these mutants were also found to be significantly less stable than the wild type enzyme preparations (data not shown). Analogously, the F732A mutation in DNMT3A likely impacts critical structural elements within DNMT3A as well as tetramerization, leaving unanswered whether dimeric DNMT3A forms retain activity.

Six residues with varying \(\Delta \Delta G\) scores and weak conservation to bacterial homologues were investigated for their contribution to tetramerization, including two residues correlated with AML and MDS syndromes (R771 and R729) (supplemental Table 1). Two conserved residues, R736 and F732, were also mutated to alanine for comparison. We found that mutation of residues with \(\Delta \Delta G\) values greater than 1.0 (the combined score from the residue of the outer and inner monomers, supplemental Table 3) resulted in the formation of homodimers on DNA, while residues with \(\Delta \Delta G\) values less than 1.0 remained tetramers on DNA. The correlation between the computed \(\Delta \Delta G\) values and our experimental results (below) supports the use of the Rosetta modeling and scoring algorithms to generate the DNMT3A tetramer model.

Disruption of the tetramer interface results in monomers in solution that dimerize on DNA— We combined size exclusion chromatography with multi-angle light scattering to evaluate the oligomeric state of wild type DNMT3A and interface mutants. Light scattering shows our preparations of wild type DNMT3A catalytic domain to have a weight-average molecular weight of 127 kD in solution, in agreement with four 36 kD monomers forming a tetramer (Fig. 2A) (9). However, the broad, tailing peak suggests the tetrameric form is in equilibrium with both the dimeric and monomeric enzyme forms on the
column. Cofactor product, AdoHcy did not affect the form of the enzyme (supplemental Fig. 4B). Surprisingly, all alanine mutations at the interface disrupted the oligomeric state of the enzyme observed by gel filtration and multi-angle light scattering (Fig. 2A, supplemental Table 3, supplemental Fig. 4). Three of the mutants had weight-average molecular weights slightly larger than the theoretical dimer mass (H739A, D768A and R736A). Four mutants (F732A, R771A, E733A and R729A) have approximate molecular masses close to the theoretical mass of a DNMT3A monomer (Fig. 2A). These monomeric mutants have very little tailing in their elution profiles and elute from the column slower than dimers and tetramers.

We also investigated if binding to DNA changes the oligomeric state of wild type and mutant enzymes. Gel-filtration of the wild type DNMT3A catalytic domain/DNA complex was attempted but resulted in large aggregates with equimolar and two-fold excess DNA concentrations (supplemental Fig. 4B) at the micromolar concentrations needed for detectable multi-angle scattering. As a control, M.HhaI was tested with and without DNA, showing distinct peaks corresponding to monomeric bound or free states, respectively (supplemental Fig. 4A).

To resolve the oligomeric states of DNMT3A with DNA at concentrations closer to cellular conditions and in vitro assay conditions, electrophoretic mobility gel shift assays using a 30 basepair (bp), fluoroceinly labeled deoxyoligonucleotide (GCbox30) were performed. The GCbox30 contains binding sites for two well-behaved proteins; a central GC GC site bound by the monomeric M.HhaI (37 kD) and two GATATC sites bound by the homodimeric EcoRV endonuclease (29 kD) (28) at either end (Fig. 2B). At equimolar concentrations of protein and DNA, one or two EcoRV dimers bind to the DNA, resulting in standards for dimers and tetramers, respectively (supplemental Fig. 5A), while M.HhaI on DNA represent the monomer standard.

The GCbox30 has a single central CG site that is the major recognition site of DNMT3A (29). Figure 2C shows the formation of discrete bands when wild type and mutant DNMT3A enzymes are incubated with GCbox30. Wild type enzyme bound to DNA as a tetramer, smearing in the higher molecular mass region and titrating enzyme concentration from 40-360 nM did not change this pattern (supplemental Fig. 5B). H739A, D768A and R736A formed tetramers or larger aggregate forms on the DNA but were dimeric by light scattering. Five mutants (R771A, E733A, R729A, F732A, and Y735A) shift the DNA closest to the dimer size standard (EcoRV dimer: 58 kD, DNMT3A dimer: 68 kD). Figure 2D shows that titration of the E733A and R771A mutant enzyme concentrations down to 40 nM resulted in the presence of 200 nM CpG sites results in only the presence of dimeric enzyme bound to the DNA (see also Supplemental Fig. 5C). These results clearly show that DNA facilitates the dimerization of DNMT3A mutants that are monomeric in the absence of DNA even at 200 fold higher concentrations (Fig. 2A). When enzyme was present at higher concentrations than DNA, bands indicative of tetramers became observable from tetramerization of the enzyme or via multiple dimers binding to a single DNA substrate (Fig. 2D). The oligomeric states of the tetramer interface mutants are summarized in Figure 2E.

Dimers are fully active with an altered mechanism— DNMT3A mutants were examined for changes in their steady state kinetic parameters by monitoring the transfer of tritiated methyl groups from cofactor AdoMet onto DNA. $k_{cat}$, $K_{m, DNA}$, $K_{m, AdoMet}$, and processivity values were obtained for all the mutants and compared to wild type enzyme (supplemental Table 5). $k_{cat}$ values were obtained at saturating DNA and AdoMet concentrations after determining the respective $K_{m}$ values. In agreement with previous data (8), wild type DNMT3A shows curvature in the plot of product formed over time (Figure 3A). This is due to the slow product release step of this processive enzyme on this multi-site DNA substrate, where the enzyme partitions between migrating along the DNA to continue carrying out catalysis and dissociation from the DNA into solution (30). This same curvature is seen for the H739A and D768A mutants that do not disrupt tetramer formation, and these mutants differ less than 10% in their $k_{cat}$ values from the wild type value of 3.2 h$^{-1}$ (Fig. 3A, supplemental Fig. 6). The other mutant that appears tetrameric by gel shift, R736A, has a 5-fold decrease in $k_{cat}$. However, R736 is homologous to the previously investigated R106 in M.HhaI that has a 2.5-fold decrease in $k_{cat}$ when mutated to alanine (supplemental Fig. 2B) (31).
The mutants that form dimers on DNA (R771A, E733A, R729A, F732A, and Y735A) all show linear product formation with time, unlike the curved product formation seen for the tetramers (Fig. 3A, supplemental Fig. 6A). Linear product formation with a y-intercept of zero is characteristic of an enzyme where product release is no longer the rate-limiting step on a substrate with multiple methylation sites (32). A less likely possibility is that the enzyme is 100% processive and does not dissociate from the DNA in the three-hour reaction. The $k_{\text{cat}}$ values for R771A and R729A were slightly greater than wild type, while E733A and Y735A have $k_{\text{cat}}$ values slightly less than wild type enzyme (supplemental Table 5). F732A, previously reported to have no activity (9), was found to have a 15.2-fold decrease in $k_{\text{cat}}$ relative to the wild type enzyme. This reduction is very similar to the 13.4-fold $k_{\text{cat}}$ reduction seen in the F102A M.HhaI mutant (supplemental Table 2). Indeed, the two mutants that have significant decreases in $k_{\text{cat}}$, F732A and R736A, are residues conserved in the monomeric bacterial homolog M.HhaI (supplemental Table 1 and supplemental Fig. 2). As mutation of the F102 and R106 residues to alanine causes reduction in the catalytic turnover in M.HhaI (supplemental Table 2), it is unlikely the kinetic consequences of mutation of these residues in DNMT3A are a result of altered oligomeric forms. The four mutants that function as dimers have $k_{\text{cat}}$ values similar to wild type, demonstrating DNMT3A is a functional dimer and tetramerization is not necessary for catalysis, although the two forms have distinct kinetic mechanisms (see below).

$K_m^{\text{AdoMet}}$ for wild type DNMT3A is 206 nM and similar values were found for D768A and H739A. Mutants with larger decreases in activity have a 5 to 6-fold increase in $K_m^{\text{AdoMet}}$ (R736A and F732A) and the other dimers showed at 2-3-fold increase (supplemental Fig. 7B and 7C). The minimal effect on cofactor binding indicates oligomerization does not result in large pleotropic affects to catalysis.

**Dimers rapidly dissociate from DNA**— Wild type DNMT3A has a $K_m^{\text{DNA}}$ of 1.2 µM basepairs (bp) on poly-dIdC (19). The four mutants are fully active as dimers and all have an increase in $K_m^{\text{DNA}}$ from wild type between 10 and 32-fold (Fig 3B, supplemental Table 5 and supplemental Fig. 7A). These results demonstrate that oligomerization at the tetramer interface has important consequences for DNA binding. The observed increase in $K_m^{\text{DNA}}$ indicates a perturbation to substrate binding and/or product dissociation in the dimer mutants. To characterize the effect of oligomerization on dissociation from DNA, dissociation rates ($k_{\text{off}}$) were determined by binding excess enzyme to 5′ 6-FAM labeled GCbox30 duplex and measuring the rate of change in anisotropy upon the addition of unlabeled GCbox30 (Fig. 3C). Change in fluorescence anisotropy was plotted against time and fit to a single exponential decay to give the $k_{\text{off}}$ value. Wild type DNMT3A has a $k_{\text{off}}$ of 0.20 min$^{-1}$ (supplemental Table 5 and supplemental Fig. 8), H739A and D768A show similar rates and R736A dissociates more quickly, 0.32 min$^{-1}$. The dimeric mutants dissociate 2-5 times faster than the wild type enzyme (Fig. 3C), indicating that the linear product formation observed for the dimer mutants (Fig. 3A) is most likely due to the increase in $k_{\text{off}}$. The increase in the rate of dissociation from product also likely contributes to the observed increase in $K_m^{\text{DNA}}$ in the dimeric mutants (Fig. 3B). Changes in processive catalysis could account for changes in in vivo methylation patterns such as those observed in AML patients with minor changes in overall methylation (28). We sought to directly relate the oligomeric form of the enzyme to changes in processivity.

**The tetramer interface controls processivity**— Both full length and the catalytic domain of wild type DNMT3A act processively on several DNA substrates, including human promoters (8). Through mathematical modeling it was determined that the enzyme carries out an average of 27 turnovers ($n^{1/2} = 27$) before 50% of the enzyme dissociates from DNA (poly-dIdC) (8). Tetrameric H739A and D768A have similar processivity values as wild type ($n^{1/2} = 22, 20$ respectfully, Supplemental Table 6), in agreement with their unchanged $k_{\text{off}}$. The highly linear formation of product over time with the dimer mutants (Fig. 3A, Supplemental Fig. 6) made fitting to the processivity equations impossible, as the multi-exponential equation cannot converge to fit the linear data. Because linear product formation is observed for highly processive enzymes like M.SssI (8), we sought to qualitatively examine processivity in the dimer mutants using a chase assay (8, 30).
The processivity assay determines the length of time the enzyme stays associated with the multiple site DNA substrate and the number of turnovers using a chase experiment. The chase assay begins with DNMT3A carrying out 1-2 turnovers, approximately 20 minutes, on poly-dIdC DNA. After this time, a 25-fold excess of pCpG, a 3872-bp plasmid lacking the recognition site (CG) for DNMT3A, is added to capture enzyme that dissociates from the substrate after carrying out catalysis. DNMT3A shows minimal methylation activity with pCpG ($k_{cat} = 0.11 \pm 0.03 \text{ h}^{-1}$), and when pre-mixed at 25-fold excess poly-dIdC, it eliminated 85% of product formation. A processive enzyme will continue to methylate the original, multisite substrate after chase DNA is added and addition of chase DNA has little, or delayed, impact on the rate of methylation (wild type DNMT3A, Fig. 3D). A non-processive enzyme will dissociate after one methyl transfer; thus, after the addition of chase DNA, the enzyme will bind the excess pCpG and cause an immediate decrease in product formation (see R771A, Fig. 3D).

Wild type DNMT3A, H739A, and D768A show unaltered methylation 90 minutes after chase DNA is added. In contrast, the mutants that exist as dimers on DNA (supplemental Table 3) show an immediate decrease in activity upon addition of the chase DNA. This indicates that these mutants no longer act processively (supplemental Fig. 10), showing that tetramerization has a profound impact on the catalytic properties of DNMT3A. The enzymes with large $k_{cat}$ changes were not tested for processivity changes.

**DNMT3L binding to DNMT3A dimers restores processivity**— The DNMT3A:DNMT3L heterotetramer complex forms a discrete band when bound to the fluorescently labeled GCbox30 substrate in the electrophoretic mobility assay (Fig. 4A). Adding an equal concentration of DNMT3L eliminates any wild type DNMT3A homotetramer band and formed a strong heterotetramer band (Fig. 4A). This discrete DNMT3A:DNMT3L band is seen with all interface mutants except F732A and Y735A. For dimer mutants R771A, E733A and R729A, there was a loss of the dimer band and a gain of a heterotetramer band; gel shift results are detailed in Figure 4B. These results are in good agreement with computational alanine screening of the heterotetramer interface that describes the central aromatic pocket (including F732 and Y735) as critical for the DNMT3A:DNMT3L interface, with more modest contributions from salt bridges flanking the hydrophobic core (supplemental Fig. 11). Like the homotetramer interface, disruption of residues with a ΔΔG value greater than 1.2 results in loss of heterotetramerization (supplemental Table 3 and 7).

DNMT3L activates DNMT3A in vitro (13) and in vivo (10) and increases processivity 3-fold (8). DNMT3L activates DNMT3A approximately five-fold at a 1:1 ratio with pre-incubation (8,13)(supplemental Fig. 9A). Pre-incubation of DNMT3L with wild type DNMT3A is needed before DNA is added for heterotetramer formation and activation, presumably to allow exchange of the outer two DNMT3A monomers for DNMT3L (13). Homotetramers H739A, D768A and R736A also showed a five-fold increase in activity when incubated with equimolar or greater concentrations of DNMT3L. Dimer mutants partitioned into two groups; activation close to 10-fold (E733A, R771A and R729A) or no activation (F732A and Y735A) (Fig. 4C, supplemental Fig. 9). F732A and Y735A are the two residues predicted to be the most important for the DNMT3L interface (supplemental Table 3) and experimentally are the only mutants that showed no activation and no tetramerization. In addition E733A, R771A and R729A did not require pre-incubation to be activated by DNMT3L and could be pre-incubated with DNA and were still activated unlike the tetramers (data not shown).

The addition of DNMT3L to wild type DNMT3A enhances $k_{cat}$, and $K_{m}^{\text{AdoMet}}$ and leaves $K_{m}^{\text{DNA}}$ unchanged (19). In contrast, DNMT3L regulates the $K_{m}^{\text{DNA}}$ for the functional dimers, which in the absence of DNMT3L show a 15-30-fold increase compared to wild type and heterotetramerization restores the $K_{m}^{\text{DNA}}$ to only 1.2 to 1.6-fold greater than wild type enzyme (Fig. 4D, supplemental Fig. 9B, supplemental Table 8). This data shows that the changes in $K_{m}^{\text{DNA}}$ did not result from these residues being directly involved in DNA binding, but rather through the disruption of homo or hetero-tetramerization.

As shown above, DNMT3A dimer mutants tested in this study have an increase in $k_{diff}$ relative to wild type (Fig. 3C, supplemental Fig. 8). The addition of DNMT3L to wild type DNMT3A
caused $k_{\text{off}}$ to be decreased 25% from 0.2 min$^{-1}$ to 0.15 min$^{-1}$ (Fig. 4E, supplemental Table 8) on a single site substrate. Previous kinetic modeling showed that DNMT3L decreases DNMT3A off rates more substantially on a multiple site substrate (8). DNMT3L decreased the dissociation rate for R771A, R729A and E733A to wild type levels.

Dimer mutants show linear product formation with time, most likely because of an increase in the off rate, resulting in $k_{\text{chemistry}}$ being the rate limiting step. Since DNMT3L decreases the off rate for DNMT3A dimer mutants (Fig. 4F, supplemental Fig. 8), we predicted that DNMT3L would rescue the curved product formation of a processive enzyme, like the homotetramer. As expected, dimer mutants with DNMT3L (E733A, R771A and R729A) all showed curved product formation with time like wild type DNMT3A (Fig. 4F, supplemental Fig. 12).

The chase assay was also used on DNMT3A mutants with DNMT3L for further qualitative assessment of processivity, with DNMT3A wild type and mutants all showing an increase in processivity. Dimer mutants R771A, E733A and R729A were all non-processive without DNMT3L. When DNMT3L was added, these three mutants become highly processive like the wild type DNMT3A:DNMT3L complex. As shown in Figure 5E the addition of chase DNA has little effect on product formation even over 100 methylation cycles for all DNMT3A:DNMT3L complexes (supplemental Fig. 12).

**DISCUSSION**

The regulatory mechanisms controlling DNMT3A targeting and activity remain poorly understood, but there is increasing evidence for the role of protein-protein interactions in establishment of 5mC patterning (33). We investigated how oligomerization contributes to DNMT3A function by testing distinct oligomers, created by mutating several residues identified via computational modeling as critical towards formation of the homotetramer complex. The DNMT3A dimer is fully functional but dissociates from DNA more rapidly, eliminating methylation on clustered CpG sites (processive catalysis) that is observed with the wild type homotetramer. Dimeric DNMT3A mutants that form heterotetramers with DNMT3L exhibit processivity like the wild type heterotetramer complex (diagramed in Fig. 5). Our results demonstrate protein binding by DNMT3A, DNMT3L, or other regulators of DNMT3A at the tetramer interface are important for methylation patterning. Recently, mutations in DNMT3A were found in AML (16, 17) or MDS (15) patients along the tetramer interface (Fig. 1D, supplemental Table 10), and we show that two of these residues (R771 and R729) are necessary for tetramerization. The clustered loss (and gain) of promoter methylation at multiple CpG sites within CpG islands in these patients (34) is consistent with our in vitro observations involving DNMT3A mutants. Differences between oligomeric states provide a possible mechanism for how DNMT3A mutations in these cancer patients create aberrant DNA methylation patterns without global decreases in methylation levels (16).

Processive catalysis and its regulation by accessory proteins are critical to the action of several DNA modifying enzymes (35-37). The control of DNMT3A processivity is relevant to the establishment of methylation patterning, as the degree of 5mC content in promoter regions has important consequences for transcriptional regulation (1). Processive methylation of DNA has been demonstrated for the maintenance DNMT1 (30, 38) and both de novo DNMT3 enzymes in vitro (8, 10), with increasing evidence for this being the predominant mechanism of methylation in vivo. Recently, whole genome methylation analysis of human cells showed promoter regions to have bimodal distribution patterns, where promoters were either mostly methylated or unmethylated (39, 40). We show the major consequence of eliminating tetramerization in DNMT3A is a dramatic change in the number of processive cycles of methylation that occur each time the enzyme associates with DNA. There is also clear evidence that interactions with accessory proteins play an essential role in the establishment of methylation patterning (41, 42). DNMT3L is critical for early developmental de novo DNA methylation (2, 12), by binding DNMT3A at the tetramer interface (9), which serves to increase cofactor affinity (13) and catalytic turnover (13). This same interface is shown here to regulate DNMT3A processivity, thereby playing an essential role in coordinating genomic methylation. Since DNMT3L is only expressed in early...
development (12), DNMT3A either acts as a homo-oligomer or interacts with unidentified binding partners (43-46) that modulate the oligomeric state and activity in differentiated cells.

We propose that protein binding at the tetramer interface controls processivity through stabilization of the catalytic loop in the closed form. Cheng and Blumenthal suggested that binding of DNMT3L or DNMT3A at the tetramer interface locks the flexible catalytic loops (707-727) of the inner dimer in the closed position (47) which increases the contact area with DNA (supplemental Fig. 12). Increasing the binding surface is established to enhance processivity in other DNA modifying enzymes (48-50) by decreasing product release and facilitating translocation along the substrate. In M.HhaI, the closure of this loop is correlated with tight binding of cofactor and DNA (18). The other human de novo methyltransferase, DNMT3B, may be regulated by a similar mechanism (10), since the tetramer interface is highly homologous with DNMT3A. Newly solved crystal structures of DNMT1 (51, 52) show a hydrophobic core at the interface analogous to the DNMT3A tetramer interface. Structures of mouse DNMT1 show the catalytically essential CXXC domain (53) forms specific interactions across this interface, including along the catalytic loop (52). We propose that oligomerization at this structurally conserved interface is a general mechanism by which cis (CXXC domain) and trans (DNMT3L) acting interactions modulate the activity of the DNMT family. Further work is needed to identify and understand how binding partners (46) (including possibly regulatory RNA (54)), mutations, and/or cellular conditions drive oligomerization at the tetramer interface of DNMT3A that have important manifestations in the cellular phenotype.

References
1. Bird, A. (2002) Genes Dev. 16, 6–21
2. Reik, W., Dean, W. & Walter, J. (2001) Science 293, 1089–1093
3. Wu, H. et al. (2010) Science 329, 444–448
4. Okano, M., Bell, D.W., Haber, D.A. & Li, E. (1999) Cell 99, 247-57
5. Chen, T., Ueda, Y., Dodge, J.E., Wang, Z. & Li, E. (2003) Mol. Cell. Biol. 23, 5594–5605
6. Robertson, K.D. (2005) Nat. Rev. Genet. 6, 597-610
7. Van Emburgh, B.O. & Robertson, K.D. (March 4, 2011) Nucleic Acids Res. doi: 10.1093/nar/gkr116
8. Holz-Schietinger, C. & Reich, N.O. (2010) J. Biol. Chem. 285, 29091–29100
9. Jia, D., Jurkowska, R.Z., Zhang, X., Jeltsch, A. & Cheng, X. (2007) Nature 449, 248-51
10. Wienholz, B.L. et al. (2010) PLoS Genet 6
11. Chedin, F., Lieber, M.R. & Hsieh, C.L. (2002) Proc. Natl. Acad. Sci. USA 99, 16916-21
12. Webster, K.E. et al. (2005) Proc. Natl. Acad. Sci. USA 102, 4068-73
13. Kareta, M.S., Botello, Z.M., Ennis, J.J., Chou, C. & Chedin, F. (2006) J. Biol. Chem. 281, 25893-902
14. Malygin, E.G., Evdokimov, A.A. & Hattman, S. (2009) Biol. Chem. 390, 835-44
15. Walter, M.J. et al. (March 18, 2011) Leukemia doi:10.1038/leu.2011.44
16. Ley, T.J. et al. (2010) N. Engl. J. Med. 363, 2424–2433
17. Shah, M.Y. & Licht, J.D. (2011) Nat. Genet. 43, 289-90
18. Matje, D.M., Coughlin, D.F., Connolly, B.A., Dahlquist, F.W. & Reich, N.O. (2011) Biochemistry 50, 1465-73
19. Purdy, M.M., Holz-Schietinger, C. & Reich, N.O. (2010) Arch. Biochem. Biophys. 498, 13-22
20. Hiller, D.A. et al. (2003) Biochemistry 42, 14375-85
21. Lyskov, S. & Gray, J.J. (2008) Nucleic Acids Res. 36, W233-8
22. Kortemme, T., Kim, D.E. & Baker, D. (2004) Sci. STKE 2004, pl2
23. Kortemme, T. & Baker, D. (2002) Proc. Natl. Acad. Sci. USA 99, 14116-21
24. Klug, M. & Rehli, M. (2006) Epigenetics 1, 127-30
25. Abramoff, M.D., Magelhaes, P.J., Ram, S.J. (2004) Biophoton. Int. 11, 36-42
26. Lindstrom, W.M., Jr., Flynn, J. & Reich, N.O. (2000) J. Biol. Chem. 275, 4912-9
27. Shieh, F.K., Youngblood, B. & Reich, N.O. (2006) J. Mol. Bio. 362, 516-27
Acknowledgments

We thank Frederick Dahlquist and John Perona for helpful discussions and critical reading of the manuscript; and Armand Vartanian and Andrew Hadd for light scattering support. We are also thankful for Matt Purdy’s work on M.HhaI F101A.

Footnote

Abbreviations: DNA Methyl Transferase (DNMT), 5-methylcytosine (5mC), acute myeloid leukemia (AML), myelodysplastic syndromes (MDS), S-Adenosyl-L-homocysteine, S-Adenosyl-methionine and dithiothreitol (DTT), wild type (WT), poly-deoxyinosinic-deoxyctydylic (poly-dldC)
Figure Legends

Fig. 1 DNMT3A homotetramer model. A. A homotetramer model was generated by aligning a DNMT3A monomer to DNMT3L (from PDB: 2QRV) followed by prediction of the lowest-energy orientation using RosettaDock. Below is a close up of the DNMT3A-DNMT3A tetramer interface showing the core residues and the predicted interactions. The interface shows an aromatic pocket in the center of the interface, with ionic interactions on each edge of the interface. B. Depiction of the interactions at the DNMT3A:DNMT3A tetramer interface identified from the model. C. DNMT3A residues for one molecule (pink) of the tetramer interface. Residues are colored based upon their contribution, in ΔΔG, to the tetramer interface compared with alanine as determined using the Rosetta interface alanine scanning module. Bright yellow residues provide the greatest contribution to the DNMT3A:DNMT3A interface. D. Multiple sequence alignment of DNMT3A tetramer interface with related DNMTs and monomeric bacterial homolog (M.Hhal). Red and green residues were mutated in this study to alanine; green and purple residues are mutated in AML or MDS patients. Shaded regions show conserved residues between DNMT3A to M.Hhal.

Fig. 2 DNMT3A biophysical characterization. A. Light scattering data of DNMT3A mutations along the tetramer interface shows that mutants in solution are either mostly monomer or dimers unlike wild type tetramers. Size-exclusion chromatography of light scattering traces of tetrameric wild type catalytic domain (black trace), representative dimeric H739A (blue trace) and representative dimeric/monomeric R771A (red trace). Molecular weights were determined from the amount of scattered light, in relation to protein concentration determined by A_280. B. Electrophoretic mobility assay of size markers, DNA (GCbox30) has binding sites for size standards, one site for M. Hhal (37 kD), a known monomer and two binding sites for EcoRV (29 kD), a known dimer, which creates a dimer and ~ tetramer band. C. DNA facilitates oligomerization for DNMT3A mutants; five mutants were dimers on DNA. The distance the fluorescently labeled duplex DNA (GCbox30) shifted with DNMT3A determined the oligomeric state. D. Electrophoretic mobility assay varying the concentration of R771A and E773A (20 -400 nM) with DNA at 200 nM demonstrating these mutants are dimers even at low concentrations, unlike the monomers seen in solution at 200x the concentration. E. Diagram of oligomeric mutants with and without DNA, showing DNA facilitates oligomerization for DNMT3A mutants via the dimer interface.

Fig. 3 DNMT3A homodimers and homotetramers are both active with different mechanisms. A. Dimers and tetramers have similar rates of reaction. Left: bar chart of k_{cat} values for oligomeric mutants, right: representative time course for tetramer (wild type) showing curved product formation and dimer (R771A) showing linear product formation. B. Eliminating tetramer formation increases K_m^{DNA}. Left: bar chart of k_{cat} values for oligomeric mutants, right: representative tetramer and dimer curves. C. DNMT3A homodimers had an increase in off rate compared to homotetramers. Left: bar chart of k_{off} values for oligomeric mutants, k_{off} values were determined by the enzyme being bound to fluorescein labeled GCbox30 DNA, then adding saturating unlabeled DNA. Right: representative tetramer and dimer curves. D. DNMT3A homotetramers are processive and dimers are non-processive. Representative tetramer and dimer processive chase assay data. ● = only substrate, 20 µM bp poly-dIdC; ■ = substrate and then 400 µM bp chase (pCpG_L) at 20 minutes; ▲ = substrate and pCpG_L at the start of the reaction. No methylation is detected after addition of chaser DNA with the dimer mutant (R771A), unlike tetramer (WT) which shows less than 10% change in activity. All error bars are at least three experiments given as SE, one-way ANOVA was used to compare wild values to each mutant, * p> 0.05, ** p > 0.01, *** p > 0.001.
**Fig. 4** DNMT3L heterotetramers restores processivity in DNMT3A dimer mutants

A. DNMT3L binds to DNMT3A tetramers, forming defined heterotetramer. Three dimer mutants become heterotetramers with DNMT3L resulting in the loss of dimers; Y735A and F732A do not bind DNMT3L.

B. Schematic of the three forms that occurs with the addition of DNMT3L, how the forms change function is below. C. DNMT3L (1:1 ratio) activates DNMT3A tetramers ~ 5-fold. Homodimers that bind DNMT3L were activated ~ 10-fold; Y735A and F732A saw no activation by DNMT3L. D. Heterotetramer formation decreased $K_m^{DNA}$ for dimer mutants, and no change for homotetramers. E. DNMT3A heterotetramers had a decrease in off rate compared to homotetramers and homodimers. All heterotetramers have similar rates. F. All DNMT3A heterotetramers are processive as demonstrated by the processive chase assay, $\bullet$ = only substrate, 20 µM bp poly-dIdC; $\blacksquare$ = substrate and then 400 µM bp chase (pCpG$^2$) at 20 minutes; $\blacktriangle$ = substrate and pCpG$^2$ at the start of the reaction. Data indicates 100 turnovers occur before the enzyme dissociates from the DNA. All error bars are at least three experiments given as SE, one-way ANOVA was used to compare wild type enzyme to mutants, * p > 0.05, ** p > 0.01, *** p > 0.001.

**Fig. 5** Diagram of DNMT3A oligomeric states altering methylation patterns through changes in processivity s.A. Homodimers (R771A, R729A, E733A, Y735A and F732A) formed by disrupting the tetramer interface resulted in enzymes that bind DNA, methylation, fast dissociation and rebinding a new piece of DNA (non-processive catalysis). B. Homotetramers (WT, H739A, D768A and R736A) bind DNA followed by methylation then translocation along the same piece of DNA to a new site that the enzyme methylates, thus carrying out processive catalysis multiple times before slow dissociation. C. DNMT3A:DNMT3L heterotetramers (WT, H739A, D768A, R736A, R771A, E733A and R729A) have increased processive catalysis compared to homotetramers and restores processive in the homodimer, carrying out processive catalysis multiple times before slow dissociation.
Figure 2
Figure 3
Figure 4
Figure 5
Oligomerization of DNMT3A controls the mechanism of de novo DNA methylation
Celeste Holz-Schietinger, Douglas M. Matje, Madeleine Flexer Harrison and Norbert O. Reich

J. Biol. Chem. published online October 6, 2011

Access the most updated version of this article at doi: 10.1074/jbc.M111.284687

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2011/10/06/M111.284687.DC1