Published in final edited form as:

Curr Opin Struct Biol. 2022 August ; 75: 102433. doi:10.1016/j.sbi.2022.102433.

Mediating and maintaining methylation while minimizing mutation: Recent advances on mammalian DNA methyltransferases

Xiaodong Cheng[1], **Robert M. Blumenthal**[2]

[1]Department of Epigenetics and Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[2]Department of Medical Microbiology and Immunology, and Program in Bioinformatics, The University of Toledo College of Medicine and Life Sciences, Toledo, OH 43614, USA

Abstract

Mammalian genomes are methylated on carbon-5 of many cytosines, mostly in CpG dinucleotides. Methylation patterns are maintained during mitosis via DNMT1, and regulatory factors involved in processes that include histone modifications. Methylation in a sequence longer than CpG can influence the binding of sequence-specific transcription factors, thus affecting gene expression. 5-Methylcytosine deamination results in C-to-T transition. While some mutations are beneficial, most are not; so boosting C-to-T transitions can be dangerous. Given the role of DNMT3A in establishing *de novo* DNA methylation during development, it is this CpG methylation and deamination that provide the major mutagenic impetus in the *DNMT3A* gene itself, including the R882H dominant-negative substitution associated with two diseases: germline mutations in *DNMT3A* overgrowth syndrome, and somatic mutations in clonal hematopoiesis that can initiate acute myeloid leukemia. We discuss recent developments in therapeutics targeting DNMT1, the role of noncatalytic isoform DNMT3B3 in regulating *de novo* methylation by DNMT3A, and structural characterization of DNMT3A in various configurations.

Introduction

Epigenetic modifications profoundly impact chromatin structure and gene expression, affecting inheritance, evolution, and human development and health [1]. *S*-adenosyl-L-methionine (SAM)-dependent methylations occur on DNA, RNA, proteins and an array of other biological molecules [2]. Here, we focus on DNA cytosine-C5 methylation – one of the most studied epigenetic modifications in mammals, and a process associated with a range of human diseases, including cancer [3].

Corresponding author: Cheng, Xiaodong (xcheng5@mdanderson.org).

Conflict of interest statement

Nothing declared.

There are three major human DNA cytosine methyltransferases (DNMTs), belonging to two structurally-and functionally-distinct families [4] (Figure 1). These DNMTs act primarily on CpG dinucleotides, generating 5-methylcytosine (5mC). DNMT3A and DNMT3B establish the initial methylation pattern *de novo* [5], in response to signals that are still not fully understood. DNMT1, in contrast, maintains this pattern during replication, by acting on the unmethylated strand in hemimethylated DNA [6]. Two additional proteins function in these processes: DNMT3L in *de novo* methylation [7], and UHRF1 in maintenance methylation [8,9], with both of them linking DNA methylation and histone modifications. DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* DNA methylation [10], whereas UHRF1 binds to both hemimethylated CpGs and to modified histone H3 tails ([11] and references therein) (Figure 1). In rodent genomes (but not in primates), Dnmt3c is a product of gene duplication of Dnmt3b, and is a catalytically-active DNA methyltransferase expressed in male germline cells [12,13]. Together with Dnmt3a, which broadly methylates the genome [14], DNMT3c methylates the promoters of young retrotransposons and is required for mouse fertility [12].

This review focuses on recent developments in cancer therapeutics targeting DNMT1, hotspot mutations of DNMT3A in clonal hematopoiesis, structural characterization of DNMT3A in various configurations, roles of DNMT3B variants in regulation of DNMT3A activity, and the deamination of 5mC within a CpG dinucleotide which can cause C-to-T mutations.

Dicyanopyridine-based DNMT1-selective, non-nucleoside inhibitors

The reversibility of epigenetic modifications makes them attractive therapeutic targets. Historically, inducing DNA hypomethylation in humans relied on the cytosine nucleoside analogs azacytidine and 5-aza-2′-deoxycytidine (decitabine) [15,16]. These FDA-approved prodrugs, after conversion to triphosphates, are incorporated into RNA or DNA, where they trap DNMTs in irreversible, covalently-bonded suicide complexes [17]. These complexes result in persistent nucleoprotein adducts throughout the genome (for example, see Ref. [18]), repair of which tends to be error-prone, contributing to the toxicity of these drugs.

Despite their potential, the toxicity of, limited patient tolerance for, and ineffectiveness in treating solid tumors of nucleoside cytosine analogs [19] has stimulated development of non-nucleoside DNMT inhibitors. GlaxoSmithKline recently described a new class of reversible, DNMT1-selective inhibitors containing a dicyanopyridine moiety [20,21]. These inhibitors, including GSK3685032 (Figure 2a), are specific for DNMT1 with respect to >300 protein kinases and 30 other methyltransferases, significantly including DNMT3A and 3B [20]. GSK3685032 inhibits cancer cell growth *in vitro*, and is superior to decitabine for tumor regression and mouse survival in models of acute myeloid leukemia (AML) [20]. This new class of DNMT1-specific inhibitors shows promising therapeutic potential. In a transgenic mouse model of sickle cell disease, oral GSK3482364 was well tolerated and boosted both fetal hemoglobin levels and the percentage of erythrocytes expressing fetal hemoglobin [22]. GSK3484862 (the purified *R*-enantiomer of GSK3482364) inhibits Dnmt1 in murine pre-implantation *Dnmt3a/3b* knockout embryos [23], and results in DNA hypomethylation nearly to the extent in *Dnmt1* knockouts, with minimal toxicity [24].

These dicyanopyridine compounds contain a planar dicyanopyridine that intercalates specifically into DNMT1-bound hemimethylated CpG dinucleotides, displacing the active-site loop (Figure 2b). It is not yet understood how the inhibitor is recruited to the CpG site, but the active-site loop also intercalates into the same CpG site, suggesting that the inhibitor collaborates and/or competes with the active-site loop.

The DNMT1 active-site loop begins a catalytic cycle in an open conformation (Figure 2c). Upon DNA binding, this loop intercalates into the DNA minor groove, flipping the target cytosine into the active-site pocket (Figure 2d). Once inside the active site, the Cys1226 nucleophile approaches cytosine C6, and the methyl-donor SAM approaches C5 from the opposite face of the cytosine ring, allowing catalysis to proceed (Figure 2e). After methyltransfer, elimination of the C5 proton breaks the covalent bond between Cys and C6, releasing the reaction products (methylated DNA and SAH) prior to the next round of catalysis. The inhibitor appears to act just *prior* to the methyltransfer, competing with or displacing the active-site loop from the DNA minor groove, restoring the open conformation without releasing the DNA (Figure 2f).

The dicyanopyridine moiety intercalates into the DNA in the minor groove, between the adjacent C:G and G:C base pairs, and interacts with DNMT1 side chains on the major groove side (Figure 2g). A planar moiety inserting between adjacent bases is a well-known process for drug intercalation into DNA or RNA ([25] and references therein). However, in this case the intercalation takes place specifically between the DNMT1-bound hemimethylated CpG dinucleotides. Although Class I methyltransferases (to which DNMT1 belongs) share a common seven-stranded Rossman-fold structure [26], the size and amino acid sequence of their active-site loops are not conserved (Figure 2h). Nevertheless, these results may also help in developing DNMT3A/3B-selective inhibitors.

Hotspot somatic mutations of DNMT3A in CpG dinucleotides

Clonal hematopoiesis (CH) is a prevalent age-related condition, associated with greatly increased risk of hematologic disease [27]. Somatic mutations in *DNMT3A* can initiate AML in some patients ([28,29] and references therein). Genome sequencing of hematopoietic cells from ~50,000 healthy individuals revealed that the top 10 most common single-nucleotide mutations *all* involve C-to-T alterations within CpG sites [30]. These mutations alter seven arginines and one proline in DNMT3A (Figure 3a). We note that 4/6 arginine codons include a CpG. Ironically, the CpG methylation activity of DNMT3A itself provides the major mutagenic impetus: 5mC deaminates to thymine [31], yielding a T:G mismatch. If not repaired in time, a round of replication generates a C:G-to-T:A transition (see below).

DNMT3A variants occur across the gene (Catalogue of somatic mutations in cancer; https://cancer.sanger.ac.uk/cosmic). Three-quarters (74%) of 253 disease-associated *DNMT3A* mutations were loss-of-function mutations, some due to reduced protein stability, and were associated with greater clonal expansion and AML development [32]. Of these top ten mutations, three are truncations (R320*, R598* and R771*) that result in loss of the histone-binding PWWP and ADD domains, the C-terminal catalytic domain, or all three

domains (Figure 3b–d). Arg326 of the PWWP domain is far from the H3K36me3 binding surface, but the R326C mutant loses an Arg–Asp electrostatic interaction (Figure 3e), which might explain its instability. Arg729 and Arg736 are in the same helix, interacting with DNMT3L (Figure 3f), which increases DNMT3A stability [33]. R729W or R736C/H could disrupt DNMT3A-DNMT3L interactions, explaining their instability. Pro904, whose codon CCG also contains CpG, curves the last helix of the catalytic domain (Figure 3g). P904L introduces a larger hydrophobic residue, that might interfere with the close packing with neighboring residues (like Pro799 in Figure 3g).

The most studied DNMT3A mutant is R882H [28]. R882 is in the DNMT3A dimer interface, and also contacts a DNA phosphate [34] (Figure 3h). R882H should not alter the DNMT3A–DNMT3L tetramer structure: aligning the DNMT3A catalytic domains (wild type R882 and mutant H882), in complex with the same DNA, yields a root-mean-square-deviation of <0.3Å over 875 aligned Cα atoms [34]. The mechanistic defects of R882H are unclear and reports seem inconsistent. In murine embryonic stem (mES) cells, the R882H equivalent mutant in mouse Dnmt3a exhibited dominant-negative effects [35]. *In vitro*, R882H did not have such effects in mixed DNMT3A/R882H complexes [36], but R882H inhibits wild-type DNMT3A by blocking formation of active tetramers [37], while R882H stabilizes the formation of large DNMT3A oligomers [38]. In the DNMT3A C-terminal catalytic domain, R882H acquires flanking sequence preference for G three nucleotides (+3) past the CpG [39], or prefers C there in another study from the same laboratory [40]. DNMT3A^{R882H}-DNMT3L complexes have reduced activity on CpGs, but higher activity on CpAs [34]. Finally, the R882H-equivalent mutation of mouse Dnmt3a is a Dnmt3b-like enzyme – the corresponding R882H mutation in mouse Dnmt3a (tested in Dnmt3a/3b double knockout mouse embryonic stem cells) impairs methylation of Dnmt3a targets at major satellite DNA, but has no clear effect on methylation of Dnmt3b targets at minor satellite DNA [41] (these authors also analyzed flanking sequence preferences of DNMT3A, DNMT3B and their variants, as discussed below). R882H-dependent CH and AML cells are particularly susceptible to azacytidine [42]. This suggests that R882H is active, at least for the initial reaction step where the covalent adduct is formed, and possibly also reflecting the changed CpG vs. CpA substrate preference.

Germline pathogenic variants (including R882H) are also found in patients with *DNMT3A* Overgrowth Syndrome (DOS) [43]. The peripheral blood DNA of DOS patients is hypomethylated (particularly in differentially-methylated regions – DMRs). Hypomethylation is more severe in patients with R882 substitutions, though still present in patients with non-R882 mutations [44]. In a mouse model ± *Dnmt3a* equivalent R882H mutation, CpG methylation levels within DMRs were ~33% (*Dnmt3a*$^{-/-}$), ~50% (*Dnmt3a*$^{+/R882H}$), and ~85% (*Dnmt3a*$^{+/-}$) of the Dnmt3a$^{+/+}$ control [44]. Since *Dnmt3a*$^{+/R882H}$ has lower methylation than *Dnmt3a*$^{+/-}$, Dnmt3a R882H protein must have dominant–negative properties *in vivo*.

The inactive DNMT3B3 isoform functions like DNMT3L in somatic cells and cancers

DNMT3L functions primarily in the germline [7], and the *DNMT3L* gene is quickly silenced upon differentiation [45,46], whereas DNMT3A and DNMT3B expression is retained. This raises a question about the molecular basis for DNMT3L-independent regulation of DNMT3A/3B. DNMT3B produces ~40 splice variants in tissue- and disease-specific manners [47], and at least 11 transcripts have been reported in the National Center for Biotechnology Information (NCBI), some of which yield catalytically inactive proteins. The variant DNMT3B3 is expressed in normal somatic cells and cancers [48,49], and this variant has the same 63-residue deletion as in Dnmt3L (Figure 4a). *In vitro* and in mES cells, Dnmt3b3, like Dnmt3L, stimulates the activities of Dnmt3a and Dnmt3b [50]. Structural characterization revealed a 3B3–3A-3A-3B3 heterotetramer, very similar to 3L-3A-3A-3L and 3L-3B-3B-3L (Figure 4b–d) [49,51,52]. One unexpected observation is that DNMT3B3 binds to the acidic patch of the nucleosome core and orients the binding of DNMT3A to the linker DNA where methylation occurs [49] (Figure 4b). The key accessory role for DNA methylation, played by DNMT3L in germ cells, may thus be played by DNMT3B3 and/or other inactive DNMT3B isoforms (*e.g.*, DNMT3B6) in somatic or cancer cells.

Flanking sequence preferences of DNMTs

Using the method of “Deep Enzymology” [53], Jeltsch and his colleagues studied CpG flanking sequence preferences for DNMT1, DNMT3A and DNMT3B [54,55]. These *in vitro* studies complemented earlier ones (*e.g.*, Ref. [41]), and used recombinant enzymes (or just their catalytic domains) and naked DNA. Nevertheless, the studies revealed distinct sequence preferences of DNMTs. For example, DNMT3A and DNMT3B strongly prefer T at the −2 flanking position (Figure 4e) [55]. Preference for a Tat −2 is present but less pronounced for DNMT1 [54]. Currently, there is no structural explanation for the preference of T at −2 for any of these DNMTs. DNMT1 strongly disfavors a C at −2 (Figure 4f) [54], while it is accepted there by DNMT3A and 3B (Figure 4e) [55]. On the 3′ side, DNMT3A prefers CpGs followed by C/T +at 1, while DNMT3B predominantly methylates CpGs followed by G/A at +1 [56], fitting structural observations [55]. DNMT1 has no (or limited) preference at the +1 position, which allows the maintenance enzyme to propagate without bias (on the 3′ side) the methylation generated by either DNMT3A or DNMT3B.

These observations raise questions about possible competing actions between *de novo* and maintenance DNMTs. For example, a C at −2, being strongly disfavored by DNMT1, could result in some CpGs preferentially losing methylation during each round of DNA replication. To avoid such bias, DNMT1 activity at replication foci is tightly regulated at multiple levels (Figure 1). Factors included in this regulation include PCNA [57], UHRF1 [8,9], DNA ligase 1 [58], and protein (histone) lysine methyltransferases G9a/GLP [59], among others, which may override DNMT1’s intrinsic flanking sequence preferences. We note that methylation of E-box sequences (CACGTG with a C at −2) or of CCAAT/enhancer sequences (TTGCGCAA with a T at −2) have profound effects on binding by

transcription factors of the bHLH (basic helix-loop-helix) and bZIP (basic leucine-zipper) families, respectively [60], so DNMT flanking sequence preferences can indirectly modulate transcription factor binding.

Transcription factor binding of a 5mC deamination product increases C-to-T mutations

Mutations in CpG-containing Arg codons correlates with genetic disease risk [61], and Arg→His mutations constitute about half of cancer-associated mutations across a range of tissues [62]. Besides the DNMT3A R882 mutations discussed here, other recurrent hotspot mutations are observed in the gene for tumor repressor p53, at R175, R213, R248, R273 and R282 [63], and in the genes for isocitrate dehydrogenase at R132 of IDH1 and R140 of IDH2 [64,65]. While spontaneous oxidative deamination of 5mC to T (or C to U) has been documented in organisms ranging from bacteria [66] to humans [67], our focus here is on the effects of 5mC methylation on the binding of specific proteins, and on how the 5mC-binding proteins modulate the mutation rate.

Deamination of 5mC-to-T generates a T:G mismatch (Figure 4g) [68]. There are two mammalian thymine DNA glycosylases, TDG [69] and MBD4 [70]. MBD4 guards against methylation-mediated deamination damage seen in CH and early-onset AML [71]. MBD4-deficient AMLs display a >30-fold higher mutation burden than AMLs generally, with nearly all (>95%) being C→T in a CpG dinucleotide [71]. In pan-cancer analysis of whole genomes, patients with rare germline *MBD4* truncations exhibited increased rates of somatic C→T mutation at CpG dinucleotide [72]. If the T:G mismatch is inaccessible to MBD4 and is not repaired in time, replication generates a C:G to T:A transition (Figure 4g). Two transcription factors, CEBPβ and ETS1, whose binding motifs contain CpG, restrict access by MBD4 to the mismatch and boost C→T mutation rates within their respective binding sites [73–75]. ETS1 binding both increases generation of UV-induced cyclopyrimidine dimers (CPDs), and blocks repair of cytosine-deaminated CPDs [75–77].

Summary and perspective

Growth of population sequencing data is now explosive, revealing the spectrum of human mutations. One example we reviewed here is *DNMT3A* mutations in DOS (germline) and CH (somatic). Similarly, p53 tumor suppressor gene (*TP53*) has germline mutations causing Li-Fraumeni syndrome, and the vast majority of mutations in diverse cancers involve CpG-containing Arg codons [78]. These observations underscore the importance of endogenous mutational mechanisms in the genesis of genetic diseases and cancer.

Knowledge on targeting p53 Arg-substitution proteins for treating cancer might be applicable to DNMT3A. Eprenetapopt (APR-246) is a thiol-reactive small molecule that covalently binds cysteines [79], stabilizing mutant p53 and restoring wild-type function in *TP53*-mutant cells [80]. The same approach might be usable in *DNMT3A*-mutant cells. The U.S. Food and Drug Administration has authorized clinical studies of a combination of APR-246 with azacitidine in the treatment of myelodysplastic syndrome (MDS) and AML, where a susceptible *TP53* mutation is present [81,82].

The new non-nucleoside and DNMT1-selective inhibitors should overcome many limitations of nucleoside analogs. Dicyanopyridine derivatives might be developed for DNMT3A/3B-specific and/or pan-DNMT inhibitors. The current use of DNA hypomethylating agents in treating the hypomethylation phenotype of patients with DNMT3A mutations seems counterproductive. Regions affected by hypomethylating agents are broader than the DMRs hypomethylated due to DNMT3A inactivation, so in some cases (such as hematopoietic stem cells) treatment with broad hypomethylating agents can have beneficial effects. It should be a high priority to identify regions of mismethylation associated with specific diseases such as AML.

Acknowledgements

We thank members of the Cheng laboratory for discussion and for the published work supported in part by the National Institutes of Health (grant R35GM134744) and Cancer Prevention and Research Institute of Texas (RR160029). X.C. is a CPRIT scholar in Cancer Research.

References

Papers of particular interest, published within the period of review, have been highlighted as:

• of special interest

•• of outstanding interest

1. Allis CD, Jenuwein T, Reinberg D: Epigenetics. 2nd ed. NY: Cold Spring Harbor Laboratory Press; 2015. Cold Spring Harbor.
2. Cheng X, Blumenthal RM: S-adenosylmethionine-dependent methyltransferases: structures and functions. New Jersey: World Scientifc Publishing Co. River Edge; 1999.
3. Gasser SM, Li E: Epigenetics and disease: pharmaceutical opportunities. Prog Drug Res 2011, 67.
4. Bestor TH: The DNA methyltransferases of mammals. Hum Mol Genet 2000, 9:2395–2402. [PubMed: 11005794]
5. Okano M, Bell DW, Haber DA, Li E: DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 1999, 99:247–257. [PubMed: 10555141]
6. Li E, Bestor TH, Jaenisch R: Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 1992, 69:915–926. [PubMed: 1606615]
7. Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH: Dnmt3L and the establishment of maternal genomic imprints. Science 2001, 294:2536–2539. [PubMed: 11719692]
8. Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S, Jacobsen SE: UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science 2007, 317:1760–1764. [PubMed: 17673620]
9. Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K, et al. : The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature 2007, 450: 908–912. [PubMed: 17994007]
10. Ooi SKT, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin S-P, Allis CD, et al. : DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 2007, 448:714–717. [PubMed: 17687327]
11. Mancini M, Magnani E, Macchi F, Bonapace IM: The multi-functionality of UHRF1: epigenome maintenance and preservation of genome integrity. Nucleic Acids Res 2021, 49: 6053–6068. [PubMed: 33939809]
12. Barau J, Teissandier A, Zamudio N, Roy S, Nalesso V, Herault Y, Guillou F, Bourc'his D: The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. Science 2016, 354:909–912. [PubMed: 27856912]

13. Jain D, Meydan C, Lange J, Claeys Bouuaert C, Lailler N, Mason CE, Anderson KV, Keeney S: Rahu is a mutant allele of Dnmt3c, encoding a DNA methyltransferase homolog required for meiosis and transposon repression in the mouse male germline. PLoS Genet 2017, 13, e1006964.

14. Dura M, Teissandier A, Armand M, Barau J, Lapoujade C, Fouchet P, Bonneville L, Schulz M, Weber M, Baudrin LG, et al. : DNMT3A-dependent DNA methylation is required for spermatogonial stem cells to commit to spermatogenesis. Nat Genet 2022, 54:469–480. [PubMed: 35410378]

15. Yoo CB, Jones PA: Epigenetic therapy of cancer: past, present and future. Nat Rev Drug Discov 2006, 5:37–50. [PubMed: 16485345]

16. Flotho C, Claus R, Batz C, Schneider M, Sandrock I, Ihde S, Plass C, Niemeyer CM, Lubbert M: The DNA methyltransferase inhibitors azacitidine, decitabine and zebularine exert differential effects on cancer gene expression in acute myeloid leukemia cells. Leukemia 2009, 23:1019–1028. [PubMed: 19194470]

17. Santi DV, Norment A, Garrett CE: Covalent bond formation between a DNA-cytosine methyltransferase and DNA containing 5-azacytosine. Proc Natl Acad Sci U S A 1984, 81: 6993–6997. [PubMed: 6209710]

18. Ferguson AT, Vertino PM, Spitzner JR, Baylin SB, Muller MT, Davidson NE: Role of estrogen receptor gene demethylation and DNA methyltransferase.DNA adduct formation in 5-aza-2'deoxycytidine-induced cytotoxicity in human breast cancer cells. J Biol Chem 1997, 272:32260–32266. [PubMed: 9405430]

19. Stomper J, Rotondo JC, Greve G, Lubbert M: Hypomethylating agents (HMA) for the treatment of acute myeloid leukemia and myelodysplastic syndromes: mechanisms of resistance and novel HMA-based therapies. Leukemia 2021, 35: 1873–1889. [PubMed: 33958699]

20••. Pappalardi MB, Keenan K, Cockerill M, Kellner WA, Stowell A, Sherk C, Wong K, Pathuri S, Briand J, Steidel M, et al. : Discovery of a first-in-class reversible DNMT1-selective inhibitor with improved tolerability and efficacy in acute myeloid leukemia. Nat Cancer 2021, 2:1002–1017. [PubMed: 34790902] A first report of non-nucleoside DNMT1-specific inhibitors having a clear translation path from *in vitro* to *in vivo* activity and possible clinical utility.

21. Horton JR, Pathuri S, Wong K, Ren R, Rueda L, Fosbenner DT, Heerding DA, McCabe MT, Pappalardi MB, Zhang X, et al. : Structural characterization of dicyanopyridine containing DNMT1-selective, non-nucleoside inhibitors. Structure 2022, 30:793–802. [PubMed: 35395178]

22. Gilmartin AG, Groy A, Gore ER, Atkins C, Long ER, Montoute MN, Wu Z, Halsey W, McNulty DE, Ennulat D, et al. : In vitro and in vivo induction of fetal hemoglobin with a reversible and selective DNMT1 inhibitor. Haematologica 2021, 106: 1979–1987. [PubMed: 32586904]

23. Haggerty C, Kretzmer H, Riemenschneider C, Kumar AS, Mattei AL, Bailly N, Gottfreund J, Giesselmann P, Weigert R, Brandl B, et al. : Dnmt1 has de novo activity targeted to transposable elements. Nat Struct Mol Biol 2021, 28:594–603. [PubMed: 34140676]

24. Azevedo Portilho N, Saini D, Hossain I, Sirois J, Moraes C, Pastor WA: The DNMT1 inhibitor GSK-3484862 mediates global demethylation in murine embryonic stem cells. Epigenet Chromatin 2021, 14:56.

25. Soni A, Khurana P, Singh T, Jayaram B: A DNA intercalation methodology for an efficient prediction of ligand binding pose and energetics. Bioinformatics 2017, 33:1488–1496. [PubMed: 28073762]

26. Schubert HL, Blumenthal RM, Cheng X: Many paths to methyltransfer: a chronicle of convergence. Trends Biochem Sci 2003, 28:329–335. [PubMed: 12826405]

27. Challen GA, Goodell MA: Clonal hematopoiesis: mechanisms driving dominance of stem cell clones. Blood 2020, 136: 1590–1598. [PubMed: 32746453]

28. Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, Kandoth C, Payton JE, Baty J, Welch J, et al. : DNMT3A mutations in acute myeloid leukemia. N Engl J Med 2010, 363:2424–2433. [PubMed: 21067377]

29. Bowman RL, Busque L, Levine RL: Clonal hematopoiesis and evolution to hematopoietic malignancies. Cell Stem Cell 2018, 22:157–170. [PubMed: 29395053]

30. Watson CJ, Papula AL, Poon GYP, Wong WH, Young AL, Druley TE, Fisher DS, Blundell JR: The evolutionary dynamics and fitness landscape of clonal hematopoiesis. Science 2020, 367:1449–1454. [PubMed: 32217721]

31. Duncan BK, Miller JH: Mutagenic deamination of cytosine residues in DNA. Nature 1980, 287:560–561. [PubMed: 6999365]

32•. Huang YH, Chen CW, Sundaramurthy V, Slabicki M, Hao D, Watson CJ, Tovy A, Reyes JM, Dakhova O, Crovetti BR, et al. : Systematic profiling of DNMT3A variants reveals protein instability mediated by the DCAF8 E3 ubiquitin ligase adaptor. Cancer Discov 2021, 12:220–235. [PubMed: 34429321] Large-scale survey of >250 disease-associated *DNMT3A* mutations, with demonstration or inferences of the molecular defects

33. Veland N, Lu Y, Hardikar S, Gaddis S, Zeng Y, Liu B, Estecio MR, Takata Y, Lin K, Tomida MW, et al. : DNMT3L facilitates DNA methylation partly by maintaining DNMT3A stability in mouse embryonic stem cells. Nucleic Acids Res 2019, 47: 152–167. [PubMed: 30321403]

34•. Anteneh H, Fang J, Song J: Structural basis for impairment of DNA methylation by the DNMT3A R882H mutation. Nat Commun 2020, 11:2294. [PubMed: 32385248] Comprehensive structural and *in vitro* activity characterizations of the pathogenic DNMT3A R882H mutant

35. Kim SJ, Zhao H, Hardikar S, Singh AK, Goodell MA, Chen T: A DNMT3A mutation common in AML exhibits dominant-negative effects in murine ES cells. Blood 2013, 122: 4086–4089. [PubMed: 24167195]

36. Emperle M, Dukatz M, Kunert S, Holzer K, Rajavelu A, Jurkowska RZ, Jeltsch A: The DNMT3A R882H mutation does not cause dominant negative effects in purified mixed DNMT3A/R882H complexes. Sci Rep 2018, 8, 13242.

37. Russler-Germain DA, Spencer DH, Young MA, Lamprecht TL, Miller CA, Fulton R, Meyer MR, Erdmann-Gilmore P, Townsend RR, Wilson RK, et al. : The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. Cancer Cell 2014, 25:442–454. [PubMed: 24656771]

38. Nguyen TV, Yao S, Wang Y, Rolfe A, Selvaraj A, Darman R, Ke J, Warmuth M, Smith PG, Larsen NA, et al. : The R882H DNMT3A hot spot mutation stabilizes the formation of large DNMT3A oligomers with low DNA methyltransferase activity. J Biol Chem 2019, 294:16966–16977. [PubMed: 31582562]

39. Emperle M, Adam S, Kunert S, Dukatz M, Baude A, Plass C, Rathert P, Bashtrykov P, Jeltsch A: Mutations of R882 change flanking sequence preferences of the DNA methyltransferase DNMT3A and cellular methylation patterns. Nucleic Acids Res 2019, 47:11355–11367. [PubMed: 31620784]

40. Mack A, Emperle M, Schnee P, Adam S, Pleiss J, Bashtrykov P, Jeltsch A: Preferential self-interaction of DNA methyltransferase DNMT3A subunits containing the R882H cancer mutation leads to dominant changes of flanking sequence preferences. J Mol Biol 2022, 434, 167482.

41. Norvil AB, AlAbdi L, Liu B, Tu YH, Forstoffer NE, Michie AR, Chen T, Gowher H: The acute myeloid leukemia variant DNMT3A Arg882His is a DNMT3B-like enzyme. Nucleic Acids Res 2020, 48:3761–3775. [PubMed: 32123902]

42. Scheller M, Ludwig AK, Gollner S, Rohde C, Kramer S, Stable S, Janssen M, Muller JA, He L, Baumer N, et al. : Hotspot DNMT3A mutations in clonal hematopoiesis and acute myeloid leukemia sensitize cells to azacytidine via viral mimicry response. Nat Cancer 2021, 2:527–544. [PubMed: 35122024]

43. Tatton-Brown K, Seal S, Ruark E, Harmer J, Ramsay E, Del Vecchio Duarte S, Zachariou A, Hanks S, O'Brien E, Aksglaede L, et al. : Mutations in the DNA methyltransferase gene DNMT3A cause an overgrowth syndrome with intellectual disability. Nat Genet 2014, 46:385–388. [PubMed: 24614070]

44•. Smith AM, LaValle TA, Shinawi M, Ramakrishnan SM, Abel HJ, Hill CA, Kirkland NM, Rettig MP, Helton NM, Heath SE, et al. : Functional and epigenetic phenotypes of humans and mice with DNMT3A Overgrowth Syndrome. Nat Commun 2021, 12: 4549. [PubMed: 34315901] Comprehensive characterization of DNMT3A R882H germline mutation effects in human patients and mouse model, establising R882H protein's dominant–negative properties *in vivo*

45. Hata K, Okano M, Lei H, Li E: Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development 2002, 129: 1983–1993. [PubMed: 11934864]

46. Guenatri M, Duffie R, Iranzo J, Fauque P, Bourc'his D: Plasticity in Dnmt3L-dependent and -independent modes of de novo methylation in the developing mouse embryo. Development 2013, 140:562–572. [PubMed: 23293288]

47. Gopalakrishnan S, Van Emburgh BO, Shan J, Su Z, Fields CR, Vieweg J, Hamazaki T, Schwartz PH, Terada N, Robertson KD: A novel DNMT3B splice variant expressed in tumor and pluripotent cells modulates genomic DNA methylation patterns and displays altered DNA binding. Mol Cancer Res 2009, 7:1622–1634. [PubMed: 19825994]

48. Weisenberger DJ, Velicescu M, Cheng JC, Gonzales FA, Liang G, Jones PA: Role of the DNA methyltransferase variant DNMT3b3 in DNA methylation. Mol Cancer Res 2004, 2:62–72. [PubMed: 14757847]

49•. Xu TH, Liu M, Zhou XE, Liang G, Zhao G, Xu HE, Melcher K, Jones PA: Structure of nucleosome-bound DNA methyltransferases DNMT3A and DNMT3B. Nature 2020, 586: 151–155. [PubMed: 32968275] Shows that DNMT3B3 binds to the acidic patch of the nucleosome core and thereby orients DNMT3A towards the target linker DNA

50•. Zeng Y, Ren R, Kaur G, Hardikar S, Ying Z, Babcock L, Gupta E, Zhang X, Chen T, Cheng X: The inactive Dnmt3b3 isoform preferentially enhances Dnmt3b-mediated DNA methylation. Genes Dev 2020, 34:1546–1558. [PubMed: 33004415] Demonstration that Dnmt3b3, and perhaps Dnmt3b6, facilitates *de novo* DNA methylation during mouse embryonic development and in somatic cells

51. Zhang ZM, Lu R, Wang P, Yu Y, Chen D, Gao L, Liu S, Ji D, Rothbart SB, Wang Y, et al. : Structural basis for DNMT3A-mediated de novo DNA methylation. Nature 2018, 554:387–391. [PubMed: 29414941]

52. Lin CC, Chen YP, Yang WZ, Shen JCK, Yuan HS: Structural insights into CpG-specific DNA methylation by human DNA methyltransferase 3B. Nucleic Acids Res 2020, 48:3949–3961. [PubMed: 32083663]

53•. Jeltsch A, Adam S, Dukatz M, Emperle M, Bashtrykov P: Deep enzymology studies on DNA methyltransferases reveal novel connections between flanking sequences and enzyme activity. J Mol Biol 2021, 433, 167186.Used high-throughput sequencing methods on CpGs within random sequences, to characterize flanking sequence preferences of several DNMTs

54. Adam S, Anteneh H, Hornisch M, Wagner V, Lu J, Radde NE, Bashtrykov P, Song J, Jeltsch A: DNA sequence-dependent activity and base flipping mechanisms of DNMT1 regulate genome-wide DNA methylation. Nat Commun 2020, 11:3723. [PubMed: 32709850]

55. Gao L, Emperle M, Guo Y, Grimm SA, Ren W, Adam S, Uryu H, Zhang ZM, Chen D, Yin J, et al. : Comprehensive structure-function characterization of DNMT3B and DNMT3A reveals distinctive de novo DNA methylation mechanisms. Nat Commun 2020, 11:3355. [PubMed: 32620778]

56. Mallona I, Ilie IM, Karemaker ID, Butz S, Manzo M, Caflisch A, Baubec T: Flanking sequence preference modulates de novo DNA methylation in the mouse genome. Nucleic Acids Res 2021, 49:145–157. [PubMed: 33290556]

57. Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, Li BF: Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science 1997, 277:1996–2000. [PubMed: 9302295]

58. Ferry L, Fournier A, Tsusaka T, Adelmant G, Shimazu T, Matano S, Kirsh O, Amouroux R, Dohmae N, Suzuki T, et al. : Methylation of DNA ligase 1 by G9a/GLP recruits UHRF1 to replicating DNA and regulates DNA methylation. Mol Cell 2017, 67:550–565 e555. [PubMed: 28803780]

59. Esteve PO, Chin HG, Smallwood A, Feehery GR, Gangisetty O, Karpf AR, Carey MF, Pradhan S: Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. Genes Dev 2006, 20:3089–3103. [PubMed: 17085482]

60. Yang J, Zhang X, Blumenthal RM, Cheng X: Detection of DNA modifications by sequence-specific transcription factors. J Mol Biol 2020, 432:1661–1673.

61. Schulze KV, Hanchard NA, Wangler MF: Biases in arginine codon usage correlate with genetic disease risk. Genet Med 2020, 22:1407–1412. [PubMed: 32371920]

62. Szpiech ZA, Strauli NB, White KA, Ruiz DG, Jacobson MP, Barber DL, Hernandez RD: Prominent features of the amino acid mutation landscape in cancer. PLoS One 2017, 12, e0183273.

63. Levine AJ: Spontaneous and inherited TP53 genetic alterations. Oncogene 2021, 40:5975–5983. [PubMed: 34389799]

64. Dang L, White DW, Gross S, Bennett BD, Bittinger MA, Driggers EM, Fantin VR, Jang HG, Jin S, Keenan MC, et al. : Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature 2009, 462:739–744. [PubMed: 19935646]

65. Ward PS, Patel J, Wise DR, Abdel-Wahab O, Bennett BD, Coller HA, Cross JR, Fantin VR, Hedvat CV, Perl AE, et al. : The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. Cancer Cell 2010, 17: 225–234. [PubMed: 20171147]

66. Cherry JL: Methylation-induced hypermutation in natural populations of bacteria. J Bacteriol 2018, 200. e00371–00318.

67. Zhou Y, He F, Pu W, Gu X, Wang J, Su Z: The impact of DNA methylation dynamics on the mutation rate during human germline development. G3 (Bethesda) 2020, 10:3337–3346. [PubMed: 32727923]

68. Chan K, Roberts SA, Klimczak LJ, Sterling JF, Saini N, Malc EP, Kim J, Kwiatkowski DJ, Fargo DC, Mieczkowski PA, et al. : An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. Nat Genet 2015, 47:1067–1072. [PubMed: 26258849]

69. Neddermann P, Gallinari P, Lettieri T, Schmid D, Truong O, Hsuan JJ, Wiebauer K, Jiricny J: Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. J Biol Chem 1996, 271:12767–12774. [PubMed: 8662714]

70. Hendrich B, Hardeland U, Ng HH, Jiricny J, Bird A: The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. Nature 1999, 401:301–304. [PubMed: 10499592]

71. Sanders MA, Chew E, Flensburg C, Zeilemaker A, Miller SE, Al Hinai AS, Bajel A, Luiken B, Rijken M, McLennan T, et al. : MBD4 guards against methylation damage and germ line deficiency predisposes to clonal hematopoiesis and early-onset AML. Blood 2018, 132:1526–1534. [PubMed: 30049810]

72. Consortium ITP-CAoWG: Pan-cancer analysis of whole genomes. Nature 2020, 578:82–93. [PubMed: 32025007]

73. Yang J, Horton JR, Akdemir KC, Li J, Huang Y, Kumar J, Blumenthal RM, Zhang X, Cheng X: Preferential CEBP binding to T:G mismatches and increased C-to-T human somatic mutations. Nucleic Acids Res 2021, 49:5084–5094. [PubMed: 33877329]

74. Ershova AS, Eliseeva IA, Nikonov OS, Fedorova AD, Vorontsov IE, Papatsenko D, Kulakovskiy IV: Enhanced C/EBP binding to G.T mismatches facilitates fixation of CpG mutations in cancer and adult stem cells. Cell Rep 2021, 36, 109365.

75. Yang J, Gupta E, Horton JR, Blumenthal RM, Zhang X, Cheng X: Differential ETS1 binding to T:G mismatches within a CpG dinucleotide contributes to C-to-T somatic mutation rate of the IDH2 hotspot at codon Arg140. DNA Repair 2022, 113, 103306.

76. Mao P, Brown AJ, Esaki S, Lockwood S, Poon GMK, Smerdon MJ, Roberts SA, Wyrick JJ: ETS transcription factors induce a unique UV damage signature that drives recurrent mutagenesis in melanoma. Nat Commun 2018, 9:2626. [PubMed: 29980679]

77. Elliott K, Bostrom M, Filges S, Lindberg M, Van den Eynden J, Stahlberg A, Clausen AR, Larsson E: Elevated pyrimidine dimer formation at distinct genomic bases underlies promoter mutation hotspots in UV-exposed cancers. PLoS Genet 2018, 14, e1007849.

78. Lozano G, Levine AJ: The p53 Protein: from cell regulation to cancer. Cold Spring Harbor Laboratory Press; 2016.

79•. Degtjarik O, Golovenko D, Diskin-Posner Y, Abrahmsen L, Rozenberg H, Shakked Z: Structural basis of reactivation of oncogenic p53 mutants by a small molecule: methylene quinuclidinone

(MQ). Nat Commun 2021, 12:7057. [PubMed: 34862374] Shows that the biologically active methylene quinuclidinone (MQ), converted from prodrug APR-246, reactivates mutant p53 by binding covalently to cysteine residues

80. Duffy MJ, Synnott NC, O'Grady S, Crown J: Targeting p53 for the treatment of cancer. Semin Cancer Biol 2022, 79:58–67. [PubMed: 32741700]
81. Sallman DA, DeZern AE, Garcia-Manero G, Steensma DP, Roboz GJ, Sekeres MA, Cluzeau T, Sweet KL, McLemore A, McGraw KL, et al. : Eprenetapopt (APR-246) and azacitidine in TP53-mutant myelodysplastic syndromes. J Clin Oncol 2021, 39:1584–1594. [PubMed: 33449813]
82. Cluzeau T, Sebert M, Rahme R, Cuzzubbo S, Lehmann-Che J, Madelaine I, Peterlin P, Beve B, Attalah H, Chermat F, et al. : Eprenetapopt Plus Azacitidine in TP53-Mutated Myelodysplastic Syndromes and Acute Myeloid Leukemia: a Phase II Study by the Groupe Francophone des Myelodysplasies (GFM). J Clin Oncol 2021, 39:1575–1583. [PubMed: 33600210]
83. Zhou L, Cheng X, Connolly BA, Dickman MJ, Hurd PJ, Hornby DP: Zebularine: a novel DNA methylation inhibitor that forms a covalent complex with DNA methyltransferases. J Mol Biol 2002, 321:591–599. [PubMed: 12206775]

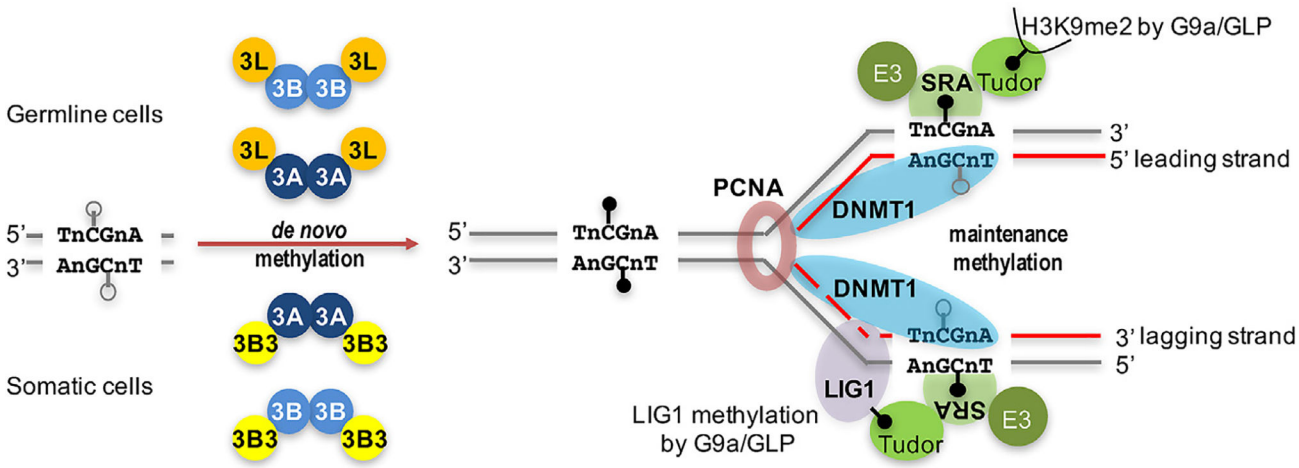

Figure 1. Schematic diagram of *de novo* and maintenance DNA methylation at CpG dinucleotide.

DNMT3A and DNMT3B establish the initial cytosine methylation pattern *de novo*, together with DNMT3L in germline and DNMT3B3 (or DNMT3B6, not shown) in somatic cells. Open lollipops represent unmethylated C, while filled ones represent 5mC. DNMT1 activity at replication foci is regulated at multiple levels to assure the faithful inheritance of genomic methylation patterns. Factors included in this regulation include PCNA, UHRF1, DNA ligase 1 (LIG1), protein lysine methyltransferases G9a/GLP, among others.

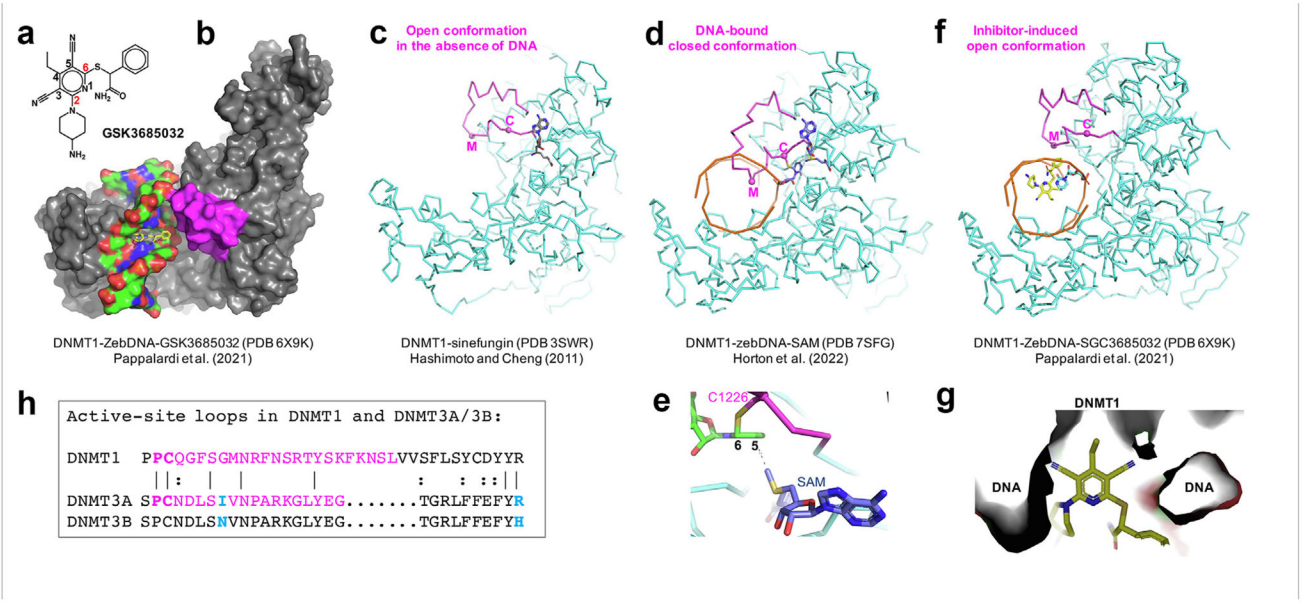

Figure 2. DNMT1-selective, non-nucleoside, dicyanopyridine containing inhibitors.

(**a**) The chemical structure of GSK3685032. The inhibitors all share a 3,5-dicyanopyridine core but vary in their chemical architecture at positions C2 and C6 of the pyridine ring. (**b**) Surface representation of inhibitor-bound DNMT1-DNA structure with the inhibitor shown in yellow and the DNMT1 active-site loop shown in magenta. (**c**) The open conformation of the DNMT1 active-site loop (magenta) in the absence of DNA. Sinefungin is shown in stick model, M, Met1232 and C, Cys1226. (**d**) The closed conformation of the DNMT1 active-site loop bound to DNA (in orange ribbon; the DNA helical axis is perpendicular to the page). The flipped-out cytosine analog zebularine and SAM are shown in stick model. (**e**) The active-site Cys1226 and SAM approach the zebularine [83] ring carbon atoms C6 and C5 from opposite directions, respectively. (**f**) The inhibitor-induced open conformation of the DNMT1 active-site loop bound to DNA. The intra-helical cytosine analog zebularine (cyan) and inhibitor (colored yellow) are shown in stick model. (**g**) The inhibitor penetrates deeply into the DNA and interacts with DNMT1 located on the major groove side. (**h**) Sequence alignment of active-site loops among DNMT1, DNMT3A and DNTM3B. The cysteine in the third position (C1226 in DNMT1) is the active site nucleophile.

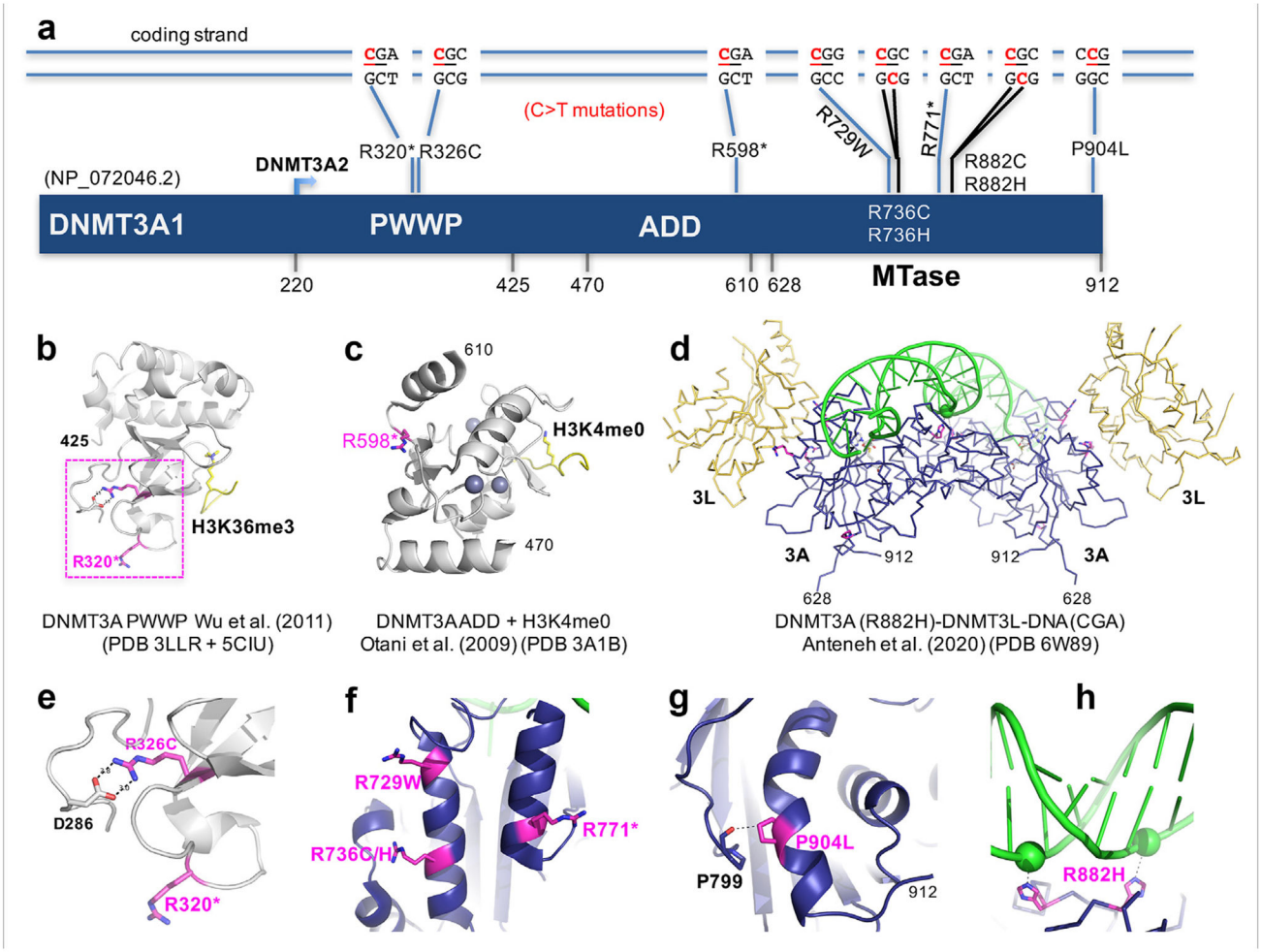

Figure 3. The top 10 most commonly-observed single-nucleotide mutations in DNMT3A

(derived from a study from ~50,000 healthy individuals [30]). (**a**) The C-to-T alterations (in red letters) within CpG sites are indicated for mutated codons. The affected amino acids are depicted in the domain structures of DNMT3A. (**b**) R320 and R326 are located within the PWWP (Pro-Trp-Trp-Pro) domain. (**c**) R598 is located in the ADD Zn-binding domain conserved in ATRX, DNMT3A and DNMT3L. Three Zn atoms are shown in silver balls. (**d**) The catalytic domain of DNMT3A (R882H)-DNMT3L tetramer in complex with DNA. (**e**) R326 forms a salt bridge with D286. (**f**) Viewing from the point of DNMT3L: R729, R736 and R771 of DNMT3A are located in the interface with DNMT3L. (**g**) P904 of DNMT3A is part of a curved helix. (**h**) R882H is located in the DNMT3A dimer interface and interacts with a DNA phosphate.

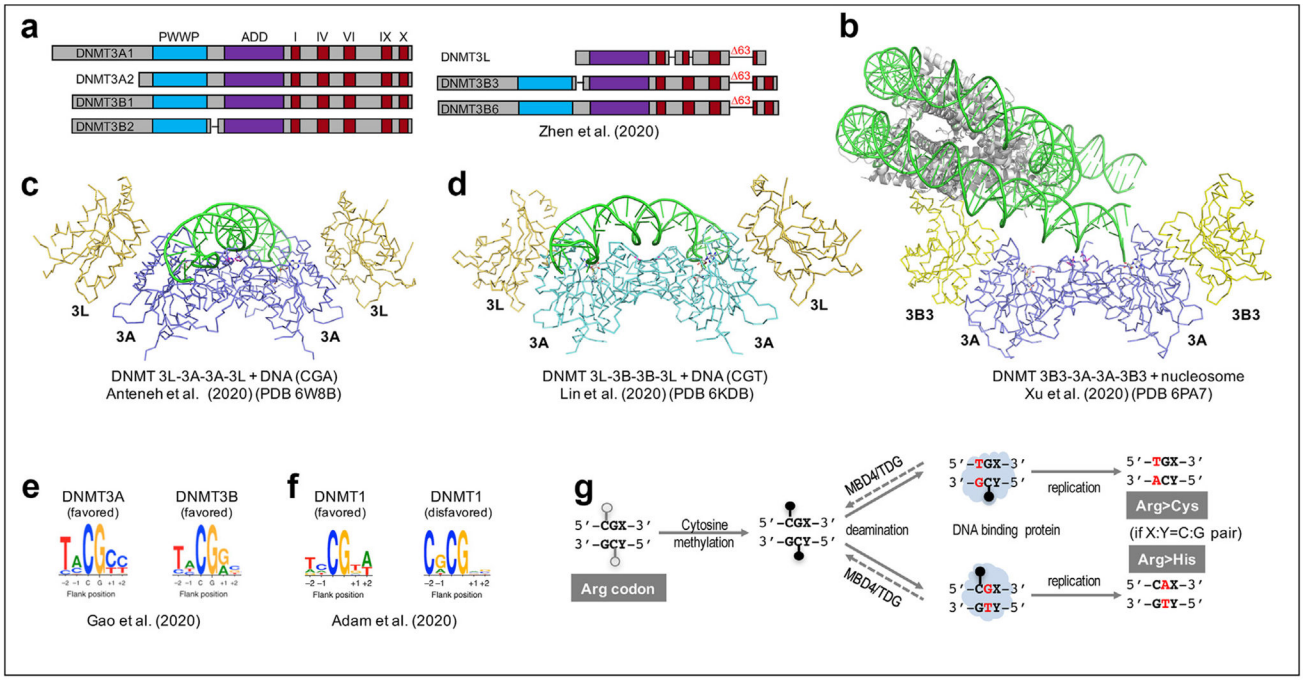

Figure 4. Structures of DNMT3A in various configurations.

(**a**) Representatives of catalytically-active members of the DNMT3 family (3A1, 3A2, 3B1 and 3B2) and inactive members (3L, 3B3 and 3B6). For a sequence comparison see figure S1B of [50]. (**b**) Structure of DNMT3B3–3A-3A-3B3 in complex with nucleosome core particle. DNA is in green, and histones are in gray. (**c**) DNMT3L-3A-3A-3L in complex with naked DNA. (**d**) DNMT3L-3B-3B-3L in complex with naked DNA. We note that the structural examples shown here are only for the C-terminal catalytic domain of 3A and 3B, or the equivalent domains of 3L and 3B3. (**e**) Flanking sequence preferences of DNMT3A and DNMT3B catalytic domains alone. (**f**) Examples of favored and disfavored DNMT1 methylation sites. (**g**) Methylation at an Arg codon. The deamination product of 5mC creates a G:T mismatch. Binding of the mismatch by transcription factors limits access to the damaged site and, if not repaired in time, a round of DNA replication converts a CpG dinucleotide to CpA/TpG.