Mammalian Cytosine DNA Methyltransferase Dnmt1: Enzymatic Mechanism, Novel Mechanism-Based Inhibitors, and RNA-directed DNA Methylation

Željko M. Svedružić*

Department of Biophysics and Biochemistry, Washington State University, Pullman, WA 99164, USA

Abstract: This is a review of the enzymatic mechanism of DNA methyltransferase Dnmt1 and analysis of its implications on regulation of DNA methylation in mammalian cells and design of novel mechanism-based inhibitors. The methylation reaction by Dnmt1 has different phases that depend on DNA substrate and allosteric regulation. Consequently, depending on the phase, the differences in catalytic rates between unmethylated and pre-methylated DNA can vary between 30-40 fold, 3-6 fold or only 1 fold. The allosteric site and the active site can bind different molecules. Allosteric activity depends on DNA sequence, methylation pattern and DNA structure (single stranded *vs*. double stranded). Dnmt1 binds poly(ADP-ribose) and some RNA molecules. The results on kinetic preferences, allosteric activity and binding preference of Dnmt1 are combined together in one comprehensive model mechanism that can address regulation of DNA methylation in cells; namely, inhibition of DNA methylation by poly(ADP-ribose), RNA-directed DNA methylation by methylated and unmethylated non-coding RNA molecules, and transient interactions between Dnmt1 and genomic DNA. Analysis of reaction intermediates showed that equilibrium between base-flipping and base-restacking events can be the key mechanism in control of enzymatic activity. The two events have equal but opposite effect on accumulation of early reaction intermediates and methylation rates. The accumulation of early reaction intermediates can be exploited to improve the current inhibitors of Dnmt1 and achieve inhibition without toxic modifications in genomic DNA. [1,2-dihydropyrimidin-2-one]-5-methylene-(methylsulfonium)-adenosyl is described as the lead compound.

Keywords: S-Adenosyl Methionine analogues, Epigenetics, non-coding RNA, Zebularine, mechanism-based inhibition, Cancer Chemotherapy, Antiviral Therapy

Cell differentiation [1], oncogenic transformation [2], viral infection [3], and long term memory [4] are some of many physiological processes that have created a large interest in DNA methylation. DNA methylation is one of the first steps in epigenetic regulation [5]. Epigenetic regulation is a sequence of molecular events that controls chromatin structure and supports functional organization of eukaryotic genome through cell divisions [6]. DNA methylation is attractive as a fundamental mechanism in functional organization of the human genome, and also as a promising target in development of new drugs for cancer chemotherapy [7, 8], suppression of viral infections [3] and possibly mental illness [9]. DNA methylation in mammalian cells is usually associated with five methyltransferases: Dnmt1, Dnmt2, Dnmt3a, Dnmt3b and Dnmt3L. This manuscript will cover Dnmt1, the most predominant DNA methyltransferase in the mammalian cells, and the major target of the current pharmaceutical interest.

Dnmt1 is crucial for keeping genomic integrity in higher eukaryotes [6] by preserving existing methylation patterns or by participating in the creation of new methylation patterns. The molecular mechanism that guides these two types of DNA methylation events represent the two most basic questions in DNA methylation research: how new methylation sites are created, and how existing DNA methylation patterns are faithfully propagated through cell divisions. In this manuscript we will analyze how enzymatic features of Dnmt1 can lead to different types of DNA methylation events in cells. We will also describe enzymatic features and experimental challenges that make Dnmt1 a unique enzyme. We will show how new insights in catalytic mechanism of Dnmt1 can lead to novel design of mechanism-based (suicide) inhibitors.

In parallel with description of enzymatic properties of Dnmt1, we are also interested in interactions between Dnmt1 and other molecules that participate in DNA methylation. The last 10 years have been devoted to the search for proteins that interact with Dnmt1 as a part of the mechanism that controls DNA methylation in cells. A number of proteins that bind to Dnmt1 has been identified [10]. Briefly, Dnmt1 interacts with: *i*) proteins found at DNA replication forks (such as PCNA [11]); *ii*) proteins involved in chromatin organization (such as HDAC1/2 [12, 13], MeCp2 [14], polycomb proteins [15], or UHRF1 protein [16]); *iii*) proteins involved in cell cycle regulation or response to DNA damage (such as p21(WAF) [11], Rb protein [17], p53 protein [18], PARP1 [19]); *iv*) with RNA polymerase II [20] and with some RNA binding proteins [21]. Interestingly, this impressive array of interacting proteins can not explain the mechanism that forms the versatile patterns of DNA methylation that can be observed in different cells.

The interaction between Dnmt1 and different non-coding RNA molecules could be the missing functional feature of Dnmt1 that can explain formation of different patterns of DNA methylation. Studies of the human genome showed that human complexity is based on a surprisingly elaborate control of gene expression and an unexpectedly small number of protein coding genes [22, 23]. In E. coli approximately 98% of the genome is coding for proteins, in humans only about 2% [22]. A large part of the human genome is coding for different types of regulatory RNA molecules. Those RNA molecules orchestrate gene expression through direct interaction with DNA, messenger RNA or chromatin proteins [22-24]. Interestingly, there are increasing evidences that non-coding RNA molecules could control DNA methylation in cells. A very stable Dnmt1-RNA complex can be found in cell extracts [25, 26], and RNA molecules can modulate catalytic activity of Dnmt1 [25]. Dnmt1 interacts in cells with several RNA binding proteins [21] and with RNA polymerase II [20]. RNA polymerase II is involved in synthesis of non-coding RNA molecules that regulate epigenetic silencing [27]. RNA mediated DNA methylation was shown on several occasions in plant cells [28, 29] and possibly in mammalian cells [27, 30, 31].

Our ability to understand how interaction between Dnmt1 and different proteins or non-coding RNA molecules affects DNA methylation depends on our ability to understand the enzymatic properties of Dnmt1. A nice general description of Dnmt1 was presented in the earlier reviews [5, 10, 32]. Briefly, Dnmt1 methylates DNA at CG sites. Dnmt1 is a large single polypeptide about 1600 amino acids long; its precise length depends on the species and the reading frame in specific cell type. The full-length Dnmt1 gene has evolved as a fusion product of at least three different genes [33].

*Address correspondence to this author at the Department of Biophysics and Biochemistry, Washington State University, Pullman, WA 99164, USA; Tel: 509-335-7457; Fax: 509-335-9688; E-mail: zeljko@wsu.edu

After December 2007: Center for Human Genetics, Campus Gasthuisberg Hospital, KU Leuven and Flanders Interuniversity Institute for Biotechnology (VIB). Herestraat 49, 3000 Leuven, Belgium; Tel: 32-16-346-371; Fax:+32-16-347-181; E-mail: Zeljko.Svedruzic@med.kuleuven.be

Dnmt1 is generally divided into a small catalytic domain at the C terminal and a large regulatory domain at the N terminal [34]. The two domains are connected by 12 Gly-Lys repeats. The catalytic domain contains sequence motifs found in other AdoMet binding proteins and sequence motifs found in the active site of other carbon 5 pyrimidine methyltransferases (generally known as motifs I-IX [35]). It is unclear to what extent the target recognition domain (i.e. target base binding domain) is a part of catalytic domain [36]. Multiple attempts to express the catalytic domain resulted in proteins that do not show enzymatic activity [10, 37]; however there is a possibility that an active catalytic domain could be prepared by a limited proteolysis of full-length Dnmt1 [34]. The large N-terminal domain contains a number of sequence motifs with specific functions [5, 10]. Briefly, the large N-terminal domain contains a poly-bromo domain generally found in chromatin binding proteins [38], a cysteine rich Zn-finger [34], a phosporylation site at Ser 514 [39] that is located within the nuclear targeting domain [40], an Rb protein binding domain [41], and a PCNA binding sequence [11]. A number of N terminal fragments have been expressed and purified [37], however at this moment we do no know how to ascertain their function. It appears that the first 580 amino acids on the N terminal domain are not crucial for catalytic activity [37, 42]. Interestingly, the fragment lacking the first 501 amino acids shows 3-10 fold faster catalytic rates than wild type Dnmt1 with all DNA substrates [42]. Thus, the N-terminal (regulatory) domain inhibits Dnmt1 activity with all DNA substrates. The enzymatic studies on mammalian Dnmt1 used murine [43-50] and human enzymes [42, 51, 52], the two enzymes share 78% sequence identity. The two enzymes have very similar catalytic rates [47], and some shared features in the allosteric regulation [47, 52], possible specific differences between the two enzymes remain to be identified.

SUBSTRATE SELECTIVITY AND ALLOSTERIC REGULATION OF DNMT1

Traditionally enzyme affinity for other molecules is described in terms of binding preference (i.e. Kd, Ka or Ki) and kinetic preference (i.e. kcat/Km) [53]. Both, binding and kinetics preferences have to be considered in Dnmt1 studies since initial results showed that Dnmt1 has a high binding affinity for some regulatory molecules that are not substrates in methylation reaction [25, 44, 49, 51]. For DNA molecules that can be methylated by Dnmt1, the highest methylation rates do not correlate with the highest binding affinity [43, 45]. It appears that the binding preference and the kinetic preference depend on a combination of DNA sequence, methylation pattern and DNA structure (i.e. single strand vs. double strand). In the next few paragraphs we will summarize results from different Dnmt1 studies that analyzed how DNA sequence, methylation pattern, and structure affect DNA binding affinity and catalytic properties of Dnmt1. We will also present a model mechanism that is most consistent with experimental results, and we will discuss possible implications of the presented model mechanism to DNA methylation in cells. In this section we will concentrate only on a descriptive analysis of Dnmt1 interaction with different DNA molecules. The quantitative analysis will be presented in the last section of this manuscript together with necessary precautions.

Binding Studies with Dnmt1

Binding affinity for specific DNA sequences was first measured in studies that used short oligonucleotides that can bind only one Dnmt1 molecule [43]. The binding studies with double strand unmethylated DNA showed that Dnmt1 has a noticeable preference for CG-rich substrates [43]. The binding preference for CG-rich substrates does not translate into kinetic preference, methylation rates are slightly slower when the target CG site is imbedded in a CG-rich sequence relative to an AT-rich sequence [45]. Dnmt1 also binds a number of other molecules that can inhibit its catalytic activity but can not be its substrate in the methylation reaction.

Namely, Dnmt1 binds poly(G), poly(dG)-poly(dC), poly(dA)-poly(dT) [25], poly(dA-dT) [48], poly (ADP-ribose) [19], fully methylated double strand DNA [49, 51], and methylated single strand oligonucleotide [44]. Most of those binding interactions have not been characterized quantitatively so we can not organize them in the terms of binding selectivity. However initial results suggested that DNA binding affinity by Dnmt1 can vary by several orders of magnitude [43-45], depending on DNA length, sequence, methylation pattern, and structure (i.e. single strand vs. double strand). Measuring interactions with low affinity substrates can be technically challenging [43].

Preexisting DNA Methylation and Dnmt1 Activity

Since the beginning of Dnmt1 research it has been known that Dnmt1 shows higher catalytic rates with pre-methylated DNA relative to unmethylated DNA [54, 55]. Dnmt1 reaction with unmethylated and different pre-methylated DNA is still analyzed to this day; discrimination between the two DNA substrates is viewed as a part of the mechanism that separates maintenance methylation from *de novo* methylation. Maintenance methylation and *de novo* methylation represent the two most basic questions in DNA methylation research: how new methylation sites are created and what is the mechanism that drives faithful propagation of existing methylation sites through cell divisions. The current challenge is to organize different Dnmt1 studies with pre-methylated and unmethylated DNA in one comprehensive model mechanism.

It is striking to see how some specific changes in DNA methylation and DNA structure have very profound effects on Dnmt1 activity (Figs. **1** and **2**). A single stranded 30 bp long oligonucleotide with one ^{5m}CG site and a specific CG rich sequence is a potent inhibitor of mouse Dnmt1 [44]. The inhibition is a genuine result of a single strand DNA, since based on its sequence the inhibitory oligonucleotide is unlikely to form higher order structures. Dnmt1 can not methylate the inhibitory oligonucleotide, rather the high affinity interaction results in a potent allosteric inhibition of methylation reaction on different DNA substrates [44, 48]. Interestingly, the high binding affinity and the potent inhibition can be observed only when the inhibitory oligonucleotide is a single stranded DNA with one ^{5m}CG site. When the ^{5m}CG site is replaced with a CG site the result is a CG rich single stranded oligonucleotide that has orders of magnitude lower binding affinity and exceptionally slow methylation rates [43, 44].

Other similar changes can convert the methylated single stranded oligonucleotide from a potent inhibitor to different double stranded oligonucleotides that actually support DNA methylation, namely hemimethylated dsDNA, unmethylated dsDNA or fully methylated dsDNA. If the inhibitory single strand oligonucleotide is paired to be a double strand DNA, the result is a hemimethylated DNA. Hemimethylated DNA has some of the highest methylation rates [45, 51], no allosteric inhibition [47, 54], high processivity [48, 50], and relatively moderate binding affinity for Dnmt1 [45]. Second, two changes can convert the inhibitor oligonucleotide to a double stranded unmethylated DNA, ^{5m}CG site can be replaced with a CG site and the single strand oligonucleotide can be paired to be a double strand DNA. Unmethylated double strand DNA shows relatively slow catalytic rates [45, 51], a partial allosteric inhibition [44, 47, 54], and a relatively moderate binding affinity [43-45, 47]. As the last option, inhibitory oligonucleotide can be paired with a complementary DNA strand carrying a ^{5m}CG site to form a double stranded fully methylated DNA. A fully methylated oligonucleotide is not a substrate for Dnmt1, rather in the presence of an excess of fully methylated oligonucleotide, Dnmt1 reaction on unmethylated DNA becomes very similar to Dnmt1 reaction on hemimethylated DNA (Fig. **1** in [49]).

The binding selectivity by Dnmt1 can be attributed to different functions of its allosteric site (Fig. **1**), though some key measure-

ments are still needed to understand the actual mechanism. First, the sequence specificity of these interactions has to be further determined. Second, based on the current data it is not clear whether higher activity on unmethylated DNA in the presence of an excess of fully methylated DNA [49, 51] is due to allosteric activation or diminished allosteric inhibition by unmethylated DNA. The later appears more likely. When an excess of fully methylated double stranded DNA is added to a methylation reaction on unmethylated DNA substrate, the resulting reaction shows close similarities to methylation reaction on hemimethylated DNA substrate (Fig. **1** in [49]). A hemimethylated DNA shows higher methylation rates than unmethylated DNA (in part?) due to lack of allosteric inhibition [44, 45, 47, 54].
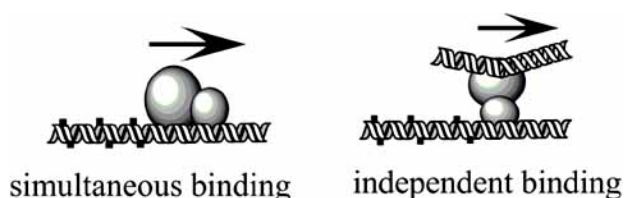

Fig. (1). DNA binding at the active site (small oval) and the allosteric site (large oval) on Dnmt1 (the arrows indicate ongoing methylation reaction, the squares indicate new methylation sites). In a case of simultaneous binding, DNA binding at the active site leads to obligatory binding of the adjacent DNA site at the enzyme's allosteric site. The active site is methylating DNA, the allosteric site is sensing existing methylation. One example of such mechanism is DNA polymerase were the catalytic domain and the proofreading domain act simultaneously in creating and proofreading nascent DNA strands. In the case of independent binding, DNA binding at the active site does not necessarily lead to DNA binding at the allosteric site. The allosteric site and the active site do not need to bind adjacent DNA sites, and allosteric regulation can come from completely different DNA molecule, RNA molecule, or some distant site on the same DNA molecule. The model mechanism with two independent DNA binding sites is most consistent with experimental data. Namely, presented model can explain regulation of Dnmt1 activity by DNA molecules than can not be substrates for Dnmt1, like single strand oligo nucleotide with one ^{5m}C site [44], double strand fully methylated DNA [49, 51], or poly(dA)-poly(dT), poly(G), poly(dC)-poly(dG) [25], or poly(dA-dT) [48]. The model mechanism with two independent DNA binding site can also explain uncompetitive inhibition with single strand oligonucleotide [44], and pulse-chase processivity experiments where different DNA molecules added as a chase have different effects on ongoing processivity at the active site [48]. Finally, the model mechanism with two independent sites can explain why changing the distance between target CG site and proximal ^{5m}CG site from 3 to 18 bases did not affect measured enzyme activity [124]. Simply, the target CG site binds at active site while ^{5m}CG site from another DNA molecule binds at the allosteric site making the distance between ^{5m}CG site and the target CG site irrelevant. Interestingly, a complex between Dnmt1 and two DNA molecules has never been observed in binding studies [43]. This is due to limited research on this problem, but also, there is a possibility that such complex is only transient and its existence can be detected only by kinetics studies [47].

The presented examples also indicated that we have to look at differences between pre-methylated and unmethylated DNA, rather than differences between hemimethylated and unmethylated substrates as it is often presented in the literature (Fig. **1**). Premethylated DNA would include hemimethylated dsDNA, proximally methylated dsDNA, fully methylated dsDNA, and methylated ssDNA. In hemimethylated DNA the target CG site is paired with ^{5m}CG site; Dnmt1 methylates hemimethylated sites during DNA replication in order to propagate methylation patterns from the old DNA strand to the new DNA strand (i.e. maintenance methylation). Proximally methylated substrates are not well defined, a loose definition could be: DNA substrates where the target CG site and ^{5m}CG site are positioned within 3 to 30 base pairs (what could be one

Dnmt1 footprint on its DNA substrate based on oligo affinity studies [45]). Dnmt1 interaction with proximally methylated dsDNA, fully methylated dsDNA and methylated ssDNA are different functions of its allosteric site (Fig. **1**). These allosteric interactions make Dnmt1 more responsive to pre-methylated DNA in comparison to the bacterial enzymes that can recognize only the difference between unmethylated and hemimethylated substrates. In summary, methylation reaction by Dnmt1 can take place on three different DNA substrates (unmethylated, hemimethylated and pre-methylated dsDNA); and each substrate can be combined with one of four different allosteric regulators (fully methylated dsDNA, hemimethylated dsDNA, unmethylated dsDNA, methylated ssDNA). We need to specify the physiological importance for each of these interactions.

Model Mechanism for DNA Methylation by Dnmt1 in Mammalian Cells

The presented mechanism (Figs. **1** and **2**) can be used to construct a model for Dnmt1 activity in cells. The best explanation for the presented experimental data is a model mechanism where the active site and the allosteric site bind DNA independently (Fig. **1**). Precisely, the active site binds the substrate DNA and the allosteric site binds the "regulatory DNA molecules". The "regulatory DNA molecules" control progress of ongoing catalysis at the active site (Fig. **2**). The progress of ongoing catalytic activity depends on DNA bound at the active site and on DNA bound at the allosteric site (Fig. **2**). In process of DNA methylation Dnmt1 is moving over genomic DNA as a part of the replication fork [11], or as a part of some other DNA-methylating complex. The catalytic domain is bound to the substrate DNA while its activity is regulated through the allosteric site that is open for interaction with different regulatory molecules (Fig. **2**). Allosteric regulation can support catalytic activity and a high processivity (Fig. **2**, complex 1b or 2b), or it can fully or partially inhibit catalytic activity (Fig. **2**, complex 1c, 2c, 1a, 2a) eventually leading to DNA dissociation and lower processivity [48]. Interestingly, the presented mechanism of allosteric regulation is consistent with the recent study showing that DNA methylation in mammalian cells is based on a series of transient interactions between Dnmt1 and genomic DNA [56].

In the introduction we summarized the arguments that in cells non-coding RNA molecules or poly(ADP-ribose) could regulate activity of Dnmt1 by binding at the allosteric site just as described for different DNA molecules (Figs. **1** and **2**). Currently we understand only allosteric regulation by DNA molecules (Fig. **2**), given that biochemical studies to this date used only DNA molecules. For example, Dnmt1 inhibition by poly(dA) and poly(dA-dT) was initially very surprising [25, 48] until it was shown that DNA methylation and Dnmt1 can be inhibited by poly(ADP-ribose) [19]. Poly(ADP-ribose) is a cellular response to DNA damage [57], and it is possible that poly(dA) and poly(dA-dT) mimic poly(ADP-ribose) in their binding to Dnmt1. In cells, Dnmt1 inhibition by poly(ADP-ribose) can prevent DNA methylation and subsequent chromatin compaction. This could assure that DNA damage is accessible and repaired by DNA repair proteins before the damage site is packed in compact chromatin structures [57-59]. Interestingly, it appears that the poly(ADP-ribose) can compete with single stranded and double stranded DNA in binding to Dnmt1 [19].

Similar to the presented control of Dnmt1 activity by DNA molecules (Fig. **2**), non-coding RNA molecules could control catalytic activity of Dnmt1 based on their sequence, methylation pattern, and structure (single strand vs. double strand). The correlation between structure and function of non-coding RNA molecules is described in the literature [27, 28]. The possible methylation of non-coding RNA molecules could be attributed to specific RNA methyltransferases that are an integral part of DNA methylation machinery. Recent studies have shown that Dnmt2 can be an RNA
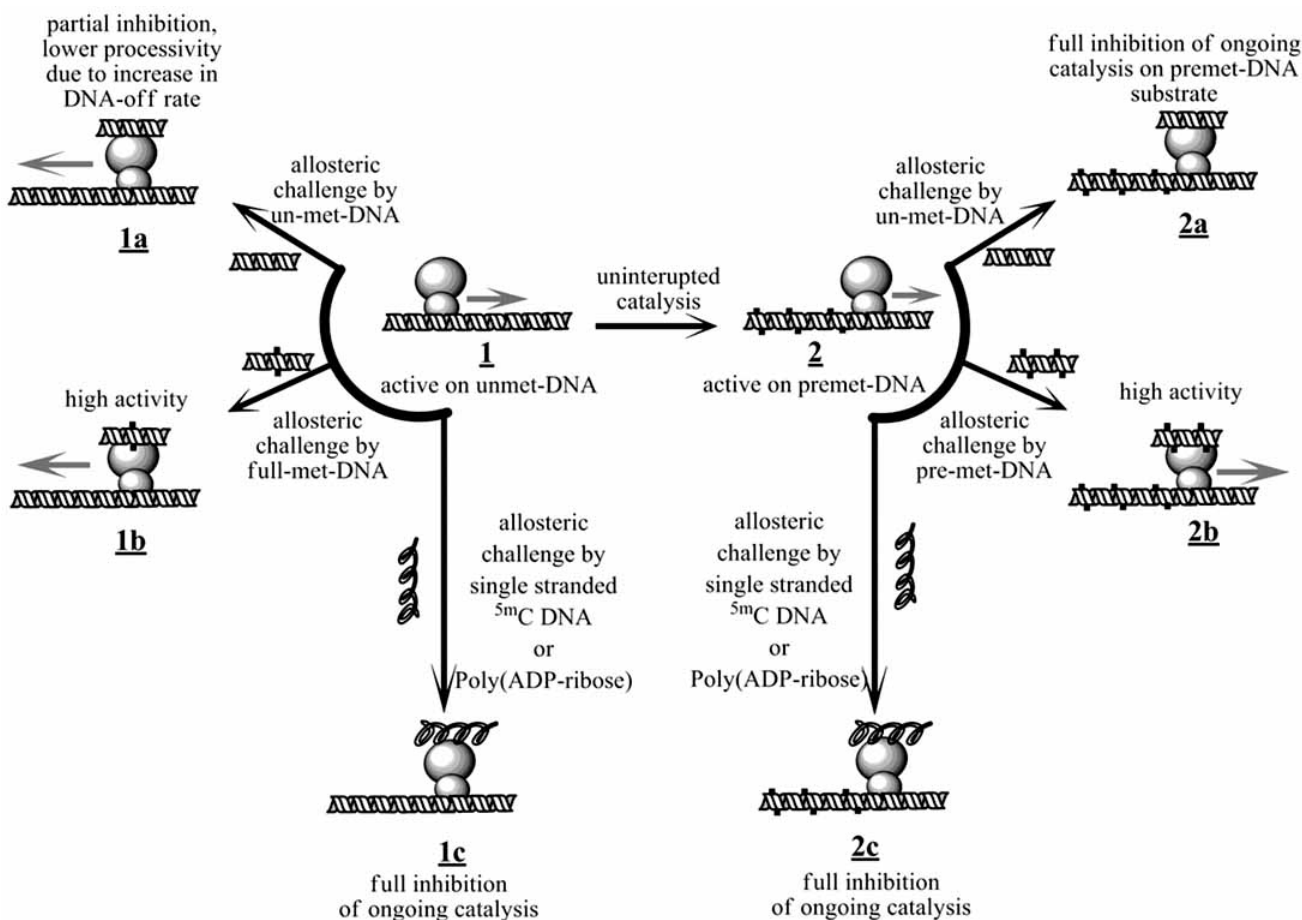
Fig. (2). Allosteric regulation of Dnmt1 (catalytic domain (small oval), regulatory domain (large oval), DNA (helix), ^{5m}C sites (■), ongoing catalysis (gray arrows)). The aim of this figure is to provide a model mechanism that can summarize interactions between Dnmt1 and different DNA molecules, different features of allosteric regulation of Dnmt1, and ultimately regulation of DNA methylation in cells. There are three types of allosteric regulation of Dnmt1;

i) fully inhibition (**1c** and **2c**), such as inhibition by single strand oligo inhibitor with one ^{5m}C site [44], poly(dA-dT) [48] and probably other inhibitory sequences such as poly(ADP-ribose) [19], poly(dA)-poly(dT), poly(dG)-poly(dC), poly(G) [25];

ii) inhibition by unmethylated DNA, such as partial inhibition of ongoing catalysis on unmethylated substrate (**1a**, [44, 47, 48, 54]) and full inhibition of ongoing catalysis on pre-methylated substrates (**2a,** [48]);

iii) allosteric regulation by fully methylated (**1b**, [49, 51]) and pre-methylated DNA (**2b**, [47, 54]).

Every allosteric complex can have different physiological function. The reaction "active on unmet-DNA" (**1**) could represents *de novo* methylation and its different forms of allosteric regulation (**1a**, **1b**, and **1c**). The reaction "active on premet-DNA" (**2**) could represent maintenance methylation and its different forms of allosteric regulation (**2a**, **2b,** and **2c**). For example, inhibition by poly(ADP-ribose) [19] can stop ongoing methylation reaction in a case of DNA damage, a potent inhibition by the single stranded DNA with one 5mC site [44] can also stop ongoing DNA methylation but its trigger mechanism is not known. Interestingly, it appears that a single stranded DNA can compete with poly(ADP-ribose) in binding to Dnmt1 [19]. Complex **2b**, has the highest activity [45, 51] and the highest processivity [50]. Most likely it is a part of maintenance methylation mechanism when the target is a hemimethylated site, or a part of methylation spreading mechanism when the target is a proximally methylated site. Complex **1b**, has activity similar to complex **2b** (Fig. **1** in [49]) and its function can be *de novo* methylation. Both complex **1a** and **2a** can inhibit DNA methylation [44, 47, 48] but not with the high potency that can be observed in the presence of ssDNA with one ^{5m}C site [43-45]. With complex **1a** there is only a partial inhibition, partitioning between two enzyme forms (see (Fig. **9**) in this text), and an increase in DNA dissociation rate [48]. Nothing is known about complex **2a**, except that initial studies suggested that it is inactive [48]. In sum, the allosteric regulation can support catalytic activity and a high processivity (complex **1b** or **2b**), or it can fully or partially inhibit catalytic activity (Fig. **2**, complex **1c**, **2c**, **1a**, **2a**), possibly leading to DNA dissociation and lower processivity [48]. Interestingly, the presented mechanism of allosteric regulation can be the molecular mechanism that is driving the series of transient interactions between Dnmt1 and genomic DNA during DNA methylation in cells [56].

methyltransferase that is active in cytosol and methylates transfer RNA for aspartate [60]. However, additional biochemical studies of substrate specificity are needed to determine if a large excess of tRNA could have overwhelmed the enzymatic activity with non-coding RNA molecules that are expressed only transiently and (or) only as a small fraction of the total RNA.

It is often difficult to identify and characterize new non-coding RNA molecules due to their "transient and subtle" function [22, 23,

27, 61]. The same challenge can face the future studies of non-coding RNA molecules that control DNA methylation. Interestingly, a very stable Dnmt1-RNA complex is routinely observed in cell extracts during Dnmt1 purification from mouse erithroleuchemia cells [26] or HeLa cells [25]. Dnmt1-RNA complex that is present in cell extracts can be used for the identification of non-coding RNA molecules that control DNA methylation. Dnmt1-RNA complex can be isolated from different cells, or from the same

cells in different stages of the cell cycle. The cells can be different cell lines grown in culture, cells isolated from specific organs, or cells isolated during cell differentiation in embryonic tissues. The prepared RNA molecules can be characterized based on their effect on Dnmt1 activity in terms of their sequence, methylation pattern and structure (single strand vs. double strand). The result of such effort will be an "RNA-methylome-epigenome", a genomic map of non-coding RNA molecules that control DNA methylation and associated epigenetic processes. Different RNA molecule that form the "RNA- methylome-epigenome" could be used in standard transfection protocols to selectively manipulate DNA methylation in cells. Similar strategy was used in the past with the single stranded DNA-oligonucleotide inhibitor of Dnmt1 [44].

TARGET BASE ATTACK, BASE FLIPPING AND NOVEL MECHANISM-BASED (SUICIDE) INHIBITORS

The majority of Dnmt1 studies have analyzed Dnmt1 activity with different DNA substrates and Dnmt1 interaction with other proteins. Very few studies have analyzed catalytic intermediates of Dnmt1. Inhibition by 5-florocytosine [62], ^{3}H exchange reaction [47], and the conserved sequence motifs [35] indicated shared catalytic steps between Dnmt1 and the previously characterized carbon 5 pyrimidine methyltransferases [63]. 5-fluorouracil is a well-known drug that is based on the catalytic mechanism of carbon 5 pyrimidine methyltransferases [64]. 5-fluorouracil is a cell growth inhibitor that specifically targets thymidylate synthase at low concentration and RNA at higher concentration [64]. Similar to 5-fluorouracil, 5-fluorocytosine, 5-aza-cytosine and 1,2-dihydro-pyrimidin-2-one (zebularine) act as mechanism-based (suicide) inhibitors that trap the covalent adduct intermediate (Fig. **3**, **2**) with all cytosine C[5] DNA methyltransferases including Dnmt1 [7, 62]. In contrast to 5-fluorouracil, the cytosine analogues are randomly incorporated in genomic DNA, which leads to excessive DNA damage and a high toxicity [7]. A challenge is to modify the existing cytosine analogues so that DNA methyltransferases can form the covalent adduct intermediate (Fig. **3**, **2**) and undergo suicide inhibition without toxic damages in genomic DNA [65].

Searching for new insights in design of suicide inhibitors recent studies have analyzed formation and stability of the covalent adduct intermediate (Fig. **3**. **2**) with Dnmt1 [47] and bacterial enzyme M.HhaI [66]. Comparative analysis of Dnmt1 and M.HhaI can be a productive approach. Unlike Dnmt1, M.HhaI has been a subject of extensive structural, enzymatic, and computational studies (over 120 studies). Dnmt1 and M.HhaI share many of the fine features in catalytic mechanism but differ in catalytic rates by one to two orders of magnitude [47]. Similar catalytic features and a large difference in catalytic rates can be explained by the initial observations indicating that for both enzymes the early catalytic intermediates accumulate as a "fast equilibrium" prior to the rate limiting methyl-transfer step (Fig. **4A**, [47, 66]).

A number of M.HhaI studies suggested that the early reaction intermediates form a fast equilibrium [66-69]. Recent studies showed that a dynamic equilibrium between base flipping and base restacking events can control specificity of other base flipping enzymes [70, 71]. Interestingly, the full implications of the "fast equilibrium mechanism" on catalytic features of DNA methyltransferases have not been described. Thus, we will describe the properties of a dynamic equilibrium mechanism using numerical simulation and extensive experimental information available for bacterial methyltransferase M.HhaI. The basic principles of "fast equilibrium" kinetics have been described in some of the well-known enzymology and enzyme kinetics textbooks (p. 177 in [53]). Briefly, "accumulation" of each intermediate is determined by the ratio between the "rate of formation" and the "rate of reversal" for all steps that participate in the equilibrium. In the case of DNA methyltransferases, the early reaction intermediates accumulate as a fast equilibrium prior to the rate limiting methyl-transfer step (Fig. **4A**) [47, 66]. Thus, methylation rate simultaneously depends on
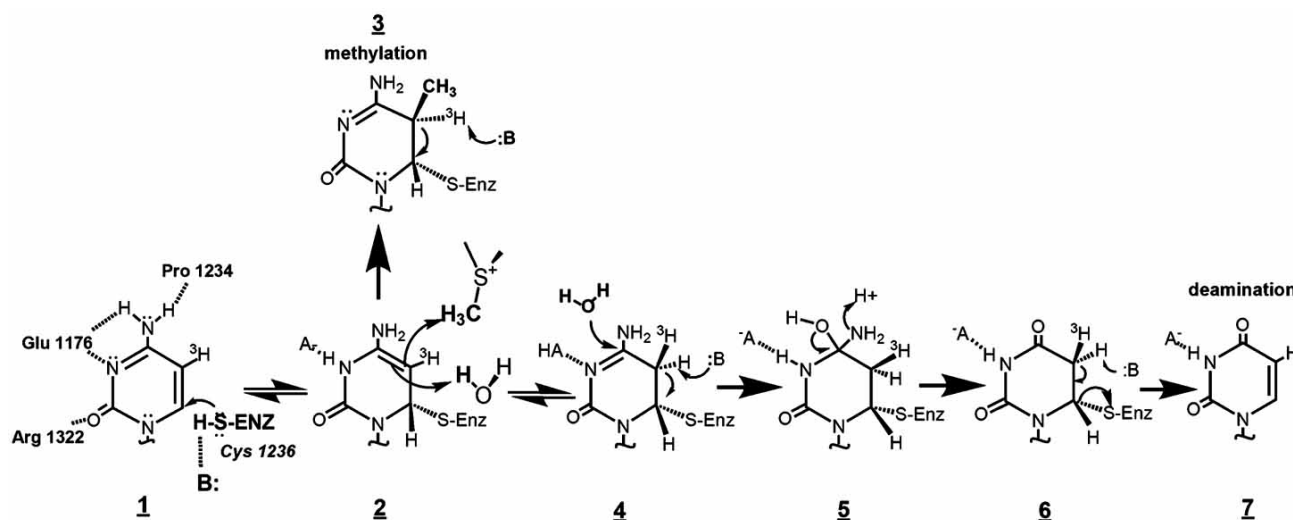

Fig. (3). Cytosine methylation and deamination by cytosine carbon 5 DNA methyltransferases. Inhibition by 5-florocytosine [62], ^{3}H exchange reaction [47], and the conserved sequence motifs [35] indicated shared catalytic steps between Dnmt1 and the previously characterized carbon 5 pyrimidine methyltransferases [63]. The target base forms asymmetric hydrogen bonds with the conserved active site residues (**1**). This facilitates polarization of the target base and a nucleophilic attack by the active site cysteine (**1** -> **2**). The result is unstable covalent adduct intermediate (**2**). The covalent adduct intermediate can rapidly revert back to cytosine (**2** -> **1**) [47], or it can be slowly methylated (**2** -> **3**), or protonated (**2** -> **3**) [47, 91]. Protonation of the target base can increase cytosine deamination (**4** -> **7**) rates by at least four orders of magnitude [63], consequently the enzyme has to protect the activated target base from proton donors to prevent mutagenic conversion of cytosine to uracil (**4** -> **7**) [66, 97]. The pKa of carbon 5 on the activated target base (**2**) can vary between 11 and 18 depending whether nitrogen 3 is protonated or not [90]. Consequently uncontrolled solvent influx in the active site can be the main cause of catalytic deamination [47, 66]. Balancing water access to the activated target base is one of the key functions for cytosine DNA methyltransferases, water is a part of the β-elimination step that follows the methyl-transfer step (**2** -> **3**), but also, water can trigger mutagenic deamination. Initial studies suggested that for both Dnmt1 and bacterial enzyme M.HhaI, carbon 5 on the activated target base is protected from solvent molecules by the cofactor and by enzyme-DNA interaction (as exemplified by the active site loop in M.HhaI) [47, 66].
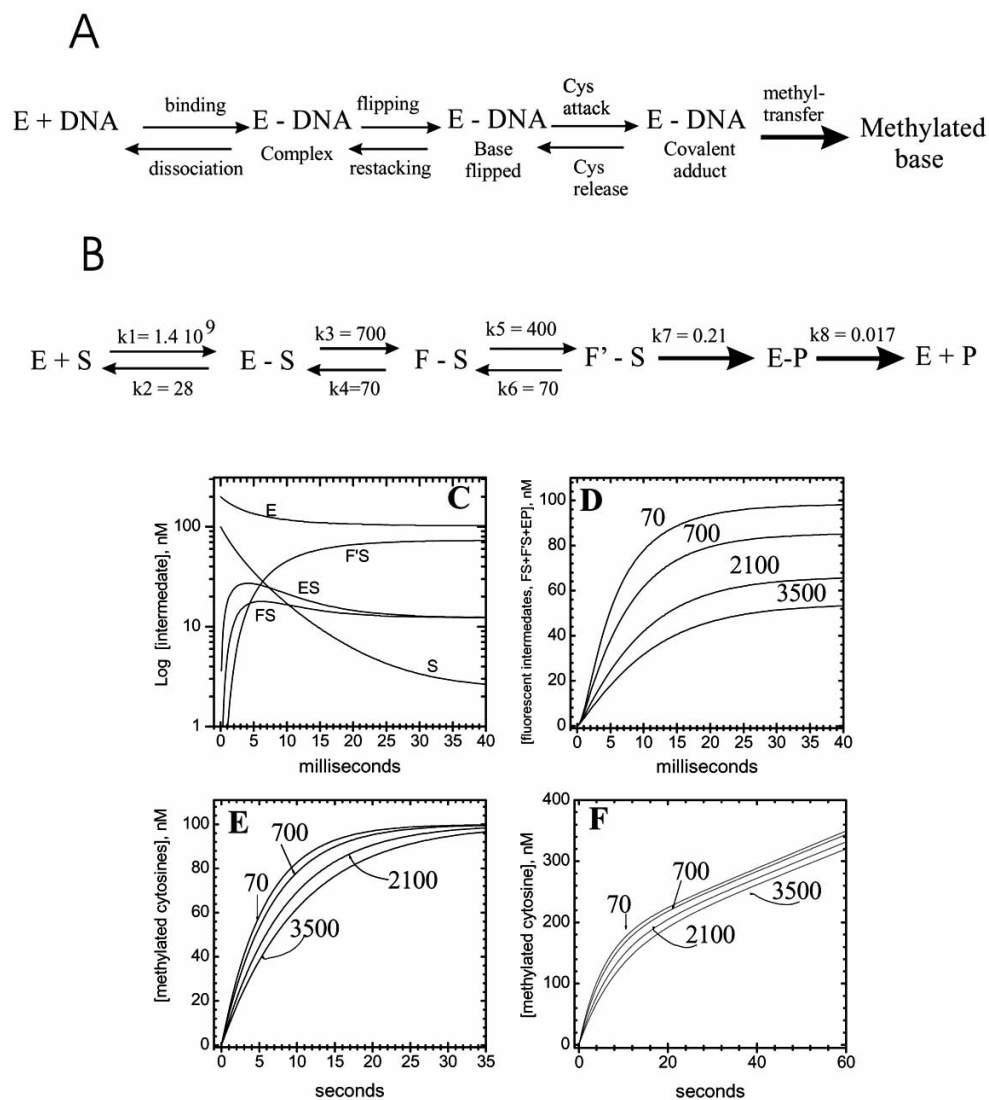
Fig. (4) <u>A</u>-<u>F</u>). The properties of a dynamic equilibrium between early catalytic intermediates can be illustrated using numerical simulations [75] and available experimental data measured with the bacterial enzyme M.HhaI [66, 69, 73, 105]. (**A**) DNA methyltransferases have four major catalytic intermediates that accumulate as a fast equilibrium prior to the slow methyl-transfer step [47, 66]. (**B**) Catalytic intermediates can be described with symbols and individual steps can be assigned rate constants in sec^{-1} (for k2 to k8) or M^{-1}sec^{-1} (for k1). The simulation used catalytic rates measured with bacterial enzyme M.HhaI [66, 69, 73, 105] to take advantage of numerous kinetic, structural and computational studies of M.HhaI. Initial studies suggested that both M.HhaI and Dnmt1 depend on the fast equilibrium mechanism [47]. We simulated time profiles for early reaction intermediates (<u>**C**</u>), and for the three most common experiments, namely: (<u>**D**</u>) base flipping studies with fluorescence base analogues [68]; (<u>**E**</u>) single turnover methylation reaction [73, 105]; and (<u>**F**</u>) pre-steady state burst in methylation reaction [66, 69, 73, 105]. In all simulations the rate constants were kept as indicated in panel **B**, except that in panels **D** to **F** the base restacking rate varied between 70, 700, 2100, 3500 sec^{-1}.

Simulation shows that the early catalytic intermediates will equilibrate within 30 milliseconds (<u>**C-D**</u>), while the pre-steady state methylation step is completed in 30 sec (<u>**E-F**</u>), or thousand times slower. Most importantly, regardless of thousand-fold difference in the relative rates, a change in base restacking rates leads to a change in methylation rates. The change in methylation rates is almost proportional to the change in base restacking rate as long as base flipping and base restacking rates are in equilibrium (i.e. changing the base restacking rate from 700 to 2100 sec-1, gives bigger response in methylation rates than changing the rates from 700 to 70 sec-1 or from 2100 to 3500 sec-1). A slower base restacking rates lead to a faster pre-steady methylation rates due to more favorable accumulation of the early reaction intermediates (FS and F'S) (<u>**D**</u>). Thus, base restacking can affect methylation rates, even though base restacking is considered to be a product release step, and even though there was no change in base flipping rate.

The rate constant for methyl-transfer step (k7) was chosen to be 0.21 sec^{-1} (750 h^{-1}) based on the average values from pre-steady state studies of M.HhaI [73, 105]. The base flipping rate constant (k3) was chosen to be similar to the rates observed in "base flipping" studies with fluorescent base analogues [68]. Notice that "base flipping" studies with fluorescent base analogues do not measure the actual base flipping rate (k3), but the rates of equilibration between intermediates with intra-helical and extra-helical base (i.e. [ES] vs. [FS]+[F'S]+[EP]). The actual "base flipping" rates can be much higher than the observed change in fluorescence and any process that can change the rate of equilibration between different intermediates can appear as an increase or a decrease in the "base-flipping" rate (<u>**D**</u>). Thus, all early rate constants (k1 to k6) were chosen to be comparable to the rates measured in fluorescence "base flipping" studies. Also, to achieve equilibrium between early reaction intermediates the difference between forward and back step has to be always less than an order of magnitude. For panels (<u>**C-E**</u>) the chosen enzyme and DNA concentration correspond to single turnover experiments; [E] = 200 nM, [DNA] =100 nM, Kd(E-DNA)= 20 nM = (k$_{-1}$/k$_1$). For panel (<u>**F**</u>) the chosen concentrations correspond to pre-steady state burst experiments [E] = 200 nM, [DNA] =1000 nM, Kd (E-DNA)= 20 nM = (k$_{-1}$/k$_1$).

"concentration of the accumulated reaction intermediates" and on the rate-limiting methyl-transfer step (i.e. methylation rate constant = methyl-transfer rate constant multiplied by concentration of intermediate **2** (Fig. **3**)). Notice the difference between methylation rate (i.e. formation of ^{5m}C in time) and methyl-transfer rate (Fig. **3** **2**->**3**). Also notice that "concentration of the accumulated reaction intermediates" and the rate limiting "methyl-transfer step" can affect the methylation rates independently.

The idea that there is a dual control of catalytic activity is supported by experimental observations. For example, different M.HhaI studies showed that small changes in base flipping mechanism result in proportional changes in methylation rate [68, 69, 72, 73]. Thus, the base flipping events can affect catalytic rates even though there is a general consensus that the base flipping is orders of magnitude faster than methylation reaction [68, 69, 74]! Furthermore, similarities between Dnmt1 and M.HhaI appear to be a result of a conserved catalytic mechanism in the active site (Fig. **3**, [47]), while large differences in catalytic rates appear to be a result of unexplored differences in accumulation of early catalytic intermediates (i.e. base flipping/base restacking steps, Fig. **4A**).

Numerical analysis [75] can be used to illustrate how changes in any of the equilibrium steps affect catalytic activity (Fig. **4**). Analysis of base restacking is a very good illustration of the basic principles of the equilibrium mechanism. Studies usually concentrate on base flipping events and base re-stacking is considered as a product release step that is not involved in the actual target base attack. However, if there is a fast equilibrium between base flipping and base restacking steps, those two steps have opposite effects on methylation rates that are equal in magnitude (Fig. **4D**-**F**). The results of numerical analysis are consistent with available experimental data (Fig. **5**). More of such examples are likely to come in the future as base flipping and base restacking steps receive equal attention.

Novel Mechanism-Based Inhibitors Can Mimic Accumulation of Early Reaction Intermediates

Accumulation of early catalytic intermediates in a fast equilibrium represents a new opportunity for design of mechanism-based (suicide) inhibitors. The new mechanism-based inhibitors can be designed to look like intermediates in the second (slow) phase of catalysis (Fig. **4A**). Thus, the suicide inhibition steps can be separated from the fast DNA binding steps and inhibition can be achieved without toxic modifications in genomic DNA. 1,2-dihydropyrimidin-2-one-5-methylene-(methylsulfonium)-adenosyl (Fig. **6**, **1**) is a structure designed with a desire to initiate suicide inhibition in the absence of DNA. When positioned in the active site, the 1,2-dihydropyrimidin-2-one ring (or shortly 2-pyrimidinone ring) can form hydrogen bonds with the conserved Glu and Asp residues and these interactions can initiate nucleophilic attack by the active site cysteine (Fig. **6**, **2** -> **3** -> **4** and [76, 77]). Unlike 5-fluorocytosine and 5-aza-cytosine, 2-pyrimidinone does not require methyl-transfer for enzyme entrapment [76-78]. Thus, the 2-pyrimidinone ring is attached at the carbon 5 to methylene-methylsulfonium-adenosyl part of AdoMet (Fig. **6**, **1**). Based on conserved sequence motifs [35] and catalytic similarities between M.HhaI and Dnmt1 [47], we can predict that the adenosyl part of AdoMet could position the 2-pyrimidinone ring in the enzyme active site. Precisely, a number of M.HhaI studies [47, 66, 79-82] and initial Dnmt1 studies [47] showed that adenosyl derivatives can interact with cytosine methyltransferases in the same orientation as the corresponding parts in AdoMet [83]. Adenosyl part of AdoMet forms 5 out of 8 conserved hydrogen bonds between AdoMet and DNA methyltransferases [83], 2-pyrimidinone can form at least 3 hydrogen bonds with the conserved active site residues (Fig. **6** and [77]). That is total of 8 hydrogen bonds between the enzyme and the inhibitor. The presented inhibitor looks like

transition state intermediate in the methyl-transfer step (Fig. **3**, **2**->**3**); i.e. the 2-pyrimidinone ring is positioned as the "flipped-out" target base and the methylsulfonium-adenosyl part is positioned as methylsulfonium-adenosyl parts of AdoMet. Interestingly, presented inhibitor is somewhat structurally similar to earlier described inhibitors of AdoMet decarboxylase [84-86].
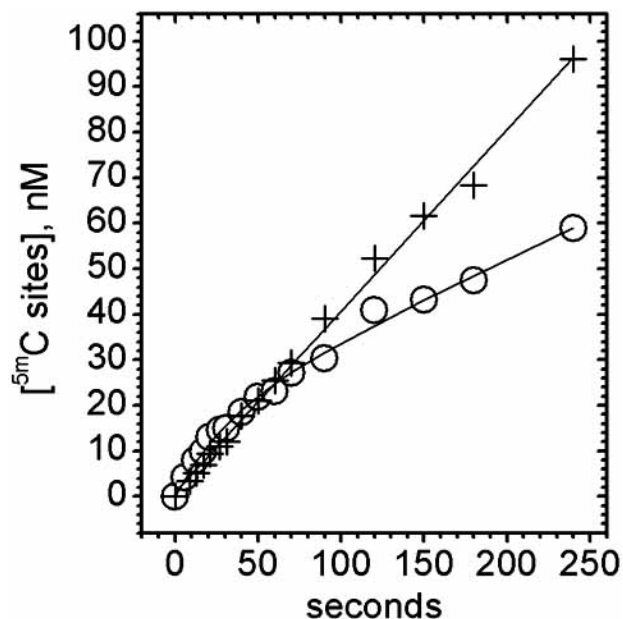

Fig. (5). Pre-steady state burst in methylation reaction with bacterial enzyme M.HhaI and poly(dG-dC) (O) **and poly(dI-dC)** (+) **substrates.** Pre-steady state burst can be used to show how base restacking and a fast equilibrium between early catalytic intermediates control methylation rates. M.HhaI shows a pre-steady state burst in methylation reaction [73, 105], the pre-steady state burst phase represents steps leading to methyl-transfer step (Fig. **4A**) and the subsequent linear steady state phase represents product release steps (i.e. AdoHcy release [108], base restacking and loop opening [66, 69]). Studies that analyzed interactions at the active site loop [66, 69] showed that a decrease in stability of the active site loop will lead to a decrease in catalytic rates; interestingly the rate decrease affects pre-steady state burst phase more than the slow product release phase [66, 69]. As one example, here we show pre-steady state burst in M.HhaI reaction with poly(dG-dC) (●) and poly(dI-dC) (o). In comparison to GCGC sequence, ICIC sequence has unstable active site loop [66] due to missing hydrogen bond between Ile86 on the active site loop and G that is in 5' position relative to target C [125]. Interestingly, when compared to GCGC sequence, ICIC sequence shows slower pre-steady state rates (65 ± 8 h^{-1} *vs.* 140 ± 20 h^{-1}) but faster steady state rates (65 ± 8 h^{-1} vs. 40 ± 4 h^{-1}). Thus, changes in the active site loop and base restacking affect the start of methylation reaction (Fig. **4A**) even though those two events are part of product release. That is one of the expected features in mechanism that is controlled by a fast equilibrium between base restacking and base flipping rates (Fig. **4C-F**). Methylation rates were measured using 21 nM M.HhaI, 10 μM ^{3}H-AdoMet (68 000 cpms/pmol) and 10 μM bp poly(dG-dC) or poly(dI-dC). Additional details about reaction conditions are given in the reference [66].

If necessary the proposed inhibitor can be modified to improve its specificity, binding affinity, reactivity and overall pharmacokinetic properties. For example, as with AdoMet decarboxylase inhibitors, methyl-sulfonium part in presented inhibitor can be replaced with sulfo (–S-), methylene (-CH$_2$-), amino (-NH-) or oxy (-O-) group. These modifications can change the binding affinity and the rate of inactivation by: *i*) changing interaction between the 2-pyrimidinone ring and the conserved residues in the active site; and by *ii*) changing electronic structure of the 2-pyrimidinone ring to alter its susceptibility to nucleophilic attack by the active site cys-

teine (Fig. **6**, **2** -> **3** -> **4**). Furthermore, short chain fragments can be attached at nitrogen 1 of 2-pyrimidinone ring to increase binding specificity and affinity. These fragments would mimic the glycosidic part(s) on DNA as suggested in a recent study of uracil-glycosylase inhibitors [87]. Precisely, AdoMet is the most widely used cofactor in cells after ATP [88], and methyl-sulfonium-adenosyl parts in the presented inhibitor could bind to every protein that has the conserved AdoMet binding domain. In the proposed inhibitor, the specificity for cytosine DNA methyltransferase comes from: *i*) the suicide inhibition steps (Fig. **6**, **2** ->**3** -> **4**) that can take place only when the 2-pyrimidinone ring forms hydrogen bonds with the conserved active site residues that are found only in cytosine methyltransferases; and from *ii*) steric hindrance that can be caused by 2-pyrimidinone ring (and its derivatives) when the adenosyl part binds to other AdoMet dependent enzymes. Adding specific short chain fragments at nitrogen 1 of 2-pyrimidinone ring can provide additional steric interference that can prevent nonspecific interactions, and also, in some cases it can further increase reactivity of the suicide inhibition steps (Fig. **6**, **2**-> **3** -> **4**). Finally, additional increase in binding specificity can be achieved by eliminating some of the 5 hydrogen bonds that could form between adenosyl part of the inhibitor and the conserved AdoMet binding domain. As indicated earlier, there could be 8 hydrogen bonds between the inhibitor and Dnmt1; three of those hydrogen bonds are specific for the active site residues found only in cytosine DNA methyltransferase, while the remaining five hydrogen bonds can be found in other enzymes that bind AdoMet.

Fast-Equilibrium as a General Mechanism of Control of Enzymatic Activity of DNA Methyltransferases

　　Interestingly, based on the current findings it appears that a fast equilibrium between the early reaction intermediates can be the principal mechanism of control of catalytic activity in the evolution of DNA methyltransferases. For example with Dnmt1, the preference for pre-methylated substrates relative to unmethylated substrate appears to be primarily due to more favorable accumulation of the early catalytic intermediates (Fig. **4A**) and thus faster target base attack [47]. Also, differences in catalytic rates between M.HhaI and Dnmt1 appear to be primarily due to differences in accumulation of the early reaction intermediates [47]. There is a very good reason why differences in catalytic activity originate from differences in accumulation of early reaction intermediates rather than from changes in the rate-limiting methyl-transfer step (Fig. **3**, **2** -> **3**). Electrophilic addition of a methyl group to a conjugated pyrimidine ring is chemically challenging process especially in an aqueous medium [89]. Target base attack by the active site cysteine (Fig. **3**, **1** -> **2**) and the subsequent methyl-transfer step (Fig. **3**, **2** -> **3**) require specific orientation between the active site residues and the target base in combination with specific electron distribution within the pyrimidine ring [90, 91]. In addition there is a permanent risk of mutagenic deamination by competing water molecules (Fig. **3**, **2**->**7**). Thus, controlling catalytic activity by modulating catalytic steps in the active site is not a flexible option in enzyme regulation or enzyme evolution. Consequently it is not surprising that pyrimidine methyltransferases have a highly conserved catalytic mechanism, conserved active site residues [35], exceptionally slow catalytic rates, and the methyl-transfer step as the rate-limiting step [47]. On the other hand, the equilibrium between early reaction intermediates is a highly flexible control mechanism. The equilibrium between base-flipping and base-restacking steps can change simply by changing the difference in free energy (ΔG) of enzyme-DNA complex with base-in and base-out of the DNA helix (i.e. the extent of enzyme-DNA interaction with base-in and base-out of DNA helix). Thus, it comes as no surprise that there is no a highly conserved base flipping mechanism
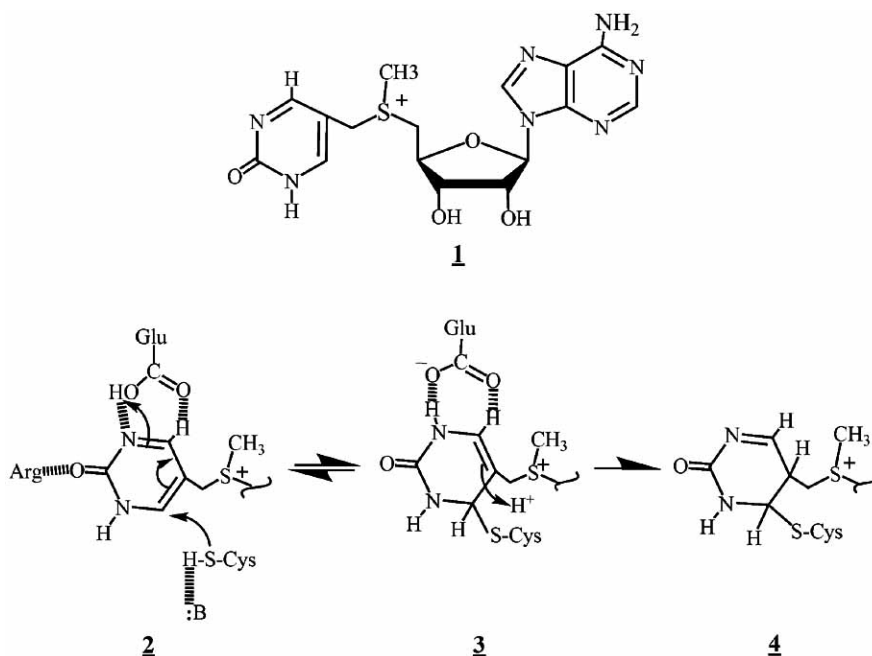

Fig. (6). 1,2-dihydropyrimidin-2-one-5-methylene-(methylsulfonium-adenosyl), a novel suicide inhibitor of DNA methyltransferases based on zebularine and AdoMet structures. 1,2-dihydropyrimidin-2-one-5-methylene-(methylsulfonium)-adenosyl (**1**) is designed with desire to initiate suicide inhibition in the absence of DNA. When positioned in the active site, 1,2-dihydropyrimidin-2-one ring can form hydrogen bonds with conserved Glu and Asp residues initiating a nucleophilic attack by the active site cysteine (**2** -> **3** -> **4** and [76, 77]). Unlike 5-fluoro-cytosine and 5-aza-cytosine, 1,2-dihydropyrimidin-2-one does not require methyl-transfer for enzyme entrapment [76-78]. Thus, 1,2-dihydropyrimidin-2-one ring is attached at carbon 5 to methylene-(methylsulfonium-adenosyl) part of AdoMet (**1**). Based on conserved sequence motifs and catalytic similarities between M.HhaI and Dnmt1 [47], we can predict that adenosyl part of AdoMet could position 1,2-dihydropyrimidin-2-one ring in the enzyme active site. Actually, the presented inhibitor looks like transition state intermediate(s) during the methyl-transfer step; i.e. 1,2-dihydropyrimidin-2-one ring is positioned as the "flipped-out" target base and sulfonium-adenosyl is positioned as sulfonium-adenosyl parts of AdoMet. Additional information about the mechanism of inhibition and inhibitor specificity is provided in the main text.

[92, 93]. Also as a general principle, in small molecule enzymology the enzyme's specificity is controlled by its binding affinity to its substrate; in nucleic acid enzymology, enzyme specificity is not controlled by its binding affinity to DNA, but by conformational changes that trigger catalysis at specific DNA site [94, 95].

Dnmt1 and Mutagenic Cytosine Deamination

Cytosine deamination at DNA methylation sites is a common mutation that can result in oncogene activation [96]. Interestingly, CG sites are significantly underrepresented in the mammalian genome, except for CG islands that are target of DNA methylation [96]. Thus, it is likely that mutagenic deamination at DNA methylation sites had a direct effect on the evolution of human genome. Bacterial DNA methyltransferases can catalyze cytosine deamination [79, 80, 97], the question is to what extent is deamination reaction catalyzed by mammalian DNA methyltransferases. Cytosine deamination depends on solvent access to the activated target base ([66] and Fig. **3**, **2** -> **7**). Crystal structures of bacterial enzyme M.HhaI suggested that solvent can penetrate in the active site where it can take part in elimination step that follows the methyl-transfer step ([98, 99], Fig. **3**, **3**). AdoMet and AdoHcy can block harmful solvent access in the active site [66] and block mutagenic deamination [79, 80, 97]. Interestingly, bacterial enzyme M.HhaI has a high rate of target base attack in the absence of cofactor. This makes the activated target base especially susceptible to solvent molecules [66, 100], and it comes as no surprise that the highest deamination rates are observed in the absence of cofactor [97]. It appears that Dnmt1 does not attack its target base in the absence of AdoMet [47]. This can make Dnmt1 much less likely to be mutagenic than the bacterial enzymes, and interestingly, the earlier cell-based studies came with the similar conclusion [101].

The difference in target base attack in the absence of cofactor might reflect physiological function of these enzymes. In a case of limiting nutrition and limiting AdoMet, mutations caused by deamination might not be lethal to bacteria yet they can increase bacterial chances to evade lethal digestion of its DNA by phage endonucleases. Thus, there could be a physiological advantage for bacterial cells if bacterial methyltransferases have evolved a parallel function as opportunistic deaminases. In difference to bacterial cells, mammalian cells are unlikely to benefit from mutagenic deamination by DNA methyltransferase.

ASSAY DESIGN AND QUANTITATIVE ANALYSIS OF CATALYTIC ACTIVITY BY DNMT1

A majority of publications in enzyme kinetics and design of enzyme assays is written for small molecule substrates. Nucleic acid enzymology is a different challenge, DNA is a complex polymer, and the differences between "substrate DNA" and "product DNA" are subtle and specific for each enzyme group. For example, polymerases, helicases, topoisomerases, ligases, nucleases, or DNA methyltransferases have different interaction with "substrate DNA" and "product DNA", consequently enzyme assays and kinetics analysis have to be adapted for each enzyme group. The challenges mostly affect quantitative analysis of assay design and data interpretation. Different experimental challenges are usually met as a "work in progress", and occasionally there is a need for a review that can serve as a reference point between different studies. It this section we would like to review different challenges in quantitative analysis of catalytic activity by Dnmt1. Also, we would like to give a brief description of some of the most common assays that can be used for Dnmt1 studies. The aim is to provide a reference point for quantitative analysis of the future and the previously published studies of Dnmt1. This section will primarily cover technical aspects of assay design and quantitative data interpretation.
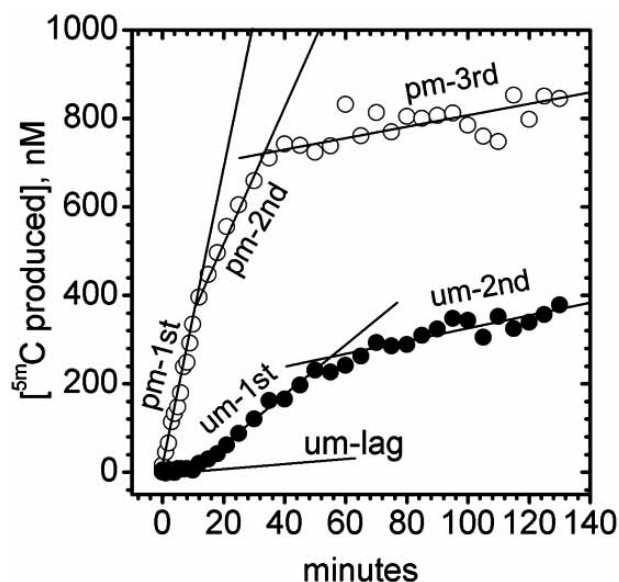

Fig. (7). Dnmt1 (350 nM), initial turnovers in methylation reaction with unmethylated (●) and pre-methyalted (○) poly(dG-dC) (10 μM base pairs). We use these two reactions as a simple model system to illustrate some of the main features that define complex catalytic mechanism of Dnmt1. Consistent with previous studies [45, 47, 50-52], both reactions have exceptionally slow catalytic turnovers that allow us to see different steps in catalytic cycle. Nonlinear reaction profiles indicate that multiple events are taking place simultaneously. First, we have to identify individual catalytic cycles (turnovers), i.e. the time it takes the reaction to produce ^{5m}C sites in one equivalent of enzyme concentration (every 350 nM). The first turnover with poly(dG-dC) (um-1st) takes almost 80 minutes! With pm-poly(dG-dC) the first turnover takes about 10 minutes (pm-1st), each subsequent turnover (pm-2nd and pm-3rd) shows a decrease in catalytic rates as a result of processivity and a gradual product dissociation that is described in the main text. The reaction with poly(dG-dC) substrate shows initial lag (um-lag) that is followed by a pre-steady state burst; i.e. a fast first turnover (um-1st) that is followed by a slow product release (um-2nd). At the start, the difference in catalytic rate between the two reactions is about 36 fold (um-lag 0.25 ± 0.04 h^{-1} vs. pm-1st 8.8 ± 0.6 h^{-1}); however these two reactions start with different catalytic processes that can not be compared directly. The reaction with unmethylated substrate starts with the initial lag (um-lag) that represents a slow relief from allosteric inhibition ([47] and Fig. **9**), while the reaction with pre-methylated substrates starts directly with methylation. Thus, the initial lag has to be separated, and the two reactions have to be compared at equivalent stages. In the first catalytic turnover the observed difference in catalytic rate is about 5 fold (um-1st 1.7 ± 0.4 h^{-1} vs. pm-1st 8.8 ± 0.6 h^{-1}). In the subsequent turnovers, when slow product release dominates the catalytic rates, the two reactions show principally identical rates (um-2nd vs. pm-3rd). Different events in the catalytic cycle have to be separated for a meaningful comparison of different reactions and for an accurate description of the underling mechanism. The presented analysis can explain variations in reported differences in catalytic rates between pre-methylated and unmethylated DNA substrate. The reported difference in catalytic rates between unmethylated and pre-methylated DNA vary considerably, namely 3-6 fold [45, 47] , 3-20 fold [52], 7-20 fold [10], 5 - 30 fold [5], or 30-40 fold [49]. However, after corrections for different phases in catalytic cycle and steady state approach (see the main text) the majority of Dnmt1 studies in the last 20 years show a very consistent difference of 3-6 fold [47]. The presented analysis is likely to be valid for Dnmt1 reaction with different DNA substrates, except that the *extent* of each feature depends on DNA sequence, preexisting methylation, enzyme saturation with its DNA substrate, or allosteric activity and assay design as described in the main text. For example, the same phases can be observed with poly(dI-dC) substrates, but duration of each phase is different due to different kinetic constants [47]. A full description of experimental procedures can be found in reference [47].

Dnmt1 is in many aspects a unique enzyme, with unique experimental challenges. Dnmt1 shows nonlinear reaction profiles (Fig. **7**), exceptionally slow catalytic rates (Fig. **7**), diverse allosteric activity (Fig. **8**) and a complex processivity [48, 50]. The nonlinear reaction profiles indicate that multiple events are taking place during catalysis. The slow catalytic rates are the result of complex catalytic mechanism as indicated earlier (Figs. **3** and **4**). For comparison, methylation rates for Dnmt1 are about one to two orders of magnitude slower than methylation rates for bacterial cytosine carbon 5 DNA methyltransferases [47], and orders of magnitude slower than the majority of metabolic enzymes. Nevertheless, the slow rates are fully reproducible in all of Dnmt1 studies that have been reported by different laboratories in the last 12 years [47]. We start analysis of different catalytic features of Dnmt1 by following the extent of enzyme saturation with its DNA substrate (Fig. **8**), and by tracing each catalytic turnover (Fig. **7**), (i.e the time it takes for an enzymatic reaction to produce ^{5m}C sites in one equivalent of enzyme concentration).

Michaelis-Menten kinetics in Dnmt1 Studies

Traditionally, "steady state approach" is used to describe enzyme affinity for its substrate (i.e. Michaelis-Menten constant, Km), its turnover rates (i.e. kcat), order of substrate addition, and enzyme inhibition [53]. "Steady state approach" is needed to facilitate quantitative analysis of enzymatic reactions, and it is based on several assumptions: *i*) there is a large excess of substrate relative to enzyme; *ii*) that measurements include multiple turnovers and linear reaction profiles; *iii*) product inhibition and substrate consumption are negligible, *iv*) catalytic process is homogenous it terms of substrate and its mechanism. The listed conditions are often impossible, or at least very difficult to achieve with slow enzymes such as Dnmt1 (or other mammalian DNA methyltransferases [102, 104]). For example, in Dnmt1 studies due to slow catalytic rates Dnmt1 concentration is comparable to the varied substrate concentration (*7-9*, *11*, *12*, *24*, *27*, *55*). Thus, competitive and noncompetitive patterns in double reciprocal plots and the calculated *K*m, *k*cat, and *K*i constants reflect enzyme/DNA ratios and assay design rather than kinetic properties of Dnmt1. Based on reported variations in kinetic constants for simple reaction like Dnmt1 methylation reaction with poly(dI-dC) substrate, it is possible to estimate that a misuse of steady-state approach in Dnmt1 studies can lead to kinetic constants that can be inaccurate by almost two orders of magnitude! Accordingly, all of Dnmt1 studies that used steady state approach to describe substrate affinity, order of substrate addition, or inhibition by potential drug candidates [103] have to be reevaluated. A satisfying consensus between different Dnmt1 studies is found when steady state kinetics is not used for data interpretation [47].

Inability to use steady-state kinetics in Dnmt1 studies requires modifications in standard experiments that can be accomplished with a help from numerical simulation (one simple example is shown in Fig. **8**). Catalytic rates in Dnmt1 reaction with different DNA substrates should be measured as a function of increasing substrate concentration until full saturation is achieved. It is important to pay attention that saturation is achieved in terms of enzyme binding sites on its DNA substrate as well as in terms of CpG sites [47]. Measuring Dnmt1 rates with increasing DNA concentration can reveal the highest catalytic rates attainable with given DNA substrate, and possibly allosteric inhibition (Fig. **8A**), allosteric activation (Fig. **8B**), or a lack of allosteric activity. The same approach can be repeated in presence of different allosteric regulators (Fig. **8C**). In a case of a competitive inhibition, the inhibitor will affect half saturation but not the maximal rates, if there is a pure non-competitive inhibition the inhibitor will affect the maximal rates, and finally an uncompetitive or a mixed inhibition will affect both. This approach is similar to the routine steady state kinetics studies, except that double reciprocal plots (i.e. Lineweaver-Burk

plots) or steady-state equations can not be used for quantitative analysis and data interpretation.

Using numerical simulations is possible to show that the ability to see allosteric response depends on experimental design; i.e. enzyme and substrate concentration, binding constants for each site, and the number and the relative distribution of experimental points. Studies of allosteric regulation by DNA or RNA molecules other than substrate DNA are the easiest to interpret (Fig. **8C**). Also, it is easy to recognize inhibition by substrate DNA when dissociation constant for the active site is lower than dissociation constant for the allosteric site (Fig. **8A**). Difficult measurements are allosteric activation by substrate DNA (Fig. **8B**), or partial allosteric inhibition by substrate DNA when the allosteric site has lower dissociation constant than the active site (Fig. **8A**). Also, profiles that do not show clear evidence of allosteric activity have to be interpreted with caution. A lack of allosteric response could be real, or a result of inadequate experimental design. To address such uncertainties allosteric activity should be analyzed using different enzyme and substrate concentrations, or using Dnmt1 fragments [37, 42], or by looking for other signs that accompany different types of allosteric activity (such as initial burst or lag, See Figs. **7** and **9**).

Pre-Steady State Burst, Initial Lag, and Allosteric Inhibition in Dnmt1 Studies

Slow catalysis by Dnmt1 allows us to see transition between different steps in the catalytic cycle as indicted by nonlinear reaction profiles (Fig. **7**). Dnmt1 can show nonlinear reaction profiles due to a pre-steady state burst, an initial lag, or processivity on its DNA substrate. Each of those features represents a different function in the catalytic mechanism, and as such they depend on pre-existing DNA methylation, allosteric regulation, or extent of enzyme saturation with AdoMet or its DNA substrate (Fig. **9**). Different functions have to be separated for a meaningful analysis of catalytic mechanism. For example, there are considerable variations in the reported difference in catalytic rates between unmethylated and pre-methylated DNA substrates in Dnmt1 studies, namely 3-6 fold [45, 47], 3-20 fold [52], 7-20 fold [10], 5-30 fold [5], or 30-40 fold [49]. Those differences are sometimes attributed to different DNA sequences and sometimes they are casually attributed to proteolytic degradation that was caused during Dnmt1 purification [32, 52]. In reality the differences can be attributed to different stages in catalytic cycle (Fig. **7**) and steady state approach as indicated in the earlier section.

A pre-steady state burst is observed with Dnmt1 and with bacterial DNA methyltransferases [45, 68, 105]. The pre-steady state burst can provide insights in enzymatic mechanism (page 195 in [53]), or it can be used to estimate often debated concentration of active enzyme (pages 143-145 in [53]). Catalytic mechanisms that can lead to a pre-steady state burst have been described in the literature (p. 195 in [53] or [106, 107]), and also, specific features of a pre-steady state burst can be explored in detail using numerical simulation like the one shown in Fig. (**4F**). Briefly, a pre-steady state burst is observed when steps leading to the detection step are faster than the subsequent steps (usually product release steps). In methylation reaction with DNA methyltransferases and a radio-labeled AdoMet, detection step is incorporation of the radio-labeled methyl-group into DNA substrate (i.e. methyl-transfer step, Fig. **3**, **2**->**3**). Thus, a pre-steady state burst is observed when methyl-transfer step, or the steps leading to the methyl-transfer step, are faster than the subsequent product release steps (i.e. target base release, DNA release or AdoHcy release [66, 69, 108]). The same reaction might not show a pre-steady state burst if different method is used for analysis. For example, in some of the coupled assays with DNA endonucleases detection step is DNA interaction with endonucleases, and that step requires the slow DNA dissociation from DNA methyltransferase. Coupled assays with DNA endonu-
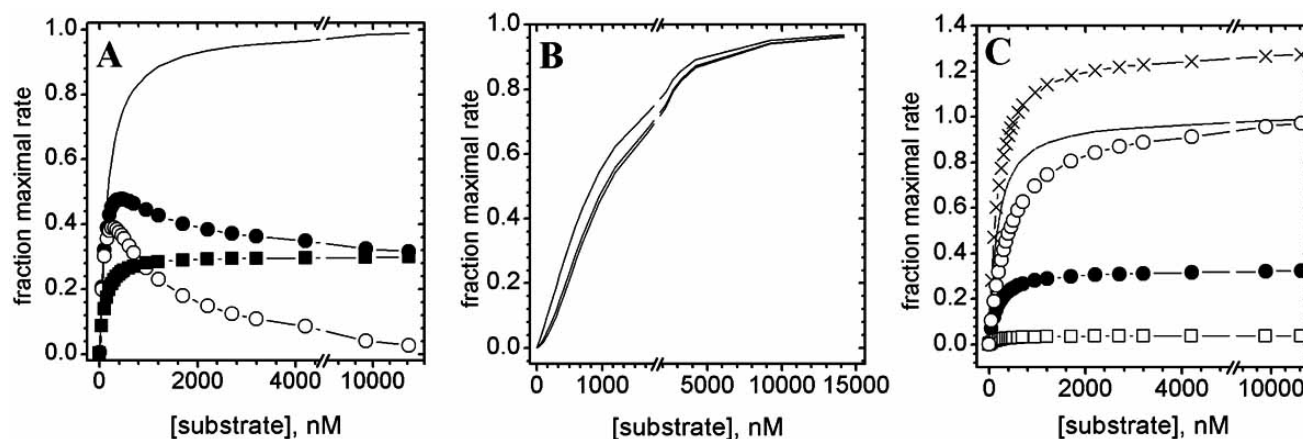
Fig. (8 <u>A</u>-<u>C</u>). Different types of allosteric responses by Dnmt1 and the associated experimental challenges can be summarized using numerical simulation. These simulations are unique because substrate DNA and Dnmt1 are present in comparable concentrations as indicated in the main text. We analyzed allosteric regulation by substrate DNA (<u>A</u>-<u>B</u>) and by DNA or RNA molecule other than substrate (<u>C</u>). Unless otherwise indicated in all panels the line without symbols indicates reaction in the absence of allosteric regulation.

(<u>A</u>) Inhibition by substrate DNA. Allosteric inhibition can be complete (o) or partial (●,■). The shape of each profile depends on binding constant for the active site and the allosteric site. When the active site has a higher binding affinity than the allosteric site, an increase in substrate concentration will lead to initial increase in catalytic rates that is followed by a decrease in catalytic rates (●, o). A less distinct curve is observed when the allosteric site has higher binding affinity than the active site (■).

(<u>B</u>) Allosteric activation by substrate DNA. A gradual increase in substrate concentration will result in a sigmoidal curve if there is allosteric activation by the substrate (right curve no activation, left two curves show activation). A sigmoidal shape can be difficult to observe when there are only two binding sites, the shape depends on binding constant, the extent of enzyme activation and on the number of allosteric sites.

(<u>C</u>). Allosteric regulation by regulatory DNA or RNA molecules that bind to Dnmt1 but can not be its substrate. Allosteric activity can be measured by following catalytic rates with increasing substrate concentration in the presence of different concentrations of regulatory DNA or RNA molecules. In the case of inhibition, the regulatory DNA or RNA molecule will decrease binding affinity for the substrate DNA (o) and (or) maximal catalytic rate (●,□). Opposite will be observed in the case of allosteric activation (x). Interestingly, with Dnmt1 the allosteric inhibition can be complete (□) or partial (●).

All curves in panels <u>A</u> and <u>B</u> were simulated using assumption that interaction between Dnmt1 and DNA is fast relative to the catalytic step (Fig. <u>4A</u>). The first binding was calculated using equation: $[ES1i]=0.5*((Et+Kd1+Si)-[(Et+Kd1+Si)^2-4*Et*Si]^{1/2})$, while the second binding was calculated using equation: $[ES2i]= 0.5*((ESi+Kd1+Sii)-[(ESi+Kd1+Sii)^2-4*ESi*Sii]^{1/2})$. ESi is concentration of the first complex at initial substrate concentration Si, initial enzyme concentration Et, and first binding constant $Kd1$. $ES2i$ is concentration of the second complex when concentration of the first complex is $ES1i$ (as calculated from the first equation), substrate concentration is $Sii = Si-ESi$, and the second binding constant is $Kd2$. Observed rates were calculated using equation: $ktot=k1*ES1 + k2*ES2$, where $ktot$ represents the actual measured rate, while $k1$ and $k2$ are catalytic rate constants for ES1 and ES2 complex respectively. (<u>A</u>) all curves were simulated using Et = 40nM and: (●) Kd1= 150 nM , Kd2= 400 nM, k1=1, k2= 0.3; (■) Kd1=400 nM, Kd2= 150 nM, k1=1, k2= 0.3; (O) Kd1= 150 nM, Kd2= 400 nM, k1= 1, k2= 0. (<u>B</u>) all curves were calculated using Et = 400 nM, Kd1=100 nM, Kd2 = 400nM, k1=1 and k2=1.05 or 1.3 for the second and the third curve from the left. (<u>C</u>) all curves were calculated using the same approach as in panels A-B, except that Sii values from the second equation were replaced with an initial concentration of a "regulatory" DNA/RNA that binds at the allosteric site (Ri value). Also, in these simulations Kd2 represents dissociation constant for regulatory DNA/RNA. All curves were simulated using Et = 40nM, Ri = 500 nM and: (x) Kd1=150nM, Kd2=100 nM, k1=1 and k2=1.3 (i.e. allosteric activation); (O) Kd1=250 nM, k1=1, k2= 1, (i.e. allosteric modulation of only binding affinity for substrate DNA); (●) Kd1= 150 nM, Kd2= 400 nM, k1=1, k2= 0.3, (i.e. allosteric modulation of maximal rates only); (□) Kd1= 150 nM, Kd2= 400 nM, k1=1, k2= 0, (i.e. binding at the allosteric site leads to full inhibition). Notice, that in the case of complete allosteric inhibition (□) a low level of catalytic activity can observed due to enzyme substrate complex that is in equilibrium with enzyme substrate complex bound to allosteric inhibitor.

cleaves are sometimes used in fluorescence FRET assays for high throughput screening studies.

The burst is most pronounced with short oligonucleotides that have one methylation site [45]. In such case product release has to include AdoHcy release and DNA release. Processivity can affect the shape of pre-steady state burst (Fig. **7** and [48]). Since the enzyme is not 100% processive only a fraction of enzyme molecules will go through the slow product release step after each turnover [48]. Thus, catalytic rates will decrease after each turnover proportionally to the fraction of enzyme molecules that goes through the slow dissociation step. The result will be a gradual decrease in catalytic rates, and each catalytic turnover will represent a "small partial burst". Control measurements can show that such gradual decrease in catalytic rates can not be attributed to other common causes of rate decline such as product inhibition, enzyme decay, or substrate depletion [48].

Initial lag is another kinetics tool that can be used to analyze enzyme mechanism [107, 109]. In general, an initial lag can be observed in enzymatic reaction as a result of assay design or as a result of a slow enzymatic process at the start of catalysis [107, 109]. With Dnmt1, an initial lag is observed in reactions with unmethylated DNA substrates that show allosteric inhibition [25, 47, 54]. Control measurements showed that the initial lag with Dnmt1 can be attributed to a slow relief from allosteric inhibition [47] that appears to be driven by AdoMet binding (Fig. **9**). The initial lag, and AdoMet driven relief from allosteric inhibition can be a mechanism that protects the Dnmt1 from attacking the target base in the absence of the cofactor, and thus protects the enzyme from catalyzing mutagenic deamination as discussed earlier in the text.

In summary, the pre-steady state burst, the initial lag, and the allosteric inhibition appear to be interconnected and a part of the mechanism that leads to the start of catalysis by Dnmt1.

Processivity Studies with Dnmt1

Several strategies have been developed to study processivity of Dnmt1, coupled assays with endonucleases [110], bisulfite sequencing assays [50], pulse-chase approach with ³H labeled DNA [48], and analysis of reaction time profiles [48]. These assays make

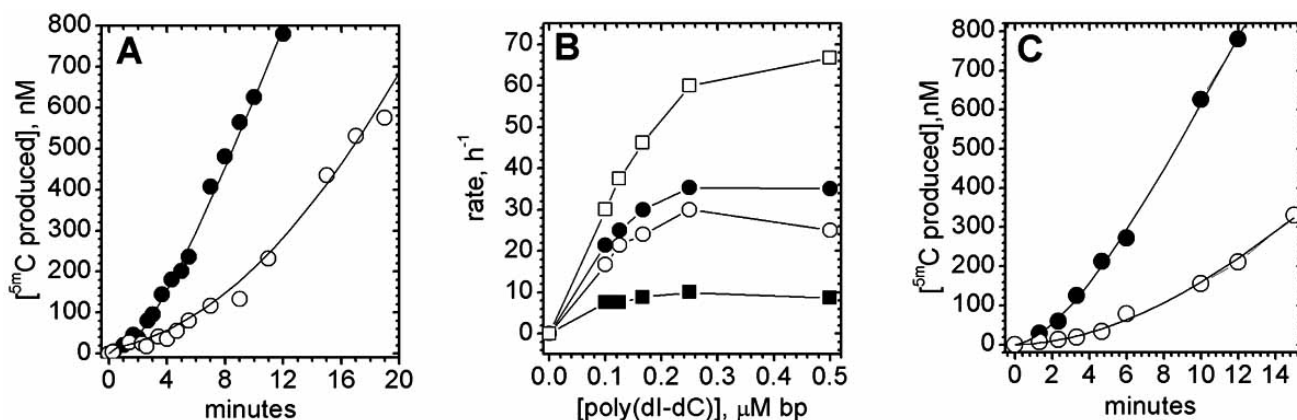**Fig. (9) <u>A</u>-<u>C</u>).** A relief from allosteric inhibition and AdoMet binding can drive slow conformational changes that trigger Dnmt1 reaction with unmethylated DNA substrate. (<u>A</u>) Previous studies [47] showed that saturation of Dnmt1 with unmethylated DNA substrate leads to allosteric inhibition and an increase in the length of the initial lag (figure shows lag with 10 μM (●) or 260 μM (O) base pairs of poly(dI-dC) substrate). The control measurements indicated that the lag can be attributed to the slow relief from allosteric inhibition at the start of reaction [47]. Interestingly, a decrease in Dnmt1 saturation with AdoMet leads to an increase in allosteric inhibition (<u>B</u>), and an increase in length of the initial lag (<u>C</u>). Combined together, these three experiments suggest that at the start of catalysis with unmethylated substrate Dnmt1 has a slow conformational change that is driven by AdoMet binding and by slow relief from allosteric inhibition. Such conformational change can represent a slow transition that drives recognition of target sites. (<u>A</u>) Presented data are taken from Fig. (**3A**) in [47]. Briefly, the enzyme activity was measured using 145 nM Dnmt1, 12.5 μM AdoMet, and 10 μM (●) or 260 μM (O) base pairs of poly(dI-dC). (<u>B</u>) Presented data are a re-plot of Fig. (**4A**) in [52]. The original data were given as a double reciprocal plot (i.e. 1/rate vs. 1/[substrate]). The enzyme activity was measured using 2 nM Dnmt1, and 1 μM (■), 2.6 μM (O), 5.0 μM (●), and 20 μM (□) of AdoMet. (<u>C</u>) presented data were measured as earlier described [47] using 110 nM Dnmt1, 3 μM bp poly(dI-dC), and 2 μM (O) or 15 μM (●) of ³H-AdoMet (7100 cpms/pmol, 72 Ci/mmol).

qualitative studies of enzyme processivity almost a routine measurement; the remaining challenge is to highlight precautions that have to be taken in processivity measurements, and to develop quantitative analysis. Several examples can illustrate these challenges. First, several studies analyzed Dnmt1 processivity on unmethylated substrates [49, 50, 110], however numerous studies showed that Dnmt1 reaction on unmethylated substrates (except for poly(dI-dC)) is so slow that it is impossible to measure multiple turnovers and processivity in those reactions ([45, 47, 50-52] and Fig. **7**). Second, catalysis by different Dnmt3 enzymes is even slower than Dnmt1 reaction on unmethylated substrates [102], and processivity studies with Dnmt3 enzymes [32, 111, 112] are not possible. Third, it has been reported that Dnmt1 inhibition by potential drug candidates can inhibit processivity [103], however inhibitors of catalytic activity can inhibit processive and non-processive reactions equally. The true factors that control enzyme processivity have to affect DNA dissociation rate with or without effect on the turnover rate (see later in the text). Finally, processivity is often evaluated and quantitatively described based on the number of adjacent methylation sites; however numerous non-processive scenarios can lead to closely spaced methylation sites, and also, a highly processive enzyme might not methylate adjacent methylation sites (Fig. **10**). A well-designed processivity study is shown by Vilkaitis and colleagues (Fig. **4** in [50]). The study showed buildup of new methylation sites caused by processivity, and more importantly, the study showed changes in the ratio between the fraction of methylated and unmethylated DNA that took place within a time-span of several turnovers.

An enzyme is processive if its turnover rate is faster than its dissociation rate for substrate DNA. For slow enzymes like Dnmt1, turnover rates are controlled by its catalytic mechanism rather than by its diffusion (sliding) on DNA substrate (Fig. **10**). Processivity is also a stochastic process, in each experiment we are looking at a large population of enzyme molecules and not all of enzyme molecules will take the same number of processive steps on given DNA molecule. The extent of enzyme processivity is controlled by the ratio between its turnover rates, and its enzyme-DNA dissociation rates:

$$P = \frac{k}{k + koff}$$

The presented ratio represents processivity probability (i.e. if a turnover rate is $k = 9$ h⁻¹, and a dissociation rate is $koff = 1$ h⁻¹, the probability that enzyme will take the next turnover on the same DNA substrate is 0.9 or 90%). The probability that enzyme will process n number of steps is equal to P^n (i.e. if processivity in the previous example is 90%, the chances that enzyme will make 5 methylation sites on the same DNA is $(0.9)^5 = 0.59$, or only 59%).

The factors that truly control processivity have to affect enzyme-DNA dissociation rates ($koff$ values) with or without an effect on turnover rates. Changes in enzyme-DNA dissociation rates can be measured directly by standard competition experiments [106], or calculated from processivity experiment by following enzyme dissociation from its initial DNA substrate [48]. With some modifications for a specific enzyme type the same equation can be used to analyze processivity on DNA by different enzymes [53]. To this day, the equation has been adopted to study processivity by DNA terminal transferase [113], DNA helicases [114, 115] and DNA methyltransferases [48]. For Dnmt1, a time course for a processive methylation reaction can be described by equation:

$$[P] = \left[\sum_{i=1}^{n} [ESo] \frac{k^i}{(k+koff)^i} \left(1 - e^{-(k+koff) \cdot t} - e^{-(k+koff) \cdot t} \cdot \sum_{l=1}^{i} \frac{(k+koff)^l \cdot t^l}{l!} \right) \right] + [ESo] \cdot kss \cdot t$$

Where [P] represents concentration of new methylation sites, [ESo] represents concentration of enzyme molecules bound to substrate DNA, t is duration of measurement, n represents the number of turnovers in a given time period t, and k and $koff$ represent the turnover rate constant and DNA dissociation rate constant respectively. The presented equation can be applied to any method that was used to study processivity of Dnmt1 [48, 50, 110], as long as the results are shown as a function of time. The first part of the equation ([positioned in brackets]) represents a gradual accumulation of new methylation sites by all enzyme molecules before their first dissociation from the original DNA substrate. The second part (*ESo·kss·t*), shows late liner phase that can be attributed to multiple
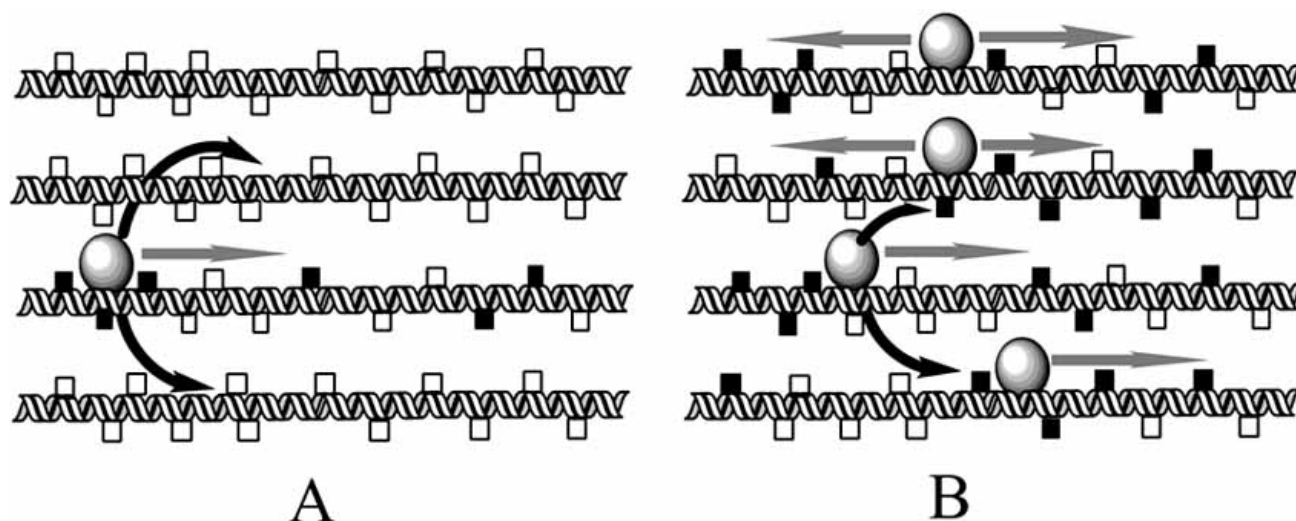
Fig. (10) <u>A</u>-<u>B</u>). Measurements of Dnmt1 processivity on its DNA substrate (enzyme (oval shape), DNA (helical chain), gray arrows (processive catalysis), black arrows (distributive catalysis), methylation sites unmethylated (open squares), methylated (closed squares)). In a processive catalysis, an enzyme goes through multiple turnovers without dissociation from the initial DNA substrate. Thus, the most basic requirement to show processivity is to show that after multiple turnovers the enzyme *did not* dissociate from the initial DNA molecule. The easiest way to accomplish this is to show that a majority of new methylation sites is concentrated to a small fraction of total DNA (Fig. <u>A</u>). Contrary to some reports, a cluster of new methylation sites that is closely spaced on one DNA molecule can not be an evidence of enzyme processivity (Fig. <u>B</u>). With long DNA substrates dissociation constant for Dnmt1-DNA complex is so low (due to numerous binding sites) that in all practically situations Dnmt1 is always bound to DNA. If Dnmt1 and substrate DNA are present in similar concentrations (which is often the case), both processive (black arrows) and distributive catalysis (gray arrows) can cause clusters of methylation sites (Fig. <u>B</u>). Diffusion on DNA substrate is never a rate-limiting step for slow enzymes such as DNA methyltransferases. The sliding on substrate DNA is relatively fast process, and the location of new methylation sites primarily depends on enzyme's preference for specific DNA site and not on their physical proximity. Thus a high processivity could easily lead to methylation sites that are randomly scattered over <u>one</u> DNA molecule. Dnmt1 shows kinetic preference for hemimethylated sites relative to unmethylated substrates, and about 3 fold kinetics preference for CG sites imbedded in AT rich sequence relative to CG rich sequences [45]. Additional explanations can be found in the main text.

association and dissociation events. This seemingly cumbersome looking equation can readily converge in nonlinear regression to give the best fit values for turnover rates (*k*) and dissociation rate (*koff*) with the low correlation coefficients between the best fit values [48, 114, 115].

All of Dnmt1 processivity studies agree that pre-methylated DNA supports enzyme processivity, and the higher the methylation density the higher processivity of Dnmt1 [50]. Interestingly, pre-methylated DNA substrates do not have allosteric inhibition [47], and initial studies showed that allosteric inhibition leads to a decrease in processivity by Dnmt1 [48].

CONCLUDING REMARKS

It is often debated how to correlate enzymatic studies with purified Dnmt1 and DNA methylation in cells. In biomedical science like in any other science, our ability to understand complex system depends on our ability to understand system components. Thus, to understand DNA methylation we have to understand enzymatic mechanism of Dnmt1, first with purified Dnmt1 then in the presence of other interacting molecules. The difference between enzyme activity in cytosol and in purified protein solution is addressed in a number of excellent review articles [116-119] and in numerous specific studies including a recent study by the author [120]. Briefly, in cells enzymes function in highly concentrated solution of macromolecules, this leads to molecular crowding and excluded volume effects. Such conditions lead to lower solubility of macromolecules, transitions to more compact macromolecular conformations, and tens of kJ/mol of free energy in support of protein-protein or protein-DNA interaction. Protein-protein and protein-DNA interactions are adapted to cytosolic conditions, and as such they are less stable in solutions with purified proteins. For example, purified Dnmt (like many other DNA

binding enzymes) is inactive at ionic strength that is comparable to the ionic strength in cytoplasm. Nevertheless, molecular crowding conditions do not affect the major enzymatic properties that were discussed in this manuscript, namely catalytic mechanism at the active site, the number of DNA binding sites, enzyme's predisposition for interaction with other molecules, or the functional consequences of those interactions. Rather, molecular crowding conditions affect *the extent* of intermolecular interactions and consequently *the extent* of functional properties that depend on those interactions. As shown in this manuscript, the final conclusion about biological process is always a combined result of cell based studies and enzyme based studies. Cell based studies and biochemical studies are complementary, one can not replace the other. Thus, enzymatic studies of Dnmt1 do not need to recreate DNA methylation in cells, such systems are already analyzed in the cell based studies. Rather, enzymatic studies need to focus on different aspects of enzyme function with a clarity that is impossible with the cell bases studies. In summary, we have outlined the basic principles that have to be taken in account when biochemical studies with purified Dnmt1 are criticized as inadequate for analysis of DNA methylation in cells [121], or when in-cell protein expression studies are suggested as a superior alternative [121]. In-cell protein expression studies can not replace the studies with purified Dnmt1 since they lack the clarity that is required to understand biomolecular mechanism (uncertainty about Dnmt2 function is a good example [122]). Also in-cell expression studies might miss crucial molecular interactions due to inability to meet stoichiometric ratios as a result of arbitrary expression levels.

In this manuscript we summarized known catalytic features of Dnmt1 as well as remaining questions and exciting studies that could come in the future. The experimental approaches that were developed in the past can meet the future challenges. When all of described precautions are taken in account, all of Dnmt1 studies

that were published in the last 20 years appear consistent and independent from research group or Dnmt1 preparation [47]. In principle, technical features in assay design and data analysis are easy to analyze if the study is based on the established methods for enzyme mechanism studies (an excellent standard textbooks is [53]). The most difficult to interpret, and the most inaccurate in their conclusions, are the studies that ignored established methods. For example, a good number of published Dnmt1 or Dnmt3 studies can not be interpreted simply because they lack information about enzyme, substrate or product concentration.

A standard assay for Dnmt1 activity can greatly facilitate our ability to compare different Dnmt1 studies. Standard assays are common in enzymology [123]. A good standard assay needs to be robust, simple, inexpensive, versatile, and widely present in literature. Historically, Dnmt1 methylation reaction with poly(dI-dC) is the best fit to this description. It appears that Dnmt1 methylation reaction with poly(dI-dC) has all of the main features that can be observed with other DNA substrates, but with different extent [47]. Poly(dI-dC) is not a natural substrate, and as such it is used only for initial quantitative analysis in parallel with other DNA substrates as shown in the past [42, 44-47, 50-52].

ACKNOWLEDGEMENTS

REFERENCES

[1]　Reik, W.; Dean, W.; Walter, J. *Science*, **2001**, *293*, 1089.
[2]　Esteller, M. *Br. J. Cancer*, **2006**, *94*, 179.
[3]　Minarovits, J. *Curr. Top. Microbiol. Immunol.*, **2006**, *310*, 61.
[4]　Miller, C. A.; Sweatt, J. D. *Neuron*, **2007**, *53*, 857.
[5]　Grace Goll, M.; Bestor, T. H. *Annu. Rev. Biochem.*, **2005**, *74*, 481.
[6]　Chen, T.; Hevi, S.; Gay, F.; Tsujimoto, N.; He, T.; Zhang, B.; Ueda, Y.; Li, E. *Nat. Genet.*, **2007**, *39*, 391.
[7]　Stresemann, C.; Brueckner, B.; Musch, T.; Stopper, H.; Lyko, F. *Cancer Res.*, **2006**, *66*, 2794.
[8]　Brueckner, B.; Lyko, F. *Trends Pharmacol. Sci.*, **2004**, *25*, 551.
[9]　Egger, G.; Liang, G.; Aparicio, A.; Jones, P. A. *Nature*, **2004**, *429*, 457.
[10]　Pradhan, S.; Esteve, P. O. *Clin. Immunol.*, **2003**, *109*, 6.
[11]　Chuang, L. S.; Ian, H. I.; Koh, T. W.; Ng, H. H.; Xu, G.; Li, B. F. *Science*, **1997**, *277*, 1996.
[12]　Fuks, F.; Burgers, W. A.; Brehm, A.; Hughes-Davies, L.; Kouzarides, T. *Nat. Genet.*, **2000**, *24*, 88.
[13]　Rountree, M. R.; Bachman, K. E.; Baylin, S. B. *Nat. Genet.*, **2000**, *25*, 269.
[14]　Kimura, H.; Shiota, K. *J. Biol. Chem.*, **2003**, *278*, 4806.
[15]　Schlesinger, Y.; Straussman, R.; Keshet, I.; Farkash, S.; Hecht, M.; Zimmerman, J.; Eden, E.; Yakhini, Z.; Ben-Shushan, E.; Reubinoff, B. E.; Bergman, Y.; Simon, I.; Cedar, H. *Nat. Genet.*, **2007**, *39*, 232.
[16]　Bostick, M.; Kim, J. K.; Esteve, P. O.; Clark, A.; Pradhan, S.; Jacobsen, S. E. *Science*, **2007**, *317*, 1760.
[17]　Pradhan, S.; Kim, G. D. *EMBO J.*, **2002**, *21*, 779.
[18]　Esteve, P. O.; Chin, H. G.; Pradhan, S. *Proc. Natl. Acad. Sci. USA*, **2005**, *102*, 1000.
[19]　Reale, A.; Matteis, G. D.; Galleazzi, G.; Zampieri, M.; Caiafa, P. *Oncogene*, **2005**, *24*, 13.
[20]　Carty, S. M.; Greenleaf, A. L. *Mol. Cell Proteomics*, **2002**, *1*, 598.
[21]　Jeffery, L.; Nakielny, S. *J. Biol. Chem.*, **2004**, *279*, 49479.
[22]　Mattick, J. S. *Nat. Rev. Genet.*, **2004**, *5*, 316.
[23]　Mattick, J. S. *Sci. Am.*, **2004**, *291*, 60.
[24]　Zamore, P. D.; Haley, B. *Science*, **2005**, *309*, 1519.
[25]　Bolden, A.; Ward, C.; Siedlecki, J. A.; Weissbach, A. *J. Biol. Chem.*, **1984**, *259*, 12437.
[26]　Glickman, J. F.; Reich, N. O. *Biochem. Biophys. Res. Commun.*, **1994**, *204*, 1003.
[27]　Zaratiegui, M.; Irvine, D. V.; Martienssen, R. A. *Cell*, **2007**, *128*, 763.
[28]　Matzke, M. A.; Birchler, J. A. *Nat. Rev. Genet.*, **2005**, *6*, 24.
[29]　Henderson, I. R.; Jacobsen, S. E. *Nature*, **2007**, *447*, 418.
[30]　Morris, K. V.; Chan, S. W.; Jacobsen, S. E.; Looney, D. J. *Science*, **2004**, *305*, 1289.
[31]　Kim, J. W.; Zhang, Y. H.; Zern, M. A.; Rossi, J. J.; Wu, J. *Biochem. Biophys. Res. Commun.*, **2007**, *359(2)*, 292.
[32]　Jeltsch, A. *Curr. Top. Microbiol. Immunol.*, **2006**, *301*, 203.
[33]　Margot, J. B.; Aguirre-Arteta, A. M.; Di Giacco, B. V.; Pradhan, S.; Roberts, R. J.; Cardoso, M. C.; Leonhardt, H. *J. Mol. Biol.*, **2000**, *297*, 293.
[34]　Bestor, T. H. *EMBO J.*, **1992**, *11*, 2611.
[35]　Lauster, R.; Trautner, T. A.; Noyer-Weidner, M. *J. Mol. Biol.*, **1989**, *206*, 305.
[36]　Araujo, F. D.; Croteau, S.; Slack, A. D.; Milutinovic, S.; Bigey, P.; Price, G. B.; Zannis-Hajopoulos, M.; Szyf, M. *J. Biol. Chem.*, **2001**, *276*, 6930.
[37]　Pradhan, S.; Esteve, P. O. *Biochemistry*, **2003**, *42*, 5321.
[38]　Callebaut, I.; Courvalin, J. C.; Mornon, J. P. *FEBS Lett.*, **1999**, *446*, 189.
[39]　Glickman, J. F.; Pavlovich, J. G.; Reich, N. O. *J. Biol. Chem.*, **1997**, *272*, 17851.
[40]　Leonhardt, H.; Page, A. W.; Weier, H. U.; Bestor, T. H. *Cell*, **1992**, *71*, 865.
[41]　Robertson, K. D.; Ait-Si-Ali, S.; Yokochi, T.; Wade, P. A.; Jones, P. L.; Wolffe, A. P. *Nat. Genet.*, **2000**, *25*, 338.
[42]　Bacolla, A.; Pradhan, S.; Larson, J. E.; Roberts, R. J.; Wells, R. D. *J. Biol. Chem.*, **2001**, *276*, 18605.
[43]　Flynn, J.; Azzam, R.; Reich, N. *J. Mol. Biol.*, **1998**, *279*, 101.
[44]　Flynn, J.; Fang, J. Y.; Mikovits, J. A.; Reich, N. O. *J. Biol. Chem.*, **2003**, *278*, 8238.
[45]　Flynn, J.; Glickman, J. F.; Reich, N. O. *Biochemistry*, **1996**, *35*, 7308.
[46]　Flynn, J.; Reich, N. *Biochemistry*, **1998**, *37*, 15162.
[47]　Svedruzic, Z. M.; Reich, N. O. *Biochemistry*, **2005**, *44*, 9472.
[48]　Svedruzic, Z. M.; Reich, N. O. *Biochemistry*, **2005**, *44*, 14977.
[49]　Goyal, R.; Reinhardt, R.; Jeltsch, A. *Nucleic Acids Res.*, **2006**, *34*, 1182.
[50]　Vilkaitis, G.; Suetake, I.; Klimasauskas, S.; Tajima, S. *J. Biol. Chem.*, **2005**, *280*, 64.
[51]　Bacolla, A.; Pradhan, S.; Roberts, R. J.; Wells, R. D. *J. Biol. Chem.*, **1999**, *274*, 33011.
[52]　Pradhan, S.; Bacolla, A.; Wells, R. D.; Roberts, R. J. *J. Biol. Chem.*, **1999**, *274*, 33002.
[53]　Fersht, A. *Enzyme Structure and Mechanism 2nd Ed.*, W.H. Freeman and Company: New York, **1985**; pp. 1-475.
[54]　Pedrali-Noy, G.; Weissbach, A. *J. Biol. Chem.*, **1986**, *261*, 7600.
[55]　Stein, R.; Gruenbaum, Y.; Pollack, Y.; Razin, A.; Cedar, H. *Proc. Natl. Acad. Sci. USA*, **1982**, *79*, 61.
[56]　Schermelleh, L.; Haemmer, A.; Spada, F.; Rosing, N.; Meilinger, D.; Rothbauer, U.; Cristina Cardoso, M.; Leonhardt, H. *Nucleic Acids Res.*, **2007**, *35(13)*, 4301.
[57]　Rouleau, M.; Aubin, R. A.; Poirier, G. G. *J. Cell Sci.*, **2004**, *117*, 815.
[58]　Gong, F.; Kwon, Y.; Smerdon, M. J. *DNA Repair*, **2005**, *4*, 884.
[59]　Svedruzic, Z. M.; Wang, C.; Kosmoski, J. V.; Smerdon, M. J. *J. Biol. Chem.*, **2005**, *280*, 40051.
[60]　Goll, M. G.; Kirpekar, F.; Maggert, K. A.; Yoder, J. A.; Hsieh, C. L.; Zhang, X.; Golic, K. G.; Jacobsen, S. E.; Bestor, T. H. *Science*, **2006**, *311*, 395.
[61]　Huttenhofer, A.; Vogel, J. *Nucleic Acids Res.*, **2006**, *34*, 635.
[62]　Yoder, J. A.; Soman, N. S.; Verdine, G. L.; Bestor, T. H. *J. Mol. Biol.*, **1997**, *270*, 385.
[63]　Ivanetich, K. M.; Santi, D. V. *Prog. Nucl. Acids Res. Mol. Biol.*, **1992**, *42*, 127.
[64]　Tanaka, F.; Fukuse, T.; Wada, H.; Fukushima, M. *Curr. Pharm. Biotechnol.*, **2000**, *1*, 137.
[65]　Jung, Y.; Park, J.; Kim, T. Y.; Park, J. H.; Jong, H. S.; Im, S. A.; Robertson, K. D.; Bang, Y. J.; Kim, T. Y. *J. Mol. Med.*, **2007**, *85(10)*, 1137.
[66]　Svedruzic, Z. M.; Reich, N. O. *Biochemistry*, **2004**, *43*, 11460.
[67]　Holz, B.; Klimasauskas, S.; Serva, S.; Weinhold, E. *Nucleic Acids Res.*, **1998**, *26*, 1076.
[68]　Vilkaitis, G.; Dong, A.; Weinhold, E.; Cheng, X.; Klimasauskas, S. *J. Biol. Chem.*, **2000**, *275*, 38722.
[69]　Youngblood, B.; Shieh, F. K.; De Los Rios, S.; Perona, J. J.; Reich, N. O. *J. Mol. Biol.*, **2006**, *362*, 334.
[70]　Min, J. H.; Pavletich, N. P. *Nature*, **2007**, *449*, 570.
[71]　Parker, J. B.; Bianchet, M. A.; Krosky, D. J.; Friedman, J. I.; Amzel, L. M.; Stivers, J. T. *Nature*, **2007**, *449*, 433.
[72]　Mi, S.; Alonso, D.; Roberts, R. J. *Nucleic Acids Res.*, **1995**, *23*, 620.
[73]　Vilkaitis, G.; Merkiene, E.; Serva, S.; Weinhold, E.; Klimasauskas, S. *J. Biol. Chem.*, **2001**, *276*, 20924.
[74]　Allan, B. W.; Beechem, J. M.; Lindstrom, W. M.; Reich, N. O. *J. Biol. Chem.*, **1998**, *273*, 2368.
[75]　Zimmerle, C. T.; Frieden, C. *Biochem. J.*, **1989**, *258*, 381.
[76]　Hurd, P. J.; Whitmarsh, A. J.; Baldwin, G. S.; Kelly, S. M.; Waltho, J. P.; Price, N. C.; Connolly, B. A.; Hornby, D. P. *J. Mol. Biol.*, **1999**, *286*, 389.
[77]　Zhou, L.; Cheng, X.; Connolly, B. A.; Dickman, M. J.; Hurd, P. J.; Hornby, D. P. *J. Mol. Biol.*, **2002**, *321*, 591.
[78]　Taylor, C.; Ford, K.; Connolly, B. A.; Hornby, D. P. *Biochem. J.*, **1993**, *291* (Pt 2), 493.
[79]　Zingg, J. M.; Shen, J. C.; Jones, P. A. *Biochem. J.*, **1998**, *332*, 223.
[80]　Zingg, J. M.; Shen, J. C.; Yang, A. S.; Rapoport, H.; Jones, P. A. *Nucleic Acids Res.*, **1996**, *24*, 3267.

[81] Dalhoff, C.; Lukinavicius, G.; Klimasauskas, S.; Weinhold, E. *Nat. Chem. Biol.*, **2006**, *2*, 31.

[82] Pignot, M.; Pljevaljcic, G.; Weinhold, E. *Eur. J. Org. Chem.*, **2000**, *3*, 549.

[83] O'Gara, M.; Zhang, X.; Roberts, R. J.; Cheng, X. *J. Mol. Biol.*, **1999**, *287*, 201.

[84] Guo, J.; Wu, Y. Q.; Rattendi, D.; Bacchi, C. J.; Woster, P. M. *J. Med. Chem.*, **1995**, *38*, 1770.

[85] Wu, Y.; Woster, P. M. *J. Med. Chem.*, **1992**, *35*, 3196.

[86] Wu, Y. Q.; Woster, P. M. *Bioorg. Med. Chem.*, **1993**, *1*, 349.

[87] Krosky, D. J.; Bianchet, M. A.; Seiple, L.; Chung, S.; Amzel, L. M.; Stivers, J. T. *Nucleic Acids Res.*, **2006**, *34*, 5872.

[88] Cantoni, G. L. *Annu. Rev. Biochem.*, **1975**, *44*, 435.

[89] Chen, L.; MacMillan, A. M.; Chang, W.; Ezaz-Nikpay, K.; Lane, W. S.; Verdine, G. L. *Biochemistry*, **1991**, *30*, 11018.

[90] Perakyla, M. *J. Am. Chem. Soc.*, **1998**, *120*, 12895.

[91] Zhang, X.; Bruice, T. C. *Proc. Natl. Acad. Sci. USA*, **2006**, *103*, 6148.

[92] Hopkins, B. B.; Reich, N. O. *J. Biol. Chem.*, **2004**, *279*, 37049.

[93] Roberts, R. J.; Cheng, X. *Annu. Rev. Biochem.*, **1998**, *67*, 181.

[94] Spolar, R. S.; Record, M. T., Jr. *Science*, **1994**, *263*, 777.

[95] Kalodimos, C. G.; Biris, N.; Bonvin, A. M.; Levandoski, M. M.; Guennuegues, M.; Boelens, R.; Kaptein, R. *Science*, **2004**, *305*, 386.

[96] Pfeifer, G. P.; Tang, M.; Denissenko, M. F. *Mutation Hotspots and DNA Methylation. In DNA Methylation and Cancer,* Johns, P. A. and Vogt, P. K., Eds.; Springer-Verlag: Berlin, **2000**, *249*, pp. 1-19.

[97] Shen, J. C.; Rideout, W. M. d.; Jones, P. A. *Cell*, **1992**, *71*, 1073.

[98] Lau, E. Y.; Bruice, T. C. *J. Mol. Biol.*, **1999**, *293*, 9.

[99] O'Gara, M.; Klimasauskas, S.; Roberts, R. J.; Cheng, X. *J. Mol. Biol.*, **1996**, *261*, 634.

[100] Wu, J. C.; Santi, D. V. *J. Biol. Chem.*, **1987**, *262*, 4778.

[101] Chan, M. F.; van Amerongen, R.; Nijjar, T.; Cuppen, E.; Jones, P. A.; Laird, P. W. *Mol. Cell. Biol.*, **2001**, *22*, 7587.

[102] Yokochi, T.; Robertson, K. D. *J. Biol. Chem.*, **2002**, *277*, 11735.

[103] Lee, B. H.; Yegnasubramanian, S.; Lin, X.; Nelson, W. G. *J. Biol. Chem.*, **2005**, *280*, 40749.

[104] Suetake, I.; Miyazaki, J.; Murakami, C.; Takeshima, H.; Tajima, S. *J. Biochem.*, **2003**, *133*, 737.

[105] Lindstrom, W. M., Jr.; Flynn, J.; Reich, N. O. *J. Biol. Chem.*, **2000**, *275*, 4912.

[106] Johnson, K. A. In *The Enzymes*, Sigman D.S., Ed.; Academic Press, Inc.: New York, **1992**; Vol. *XX*, pp. 1-546.

[107] Tipton, K. F. In *Enzyme Assays Practical Approach,* Eisenthal, R. and Danson, M. J., Eds.; Oxford University Press: Oxford, **2002**; pp. 1-282.

[108] Merkiene, E.; Klimasauskas, S. *Nucleic Acids Res.*, **2005**, *33*, 307.

[109] Frieden, C. *J. Biol. Chem.*, **1970**, *245*, 5788.

[110] Hermann, A.; Goyal, R.; Jeltsch, A. *J. Biol. Chem.*, **2004**, *279*, 48350.

[111] Gowher, H.; Jeltsch, A. *J. Mol. Biol.*, **2001**, *309*, 1201.

[112] Gowher, H.; Jeltsch, A. *J. Biol. Chem.*, **2002**, *277*, 20409.

[113] Hayes, F. N.; Mitchell, V. E.; Ratliff, R. L.; Williams, D. L. *Biochemistry*, **1967**, *6*, 2488.

[114] Ali, J. A.; Lohman, T. M. *Science*, **1997**, *275*, 377.

[115] Porter, D. J.; Short, S. A.; Hanlon, M. H.; Preugschat, F.; Wilson, J. E.; Willard, D. H., Jr.; Consler, T. G. *J. Biol. Chem.*, **1998**, *273*, 18906.

[116] Goodsell, D. S. *Trends Biochem. Sci.*, **1991**, *16*, 203.

[117] Minton, A. P. *Curr. Opin. Struct. Biol.*, **2000**, *10*, 34.

[118] Srere, P. A. *Trends Biochem. Sci.*, **2000**, *25*, 150.

[119] Zimmerman, S. B.; Minton, A. P. *Annu. Rev. Biophysic. Biomol. Struct.*, **1993**, *22*, 27.

[120] Svedruzic, Z. M.; Spivey, H. O. *Proteins Struct. Funct. Bioinform.*, **2006**, *63*, 501.

[121] Spada, F.; Rothbauer, U.; Zolghadr, K.; Schermelleh, L.; Leonhardt, H. *Adv. Enzyme Regul.*, **2006**, *46*, 224.

[122] Jeltsch, A.; Nellen, W.; Lyko, F. *Trends Biochem. Sci.*, **2006**, *31*, 306.

[123] Bergmeyer, H. U. In *Enzyme Assays Methods of Enzymatic Analysis 3rd ed.,* Bergmeyer, H. U., Ed.; John Wiley and Sons Inc.: Hoboken NJ, **1983**; Vol. 2, pp. 1-539.

[124] Aubol, B. E.; Reich, N. O. *Biochem. Biophys. Res. Commun.*, **2003**, *310*, 209.

[125] Klimasauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. *Cell*, **1994**, *76*, 357.