Article

Single-Molecule Nanopore Sequencing of the CpG Island from the Promoter of O6-Methylguanine-DNA Methyltransferase Provides Insights into the Mechanism of De Novo Methylation of G/C-Rich Regions

Alexander V. Sergeev [1,*], Daniil P. Malyshev [1], Adelya I. Genatullina [1], Galina V. Pavlova [2,3], Elizaveta S. Gromova [1] and Maria I. Zvereva [1]

[1] Department of Chemistry, M.V. Lomonosov Moscow State University, 119991 Moscow, Russia; daniilmalyshev0@gmail.com (D.P.M.); delyaaa.smile@mail.ru (A.I.G.); gromova@belozersky.msu.ru (E.S.G.); maria.i.zvereva@yandex.ru (M.I.Z.)
[2] Institute of Higher Nervous Activity and Neurophysiology of the Russian Academy of Sciences, 117485 Moscow, Russia; lkorochkin@mail.ru
[3] Burdenko National Medical Research Institute for Neurosurgery, 125047 Moscow, Russia
* Correspondence: avsergeev@belozersky.msu.ru

Abstract: Background: The methylation of cytosine residues at CpG sites within the O6-methylguanine-DNA methyltransferase (*MGMT*) promoter is a key biomarker in glioblastoma therapy. The *MGMT* promoter (MGMTp) contains multiple guanine-rich sequences capable of folding into G-quadruplexes (G4s), but their relevance for MGMTp methylation is poorly understood. Objectives: Our study explores the impact of potential G-quadruplex-forming sequences (PQS) in the *MGMT* promoter CpG island on the activity of de novo DNA methyltransferase Dnmt3a. Additionally, we investigate their influence on the accuracy of methylation pattern detection using nanopore sequencing. Methods: Nanopore sequencing was employed to analyze the methylation of 94 clinically significant CpG sites in the human MGMTp using an in vitro de novo methylation system. Circular dichroism spectroscopy was used to identify G4 structures within the MGMTp CpG island. Interactions between the catalytic domain of Dnmt3a and the PQS from the MGMTp were examined by biolayer interferometry. Results: Guanine-rich DNA strands of the PQSs in the MGMTp were hypomethylated, while the complementary cytosine-rich strands were methylated by DNA methyltransferase Dnmt3a with higher efficiency. The accuracy of detecting modified bases in the PQS was significantly lower compared to surrounding sequences. Single-stranded guanine-rich DNA sequences from the MGMTp exhibited strong binding to Dnmt3a-CD, with an affinity approximately 10 times higher than their cytosine-rich complements ($K_d = 3 \times 10^{-8}$ M and 3×10^{-7} M, respectively). By binding to Dnmt3a, G4-forming oligonucleotides from MGMTp effectively inhibited the methylation reaction (IC$_{50}$ 6×10^{-7} M). Conclusions: The obtained data indicate the role of PQSs in establishing de novo methylation of the *MGMT* promoter. They also highlight the challenges of sequencing guanine-rich regions and the impact of specific de novo methylation patterns on clinical data interpretation.

Keywords: G-quadruplexes; *MGMT* promoter; DNA methylation; Dnmt3a; nanopore sequencing

1. Introduction

Brain cancer is among the most aggressive cancer types, with a 5-year survival rate of less than 30% [1]. Gliomas are particularly prevalent within this category, originating from glial cells that support and surround neurons. Glioblastoma multiforme (GBM), the most aggressive glioma subtype, has a five-year survival rate of just 6%. GBM accounts for 45.2% of all malignant central nervous system (CNS) tumors and 80% of primary malignant CNS tumors [2]. These statistics highlight the critical need for molecular biomarkers to enhance tumor classification and inform more effective therapeutic strategies. Well-established molecular biomarkers for CNS tumors include mutations in telomerase reverse transcriptase promoter, tumor suppressor protein 53, isocitrate dehydrogenase 1 and 2, and epidermal growth factor receptor [3]. Additionally, several promising prognostic biomarkers for glioma have been identified in recent years, such as the voltage-gated sodium channel β3 subunit, cyclin-dependent kinase 2, and insulin-like growth factor binding proteins [4–6].

Over the past three decades, the primary treatments for high-grade gliomas have remained consistent, comprising maximal surgical resection, external beam radiation therapy, and chemotherapy [2,7]. However, ongoing research is exploring innovative approaches, including gene therapy and immunotherapy. Currently, the standard of care involves the use of temozolomide, an oral cytotoxic DNA-alkylating agent, combined with radiation therapy. Notably, cytosine methylation at CpG sites in the O-6-methylguanine-DNA methyltransferase (*MGMT*) promoter has been shown to significantly improve survival outcomes in patients undergoing this combined treatment.

DNA methylation at CpG sites in mammalian cells is a crucial epigenetic modification that underpins the regulation of gene expression, development, genomic imprinting, and other cellular processes [8–10]. DNA methylation has a regulatory role in mammalian development from a zygote into a complex, multicellular adult organism [8]. The cytosine methylation pattern is bimodal: highly methylated intragenic regions with high CpG methylation level coexist with CpG islands—methylation-depleted regions of high CpG density—typically located in the promoters of actively transcribed genes [11–13]. This pattern is maintained through the dynamic balance of enzymatic DNA methylation and active or passive demethylation processes [14]. The initial establishment of DNA methylation patterns occurs during early development and is mediated by the de novo DNA methyltransferases (MTases) Dnmt3a and Dnmt3b [15]. Maintenance of these patterns during DNA replication is primarily carried out by Dnmt1.

In mammals, DNA methylation occurs at the C5-position of the cytosine residues, primarily in CpG dinucleotide sequences. All MTases use S-adenosyl-L-methionine (AdoMet) as a methyl group donor. This reaction is catalyzed by the C-terminal catalytic domain of DNMT3A/3B, which form linear tetramers with two active sites in association with the catalytically inactive DNMT3L [15,16]. In addition to its catalytic domain, DNMT3a contains an N-terminal regulatory region composed of PWWP (Pro-Trp-Trp-Pro) [17] and ADD (ATRX-DNMT3-DNMT3L) [18] domains. These domains facilitate specific recruitment of the MTase tetramer to distinct genomic regions by recognizing histone marks, such as methylated (H3K36me2/3 [19]), unmethylated (H3K4me0 [20]), or ubiquitinylated (H2AK119ub [21]) residues [22].

Simultaneously, an increasing body of evidence highlights an additional layer of regulation in MTase activity, mediated by interactions with non-canonical DNA structures, particularly G-quadruplexes (G4s) [23–29]. G4s are formed through the stacking of G-tetrads—planar arrangements of four guanines stabilized by Hoogsteen hydrogen bonds and centrally located monovalent cations, such as K^+ or Na^+ [30,31]. Using the G4-seq technique, over 700,000 G4 structures have been identified in the human genome, with most

formations associated with oncogenes, tumor suppressors, and other genes related to cancer development [32]. Additionally, G4-specific antibodies have been utilized to visualize G4 formation within the genomic DNA of human osteosarcoma cells [33], highlighting a potentially important biological role of G4 structures.

G4 structures are classified based on strand polarity: (i) parallel, where all four strands run in the same direction; (ii) antiparallel, where two pairs of strands run in opposite directions; and (iii) hybrid, where three strands run in one direction and the fourth strand runs in the opposite direction [31]. Thus, the activity of MTases in a given genomic region is influenced by a complex integration of signals from histone modifications, the enzyme's intrinsic sequence specificity [34], protein multimerization [35], and non-canonical DNA structures. This interplay makes it challenging to fully elucidate the mechanisms underlying the de novo establishment of DNA methylation pattern.

Whole-genome sequencing studies of the epigenome have further underscored the role of non-canonical DNA structures in shaping DNA methylation landscapes [27]. A notable correlation has been observed across various tissues: genomic regions enriched in G4 structures tend to exhibit low CpG methylation, whereas regions with high CpG methylation are relatively depleted of G4s. These data were derived from the analysis of 2.1 million CpG sites in humans as part of the Human Epigenome Project. Additionally, the authors experimentally measured methylation at over 600,000 CpG sites across 18 individuals using bisulfite mapping, revealing significantly lower methylation levels within quadruplex-forming regions [27]. These findings suggest that G4 structures act as a genome-wide impediment to CpG site methylation. This observation is particularly significant, given that cytosine methylation plays a critical role in the epigenetic regulation of gene expression in mammals. It should be noted that the role of G4 structures in gene regulatory regions is not exclusively repressive. For instance, a recent study demonstrated that a G4 structure in the MYC proto-oncogene promoter facilitates the recruitment of transcription factors and actively enhances transcription [36]. The specific effect of a G4 structure likely depends on its conformation and its position within the promoter. The coexistence of multiple regulatory mechanisms, such as G4 structures and CpG island methylation, offers enhanced adaptability. If one mechanism fails or operates less efficiently under certain conditions, the other can compensate, ensuring robust gene regulation.

Aberrant methylation patterns are frequently associated with cancer and other diseases [37–39], highlighting the importance of understanding the interplay between DNA methyltransferase activity and alternative DNA structures. Recent studies have shown a high density of potential G-quadruplex-forming sequences (PQS) in the promoters of human DNA repair genes [40]. Using various experimental approaches, it has been demonstrated that some of the identified PQSs indeed fold into G4 structures both in vitro and in vivo. Among these, the *MGMT* gene has been studied, which encodes an enzyme responsible for the repair of alkylated guanine residues [41,42]. This enzyme restores damaged (alkylated) guanine by transferring the methyl group from the O6 position of guanine to a cysteine residue in the protein. This mechanism prevents gene mutations, cell death, and oncogenesis caused by alkylating agents [41]. The expression of the *MGMT* gene is primarily regulated by epigenetic modifications, specifically the methylation of the CpG island within the *MGMT* promoter (MGMTp). When the promoter is methylated at CpG sites, the synthesis of the MGMT repair enzyme is significantly reduced, leaving alkylation-induced damage unrepaired. Under these conditions, chemotherapy with alkylating agents targeting cancer cells becomes a viable treatment option. Therefore, determining the methylation status of the *MGMT* promoter is critically important [43].

MGMTp is considered a marker of precancerous lesions and a biomarker for the early diagnosis of various tumors, including gastric cancer, colorectal cancer, breast cancer,

squamous cell carcinoma of the oral cavity, and cervical carcinoma [44–48]. It has also acquired a key diagnostic role for brain tumor lesions, serving as a molecular biomarker for selecting anti-cancer therapies [49]. However, several challenges remain for its clinical application, such as achieving consensus on MGMTp methylation detection methods, as these methods vary significantly across laboratories. Additionally, optimal MGMTp methylation thresholds for glioblastoma diagnostics are still lacking [50]. The CpG island in the MGMTp region contains several PQSs, raising the question of their potential influence on Dnmt3a activity, which is responsible for de novo DNA methylation and establishing the MGMTp methylation pattern. Previously, we reported a crosstalk between G4 structures and Dnmt3a-mediated methylation of the *c-MYC* oncogene promoter [26].

In this study, we explored the mechanistic aspects of Dnmt3a function in MGMTp. We focused on the influence of DNA sequence context and non-canonical structures on the methylation process. Using an in vitro methylation model, we examined how G4 structures affect DNA methyltransferase activity. Understanding the mechanisms that shape methylation patterns in MGMTp not only provides a deeper insight into the biology of de novo DNA methylation but also holds the potential to explain variability in clinical data. Additionally, it may enhance the utility of MGMTp methylation as a biomarker for disease prognosis and treatment.

2. Results

2.1. MGMTp PQS Reduces the Accuracy of Modified Base Identification in Nanopore Sequencing Data

In this work, we focused on studying the methylation of a G/C-rich region of MGMTp that contains a CpG island overlapping with the first exon (Figure 1A). The sequence of interest included CpG sites 3–97 out of 97 CpG sites within the island. The 752-base pair (bp) long sequence was amplified from Raji human lymphoblast-like cell line genomic DNA, and the resulting PCR product was referred to as MGMT-752 (Table 1).

Table 1. MGMTp DNA duplexes and oligonucleotides that are used in this study.

Description	Name	Sequence *
MGMT promoter CpG island amplicon	MGMT-752	5′-TGACTAGGGGAGCGGCACCAGGAGGGGAGAGACTCGCG CTCCGGG CTCAGCGTAGCCGCCCCGAGCAGGACCGGGATTC TCACTAAGCGGG CG CCGTCCTACGACCCCCGCGCGCTTTC A GGACCACTCGGGCACGTGGCAGGTCGCTTGCACGCCCGCGG ACTATCCCTGTGACAGGAAAAGGTACGGGCCATTTGGCAAA CTAAGGCACAGAGCCTCAGGCGGAAGCTGGGAAGGCGCCGC CCGGCTTGTACCGGCCGAAGGGCCATCCGGGTCAGGCGCAC AGGGCAGCGGCGCTGCCGGAGGACCAGGGCCGGCGTGCCGG CGTCCAGCGAGGATGCGCAGACTGCCTCAGGCCCGGCGCCG CCGCACAGGGCATGCGCCGACCCGGT CGGGCGGGAACACCC CGCCCCTCCCGGGCTCCGCCCCAGCTCCGCCCCCGCGCGCC CCGGCCCCGCCCCCGCGCGCTCTCTTGCTTTT CTCAGG TCC TCGGCTCCGCCCCGCTCTAGACCCCGCCCCACGCCGCCATC CCCGTGCCCCTCGGCCCCGCCCCCGCGCCCCGGATATGCTG G GACAGCCCGCGCCCCTAG AACGCTTTGCGTCCCGACGCCC GCAGGTCCTCGCGGTGCGCACCGTTTGCGACTTG GTGAGTG TCTGGGTCGCCTCGCTCCCGGAAGAGTGCGGAGCTCTCCCT CGGGACGGTGGCAGCCTCGAGTGGTCCTGCAGGCGCCCTCA CTTCGCCGTCGGGTGTG
G4-forming oligonucleotides	MGMT-G4	5′-GGGCCGGGGCGCGCGGGGGCGGAG
	MGMT-G4-mut	5′-GTGCCGGAGCGCGCGGAGGCGGAG
	MGMT-G4-bio	5′-bio-GGGCCGGGGCGCGCGGGGGCGGAG

Table 1. *Cont.*

Description	Name	Sequence *
Controls	MGMT-C	5′-CTCCGCCCCCGCGCGCCCCGGCCC
	MGMT-C-bio	5′-bio-CTCCGCCCCCGCGCGCCCCGGCCC
	non-specific control (no CpG)	5′-CTGAATACTACTTCCTACCCCTTACCTGAT
DNA duplexes	MGMT-ds1-f	5′-FAM-GGTCCTGCAGGCGCCCTCACTT 3′-CCAGGACGTCCGCGGGAGTGAA
	MGMT-ds1-bio	5′-bio-GGTCCTGCAGGCGCCCTCACTT 3′-CCAGGACGTCCGCGGGAGTGAA
	MGMT-ds2-bio	5′-bio-CTCCGCCCCCGCGCGCCCCGGCCC 3′-GAGGCGGGGCGCGCGGGGCCGGG

* In the case of MGMT-752, only the sequence of the cytosine-rich (coding) strand is shown. bio—biotin, f—5-carboxyfluorescein. MGMTp PQSs are highlighted in gray. Exon 1 sequence is highlighted in black. CpG sites are shown in bold; G4-disrupting substitutions are highlighted in red.

Then, we examined the efficiency of de novo methylation of MGMT-752 by the catalytic domain of mouse MTase Dnmt3a (Dnmt3a-CD), which is identical in amino acid sequence to the human enzyme, and is catalytically active in absence of its N-terminal chromatin-targeting part [51]. To this end, MGMT-752 duplex was enzymatically methylated by Dnmt3a-CD in the presence of methyl group donor AdoMet (Figure 1B). For comparison, a portion of MGMT-752 was methylated by a procaryotic monomeric C5 MTase M.SssI from *Spiroplasma* sp. strain MQ1 that targets CpG sites, while the control sample was left unmethylated.

In order to determine the enzymatic CpG methylation patterns of the amplicons, we employed nanopore sequencing. Purified DNA samples were analyzed using the Mk1B sequencing device equipped with a MinION Flow Cell (R10.4.1) (Figure 1C). The sequencing run generated 140,670 reads, yielding 117.69 megabases (Mb) of nucleotide data, which were subsequently analyzed using Dorado Basecaller software (version 0.6.2) and mapped to the GRCh38.p14 reference human genome assembly. Here, base modification calling was performed concurrently with basecalling using the supported Dorado model. The analysis utilized a model capable of detecting both 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC). Although 5hmC was not present in the tested samples, its detection served as a quality control measure for assessing the accuracy of the modification model. Using the unmethylated control amplicon as a reference, the 5mC probability distribution was plotted to evaluate modification calling accuracy across the region of interest. The accuracy distribution, generated with the Modkit Pileup tool, was separately visualized for the C-rich (coding) and G-rich (template) DNA strands of the control amplicon (Figure 2A). Unexpectedly, an anomalous region between CpG sites 40 and 65 exhibited a marked reduction in modification calling accuracy for both DNA strands, with the probability of correctly identifying unmethylated cytosines reaching as low as 25%. In contrast, accuracy across the remainder of the region consistently exceeded 75%, except at the 5′- and 3′- ends of the amplicon, where lower sequencing quality is a well-documented artifact [53].

To quantify the extent of false positive modification calls in the unmethylated sample, the probability distributions for all three cytosine variants—canonical, 5mC, and 5hmC—were plotted for control, Dnmt3a-CD-treated amplicons, and M.SssI-treated amplicons (Figure 2B). A significant amount of low-confidence, erroneous 5mC and 5hmC calls were detected in the control amplicon, likely corresponding to the low accuracy region between CpG sites 47–72. A number of 5hmC calls were also present in the methylated samples in the same low confidence interval. To account for these inaccuracies and exclude the erroneous

modification calls, a confidence threshold of 0.65 was used in the following analyses. Next, total amount of modified cytosines was calculated for all three studied amplicons. The 5mC levels were 0.04, 0.505, and 0.622 for control, Dnmt3a-CD-treated amplicons, and M.SssI-treated amplicons, respectively, with all three samples exhibiting a low level (<0.03) of high-confidence 5hmC calls. These calls were considered insignificant in the following analyses, and raising the confidence threshold further would invalidate a large number of correct modification calls. Surprisingly, the anomalous region of low modification calling accuracy closely coincided with the PQSs in the MGMTp CpG island, predicted using G4Hunter® software (Figure 2D). Therefore, the accuracy of detecting modified bases in PQS was significantly reduced compared to the surrounding sequences (Figure 2A,D). Notably, modification calling with the open-source 5mC analysis tool DeepMod 2 demonstrated greater accuracy in detecting canonical cytosines on the control amplicon compared to Oxford Nanopore's Dorado (Figure S1A).

A

B

C
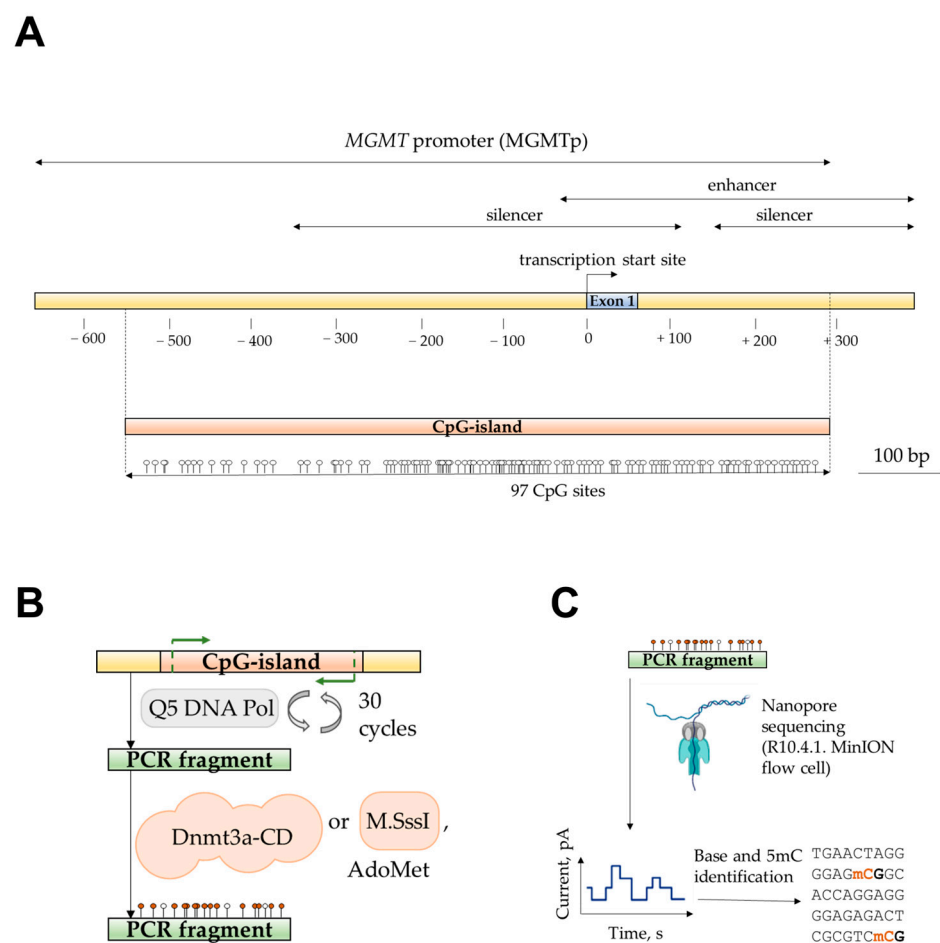
Figure 1. Detection of MGMTp CpG island methylation by nanopore sequencing. (**A**). Schematic representation of MGMTp; the region containing the CpG island is shown. Figure based on [52] and NCBI (Gene ID: 4255). (**B**). PCR amplification and in vitro methylation of a MGMTp region containing CpG sites 3–97 using Dnmt3a-CD or M.SssI. (**C**). Schematic representation of nanopore sequencing of methylated PCR product and detection of 5mC.
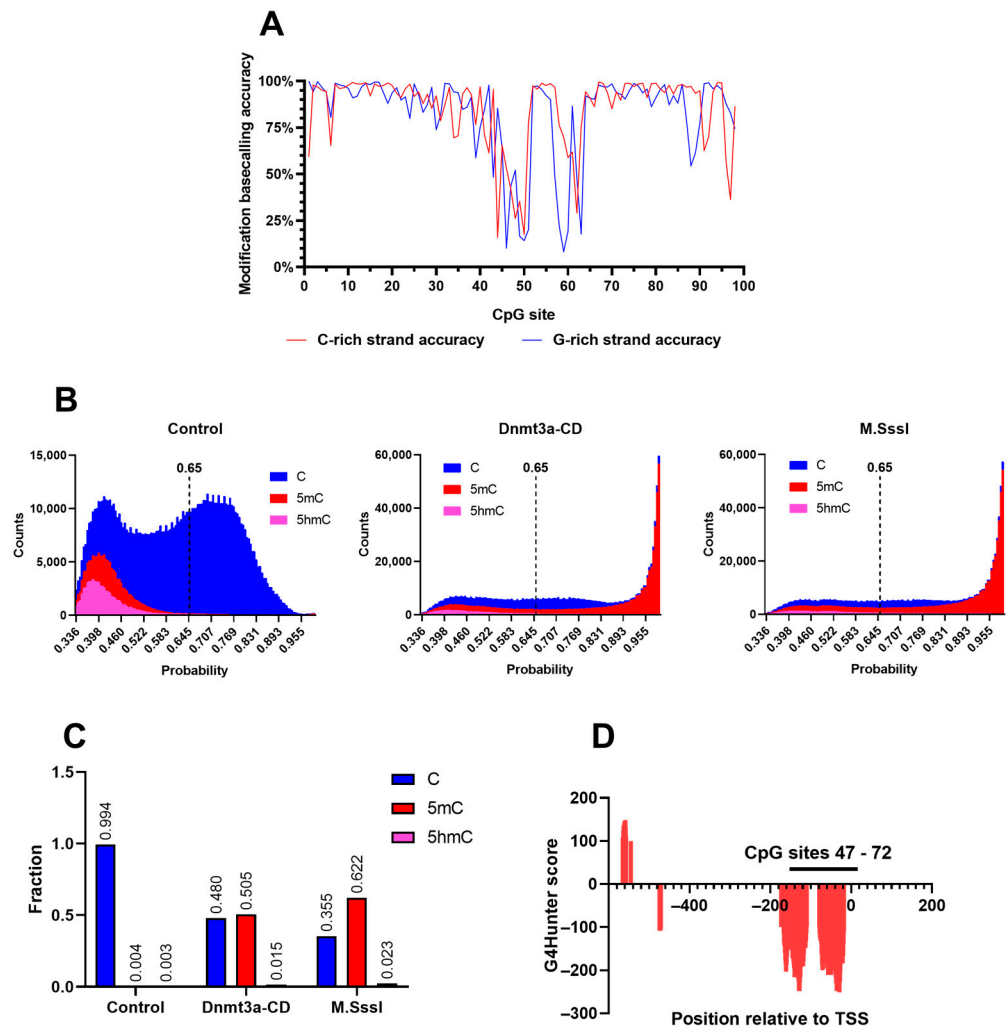
Figure 2. Modification calling of methylated and control PCR fragments of MGMTp CpG island (MGMT-752). (**A**). Distribution of base modification calling accuracy scores. Cytosine- and guanine-rich strands of the PCR fragment (C-rich and G-rich strand, respectively) were analyzed separately. (**B**). Modification probability distribution in unmethylated control and MGMT-752 methylated by Dnmt3a-CD or M.SssI. A confidence threshold of 0.65 (dashed line) was used to filter out erroneous methylation calls. (**C**). Total cytosine modification levels in unmethylated control and MGMT-752 methylated by Dnmt3a-CD or M.SssI. (**D**). PQS coverage of MGMTp CpG island predicted using G4Hunter software. TSS: transcription start site.

2.2. Differential Methylation of MGMTp DNA Strands by Dnmt3a-CD

The cytosine methylation patterns of the analyzed amplicons were next examined (Figure 3). Single-molecule sequencing allowed to differentiate the methylation patterns of the DNA strands. Control amplicons were uniformly unmethylated, whereas MTase-treated molecules exhibited methylation patterns that varied depending on the DNA strand and the enzyme used.

Unexpectedly, both Dnmt3a-CD and M.SssI exhibited differential methylation patterns between the C- and G-strands of the amplicon, with the C-strand generally displaying higher 5mC levels. Notably, amplicons treated with Dnmt3a-CD revealed a region of pronounced strand-specific methylation differences spanning CpG sites 47 to 72 (Figure 3, bottom panel). Interestingly, this region corresponded to an area previously identified as having reduced modification calling accuracy and overlapped with the PQS in MGMTp, as predicted by G4Hunter software (Figure S2A). Specifically, the G-rich DNA strand in

the primary PQS region displayed significantly lower methylation levels compared to its complementary C-rich strand (33% ± 3% vs. 58% ± 3%, respectively; *p* value < 0.0001) (Figure S2B). On the other hand, the difference of methylation efficiency between DNA strands outside of the primary PQS was not significant (47% ± 3% vs. 49% ± 3%, for G-rich and C-rich strands, respectively; *p* value 0.40) (Figure S2C). Additionally, a separate analysis using the DeepMod 2 tool produced a comparable methylation pattern, revealing significantly lower methylation levels in the PQS region of the G-rich strand (Figure S1B). Publicly available datasets from Oxford Nanopore Technologies Open Data were analyzed to validate the methylation calls within the MGMTp region of the GM24385 cell line genomic DNA. The resulting DNA methylation heatmaps demonstrate that methylation calls derived from nanopore data are largely consistent with those obtained through bisulfite sequencing of the same region (Figure S1C).

Figure 3. Heatmaps of methylation of MGMT-752 by Dnmt3a-CD or M.SssI. Unmethylated amplicon was used as a negative control. Methylation level represents the probability that the CpG site contains 5mC, averaged from all sequencing reads. **Top panel**: C-strand methylation heatmaps. **Middle panel**: G-strand methylation heatmaps. **Bottom panel**: combined heatmap of C- and G-strand methylation by Dnmt3a-CD. G4 coverage was calculated using G4Hunter software.

These findings suggest a possible relationship between the presence of PQS and the regulation of de novo methylation at CpG sites within these sequences. While the PQSs in the MGMT-752 amplicon are unlikely to form stable G4 structures due to competition with the energetically favorable B-form DNA duplex, MGMTp PQSs have been shown to adopt G4 folds in human cells [40]. Mao et al. [25] provided evidence that G4 structures which

were observed in human using a G4-specific antibodies influence the methylation at CpG islands. The authors proved the ability of maintenance DNMT1 MTase to colocalize at sites of G4 formation and proposed a mechanism for protecting CpG islands from methylation by G4 structures. Consequently, the next objective of this study was to assess the effect of MGMTp G4 structures on Dnmt3a-CD methylation efficiency and to determine the extent to which these differences could be attributed to the presence of G4s [23] or MTase sequence preferences [34].

2.3. MGMTp PQS and G4 Structures Form Stable Complexes with Dnmt3a-CD and Inhibit Its Methylation Activity

For this study, we selected a range of short DNA duplexes and oligonucleotides derived from the MGMTp CpG island (Figure 4A). These included a 24-bp double-stranded DNA fragment, MGMT-ds2, from the MGMTp PQS; its G-rich strand, MGMT-G4; and an analog, MGMT-G4-mut, which contained guanine substitutions designed to inhibit G4 formation (Table 1). Additionally, a C-rich oligonucleotide, MGMT-C, and a 24-bp DNA duplex from outside the PQS region (MGMT-ds1) were included. As a reference, a non-specific control oligonucleotide lacking CpG sites (non-sp. control) was also utilized.

The secondary structures of the single-stranded DNA fragments were analyzed using circular dichroism (CD) spectroscopy (Figure 4B). The MGMT-G4 oligonucleotide displayed a positive peak at 295 nm, indicative of a hybrid quadruplex structure characterized by a combination of parallel and antiparallel DNA strands [54]. In contrast, the MGMT-G4-mut oligonucleotide exhibited no evidence of non-canonical structures; its CD spectrum closely resembled that of the single-stranded non-specific control. Earlier, it was shown that a 24-bp oligonucleotide representing the MGMT promoter PQS exhibited CD spectra characteristic of a parallel G4 structure [40]. In contrast, our findings suggest a mixed-hybrid G4 folding conformation for the MGMT-G4 oligonucleotide (Figure 4A). This difference may result from our oligonucleotide covering a slightly different region of the same PQS, incorporating more downstream and fewer upstream nucleotides in the sequence (Figure 4A).

To evaluate the impact of the G4 structure on Dnmt3a-CD function, we investigated the inhibition of the methylation reaction of the fluorescently tagged MGMT-ds1-f duplex by oligonucleotides MGMT-G4 and controls (Table 1), following a previously established methodology (Figure 4C) [26]. MGMT-ds1-f was a 22-bp DNA duplex containing a single central CpG site, derived from a region of the MGMTp CpG island located outside the PQS. Methylation efficiency was measured by protection of methylated DNA from digestion by restriction endonuclease Hin6I. In the absence of MGMT-G4, the MGMT-ds1-f duplex was almost fully methylated by Dnmt3a-CD. However, as the concentration of MGMT-G4 increased, the degree of MGMT-ds1-f methylation decreased, indicating inhibition of Dnmt3a-CD activity (Figure S3). The IC_{50} value, derived from the plot of methylation fraction as a function of MGMT-G4 concentration (Figure 4C), was 0.61 ± 0.03 μM. In contrast, the control oligonucleotide without CpG sites (non-sp. control, Table 1) demonstrated significantly weaker inhibition of Dnmt3a-CD activity (IC_{50} 2.4 ± 0.7 μM). Interestingly, the MGMT-G4-mut, which, according to CD data, lacks a G4 structure, exhibited a similar level of inhibition as MGMT-G4 (IC_{50} 0.59 ± 0.09 μM). In this case, the presence of multiple CpG repeats in MGMT-G4-mut may promote the formation of transient double-stranded structures, resulting in a high affinity for Dnmt3a-CD. Therefore, the G4-forming oligonucleotide MGMT-G4, derived from the MGMTp PQS, effectively inhibited the methylation reaction by tightly binding to Dnmt3a-CD and preventing the MTase from binding to its regular double-stranded substrate.

Figure 4. Molecular characterization of the interaction of MGMTp G4 oligonucleotides with Dnmt3a-CD. (**A**). Short oligonucleotides and DNA duplexes based on MGMTp CpG island. (**B**). Circular dichroism studies of MGMT-G4 oligonucleotide and non-G4-forming controls (Table 1). (**C**). Inhibition of methylation of a double-stranded Dnmt3a substrate MGMT-ds1-f by Dnmt3a-CD in the presence of MGMT-G4. Error bars represent the SD from at least two independent experiments. (**D**). Displacement of MGMT-ds1-f from the complex with Dnmt3a-CD in the presence of MGMT-G4 studied by fluorescence polarization. (**E**). Binding of MGMT-G4 to Dnmt3a-CD studied by biolayer interferometry. The experiments were conducted in buffer A.

The mechanism of Dnmt3a–CD interaction with guanine-rich DNA sequences is supported by the results of a DNA displacement experiment involving the Dnmt3a-CD/MGMT-ds1-f complex. Fluorescence polarization demonstrated that adding the unlabeled guanine-rich sequence MGMT-G4, which can form a G4, to the Dnmt3a-CD/MGMT-ds1-f complex decreases the fluorescence polarization signal (Figure 4D). This suggests that the MGMT-G4 oligonucleotide displaces the FAM-labeled DNA duplex from the enzyme's binding site.

Next, we used bio-layer interferometry (BLI) to investigate the interaction between MGMTp G4 structures or PQS sequences and Dnmt3a-CD and to evaluate their properties as MTase substrates. Biotinylated DNA substrates were immobilized on streptavidin-coated biosensors. The experiment consisted of an association phase, during which the DNA substrates interacted with Dnmt3a-CD, followed by a dissociation phase to observe the stability of the complexes.

The BLI results demonstrate that Dnmt3a-CD binds to MGMT-ds1-bio and MGMT-ds2-bio with high affinity (Figure S4). The dissociation constants (K_d) of the DNA–protein complexes varied by a factor of 4, depending on the number of CpG-sites and the presence of G clusters (Table 2), with the guanine-rich duplex MGMT-ds2-bio binding to Dnmt3a-CD more strongly than the MGMT-ds1-bio duplex. Furthermore, we observed notable differences in the binding affinity of Dnmt3a-CD to complementary cytosine- and guanine-rich DNA strands of MGMT-ds2-bio (Figure 4E). The guanine-rich strand MGMT-G4-bio binds to the protein similarly to G/C-rich MGMT-ds2-bio and is 14 times more strongly than its complementary cytosine-rich strand MGMT-C-bio, confirming the high affinity of Dnmt3a-CD for G4 structures [23]. These findings are supported by the results of a DNA displacement experiment (Figure 4D).

Table 2. Dissociation constants of Dnmt3a-CD complexes with MGMTp DNA duplexes and oligonucleotides determined using BLI.

DNA Substrate	K_d, nM
MGMT-ds1-bio	65 ± 2
MGMT-ds2-bio	15 ± 0.5
MGMT-G4-bio	14 ± 0.4
MGMT-C-bio	211 ± 21

3. Discussion

In this study, we characterized, for the first time, the de novo methylation pattern of a long model DNA fragment from MGMTp using a Dnmt3a-CD-based in vitro methylation system (Figure 1). Utilizing nanopore sequencing for the direct detection of methylation products (Figure 2), we identified strand-specific effects and a distinct methylation pattern for Dnmt3a (Figure 3). Specifically, the region of strand-dependent Dnmt3a-mediated methylation coincided with the MGMTp PQS identified by G4Hunter algorithm, leading to highly uneven methylation across the 97 CpG positions of MGMTp.

The observed preferential methylation of the C-rich strand could be attributed to Dnmt3a's known strong flanking sequence preference [55], favoring CpG sites followed by cytosines and thymines. Bisulfite sequencing experiments have demonstrated that Dnmt3a exhibits preferences for bases flanking the CpG site at more distal positions [56]. These preferences may arise from the influence of the flanking sequence on protein–DNA complex formation or specific steps of the enzymatic reaction. In general, CpG sites flanked by cytosines within three bases in either the 3′ or 5′ direction are expected to be more favorable substrates for Dnmt3a-CD compared to CpG sites flanked by guanines.

Dnmt3a's distributive mechanism necessitates either dissociation and rebinding to hemimethylated CpG sites to methylate the complementary strand or the binding of a second tetramer to the same site [35]. In contrast, the effects were less pronounced with the monomeric CpG recognizing MTase M.SssI, which in spite of strong homology to ten conserved regions of the tetramerDnmt3a-CD, exhibits a differently arranged complex with DNA [57,58]. Thus, one can suggest that in the context of a PQS with a high density of CpG sites, Dnmt3a tetramers may preferentially methylate CpGs with favorable nucleotide contexts on the C-rich strand, resulting in the accumulation of hemimethylated CpG sites.

The accuracy of 5mC detection within the MGMTp PQS using nanopore sequencing data was significantly lower than for surrounding sequences, underscoring the need for improved modified basecalling models. This limitation, coupled with known challenges in nanopore sequencing of homopolymer regions—such as cytosine and guanine repeats in PQSs—highlights the importance of considering such sequences in the development and

refinement of basecalling algorithms. High-GC and homopolymeric sequences are known to be susceptible to basecalling errors, including deletions, insertions, and mismatches [53]. Simultaneously, ONT technology demonstrated a lower overall error rate in non-B DNA motifs compared to Illumina and PacBio, making it particularly advantageous for studying regions such as telomeres and CpG islands [59]. ONT methylation calls have undergone extensive validation in recent years, as evidenced by several studies [60,61], which include the analysis of PQS-containing promoter regions. Notably, nanopore methylation calls demonstrated a strong correlation with Sanger bisulfite sequencing data at the TRPA1 promoter, showing discrepancies primarily in the methylation range below 20%. Moreover, a recent study specifically validated ONT methylation calls for four CpGs within the MGMT promoter [62] using pyrosequencing, revealing a significant correlation between the results. Our findings indicate that the distinctive sequence composition of the MGMTp CpG island, characterized by clusters of homopolymeric cytosine and guanine tracts, introduces a different type of error—an increase in low-confidence erroneous methylation and hydroxymethylation calls.

The alignment of strand-specific methylation sites with PQS regions (Figure 3) suggested an influence of these structures on methylation in human cells. G4 structures might form locally in the G-rich strand, potentially affecting the catalytic activity of Dnmt3a-CD (Figure 4). The differences in the dissociation constants of Dnmt3a-CD complexes with MGMTp DNA duplexes and oligonucleotides were also revealed (Table 2). One possible explanation is that G4 structures, formed locally on the G-rich strand, may provide an independent interaction site for DNMT3a-CD, attracting the enzyme to G4s and resulting in hypomethylation of the surrounding regions. This result is consistent with the high binding affinity of recombinant human DNMT3A to G4 containing oligonucleotides derived from promoters of various human genes (*CDKN1C*, *c-MYC*, and others) [23]. At the same time, if the double-stranded form is preserved, increased methylation can be expected in PQS.

The characteristic methylation pattern observed in brain tumor samples aligns with our findings. This pattern, as reported by Shah et al. [63] in their comprehensive analysis of MGMT promoter methylation, correlates with MGMT mRNA expression and patient responses to primary GBM therapy. Our data may provide insight into the specific G4-forming sequence of the MGMT promoter that is predictive of gene silencing and clinical response.

The ability of PQS from MGMTp to fold into G4 structures upon interaction with G4-specific ligands in human glioblastoma cells was demonstrated by Fleming et al. in 2018 [40]. This finding was supported by bioinformatic analysis and circular dichroism (CD) experiments. However, these results and our CD experiments using short DNA oligonucleotides may not directly reflect the precise topology of MGMT promoter PQS G4 structures within human cells. Instead, they provide evidence supporting the potential for G4 formation in the studied region. Further studies are needed to clarify the interplay between strong flanking sequence preference and interaction with histone and non-canonical structures in de novo methylation of MGMTp CpG island by Dnmt3a.

Our findings also necessitate a reevaluation of research data obtained using methods such as the OneStep qMethyl™ Kit from Zymo Research [64]. This kit detects locus-specific DNA methylation by selectively amplifying methylated regions of DNA and comparing them to human non-methylated DNA. The reference DNA is purified from cells with genetic knockouts of both DNMT1 and DNMT3b DNA methyltransferases, resulting in a low level of DNA methylation (~5%) [65]. In the context of MGMTp, the application of this kit must be reconsidered, as the presence of PQS could influence the interpretation of methylation levels and compromise the accuracy of the results.

4. Materials and Methods

Oligonucleotides (Table 1) were commercial products (Genterra, Moscow, Russia). Some of the oligonucleotides contained biotin or the fluorescent label 6-carboxyfluorescein (FAM). The concentrations of the oligonucleotides were determined spectrophotometrically. DNA duplexes and G4 structures were formed using by heating at 95 °C for 3 min and slow cooling to 4 °C in buffer A: 20 mM HEPES-NaOH (pH 7.5), 100 mM KCl, 1 mM EDTA, 1 mM 1,4-dithiothreitol.

MGMTp amplification. MGMTp CpG island PCR product (MGMT-752; Table 1) was amplified using 1 μg of genomic DNA from the Raji human lymphoblast-like cell line (Eurogen, Moscow, Russia) as the template, along with NEB Q5® High-Fidelity DNA Polymerase (New England Biolabs, MA, USA) and the primers 5′-GGGATTCTCACTAAGCGGGC and 5′-CTGGCACCTAGAGGTAAGGC. The thermocycling conditions were as follows: initial denaturation at 98 °C for 30 s, followed by 30 cycles of denaturation at 98 °C for 5 s, primer annealing at 72 °C for 10 s, and extension at 72 °C for 30 s. A final extension step was performed at 72 °C for 2 min. The PCR products were analyzed by 1% agarose gel electrophoresis, and bands corresponding to the 752-bp product were excised and purified using the Cleanup S-Cap DNA Purification Kit (Eurogen, Moscow, Russia) following the manufacturer's instructions.

Enzymes. To obtain Dnmt3a-CD, Escherichia coli BL21 (DE3) cells were transformed with plasmid pET-28a(+) carrying the gene encoding Dnmt3a-CD with an N-terminal $6 \times$ His tag. Subsequently, Dnmt3a-CD was isolated and purified using metal-affinity chromatography on Co^{2+}-containing TALON® resin (GE Healthcare, Chicago, IL, USA). Plasmid pET-28a(+) encoding Dnmt3a-CD was provided by Prof. A. Jeltsch (University of Stuttgart, Stuttgart, Germany). R.Hin6I and M.SssI were commercial products (SibEnzyme, Novosibirsk, Russia). The purity of the protein samples was evaluated using electrophoresis in a 12% SDS-polyacrylamide gel. The protein concentrations per protein monomer were determined using the Bradford assay. The proteins were stored at -80 °C.

DNA methylation. 300 ng of the MGMT-752 amplicons were methylated using either Dnmt3a-CD (2 μM) or M.SssI (5 U) in the presence of AdoMet (25 μM) (Sigma, Steinheim, Germany) for 1 h at 37 °C in buffer A. The methylated amplicons were purified using the Cleanup S-Cap DNA Purification Kit.

Sequencing Library Preparation and Nanopore Sequencing. Sequencing libraries were prepared using 130 ng of either Dnmt3a-CD/M.SssI-treated, or unmethylated MGMT-752 amplicons following the Ligation Sequencing Amplicons—Native Barcoding Kit 24 V14 (SQK-NBD114.24) protocol (Oxford Nanopore, Oxford, UK). The sequencing mix was prepared with 30 μl of the barcoded DNA library. The mix was loaded onto a MinION R10.4.1 flow cell and run on an Mk1-B sequencing unit (Oxford Nanopore, Oxford, UK). The sequencing run lasted 1 h and 56 min, yielding 141,000 raw reads and 177 Mb of sequence data in POD5 format. The run was monitored using MinKNOW software (version 5.7.2). Offline basecalling of raw POD5 data was performed with the Dorado Basecaller tool (version 0.6.2) [66] using the supported dna_r10.4.1_e8.2_400bps_sup@v5.0.0 model. The following parameters were used: –modified-bases 5mCG_5hmCG; --kit-name SQK-NBD114-24. The resulting .BAM files containing basecalled reads with base modification calls were mapped to the GRCh38.p14 reference human genome assembly using Dorado aligner tool. DNA methylation levels and profiles were then analyzed using the Modkit Pileup tool (version 0.2.7) [67]. The following parameters were used: --motif CG 0; --with-header; --ref NC_000010.11\[129465781..129468355\].fa. In methylation calling accuracy analysis, '--no-filtering' parameter was included. In methylation pattern analysis, parameters '--filter-threshold 0.65 --mod-thresholds m:0.65 --mod-thresholds h:0.65' were used instead.

The methylation levels were visualized with Prism GraphPad software (version 8.0.1). The modification calling accuracy was defined as a probability that the cytosine is correctly identified as unmethylated. Therefore, an accuracy of 100% represented a 100% average probability among all sequencing reads that the position contained an unmodified cytosine, while an accuracy of 0% meant that there were equal probabilities of cytosine, 5mC, and 5hmC being present in the analyzed CpG site.

For comparison, 5mC calls in the unmethylated MGMT-752 amplicon were generated using DeepMod2 tool (version 0.3.0) [68]. Here, base modification calling was conducted after basecalling by Dorado Basecaller. The following parameters were used with DeepMod2 Detect command: --bam eb99195cdaae85326c878417bf262ed568a6d746_SQK-NBD114-24_barcode05.bam; --input pod5/; --model bilstm_r10.4.1_5khz_v4.3; --file_type pod5; --ref NC_000010.11\[129465781..129468355\].fa; --output deepmod2/--threads 8 --seq_type dna. The input .BAM file for DeepMod2 containing basecalled reads without base modification information was generated using Dorado Basecaller tool as detailed above, excluding the –modified-bases parameter. DNA methylation profiles were analyzed with the Modkit Pileup tool as described above. Methylation accuracy distributions were plotted with Prism GraphPad software (version 8.0.1).

Datasets used to validate 5mC calls were sourced from the Oxford Nanopore Technologies Open Data repository, available on Amazon Web Services S3 (https://labs.epi2 me.io/dataindex/, accessed on 2 January 2025). BED tables containing methylation frequencies, derived from either bisulfite sequencing or nanopore sequencing of GM24385 cell line genomic DNA, were used to visualize and compare the methylation patterns in the MGMTp region.

Statistical Analysis. The significance of the observed effects was assessed using a two-sided *t*-test with unequal variances, performed in Prism GraphPad software (version 8.0.1). Methylation efficiencies within or outside the primary PQS region of MGMTp (CpG sites 47–72) were compared between the C-rich and G-rich DNA strands, and two-tailed *p* values were calculated. Results were reported as mean ± SEM. PQS regions were identified with G4Hunter web application (https://bioinformatics.ibp.cz/#/analyse/quadruplex, accessed on 2 January 2025).

CD Measurements. CD spectra of oligonucleotides were recorded in a quartz cuvette of 10 mm optical path length at room temperature in buffer A on a Chirascan CD spectrometer (Applied Photophysics Ltd., Surrey, UK) equipped with a thermoelectric controller. The DNA concentration (~2 µM concentration per oligonucleotide strand) was chosen to attain an absorption of 0.6–0.8 at 260 nm, which gives an optimum signal-to-noise ratio. The measurements were performed in the 230–350 nm wavelength range at a scanning speed of 30 nm/min and a signal averaging time of 2 s with a constant flow of dry nitrogen. The CD spectra were normalized to molar circular dichroism ($\Delta\varepsilon$) using molar strand concentration as a reference. Spectra were baseline-corrected for signal contributions caused by the buffer and processed with GraphPad Prism 8.0.1 (GraphPad Software, San Diego, CA, USA).

Inhibition of DNA methylation by G4-forming oligonucleotides. MGMT-ds1-f (300 nM) was methylated by 2 µM Dnmt3a-CD in the presence of various concentrations of MGMT-G4, MGMT G4-mut or non-specific control oligonucleotide for 1.5 h at 37 °C in buffer A containing AdoMet (25 µM). Methylation efficiency was analyzed by the protection of methylated DNA from cleavage by R.Hin6I (G↓CGC). After digestion with R.Hin6I, the mixtures were analyzed on 20% polyacrylamide gel containing 7 M urea, with the determination of the extent of methylation as described in [26]. The gels were visualized using a Typhoon FLA 9500 scanner (GE Healthcare Life Sciences, Buckinghamshire, UK),

and the fluorescence intensities of intact DNA and cleavage products were determined. The methylation efficiency was calculated using GelQuantNET version 1.7.8. using the equation:

$$Methylation\ efficiency = \frac{w_0 - w_{Dnmt3a}}{w_0}$$

where w_0 is the extent of DNA cleavage before methylation and w_{Dnmt3a} is the extent of DNA cleavage after methylation by Dnmt3a-CD. IC_{50} values were calculated via fitting the dependence of the extent of methylation on the concentration of an oligonucleotide using GraphPad Prism 8.0.1 (GraphPad Software, San Diego, CA, USA).

Bio-layer interferometry. DNA duplexes or oligonucleotides (690 nM) were used for the experiment. The binding buffer contained 20 mM HEPES-NaOH (pH 7.5), 100 mM KCl, 1 mM EDTA, 0.2 mM DTT, 100 μM AdoHcy, 5% glycerol, 0.02% Tween-20, and 0.5 mg/mL BSA. Bio-layer interferometry (BLI) analyses were performed using the BLItz instrument (ForteBio, Fremont, CA, USA) in extended kinetics mode with stirring at 2200 rpm in triplicate. Streptavidin biosensors (ForteBio, Fremont, CA, USA) were hydrated in the binding buffer for 10 min. prior to measurements. The optimized BLI protocol included the following steps: (i) incubation for 30 s; (ii) biotinylated DNA substrate immobilization for 120 s; (iii) sensor wash for 30 s; (iv) DNA binding to Dnmt3a-CD (690 nM in binding buffer) for 300 s; (v) dissociation of the DNA-protein complex in binding buffer for 120 s. All measurements were conducted in black microtubes (Sigma-Aldrich, New York, NY, USA) using a minimum of 300 μL of the appropriate solution. The resulting binding curves were fitted to a 1:1 DNA:protein binding model using an exponential approximation using GraphPad Prism 8.0.1 (GraphPad Software, San Diego, CA, USA).

5. Conclusions

We demonstrated that PQS within the MGMT promoter influence the activity of MTase Dnmt3a, a key enzyme responsible for de novo methylation and establishing the MGMT promoter's methylation pattern. Our findings underscore the role of PQS in regulating MGMT promoter methylation and highlight the technical challenges posed by guanine-rich sequences in nanopore sequencing. Specifically, the presence of PQS in the sequenced region impacts basecalling accuracy in nanopore data, emphasizing the need to account for such structures during analysis. We found the strong specific binding of Dnmt3a-CD to the G4-forming oligonucleotide MGMT-G4, derived from the MGMTp PQS, as well as inhibition of Dnmt3a-CD activity by this oligonucleotide. Also, effective binding of guanine-rich duplex MGMT-ds2 to Dnmt3a-CD was observed. These findings suggest that Dnmt3a-CD recognizes MGMTp CpG island G4 structures and competes with the DNA duplex substrate, thereby modulating the methylation efficiency at nearby CpG sites.

Author Contributions: Conceptualization, M.I.Z.; Funding acquisition, G.V.P. and M.I.Z.; Investigation, A.V.S., D.P.M. and A.I.G.; Project administration, M.I.Z.; Resources, M.I.Z.; Supervision, A.V.S., E.S.G. and M.I.Z.; Visualization, A.V.S.; Writing—original draft, A.V.S.; Writing—review and editing, E.S.G. and M.I.Z. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Siegel, R.L.; Giaquinto, A.N.; Jemal, A. Cancer Statistics, 2024. *CA Cancer J. Clin.* **2024**, *74*, 12–49. [CrossRef] [PubMed]
2. Bush, N.A.O.; Chang, S.M.; Berger, M.S. Current and Future Strategies for Treatment of Glioma. *Neurosurg. Rev.* **2017**, *40*, 1–14. [CrossRef] [PubMed]
3. Hasanau, T.; Pisarev, E.; Kisil, O.; Nonoguchi, N.; Le Calvez-Kelm, F.; Zvereva, M. Detection of TERT Promoter Mutations as a Prognostic Biomarker in Gliomas: Methodology, Prospects, and Advances. *Biomedicines* **2022**, *10*, 728. [CrossRef]
4. Liu, H.; Weng, J.; Huang, C.L.H.; Jackson, A.P. Is the Voltage-Gated Sodium Channel B3 Subunit (SCN3B) a Biomarker for Glioma? *Funct. Integr. Genomics* **2024**, *24*, 162. [CrossRef] [PubMed]
5. Liu, H.; Weng, J. A Comprehensive Bioinformatic Analysis of Cyclin-Dependent Kinase 2 (CDK2) in Glioma. *Gene* **2022**, *822*, 146325. [CrossRef] [PubMed]
6. Liu, H.; Tang, T. A Bioinformatic Study of IGFBPs in Glioma Regarding Their Diagnostic, Prognostic, and Therapeutic Prediction Value. *Am. J. Transl. Res.* **2023**, *15*, 2140–2155.
7. Sonkin, D.; Thomas, A.; Teicher, B.A. Cancer Treatments: Past, Present, and Future. *Cancer Genet.* **2024**, *286–287*, 18–24. [CrossRef]
8. Smith, Z.D.; Meissner, A. DNA Methylation: Roles in Mammalian Development. *Nat. Rev. Genet.* **2013**, *14*, 204–220. [CrossRef]
9. Bird, A. DNA Methylation de Novo. *Science* **1999**, *286*, 2287–2288. [CrossRef]
10. Smith, Z.D.; Hetzel, S.; Meissner, A. DNA Methylation in Mammalian Development and Disease. *Nat. Rev. Genet.* **2024**, *26*, 7–30. [CrossRef]
11. Deaton, A.M.; Bird, A. CpG Islands and the Regulation of Transcription. *Genes Dev.* **2011**, *25*, 1010–1022. [CrossRef] [PubMed]
12. Jones, P.A. Functions of DNA Methylation: Islands, Start Sites, Gene Bodies and Beyond. *Nat. Rev. Genet.* **2012**, *13*, 484–492. [CrossRef] [PubMed]
13. Shen, L.; Kondo, Y.; Guo, Y.; Zhang, J.; Zhang, L.; Ahmed, S.; Shu, J.; Chen, X.; Waterland, R.A.; Issa, J.-P.J. Genome-Wide Profiling of DNA Methylation Reveals a Class of Normally Methylated CpG Island Promoters. *PLoS Genet.* **2007**, *3*, e181. [CrossRef] [PubMed]
14. Jeltsch, A.; Jurkowska, R.Z. New Concepts in DNA Methylation. *Trends Biochem. Sci.* **2014**, *39*, 310–318. [CrossRef]
15. Jurkowska, R.Z.; Jurkowski, T.P.; Jeltsch, A. Structure and Function of Mammalian DNA Methyltransferases. *ChemBioChem* **2011**, *12*, 206–222. [CrossRef]
16. Xu, T.-H.; Liu, M.; Zhou, X.E.; Liang, G.; Zhao, G.; Xu, H.E.; Melcher, K.; Jones, P.A. Structure of Nucleosome-Bound DNA Methyltransferases DNMT3A and DNMT3B. *Nature* **2020**, *586*, 151–155. [CrossRef]
17. Chen, T.; Tsujimoto, N.; Li, E. The PWWP Domain of Dnmt3a and Dnmt3b Is Required for Directing DNA Methylation to the Major Satellite Repeats at Pericentric Heterochromatin. *Mol. Cell Biol.* **2004**, *24*, 9048–9058. [CrossRef]
18. Guo, X.; Wang, L.; Li, J.; Ding, Z.; Xiao, J.; Yin, X.; He, S.; Shi, P.; Dong, L.; Li, G.; et al. Structural Insight into Autoinhibition and Histone H3-Induced Activation of DNMT3A. *Nature* **2014**, *517*, 640–644. [CrossRef]
19. Dukatz, M.; Holzer, K.; Choudalakis, M.; Emperle, M.; Lungu, C.; Bashtrykov, P.; Jeltsch, A. H3K36me2/3 Binding and DNA Binding of the DNA Methyltransferase DNMT3A PWWP Domain Both Contribute to Its Chromatin Interaction. *J. Mol. Biol.* **2019**, *431*, 5063–5074. [CrossRef]
20. Kubo, N.; Uehara, R.; Uemura, S.; Ohishi, H.; Shirane, K.; Sasaki, H. Combined and Differential Roles of ADD Domains of DNMT3A and DNMT3L on DNA Methylation Landscapes in Mouse Germ Cells. *Nat. Commun.* **2024**, *15*, 3266. [CrossRef]
21. Wapenaar, H.; Clifford, G.; Rolls, W.; Pasquier, M.; Burdett, H.; Zhang, Y.; Deák, G.; Zou, J.; Spanos, C.; Taylor, M.R.D.; et al. The N-Terminal Region of DNMT3A Engages the Nucleosome Surface to Aid Chromatin Recruitment. *EMBO Rep.* **2024**, *25*, 5743–5779. [CrossRef] [PubMed]
22. Bannister, A.J.; Kouzarides, T. Regulation of Chromatin by Histone Modifications. *Cell Res.* **2011**, *21*, 381–395. [CrossRef] [PubMed]
23. Cree, S.L.; Fredericks, R.; Miller, A.; Pearce, F.G.; Filichev, V.; Fee, C.; Kennedy, M.A. DNA G-Quadruplexes Show Strong Interaction with DNA Methyltransferases in Vitro. *FEBS Lett.* **2016**, *590*, 2870–2883. [CrossRef] [PubMed]
24. Rauchhaus, J.; Robinson, J.; Monti, L.; Di Antonio, M. G-Quadruplexes Mark Sites of Methylation Instability Associated with Ageing and Cancer. *Genes* **2022**, *13*, 1665. [CrossRef]

25. Mao, S.-Q.; Ghanbarian, A.T.; Spiegel, J.; Martínez Cuesta, S.; Beraldi, D.; Di Antonio, M.; Marsico, G.; Hänsel-Hertsch, R.; Tannahill, D.; Balasubramanian, S. DNA G-Quadruplex Structures Mold the DNA Methylome. *Nat. Struct. Mol. Biol.* **2018**, *25*, 951–957. [CrossRef]

26. Sergeev, A.V.; Loiko, A.G.; Genatullina, A.I.; Petrov, A.S.; Kubareva, E.A.; Dolinnaya, N.G.; Gromova, E.S. Crosstalk between G-Quadruplexes and Dnmt3a-Mediated Methylation of the c-MYC Oncogene Promoter. *Int. J. Mol. Sci.* **2023**, *25*, 45. [CrossRef]

27. Halder, R.; Halder, K.; Sharma, P.; Garg, G.; Sengupta, S.; Chowdhury, S. Guanine Quadruplex DNA Structure Restricts Methylation of CpG Dinucleotides Genome-Wide. *Mol. Biosyst.* **2010**, *6*, 2439–2447. [CrossRef]

28. Jara-Espejo, M.; Line, S.R. DNA G-quadruplex Stability, Position and Chromatin Accessibility Are Associated with CpG Island Methylation. *FEBS J.* **2020**, *287*, 483–495. [CrossRef]

29. Varizhuk, A.; Isaakova, E.; Pozmogova, G. DNA G-Quadruplexes (G4s) Modulate Epigenetic (Re)Programming and Chromatin Remodeling. *BioEssays* **2019**, *41*, 1900091. [CrossRef]

30. Qin, Y.; Hurley, L.H. Structures, Folding Patterns, and Functions of Intramolecular DNA G-Quadruplexes Found in Eukaryotic Promoter Regions. *Biochimie* **2008**, *90*, 1149–1171. [CrossRef]

31. Ngo, K.H.; Liew, C.W.; Heddi, B.; Phan, A.T. Structural Basis for Parallel G-Quadruplex Recognition by an Ankyrin Protein. *J. Am. Chem. Soc.* **2024**, *146*, 13709–13713. [CrossRef] [PubMed]

32. Chambers, V.S.; Marsico, G.; Boutell, J.M.; Di Antonio, M.; Smith, G.P.; Balasubramanian, S. High-Throughput Sequencing of DNA G-Quadruplex Structures in the Human Genome. *Nat. Biotechnol.* **2015**, *33*, 877–881. [CrossRef] [PubMed]

33. Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative Visualization of DNA G-Quadruplex Structures in Human Cells. *Nat. Chem.* **2013**, *5*, 182–186. [CrossRef] [PubMed]

34. Dukatz, M.; Dittrich, M.; Stahl, E.; Adam, S.; de Mendoza, A.; Bashtrykov, P.; Jeltsch, A. DNA Methyltransferase DNMT3A Forms Interaction Networks with the CpG Site and Flanking Sequence Elements for Efficient Methylation. *J. Biol. Chem.* **2022**, *298*, 102462. [CrossRef] [PubMed]

35. Rajavelu, A.; Jurkowska, R.Z.; Fritz, J.; Jeltsch, A. Function and Disruption of DNA Methyltransferase 3a Cooperative DNA Binding and Nucleoprotein Filament Formation. *Nucleic Acids Res.* **2012**, *40*, 569–580. [CrossRef]

36. Esain-Garcia, I.; Kirchner, A.; Melidis, L.; de Cesaris Araujo Tavares, R.; Dhir, S.; Simeone, A.; Yu, Z.; Madden, S.K.; Hermann, R.; Tannahill, D.; et al. G-Quadruplex DNA Structure Is a Positive Regulator of MYC Transcription. *Proc. Natl. Acad. Sci. USA* **2024**, *121*, e2320240121. [CrossRef]

37. Kulis, M.; Esteller, M. DNA Methylation and Cancer. *Adv. Genet.* **2010**, *70*, 27–56. [CrossRef]

38. Zhang, J.; Yang, C.; Wu, C.; Cui, W.; Wang, L. DNA Methyltransferases in Cancer: Biology, Paradox, Aberrations, and Targeted Therapy. *Cancers* **2020**, *12*, 2123. [CrossRef]

39. Berdasco, M.; Esteller, M. Aberrant Epigenetic Landscape in Cancer: How Cellular Identity Goes Awry. *Dev. Cell* **2010**, *19*, 698–711. [CrossRef]

40. Fleming, A.M.; Zhu, J.; Ding, Y.; Visser, J.A.; Zhu, J.; Burrows, C.J. Human DNA Repair Genes Possess Potential G-Quadruplex Sequences in Their Promoters and 5′-Untranslated Regions. *Biochemistry* **2018**, *57*, 991–1002. [CrossRef]

41. Yu, W.; Zhang, L.; Wei, Q.; Shao, A. O6-Methylguanine-DNA Methyltransferase (MGMT): Challenges and New Opportunities in Glioma Chemotherapy. *Front. Oncol.* **2020**, *9*, 1547. [CrossRef] [PubMed]

42. Chen, Y.; Simeone, A.; Melidis, L.; Cuesta, S.M.; Tannahill, D.; Balasubramanian, S. An Upstream G-Quadruplex DNA Structure Can Stimulate Gene Transcription. *ACS Chem. Biol.* **2024**, *19*, 736–742. [CrossRef] [PubMed]

43. Jiang, J.; Xu, J.; Ji, S.; Yu, X.; Chen, J. Unraveling the Mysteries of MGMT: Implications for Neuroendocrine Tumors. *Biochim. Biophys. Acta (BBA)—Rev. Cancer* **2024**, *1879*, 189184. [CrossRef] [PubMed]

44. Zhang, Z.; Xin, S.; Gao, M.; Cai, Y. Promoter Hypermethylation of MGMT Gene May Contribute to the Pathogenesis of Gastric Cancer. *Medicine* **2017**, *96*, e6708. [CrossRef] [PubMed]

45. Inno, A. Role of MGMT as Biomarker in Colorectal Cancer. *World J. Clin. Cases* **2014**, *2*, 835–839. [CrossRef]

46. An, N.; Shi, Y.; Ye, P.; Pan, Z.; Long, X. Association Between MGMT Promoter Methylation and Breast Cancer: A Meta-Analysis. *Cell. Physiol. Biochem.* **2017**, *42*, 2430–2440. [CrossRef]

47. Kordi-Tamandani, D.M.; Moazeni-Roodi, A.-K.; Rigi-Ladiz, M.-A.; Hashemi, M.; Birjandian, E.; Torkamanzehi, A. Promoter Hypermethylation and Expression Profile of MGMT and CDH1 Genes in Oral Cavity Cancer. *Arch. Oral. Biol.* **2010**, *55*, 809–814. [CrossRef]

48. Huang, J.; Luo, J.-Y.; Tan, H.-Z. Associations of MGMT Promoter Hypermethylation with Squamous Intraepithelial Lesion and Cervical Carcinoma: A Meta-Analysis. *PLoS ONE* **2019**, *14*, e0222772. [CrossRef]

49. Chen, Y.; Qu, W.; Tu, J.; Qi, H. Implications of Advances in Studies of O6-Methylguanine-DNA- Methyltransferase for Tumor Prognosis and Treatment. *Front. Biosci.-Landmark* **2023**, *28*, 197. [CrossRef]

50. Della Monica, R.; Cuomo, M.; Buonaiuto, M.; Costabile, D.; Franca, R.A.; Del Basso De Caro, M.; Catapano, G.; Chiariotti, L.; Visconti, R. MGMT and Whole-Genome DNA Methylation Impacts on Diagnosis, Prognosis and Therapy of Glioblastoma Multiforme. *Int. J. Mol. Sci.* **2022**, *23*, 7148. [CrossRef]

51. Gowher, H.; Jeltsch, A. Molecular Enzymology of the Catalytic Domains of the Dnmt3a and Dnmt3b DNA Methyltransferases. *J. Biol. Chem.* **2002**, *277*, 20409–20414. [CrossRef] [PubMed]

52. Ji, W.; Yang, L.; Yu, L.; Yuan, J.; Hu, D.; Zhang, W.; Yang, J.; Pang, Y.; Li, W.; Lu, J.; et al. Epigenetic Silencing of O6 -Methylguanine DNA Methyltransferase Gene in NiS-Transformed Cells. *Carcinogenesis* **2008**, *29*, 1267–1275. [CrossRef] [PubMed]

53. Delahaye, C.; Nicolas, J. Sequencing DNA with Nanopores: Troubles and Biases. *PLoS ONE* **2021**, *16*, e0257521. [CrossRef] [PubMed]

54. Kejnovská, I.; Renčiuk, D.; Palacký, J.; Vorlíčková, M. CD Study of the G-Quadruplex Conformation. In *Methods in Molecular Biology (Clifton, N.J.)*; Humana Press Inc.: Totowa, NJ, USA, 2019; Volume 2035, pp. 25–44.

55. Mallona, I.; Ilie, I.M.; Karemaker, I.D.; Butz, S.; Manzo, M.; Caflisch, A.; Baubec, T. Flanking Sequence Preference Modulates *de Novo* DNA Methylation in the Mouse Genome. *Nucleic Acids Res.* **2021**, *49*, 145–157. [CrossRef]

56. Jurkowska, R.Z.; Siddique, A.N.; Jurkowski, T.P.; Jeltsch, A. Approaches to Enzyme and Substrate Design of the Murine Dnmt3a DNA Methyltransferase. *ChemBioChem* **2011**, *12*, 1589–1594. [CrossRef]

57. Wojciechowski, M.; Czapinska, H.; Bochtler, M. CpG Underrepresentation and the Bacterial CpG-Specific DNA Methyltransferase M.MpeI. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 105–110. [CrossRef]

58. Zhang, Z.M.; Lu, R.; Wang, P.; Yu, Y.; Chen, D.; Gao, L.; Liu, S.; Ji, D.; Rothbart, S.B.; Wang, Y.; et al. Structural Basis for DNMT3A-Mediated de Novo DNA Methylation. *Nature* **2018**, *554*, 387–391. [CrossRef]

59. Weissensteiner, M.H.; Cremona, M.A.; Guiblet, W.M.; Stoler, N.; Harris, R.S.; Cechova, M.; Eckert, K.A.; Chiaromonte, F.; Huang, Y.F.; Makova, K.D. Accurate Sequencing of DNA Motifs Able to Form Alternative (Non-B) Structures. *Genome Res.* **2023**, *33*, 907–923. [CrossRef]

60. Nyaga, D.M.; Tsai, P.; Gebbie, C.; Phua, H.H.; Yap, P.; Le Quesne Stabej, P.; Farrow, S.; Rong, J.; Toldi, G.; Thorstensen, E.; et al. Benchmarking Nanopore Sequencing and Rapid Genomics Feasibility: Validation at a Quaternary Hospital in New Zealand. *npj Genom. Med.* **2024**, *9*, 57. [CrossRef]

61. Gombert, S.; Jahn, K.; Pathak, H.; Burkert, A.; Schmidt, G.; Wiehlmann, L.; Davenport, C.; Brändl, B.; Müller, F.J.; Leffler, A.; et al. Comparison of Methylation Estimates Obtained via MinION Nanopore Sequencing and Sanger Bisulfite Sequencing in the TRPA1 Promoter Region. *BMC Med. Genomics* **2023**, *16*, 257. [CrossRef]

62. Halldorsson, S.; Nagymihaly, R.M.; Patel, A.; Brandal, P.; Panagopoulos, I.; Leske, H.; Lund-Iversen, M.; Sahm, F.; Vik-Mo, E.O. Accurate and Comprehensive Evaluation of O6-Methylguanine-DNA Methyltransferase Promoter Methylation by Nanopore Sequencing. *Neuropathol. Appl. Neurobiol.* **2024**, *50*, e12984. [CrossRef]

63. Shah, N.; Lin, B.; Sibenaller, Z.; Ryken, T.; Lee, H.; Yoon, J.-G.; Rostad, S.; Foltz, G. Comprehensive Analysis of MGMT Promoter Methylation: Correlation with MGMT Expression and Clinical Response in GBM. *PLoS ONE* **2011**, *6*, e16146. [CrossRef]

64. OneStep QMethylTM Kit. Available online: https://epigenie.com/products/onestep-qmethyl-kit/ (accessed on 21 November 2024).

65. OneStep PLUS QMethylTM PCR Kit. Available online: https://files.zymoresearch.com/protocols/_d5312_onestep_plus_q_methyl_pcr_kit.pdf (accessed on 22 November 2024).

66. Dorado—Oxford Nanopore's Basecaller. Available online: https://github.com/nanoporetech/dorado (accessed on 28 November 2024).

67. Modkit—A Bioinformatics Tool for Working with Modified Bases from Oxford Nanopore. Available online: https://github.com/nanoporetech/modkit (accessed on 31 December 2024).

68. DeepMod2—A Computational Tool for Detecting DNA 5mC Methylation from Oxford Nanopore Reads. Available online: https://github.com/WGLab/DeepMod2 (accessed on 5 January 2025).