Convergence of DNA methylation and phosphorothioation epigenetics in bacterial genomes

Chao Chen,a,b,1 Lianrong Wang,a,1 Si Chen,a Xiaolin Wu,a,b Meijia Gu,a Xi Chen,a Susu Jiang,a Yunfu Wang,b Zixin Deng,a Peter C. Dedon,c and Shi Chen,a,2

*Key Laboratory of Combinatorial Biosynthesis and Drug Discovery, Ministry of Education, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China; †Taihe Hospital, Hubei University of Medicine, Shiyan 442000, Hubei, China; ‡Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139; and §Medical Research Institute, Wuhan University, Wuhan 430071, China

Edited by Philip C. Hanawalt, Stanford University, Stanford, CA, and approved March 14, 2017 (received for review February 13, 2017)

Explosive growth in the study of microbial epigenetics has revealed a diversity of chemical structures and biological functions of DNA modifications in restriction–modification (R-M) and basic genetic processes. Here, we describe the discovery of shared consensus sequences for two seemingly unrelated DNA modification systems, 6mA methylation and phosphorothioation (PT), in which sulfur replaces a nonbridging oxygen in the DNA backbone. Mass spectrometric analysis of DNA from Escherichia coli B7A and Salmonella enterica serovar Cerro 87, strains possessing PT-based R-M genes, revealed d(Gp56mA) dinucleotides in the Gp56mA/AAC consensus representing ~5% of the 1,100 to 1,300 PT-modified d(Gp5A) motifs per genome, with 6mA arising from a yet-to-be-identified methyltransferase. To further explore PT and 6mA in another consensus sequence, Gp56mATC, we engineered a strain of E. coli HST04 to express Dnd genes from Hahella chejuensis KCTC23936 (PT in Gp5ATC) and Dam methyltransferase from E. coli DH10B (6mA in Gp5ATC). Based on this model, in vitro studies revealed reduced Dam activity in Gp5ATC-containing oligonucleotides whereas single-molecule real-time sequencing of HST04 DNA revealed 6mA in all 2,058 Gp5ATC sites (5% of total GATC sites). This model system also revealed temperature-sensitive restriction by DndFGH in KCTC2396 and B7A, which was exploited to discover that 6mA can substitute for PT to confer resistance to restriction by the DndFGH system. These results point to complex but unappreciated interactions between DNA modification systems and raise the possibility of coevolution of interacting systems to facilitate the function of each.

DNA phosphorothiation | DNA methylation | epigenetics | restriction–modification | bioanalytical chemistry

The emergence of convergent technologies has led to a growing appreciation for the diversity of DNA modifications in microbial epigenetics and restriction–modification (R-M) systems (1–3). DNA methylation, the most extensively studied genetic modification, was originally discovered in bacteria in the context of R-M systems involving a methyltransferase (MTase) that modifies “self” DNA at specific target sites and a cognate restriction endonuclease (REase) that discriminates and destroys unmodified invading DNA (3–5). R-M systems are ubiquitous in prokaryotes and are classified into four types (I, II, III, and IV) based on their molecular structure, sequence recognition, cleavage position, and cofactor requirements (3, 6, 7). However, some MTases exist alone, without an apparent cognate REase partner. These so-called “orphan” MTases include DNA adenine methylase (Dam), which modifies the adenine N-6 in the GATC motif, DNA cytosine methylase (Dcm), which methylates C-5 of the second cytosine in CC(A/T) GG sequences, and cell cycle-regulated methylase (CcrM), which methylates the N-6 of adenine in GANTC (N = A, T, C, or G) (8). Despite the absence of cognate REases, orphan MTases still confer immunity against the virulence of a parasitic R-M complex with the same target sites (9). In addition to defense against bacteriophages and transposons, DNA methylation is involved in various cellular functions, including chromosome replication, DNA mismatch repair, transcriptional regulation, transposition, and pathogenesis (1, 10, 11).

The chemical diversity of DNA modifications expanded with our discovery of phosphorothioates (PTs) as physiological modifications of the DNA sugar-phosphate backbone in which a nonbridging oxygen is replaced by sulfur (12). The PT system is more complicated than methylation and entails sequence-selective and R-M-specific incorporation of sulfur by five Dnd proteins (12, 13). Since the original report in Streptomyces lividans, the five-member dndABCDE gene cluster has been found in diverse, taxonomically unrelated bacteria and archaea (14, 15). dndA can be located either adjacent to clustered dndBCDE or elsewhere in the genome, consistent with the function of DndA as a cysteine desulfurase; DndA can also be functionally replaced by an iron-sulfur cluster assembly protein (IssS) homolog in some bacterial strains (16). DndB binds to the promoter of the dnd operon to regulate dndBCDE transcription and affect PT frequency (17, 18). DndC possesses ATP pyrophosphatase activity and shows significant homology to PAPS reductase family proteins (13, 16). By comparison, DndD has ATPase activity that is possibly related to DNA nicking during sulfur incorporation, and structural analysis suggests that DndE binds to nicked double-stranded DNA (13, 19, 20). In some strains, a divergently transcribed cluster, dndFGH, occurs in the vicinity of the dndBCDE operon. Without PT protection, unrestrained DndFGH restriction activity leads to double-stranded DNA damage that triggers the SOS response, cell filamentation, and prophage

Significance

New technologies have fostered renewed interest in bacterial epigenetics, with DNA modifications defending against other microbes and controlling gene expression. Here, we describe how two diverse modifications, adenine methylation (6mA) and phosphorothiation (PT), have evolved to occupy the same genomic sites in bacteria. Using mass spectrometry, single-molecule real-time sequencing, and genetic engineering, we discovered coincident PT and 6mA in Gp56mA/AAC and Gp56mATC, representing PT consensus sequences, a GA-containing consensus for an unknown methyltransferase, and the GATC consensus for Dam methyltransferase. 6mA was found in all Gp5ATC sites in vivo, with 6mA substituting for PT in preventing DNA cleavage by PT restriction enzymes. These findings show how different DNA modification systems have evolved to avoid conflict.

Author contributions: L.W., P.C.D., and Shi Chen designed research; C.C., X.W., M.G., X.C., and S.J. conducted experiments; C.C., L.W., Si Chen, X.W., Y.W., Z.D., P.C.D., and Shi Chen analyzed data; C.C., L.W., P.C.D., and Shi Chen wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. Freely available online through the PNAS open access option.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1702450114/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1702450114

PNAS | April 25, 2017 | vol. 114 | no. 17 | 4501–4506
induction (21, 22). Dnd is thus regarded as an unusual bacterial R-M system.

The complexities of different DNA modification systems have begun to come to light with the application of new technologies. For example, the application of liquid chromatography-coupled tandem quadrupole mass spectrometry (LC-MS/MS) to quantify nucleic-acid-resistant PT-modified dinucleotides has revealed significant diversity in the quantity, sequence setting, and bacterial distribution of PT modifications, such as the d(GpsA) and d(GpsT) motifs in 1:1 proportions in the enteric pathogen Escherichia coli B7A (15). Subsequent application of single-molecule, real-time (SMRT) sequencing technology to map genomic PT sites in E. coli B7A revealed that these d(GpsA) and d(GpsT) motifs are located within conserved 4-bp complementary GpsAAC/GpsTTT contexts, with no apparent wider consensus sequence (23). However, only 12% of the 40,701 possible GAAC/GTTC sites on the context with no apparent wider consensus sequence (23). How ever, only 12% of the 40,701 possible GAAC/GTTC sites on the chromosome were found to be PT-modified, even in the presence of active DndFGH (23). These unique features differentiate dndBCDE-dndFGH from typical methylation-based R-M systems and suggest that it represents a type of bacterial defense system with highly unusual target selection by Dnd proteins.

Despite having similar functional roles, DNA methylation and PT modification are considered to be two R-M defense systems that have evolved independently in prokaryotes. Here, we report the discovery of a potential conflict among the two modification systems, which were found to share DNA substrates to generate a hybrid d(Gps6mA) modification motif. This commonality is associated with significant rules of interaction and interference: sulfur replacement on the DNA backbone remarkably inhibits methylation, and methylation can substitute for PT to confer resistance to restriction by DndFGH. This interaction between two distinct endogenous DNA modification systems suggests that such “conflict” drives the evolution of different consensus sequences to avoid interference and ensure the protective function of each.

Results

Discovery of the Doubly Modified d(Gps6mA) Consensus in Bacteria. PT modifications occur sequence-selectively in DNA, with all 16 possible dinucleotide contexts represented in archaeobacterial and eubacterial organisms (15). The PT context of d(GpsA), which occurs in E. coli B7A and Salmonella enterica serovar Cerro 87 (15), drew our attention because adenine is the substrate of N6-methyladenosine-specific methyltransferases (MTases). This enzyme class catalyzes the 6mA modification widespread in R-M systems, such as M.TagI, M.ClaI, M.EcoBI, and M.EcoRI, as well as solitary Dam and CcrM MTases. Adenine MTases modify a variety of sequence contexts such as TCAG, ATCGAT, TGA(N)8TGCT, GAATTC, GATC, and GANTC (underlined bases indicate methylation).

What was notable here is the coincidence of recognition sequences of some adenine MTases with those for PT modifications, such as TGA(N)8TGCT for M.EcoBI and others (24) and GAAC/GTTC for Dnd enzymes.

These coincidences prompted us to search for PT and 6mA in the same sequence motifs. Here, we searched first for d(Gps6mA)-containing dinucleotides in E. coli B7A and S. enterica serovar Cerro 87 possessing d(GpsA) motifs, using an LC-MS/MS method that exploits the nucleoside resistance of PT-containing dinucleotides. Genomic DNA is hydrolyzed and persistent PT-containing dinucleotides are resolved from the bulk 2′-deoxyribonucleosides by HPLC, which allows sensitive MS identification and detection of the dinucleotides. Systematic MS scanning of the HPLC eluent of hydrolyzed genomic DNA from these bacteria revealed a molecule with m/z 611 (Fig. 1A). Collision-induced dissociation (CID) analysis yielded fragment ions with m/z values of 150, 362, and 460 (Fig. 1A), which are consistent with the structure of d(Gps6mA) shown in Fig. 1B. This structure was validated in several ways. First, the m/z 611 species disappeared in DNA from PT-deficient mutant strains CX3 and XTG103 of E. coli B7A and S. enterica serovar Cerro 87, respectively, which points to the presence of a PT structure in the dinucleotide. Second, the CID fragmentation properties parallel those of d(GpsA), which is 14 atomic mass units (amu) lighter due to the absence of the N6-methyl group: a protonated molecular ion [M+H]+/m/z of 597 (vs. 611), a protonated adenine base (m/z 136 vs. 150), and other fragments (m/z 348, 446) lighter by 14 amu (12). Finally, the structure in Fig. 1B was confirmed by LC-MS/MS analysis of a synthetic d(Gps6mA) standard that possessed PT as a mixture of Rp and Sp stereoisomers (Fig. 1C), with Rp as the naturally occurring configuration (Fig. 1A). These results established the cooccurrence of a PT and a methyl group in the same d(GA) context to yield a hybrid d(Gps6mA) in Rp configuration.

We next performed quantitative analysis to establish the frequency of d(Gps6mA) in the bacterial genomes. LC-MS/MS analysis using calibration curves based on chemical standards revealed that, compared with the frequency of d(GpsA) at 301 ± 2 and 258 ± 17 per 10⁶ nt, the methylated d(Gps6mA) occurred at much lower levels of 15 ± 1 and 2 ± 0 per 10⁶ nt in E. coli B7A and S. enterica serovar Cerro 87, respectively. Interestingly, we are unaware of adenine MTases with GAAC/GTTC as the precise recognition sequence, so it is possible that the frequency of d(Gps6mA), low as it is, is resulted from the overlap of GA in the consensus sequences of a MTase and the PT-inserting Dnd proteins in a subset of GAAC/GTTC sites. To further analyze the intersecting DNA modifications, we engineered bacteria with PT and methylation systems, as discussed next.

Interaction Between PT Modification and Methylation in Bacterial Genomes. This coincidence of PT and 6mA modifications at the same sites in bacterial genomics raises questions about both the target selection mechanisms underlying the modification proximity and the implications of proximity for modification function. Here, we explored this issue with PTs and Dam-mediated methylation, both of which are considered to be postreplicative modifications, with hemimethylated sites created by passage of the replication fork later fully modified (25, 26). We used the PT modification system from Hahella chejuensis KCTC2396 and dam from E. coli DH10B, both of which recognize GATC as a consensus sequence and generate GpsATC and G6mAATC, respectively. To assess how Dam reacts with PT-modified DNA, oligonucleotide substrates harboring GATC or GpsATC (Table S1) were reacted with Dam and S-adenosylmethionine (SAM) in vitro. As shown in Fig. 2A, Dam is able to use GpsATC with both Rp and Sp stereoisomers of PT as a substrate to yield d(Gps6mA). However, a time course of the Dam methylation reaction measured by LC-MS/MS (Fig. 2B) revealed that both the rate and extent of Dam activity were significantly reduced in the presence of PTs, regardless of whether the substrate was single- or double-stranded. The sulfur substitution of GpsATC thus interfered with either recognition or catalysis by Dam in an in vitro setting.

We next assessed the interaction of Dam and the PT-modifying enzymes DndBCDE in vivo. Here, we used E. coli HST04, which lacks dam, dcm, and hsd MTases, as well as PT modification genes, as a host for vectors expressing dnd genes B, C, D, and E (dndBCDE) and dam genes (pWHU701 and pWHU702, respectively); E. coli HST04 possesses the iscS gene that substitutes for dndA. LC-MS/MS analysis revealed the formation of d(Gps6mA) in exponentially growing E. coli HST04(dndBCDE, dam) at a frequency of 219 ± 2 per 10⁶ nt, compared with 228 ± 4 (d(GpsA)) per 10⁶ nt in E. coli HST04(dndBCDE) lacking dam (Table S2). This finding stands in contrast to the observation of PT inhibition of Dam activity in vitro and suggests either that Dam activity in vivo is not affected by the presence of a PT or that 6mA forms first followed by PT insertion, with Dnd proteins unaffected by 6mA. As described next, however, only a fraction of the 37,610 GATC sites containing 6mA also contained PT, which points to different target selection mechanisms by Dnd and Dam proteins.

Chen et al.
Comparative Genome Maps of PT and 6mA Sites by SMRT Sequencing.

The evidence for coincidence motivated an attempt to apply the SMRT sequencing platform to compare the locations of PT and 6mA across the genome of engineered E. coli HST04. SMRT sequencing of DNA modifications exploits the kinetic signatures, defined as pulse width and interpulse duration (IPD), arising when the DNA polymerase copies past the modification (27). Building on SMRT applications to nucleobase methylation (27), the evidence for coincidence motivated an attempt to apply the SMRT sequencing platform to compare the locations of PT and 6mA sites detected in E. coli HST04 genome.

To this end, we used SMRT sequencing to map GpsATC in E. coli HST04(dndBCDE) and G6mATC and hybrid Gps6mAATC in E. coli HST04(dndBCDE, dam). In the 4.57 × 10^6 bp in the E. coli HST04(dndBCDE) genome, 2,058 of the 37,698 GATC sites (5%) were found to be PT-modified (Fig. 3), which is consistent with the d(GpsATC) quantitation by LC-MS/MS (22). Among the 2,058 GpsATC sites, 1,019 are located on the (+) strand and 1,039 on the (−) strand, with PT occurring on both strands of the same GATC site 517 times. The remaining 1,024 PT modification sites are single-stranded GpsATC sites.

The map of G6mATC and Gps6mA sites across the E. coli HST04(dndBCDE, dam) genome revealed modification of nearly 100% of all GATC sites (37,610 out of 37,698 GATC sites) (Fig. 3). However, the similarity of the SMRT sequencing IPD ratios for GpsATC and G6mATC makes it difficult to distinguish between them in the genome modification maps in E. coli HST04(dndBCDE, dam). Knowing that nearly all of the 2,058 PT-modified GATC sites detected in E. coli HST04(dndBCDE) contained 6mA, and that the PT-modified GATC sites in E. coli HST04(dndBCDE, dam) also contain 6mA, which leads to the observation that IPD ratios for Gps6mAATC sites were 2.8-fold larger than those for GpsATC, with 89% of the 6mA ATC sites having an IPD ratio ≥ 2x larger than GpsATC. Using a Gaussian distribution as the smoothing kernel, we estimated the probability density function of the IPD ratio values for 2,058 GpsATC, 2,058 Gps6mAATC, and 35,552 G6mATC on chromosomes. As shown in Fig. S1, the immediate observation was that the average IPD ratios of both Gps6mAATC and G6mATC were larger than that of GpsATC and that the differences were significant (both P values < 0.0001). The IPD ratios of GpsATC were largely concentrated within a small range of [1.3, 3.5] and formed a sharp peak; in contrast, the Gps6mAATC and G6mATC sites exhibited more widely ranging IPD values.
DNA Methylation Protects Against PT-Associated Restriction. The temperature-sensitive DndFGH system in *E. coli* HST04 provided an opportunity to test the effects of 6mA on PT-dependent restriction. Here, we compared the bactericidal activity of DndFGH in the presence of PT or 6mA by quantifying viability (colony-forming units [cfu]) after exponentially growing cells were plated and cultured at 37 °C or 15 °C. In the absence of either PT or 6mA, the viability of *E. coli* HST04 cells at 15 °C was significantly decreased in the presence of DndFGH (Fig. 4C). Cotransformation of *E. coli* HST04 with pWHU705 and pWHU701, which express *H. chejuensis* KCTC2396 DndFGH and DndBCDE, respectively, restored viability at 15 °C to 66% of the level at 37 °C (Fig. 4C). This partial recovery of cell viability could result from the amount of DndBCDE expressed from pWHU701 under these growth conditions, which may not be fully sufficient to counteract the temperature-dependent DndFGH. Remarkably, methylation of GATC sites conferred protection against DndFGH restriction activity to almost the same extent as PT modification (64%) (Fig. 4C), which suggests that the A methylation of GATC sites hampered DNA cleavage activity or interfered with the ability of DndFGH to bind to target sites. The experiment to assess the mechanism of A methylation on DndFGH is not yet possible due to the lack of a system for in vitro reconstitution of DndFGH activity.

**Discussion**

The advent of SMRT sequencing has accelerated the pace of discovery and characterization of prokaryotic epigenetics. In particular, methylation-based R-M systems and epigenetic marks have been detected in the vast majority of bacterial and archaeal genomes (3, 30). The chemical diversity of DNA modifications in bacteria has recently grown beyond methylation with the discovery of 7-deazaguanine modifications (2) and PT modifications of the DNA backbone (12, 15). We have shown that PTs function in R-M systems with DndFGH (22, 23, 31) but that PTs are also present in bacteria that lack restriction genes. Although it is well-established that many bacteria possess multiple R-M systems and MTases, each with unique methylation consensus sequences (3, 32), we have discovered a proximity of two diverse DNA modifications in the same core consensus sequence, with interactions that point to a functional cooperation of the two modification systems.

Although SMRT sequencing is not able to resolve closely spaced DNA modifications, including PT and 6mA in *Gps* 6mA ATC consensus sequences as observed here, we were able to detect d(GPs 6mA) dinucleotides by LC-MS/MS in two unrelated bacteria, *E. coli* B7A and *S. enterica* serovar Cerro 87, by virtue of the PT-induced nuclease inhibition that produces a limit digest of PT-containing dinucleotides. The fact that d(GPs 6mA) occurred at much lower ratios, and the corresponding probability density curves were flatter (Fig. S1). In addition, the GPs 6mA ATC and G 6mA ATC probability density curves were similar in shape: (i) the curves were symmetric about their means; (ii) the highest probability density of each did not differ much between the curves (0.405 for GPs 6mA ATC and 0.437 for G 6mA ATC); and (iii) the ranges covering the middle 90% IPD ratios for the curves did not differ much ([3.98, 7.32] for GPs 6mA ATC and [3.58, 6.67] for G 6mA ATC). Therefore, the similar IPD ratios of GPs 6mA ATC and G 6mA ATC made it difficult to distinguish between them, consistent with the SMRT result that the 2,058 GPs 6mA ATC sites were grouped with the remaining 35,552 methylation G 6mA ATC sites by the SMRT analysis portal (Fig. 3). We nonetheless conclude that PT and 6mA did not interfere with each other in terms of modification location or frequency in vivo, with the protection of PT modifications, the DndFGH restriction proteins exerted no apparent toxicity or growth inhibition in *E. coli* HST04 at 37 °C (Fig. 4A). However, a DndFGH-mediated bactericidal effect emerged when the cells were grown at 15 °C (Fig. 4B). Similar thermoregulated restriction was also observed for WXY-1, a *dndBCDE*-deficient mutant of *E. coli* B7A. Without PT protection, WXY-1 demonstrated a slight growth defect in liquid culture at 37 °C, but restriction became more evident when the cells were cultured at 15 °C (Fig. S2). To characterize this effect of growth temperature, we examined *dndFGH* transcription in *E. coli* HST04 (pWHU705) and WXY-1 at different temperatures using real-time quantitative RT-PCR (qRT-PCR) with GAPDH as an internal control. Levels of *dndFGH* transcript increased significantly when cells growing at 37 °C were shifted to 15 °C for 1 h (Table S3). These results reveal that DndFGH restriction systems in *E. coli* B7A and *H. chejuensis* KCTC2396 are thermoregulated, with DndBCDE-DndFGH possibly functioning as an unusual temperature-dependent defense system. The thermoregulated behavior of DndFGH proved useful for studying the effects of 6mA on PT restriction activity.

**Development of a Temperature-Dependent PT-Associated Restriction System.** With evidence for proximal PT and 6mA modifications, we next sought to assess the effects of 6mA on PT-dependent restriction. Here, we engineered *E. coli* HST04 to express the DndFGH restriction proteins from *H. chejuensis* KCTC2396, which, as noted for *H. chejuensis* Dnd modification proteins described earlier, lacks Dam MTase and possesses PTs as d(GPs 6mA) in the Dam methylation consensus GATC. To this end, we first cloned the *dndFGH* gene from *H. chejuensis* KCTC2396 into the low-copy plasmid pACYCDuet-1 to generate pWHU705. Without the protection of PT modifications, the DndFGH restriction endonuclease activity should be lethal or at least toxic to bacteria as a result of DNA damage and the SOS response (5, 21, 22). Surprisingly, we were able to prepare the pWHU705 plasmid, and the restriction proteins DndFGH exerted no apparent toxicity or growth inhibition in *E. coli* HST04 at 37 °C (Fig. 4A). However, a DndFGH-mediated bactericidal effect emerged when the cells were grown at 15 °C (Fig. 4B). Similar thermoregulated restriction was also observed for WXY-1, a *dndBCDE*-deficient mutant of *E. coli* B7A. Without PT protection, WXY-1 demonstrated a slight growth defect in liquid culture at 37 °C, but restriction became more evident when the cells were cultured at 15 °C (Fig. S2). To characterize this effect of growth temperature, we examined *dndFGH* transcription in *E. coli* HST04 (pWHU705) and WXY-1 at different temperatures using real-time quantitative RT-PCR (qRT-PCR) with GAPDH as an internal control. Levels of *dndFGH* transcript increased significantly when cells growing at 37 °C were shifted to 15 °C for 1 h (Table S3). These results reveal that DndFGH restriction systems in *E. coli* B7A and *H. chejuensis* KCTC2396 are thermoregulated, with DndBCDE-DndFGH possibly functioning as an unusual temperature-dependent defense system. The thermoregulated behavior of DndFGH proved useful for studying the effects of 6mA on PT restriction activity.
frequencies than the d(GPSA) dinucleotide (2 to 15 per 10^6 nt versus 250 to 300 per 10^6 nt) supports the idea that d(GPSA 6mA) is occurring in the larger consensus sequence of a DNA methyltransferase. Examples of such sequences in other E. coli strains include the GAACC consensus of many E. coli R-M systems (33) and the larger recognition sequences for EcoBI/EcoI1581 (TGAN_nTGCT), Eco377I (GGAN_nATGC), and Eco777I (GGAN_nTATC) (24). However, E. coli BTA does not possess genes for these enzymes, so the source of d(GPSA 6mA) in BTA must be another MTase. Similarly, d(GPSA 6mA) dinucleotides must occur in the GAAC consensus sequence for Dnd proteins in S. enterica serovar Cerro 87, so DNA N5-adenine-methyltransferase from phage origin (M.SenCer87ORF144P) and methyl-directed repair DNA adenine methylase (M.SenCer87DamP) (33), which target GATC, cannot account for the d(GPSA 6mA) modification. Interestingly, the GpsATC sites may also contain 5mC due to the activity of C5-cytosine-specific DNA methylase (M1.SenCer87ORF144P) (33), which raises the possibility of sites containing all three modifications, GpsA 6mA 5mC, further complicating the interplay among DNA modification systems.

The discovery of proximate PT and 6mA motivated the development of a model system to further characterize coincident DNA modification consensus sequences. Here, we focused on the GATC recognition sequence for DndBCDE and Dam MTase. Although the in vitro activity of Dam was partially inhibited by a preexisting PT modification, the SMRT sequencing map of 6mA in GATC sites in E. coli HST04(dndBCDE, dam) showed complete modification of all GATC sites (37,610 of 37,698 sites), which suggests either that the Dam inhibition is an artifact of the in vitro situation, or that Dam methylation occurs in synchonry with or before PT modification. As a double-stranded DNA modification at GATC, 6mA methylation by Dam does indeed occur quickly after DNA replication (34) so it is possible that DT and 6mA are inserted in tandem at hemimodified sites postreplication.

The E. coli HST04 model system provided an opportunity to test these in vitro observations in vivo. The observation of PT modification in 2,058 of the 37,698 GATC sites (5%) in E. coli HST04 (dndBCDE) is similar to previous SMRT analyses in E. coli B7A (12% of GAAC/GTTC) and V. cholitrophus FF75 (14% of CCA) (23), with the 228 ± 4 PT per 10^6 nt measured by LC-MS/MS analysis corroborating the quantitative nature of SMRT mapping. However, the map of the GpsATC sites across the genome (GenBank accession no. CP013952) showed PTs in 939 ORFs, 1,119 in noncoding regions, and none in tRNAs or rRNA genes (Fig. 3), which stands in contrast to the slight enrichment of PTs in tRNA genes in E. coli B7A (23), in which dnd genes are native.

We also made the surprising observation of the temperature dependence of the DndFGH restriction system in two bacteria. Thermoregulated R-M systems are relatively rare, with two recent studies describing this phenotype for the restriction endonucleases LlaJI and LmoH7 (35, 36). The first is a plasmid-associated R-M system that mediates temperature-dependent resistance to phages in Lactococcus lactis, with LlaJI activity most pronounced when plaque assays are performed at 19 °C and completely abolished at 37 °C (35). The other reported temperature-sensitive restriction phenomenon involves the endonuclease LmoH7, which shares no sequence similarity, genetic components, or recognition sites with LjaJ1, but it can confer phage resistance in Listeria monocytogenes EC1 strains at low temperature (36). Similar to LlaJI and LmoH7, the mechanism responsible for dndFGH induction may be due to the instability of mRNA at high temperature or the involvement of an unknown temperature-sensitive transcriptional regulator. At this point, we do not know whether the temperature sensitivity occurs at the level of biochemical activity or protein levels, but the phenomenon points to the potential for an intriguing fitness advantage for bacteria in terms of resisting phage infection, for example, in moderate-temperature environments outside the core temperature of homeothermic animals (e.g., soil, water).

Although the coincidence of PT and 6mA in the same DNA sequence allows an organism to repurpose consensus sequences, the phenomenon has implications for the effectiveness of R-M systems and the real and potential epigenetic functions of both PT and 6mA. The artificial GpsA 6mA ATC system used here ignores the potential disruptive effects of PT on the binding of 6mA-sensitive protein binding in Dam-regulated epigenetic functions so it will be important to explore the coincidence of PT and DNA methylation in bacteria with both native modification systems. There are also interesting implications for R-M activity against phage infection. Many phages possess orphan MTases that modify and protect the phage genome against host restriction enzymes. For example, several phages have been found to possess GATC-specific Dam MTase activities, which not only modulate phage propagation but also protect against hundreds of R-M systems (37). Based on our observation of the interchangeability of 6mA and PT against PT-dependent restriction, this protection would also be extended against
GATC-specific PT-R M systems in diverse bacteria. Interestingly, phages such as *Staphylococcus* phage K have evolved to lack 5′-GATC-3′ motifs in their DNA genome (38).

In summary, the discovery of proximate PT and 5′-A DNA modifications in several bacteria point to the potential for widespread coincidence of epigenetics systems (Fig. S3). The results reveal complex but unappreciated interactions between DNA modification systems and raise the possibility of coevolution of interacting systems to facilitate the function of each.

**Materials and Methods**

Detailed descriptions of materials and methods are provided in SI Materials and Methods.

**Bacterial Strains, Plasmids, and Growth Conditions.** All strains, plasmids and primers used in this study are listed in Table S1. In-frame deletion mutants of E. coli BTA (WXY-1, CX2 and CX3, dndBCDE, dndFGH, and dndBCDEFGH) were constructed by homologous recombination using the thermolytic and suicide-sensitive plasmid pKOV-Kan, as previously described (31). Cell viability was determined by colony counting on LB-agar plates, plates placed at 15 °C for 24 h and shifted to 37 °C for colony counting for assessing low-temperature viability. qRT-PCR analysis of gene transcription was performed as described in detail in SI Materials and Methods.

**DNA Modification Analysis.** For in vitro DNA methylation, duplex oligonucleotides containing PT modifications were reacted with Dam MTase, and modifications were analyzed by LC-MS/MS as described in SI Materials and Methods. PT modifications were detected by LC-MS/MS analysis of enzymatically hydrolyzed DNA as previously described (39); see SI Materials and Methods for details. SMRT sequencing to map PT and methylation modifications was performed as described previously (23) and in SI Materials and Methods.

**ACKNOWLEDGMENTS.** We thank Bochen Zhu and Yuanjie Ge for research assistance. This work was supported by grants from the National Science Foundation of China (31520103902, 31670086, and 31300038) and the 973 program of the Ministry of Science and Technology (2013CB734003), Hubei Province’s Outstanding Medical Academic Leader Program, the Young One Thousand Talent program of China, the US National Science Foundation (CHE-1019990), the US National Institute of Allergy and Infectious Diseases (AI12711), and the Singapore-MIT Alliance for Research and Technology sponsored by the National Research Foundation of Singapore.


