JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Replacement of Thymidine by a Modified Base in the *Escherichia coli* Genome

Angad P. Mehta, Han Li, Sean A. Reed, Lubica Supekova, Tsotne Javahishvili, and Peter G. Schultz*

The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, United States

Supporting Information

ABSTRACT: Prokaryotic and eukaryotic genomic DNA is comprised of the four building blocks A, G, C, and T. We have begun to explore the consequences of replacing a large fraction or all of a nucleoside in genomic DNA with a modified nucleoside. As a first step we have investigated the possibility of replacement of T by 2′-deoxy-5-(hydroxymethyl)uridine (5hmU) in the genomic DNA of *Escherichia coli*. Metabolic engineering with phage genes followed by random mutagenesis enabled us to achieve approximately 75% replacement of T by 5hmU in the *E. coli* genome and in plasmids.

Prokaryotic and eukaryotic genomic DNA is comprised of the four building blocks A, G, C, and T. The A-T and G-C hydrogen-bonding and base-pair stacking interactions are important stabilizing non-covalent interactions in double-helical DNA and play a central role in template-dependent replication and transcription. Modified bases have been observed in genomic DNA, including N6-methyladenine, 5-methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), 5-carboxycytosine, and 2′-deoxy-5-(hydroxymethyl)-uridine (5hmU).[1,2] Of these, 5mC and 5hmC have been established as stable epigenetic modifications in eukaryotes, including mammals, and play a key role in regulating transcription and organization of chromatin structure.[3] In bacteria, 5mC also plays a role in restriction modification systems.[4] The roles of all of these DNA modifications are an active area of research. In addition to studies of natural nucleoside modifications, a number of synthetic efforts have been made to introduce unnatural base pairs (UBPs) into a template DNA and utilize them for *in vitro* and *in vivo* replication, transcription, and translation reactions. Leonard and co-workers developed several extended purines to investigate enzyme recognition and base pairing.[5−7] Benner and co-workers have generated an isoG-isoC base pair (Figure 1A) and examined its in vitro replication, transcription, and translation reactions.[8−10] They further developed an unnatural P_B-Z_B base pair (Figure 1B) that functions in PCR with high efficiency.[11−13] Kool and co-workers have developed a Z_K-F base pair based on hydrophobic interactions rather than hydrogen bonding interactions (Figure 1C), and shown it to be selective during the replication by the Klenow fragment of *E. coli* DNA polymerase I.[14−16] They also developed an xDNA based system for plasmid based expression system.[17] Hirao and co-workers have developed an unnatural base pair between **s**-**y** (Figure 1E) that can be transcribed by T7 polymerase,[18,19] and

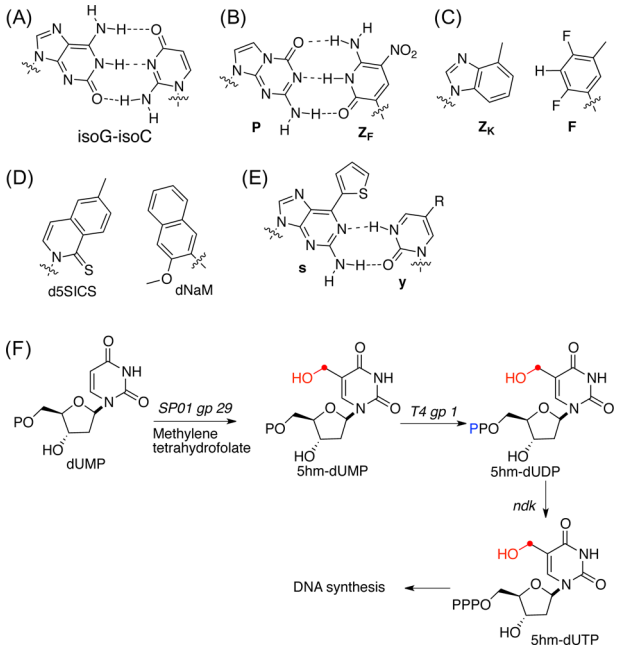

Figure 1. (A−E) Examples of unnatural base pairs. (F) Metabolic pathway to biosynthesize and incorporate 5hmU into genomic templates.

Romesberg, Schultz, and co-workers developed a series of selective, stable hydrophobic and metallo-base pairs.[20,21] Moreover, Romesberg and co-workers demonstrated that the *E. coli* replication machinery could accurately replicate plasmids containing a d5SICS-dNaM base pair (Figure 1D).[22−25]

We have begun to explore the consequences of replacing a large fraction or even all of a nucleoside in genomic DNA with a modified nucleoside. As a first step we have investigated the possibility of replacement of T by 2′-deoxy-5-(hydroxymethyl)-uridine (5hmU) in the genomic DNA of *E. coli*. Studies have detected 5hmU in mammalian DNA, and its role as an epigenetic regulator is currently an active area of investigation.[2] Thymidine replacement by the unnatural thymidine analogue, 5-bromouridine, has been previously studied by Hanawalt and co-workers,[26] and more recently, Mutzel and co-workers studied thymidine replacement by 5-chlorouridine.[27] Here we focused on 5hmU incorporation since a complete replacement of thymidine with 5hmU is observed in the SPO1 bacillus

Received: April 15, 2016
Published: May 23, 2016

DOI: 10.1021/jacs.6b03904
J. Am. Chem. Soc. 2016, 138, 7272−7275

phage genome.[28] In contrast to the previous studies, our goal is to engineer an autonomous *E. coli* strain that both biosynthesizes the modified base and efficiently incorporates it into the genome without addition of exogenous nucleoside analogs. This system also allows us to explore in an easily manipulatable system (*E. coli*) how 5hmU affects replication, transcription, and repair.

We used a similar metabolic approach to that used by SPO1 bacillus phage to engineer the *E. coli* pyrimidine triphosphate biosynthetic pathway to incorporate 5hmU into the *E. coli* genome (Figure 2).[29] Briefly, deoxyuridine monophosphate

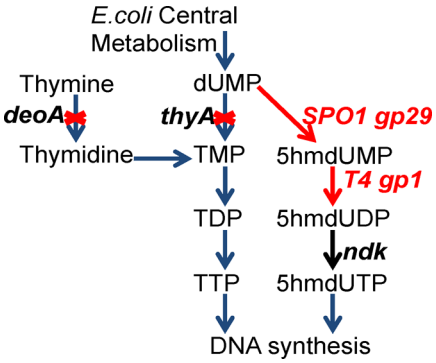

Figure 2. Metabolic engineering of the *E. coli* pyrimidine triphosphate biosynthetic pathway and incorporation of 5hmU into *E. coli* DNA.

(dUMP), the intermediate for thymidine triphosphate (TTP) biosynthesis in *E. coli*, was converted to 2′-deoxy-5-(hydroxymethyl)uridine triphosphate (5hmdUTP) by a series of reactions catalyzed by a hydroxymethylase, a mononucleotide kinase, and a nucleotide diphosphate kinase, which are encoded by SPO1 bacillus phage *gp29*, T4 bacteriophage *gp1*, and native *E. coli ndk* genes, respectively. The resulting 5hmdUTP is then incorporated into DNA by the endogenous DNA polymerase machinery. Since TTP competes with 5hmdUTP for base pairing in DNA synthesis, we also disrupted the genes *thyA*, *deoA*, and *trmA* of *E. coli* to lower the intracellular concentration of TTP.

The genes SPO1 gp29 and T4 gp1 were codon optimized and introduced as a part of a synthetic operon under the pBAD promoter with tetracycline resistance (pAM22). This operon together with native *E. coli* genes was expected to drive the biosynthesis of 5hmU (Figure 2). pAM22 was transformed into 3KO (BW25113 Δ*thyA::spec*^{R} Δ*deoA::FRT* Δ*trmA::FRT-Kan*^{R}) to give *E. coli* hT-1 and grown in the presence of thymidine supplementation. Individual colonies were picked, and an overnight culture (5%) was inoculated into LB medium containing tetracycline, thymidine (0.5 mg/L), and arabinose. Cells were harvested, and genomic DNA was isolated and digested to single nucleosides with phosphodiesterase I, calf intestinal phosphatase, and benzonase.

The resulting pool was then analyzed by high-performance liquid chromatography (HPLC, 254 nm). Along with the standard nucleosides in the genomic DNA [2′-deoxycytidine, 2′-deoxyadenosine, 2′-deoxyguanosine, and thymidine (T)], 5hmU was detected. We calculated 5hmU levels by HPLC peak area at 254 nm with standards. Under the above growth conditions, we observed that approximately half of genomic thymidine was replaced by 5hmU [%5hmU/(5hmU+T) ∼48%]. Moreover, plasmid DNA isolated from these strains showed 5hmU incorporation at somewhat higher levels than the genomic templates [%5hmU/(5hmU+T) ∼62%, Figure S4].

Because 5hmU and TTP compete at the polymerase level for incorporation into genomic DNA, we investigated the effects of decreasing TTP levels by replacing the richer LB medium with defined M9 minimal media containing L-arabinose, tetracycline, and defined levels of thymidine. No growth was observed for strain hT-1 in the absence of exogenous thymidine, and very poor growth was observed (OD_{600} ≈ 0.2) for strain hT-1 grown in the presence of 0.5 mg/L thymidine. On the other hand, better growth was observed with 2 mg/L thymidine (OD_{600} ≈ 0.6−0.8). However, genomic DNA analysis indicated that %5hmU/(5hmU+T) was ∼52%, indicating no significant change in genomic 5hmU incorporation. To further lower the level of TTP, we tested the effect of expression of thymidylate 5′-phosphatase (TMPase, from *Lactobacillus plantarum*), which converts TMP to thymidine. Introduction of a construct containing *TMPase* (pAM26) into strain hT-1 to afford strain hT-3 strain (Table 1) resulted in %5hmU/(5hmU+T) of ∼51%, again indicating no significant effect on the levels of 5hmU in the genomic templates.

Table 1. Strain Names and Properties

strain	properties/plasmid
3KO	thymidine auxotroph from *E. coli* BW25113
hT-1	3KO transformed with pAM22
hT-2	3KO transformed with pAM22 and pAM23
hT-3	3KO transformed with pAM22 and pAM26
hT-4	3KO transformed with pAM43
2hTC strains	NTG mutagenesis on hT-4 in the absence of exogenous thymidine

We then attempted to increase the pool size of 5hmdUTP in order to increase 5hmU genomic content. During the biosynthesis of 5hmUTP, SPO1 gp29 converts dUMP to 5-hydroxythymidine-5′-monophosphate (5hmdUMP) using methylenetetrahydrofolate (MFH4), which is converted to tetrahydrofolate (FH4) in this process. *E. coli glyA* regenerates MFH4 from FH4 by converting serine to glycine. In order to maintain high levels of MFH4 to drive the formation of 5hmUTP, we tested the effect of glyA overexpression and serine supplementation. Serine supplementation of the overnight culture under induction conditions (hT-1 + serine) had little effect on the net 5hmU content [%5hmU/(5hmU+T) increased to ∼53% from ∼48%; Figure S6]. Overexpression of *glyA* was carried out by introducing plasmid pAM23 into hT-1, resulting in strain hT-2. The hT-2 cultures were grown as described above with supplementation by serine (10 mg/L) and thymidine (0.5 mg/L); however, the %5hmU/(5hmU+T) was again ∼48%, indicating no significant effect (Figure S5).

We next increased the copy numbers of the phage genes encoding the 5hmUTP biosynthetic pathway, which might also increase the 5hmUTP pool. A new construct was designed where SPO1 *gp29* and T4 *gp1* were introduced into a high copy number vector (Cdf ori and chloramphenicol resistance, pAM43) under the pBAD promoter. We transformed pAM43 into 3KO to generate strain hT-4, and this strain was grown as described above in LB medium in the presence of thymidine (0.5 mg/L), serine (10 mg/L), and L-arabinose. This modification increased the %5hmU/(5hmU+T) to ∼62% as compared to 48% for hT-1. Finally, there was no significant effect when we incorporated the phage polymerase (SPO1

DOI: 10.1021/jacs.6b03904
J. Am. Chem. Soc. 2016, 138, 7272−7275

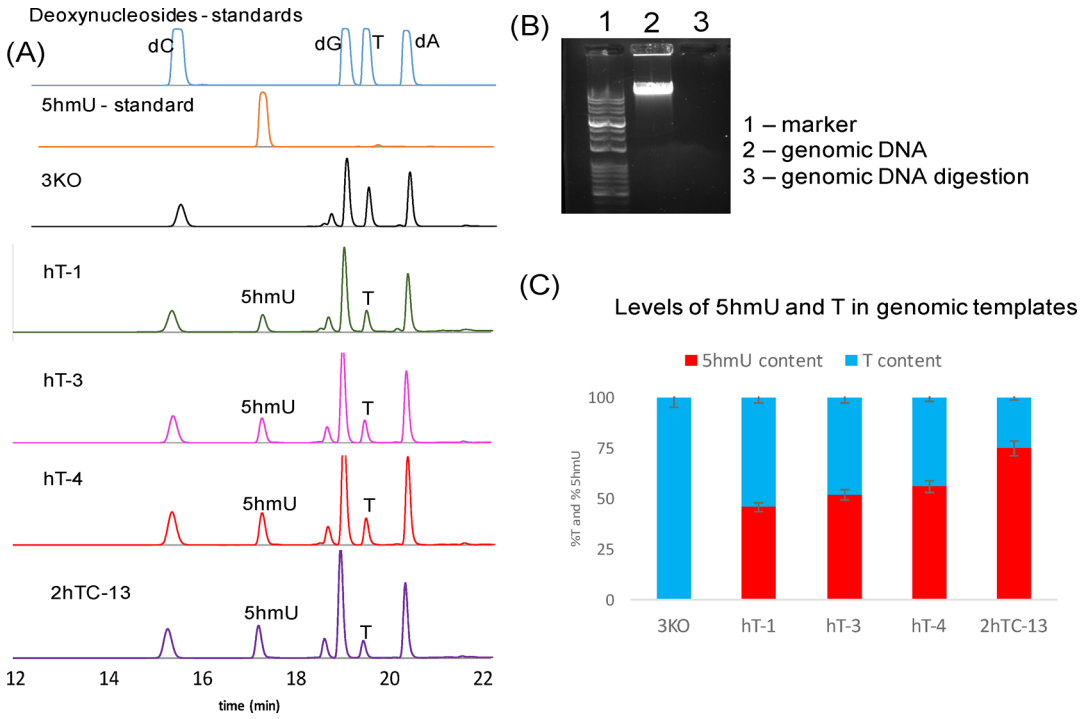

Figure 3. (A) HPLC traces (254 nm) of genomic DNA after digestion to single nucleosides. (B) Representative gel showing the complete digestion of the isolated genomic templates with phosphodiesterase I, calf intestinal phosphatase, and benzonase. (C) Levels of 5hmU content in various strains. For strain details, see Table 1.

gp31) into our engineered metabolic network [%5hmU/(5hmU+T) ~50%, Figure S10].

In all of the above experiments we observed that cells containing significant amounts of 5hmU in their genomes did not grow above $OD_{600} \approx 0.7$ (growth curve, Figure S11). These observations suggested that 5hmU incorporation into *E. coli* genomic DNA might limit growth by impairing processes such as replication, transcription, and repair, and as a consequence, alterations may be necessary in these processes to increase genomic 5hmU content. Because such changes may be difficult to predict and target by rational design, we decided to introduce genome-wide random mutations into strain hT-4 under selective pressure for 5hmU incorporation (thymidine starvation). The strain hT-4 was grown as described above in LB medium in the presence of serine (10 mg/L), tetracycline, and L-arabinose. Cells were then subjected to methylnitronitrosoguanidine (NTG) treatment (see Supporting Information for mutagenesis procedure) and plated on LB-agar plates containing arabinose, tetracycline and serine with no exogenous thymidine. Of the 19 colonies isolated, 14 survived in liquid medium containing LB, serine, arabinose, chloramphenicol, and no exogenous thymidine. All of these colonies showed higher 5hmU incorporation, and one of these strains (2hTC-13) showed %5hmU/(5hmU+T) = 73 ± 2% (Figure 3). This phenotype as well as growth was stable for at least three generations (Figures S12–S14). Similar levels of 5hmU incorporation were also observed in plasmids [%5hmU/(5hmU+T) = 67 ± 4%, Figure S8] isolated from these strains. In contrast to strain 3KO, strain 2hTC-13 took around 48 h to reach a maximum $OD_{600} \approx 0.6–0.75$ (growth curves, see Figures S12–S14). We also tested strain 2hTC-13 growth in a defined minimal medium in the absence or 0.05 mg/L of thymidine, but no growth was observed in 48 h. We next attempted to perform a second round of mutagenesis on 2hTC-13 to further increase 5hmU content. The recovered cells in this case were plated on M9 minimal medium containing casamino acids and thymidine (0.05 mg/L). Unfortunately, no colonies were observed. Thus, it appears under these experimental conditions, we seemed to have reached a maximum level of 5hmU incorporation in the genomic DNA of *E. coli* [%5hmU/(5hmU+T)] of ~75%.

To summarize, we have used a metabolic engineering approach to incorporate 5hmU into the *E. coli* genome. We further performed random mutagenesis to increase 5hmU content [%5hmU/(5hmU+T)] to ~75%. At present we do not know if 75% is the maximum possible replacement of T by 5hmU. Further mutagenesis studies and characterization of the mutations and mutant strains with increased 5hmU incorporation will likely provide insights into the key physiological regulators in *E. coli* that limit global genomic incorporation of modified nucleosides as well as the effect of 5hmU on replication, transcription, and repair. At the same time we are investigating if the 5-hydroxylmethyl group of 5hmU can be modified enzymatically[30] and whether other base or sugar modifications of the natural deoxynucleosides can be introduced into the genome at high levels, e.g., replacement of dC by 5hmdC, or dC by rC.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b03904.

HPLC data, experimental procedure, and design of the constructs, including Figures S1–S14 (PDF)

DOI: 10.1021/jacs.6b03904
J. Am. Chem. Soc. 2016, 138, 7272–7275

AUTHOR INFORMATION

Corresponding Author

*schultz@scripps.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We would like to acknowledge Kristen Williams for her assistance in manuscript preparation. This work was supported by NIH R01 GM062159-14.

REFERENCES

(1) Nabel, C. S.; Manning, S. A.; Kohli, R. M. *ACS Chem. Biol.* **2012**, *7*, 20.
(2) Olinski, R.; Starczak, M.; Gackowski, D. *Mutat. Res., Rev. Mutat. Res.* **2016**, *767*, 59.
(3) Münzel, M.; Globisch, D.; Carell, T. *Angew. Chem., Int. Ed.* **2011**, *50*, 6460.
(4) Sanchez-Romero, M. A.; Cota, I.; Casadesus, J. *Curr. Opin. Microbiol.* **2015**, *25*, 9.
(5) Leonard, N. J. *Biopolymers* **1985**, *24*, 9.
(6) Leonard, N. J.; Hiremath, S. P. *Tetrahedron* **1986**, *42*, 1917.
(7) Lessor, R. A.; Gibson, K. J.; Leonard, N. J. *Biochemistry* **1984**, *23*, 3868.
(8) Bain, J. D.; Switzer, C.; Chamberlin, A. R.; Bennert, S. A. *Nature (London, U. K.)* **1992**, *356*, 537.
(9) Horlacher, J.; Hottiger, M.; Podust, V. N.; Huebscher, U.; Benner, S. A. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 6329.
(10) Switzer, C.; Moroney, S. E.; Benner, S. A. *J. Am. Chem. Soc.* **1989**, *111*, 8322.
(11) Kim, H.-J.; Chen, F.; Benner, S. A. *J. Org. Chem.* **2012**, *77*, 3664.
(12) Yang, Z.; Sismour, A. M.; Sheng, P.; Puskar, N. L.; Benner, S. A. *Nucleic Acids Res.* **2007**, *35*, 4238.
(13) Yang, Z.-Y.; Chen, F.; Alvarado, J. B.; Benner, S. A. *J. Am. Chem. Soc.* **2011**, *133*, 15105.
(14) Morales, J. C.; Kool, E. T. *Nat. Struct. Biol.* **1998**, *5*, 950.
(15) Moran, S.; Ren, R. X. F.; Kool, E. T. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 10506.
(16) Moran, S.; Ren, R. X. F.; Rumney, S. I. V.; Kool, E. T. *J. Am. Chem. Soc.* **1997**, *119*, 2056.
(17) Krueger, A. T.; Peterson, L. W.; Chelliserry, J.; Kleinbaum, D. J.; Kool, E. T. *J. Am. Chem. Soc.* **2011**, *133*, 18447.
(18) Fujiwara, T.; Kimoto, M.; Sugiyama, H.; Hirao, I.; Yokoyama, S. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2221.
(19) Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T.; Kodama, K.; Yokogawa, T.; Nishikawa, K.; Yokoyama, S. *Nat. Biotechnol.* **2002**, *20*, 177.
(20) Matsuda, S.; Henry, A. A.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2003**, *125*, 6134.
(21) Atwell, S.; Meggers, E.; Spraggon, G.; Schultz, P. G. *J. Am. Chem. Soc.* **2001**, *123*, 12364.
(22) Henry, A. A.; Romesberg, F. E. *Curr. Opin. Chem. Biol.* **2003**, *7*, 727.
(23) Li, L.; Degardin, M.; Lavergne, T.; Malyshev, D. A.; Dhami, K.; Ordoukhanian, P.; Romesberg, F. E. *J. Am. Chem. Soc.* **2014**, *136*, 826.
(24) Malyshev, D. A.; Dhami, K.; Lavergne, T.; Chen, T.; Dai, N.; Foster, J. M.; Correa, I. R.; Romesberg, F. E. *Nature (London, U. K.)* **2014**, *509*, 385.
(25) Ogawa, A. K.; Wu, Y.; McMinn, D. L.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 3274.
(26) Hanawalt, P. C. In *Methods in Enzymology*; Academic Press: San Diego, CA, 1967; Vol. *12*, Part A, p 702.
(27) Marlière, P.; Patrouix, J.; Döring, V.; Herdewijn, P.; Tricot, S.; Cruveiller, S.; Bouzon, M.; Mutzel, R. *Angew. Chem., Int. Ed.* **2011**, *50*, 7109.
(28) Stewart, C. R.; Casjens, S. R.; Cresawn, S. G.; Houtz, J. M.; Smith, A. L.; Ford, M. E.; Peebles, C. L.; Hatfull, G. F.; Hendrix, R. W.; Huang, W. M.; Pedulla, M. L. *J. Mol. Biol.* **2009**, *388*, 48.
(29) Wilhelm, K.; Rueger, W. *Virology* **1992**, *189*, 640.
(30) Terragni, J.; Bitinaite, J.; Zheng, Y.; Pradhan, S. *Biochemistry* **2012**, *51*, 1009.

DOI: 10.1021/jacs.6b03904
J. Am. Chem. Soc. 2016, 138, 7272−7275