A semisynthetic organism engineered for the stable expansion of the genetic alphabet

Yorke Zhang1,2, Brian M. Lamb3,4, Aaron W. Feldman5, Anne Xiaozhou Zhou6, Thomas Lavergne6, Lingjun Li7, and Floyd E. Romesberg8,2

*Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037; 1Département de Chimie Moléculaire, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 5250, Université Grenoble Alpes, F-38000 Grenoble, France; and 2School of Chemistry and Chemical Engineering, Henan Normal University, Henan 453007, People’s Republic of China

All natural organisms store genetic information in a four-letter, two-base-pair genetic alphabet. The expansion of the genetic alphabet with two synthetic unnatural nucleotides that selectively pair to form an unnatural base pair (UBP) would increase the information storage potential of DNA, and semisynthetic organisms (SSOs) that stably harbor this expanded alphabet would thereby have the potential to store and retrieve increased information. Toward this goal, we previously reported that Escherichia coli grown in the presence of the unnatural nucleotides dNaMTP and d5SICSTP, and provided with the means to import them via expression of a plasmid-borne nucleoside triphosphate transporter, replicates DNA containing a single dNaM-d5SICS UBP. Although this represented an important proof-of-concept, the nascent SSO grew poorly and, more problematically, required growth under controlled conditions and even then was unable to indefinitely store the unnatural information, which is clearly a prerequisite for true semisynthetic life. Here, to fortify and vivify the nascent SSO, we engineered the transporter, used a more chemically optimized UBP, and harnessed the power of the bacterial immune response by using Cas9 to eliminate DNA that had lost the UBP. The optimized SSO grows robustly, constitutively imports the unnatural triphosphates, and is able to indefinitely retain multiple UBPs in virtually any sequence context. This SSO is thus a form of life that can stably store genetic information using a six-letter, three-base-pair alphabet.

The natural genetic alphabet is composed of four letters whose selective pairing to form two base pairs underlies the storage and retrieval of virtually all biological information. This alphabet is essentially conserved throughout nature, and has been since the last common ancestor of all life on Earth. Significant effort has been directed toward the development of an unnatural base pair (UBP), formed between two synthetic nucleotides, that functions alongside its natural counterparts (1–3), which would represent a remarkable integration of man-made, synthetic component into one of life’s most central processes. Moreover, semisynthetic organisms (SSOs) that stably harbor such a UBP in their DNA could store and potentially retrieve the increased information, and thereby lay the foundation for achieving the central goal of synthetic biology: the creation of new life forms and functions (4).

For over 15 years, we have sought to develop such a UBP (1), and these efforts eventually yielded a family of predominantly hydrophobic UBPs, with that formed between dNaM and d5SICS (dNaM-d5SICS; Fig. 1A) being a particularly promising example (5–7). Despite lacking complementary hydrogen bonding, we demonstrated that the dNaM-d5SICS UBP is well replicated by a variety of DNA polymerases in vitro (7–10), and that this efficient replication is mediated by a unique mechanism that draws upon interbase hydrophobic and packing interactions (11, 12). These efforts then culminated in the first progress toward the creation of an SSO in 2014, when we reported that Escherichia coli grown in the presence of the corresponding unnatural nucleoside triphosphates (dNaMTP and d5SICSTP, and provided with a plasmid-encoded nucleoside triphosphate transporter (NTT2) from Pseudomonas putida (which we denote as PtNTT2) (13), is able to import the unnatural triphosphates and replicate a single dNaM-d5SICS UBP) on a second plasmid (14).

Although this first SSO represented an important proof-of-concept, the generality of the expanded genetic alphabet remained unclear, as retention of the UBP was explored at only a single locus and in only a single sequence context. True expansion of the genetic alphabet requires the unrestricted retention of multiple UBPs at any loci and in any sequence context. Moreover, several limitations were already apparent with the nascent SSO (14). First, although expression of the nucleoside triphosphate transporter enabled E. coli to import dNaMTP and d5SICSTP, its expression caused the SSO to grow poorly, with doubling times twice that of the parental strain. Second, the UBP was not well retained during high-density liquid growth or during growth on solid media, presumably due to the secretion of phosphatases that degrade the unnatural triphosphates. Finally, even under optimal conditions, the nascent SSO was unable to retain the UBP with extended growth. Clearly, the ability to robustly grow under the standard repertoire of culture conditions and indefinitely retain the UBP is a prerequisite for true semisynthetic life. Here, we used genetic and chemical approaches to optimize different components of the SSO, ultimately resulting in a simplified and optimized SSO that grows robustly and is capable of the virtually unrestricted storage of increased information.

Significance

The genetic alphabet encodes all biological information, but it is limited to four letters that form two base pairs. To expand the alphabet, we developed synthetic nucleotides that pair to form an unnatural base pair (UBP), and used it as the basis of a semisynthetic organism (SSO) that stores increased information. However, the SSO grew poorly and lost the UBP under a variety of standard growth conditions. Here, using chemical and genetic approaches, we report the optimization of the SSO so that it is healthy, more autonomous, and able to store the increased information indefinitely. This SSO constitutes a stable form of semisynthetic life and lays the foundation for efforts to impart life with new forms and functions.


Conflict of interest statement: Y.Z., B.M.L., and F.E.R. have filed a patent application based on the use of Cas9 for enforced retention of the UBP. Y.Z. and F.E.R. have filed a patent application for the truncated transporter. F.E.R. has a financial interest (shares) in Synthorx Inc., a company that has commercial interests in the UBP. The other authors declare no other competing financial interests.

This article is a PNAS Direct Submission.

1Y.Z. and B.M.L. contributed equally to this work.

2To whom correspondence should be addressed. Email: floyd@scripps.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1616443114/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1616443114

Edited by Clyde A. Hutchison III, The J. Craig Venter Institute, San Diego, CA, and approved December 21, 2016 (received for review October 6, 2016)
triphosphate uptake across a population of cells, which we reasoned from a low-copy plasmid or a chromosomal locus, which we an-

tegrated and constitutive PtNTT2(66-575) (green). The promoters from which PtNTT2 is expressed are indicated by the labels next to their corresponding markers. Open triangles denote corresponding control strains without Pt

-32P]-dATP uptake in E. coli C41(DE3); pSC plasmids and integrants are in E. coli BL21(DE3). All PtNTT2 strains are non-codon-optimized for plasmid-based expression and codon-optimized for chromosomal expression unless otherwise indicated; r.d.u. is relative decay units, which corresponds to the total number of radioactive counts per minute normalized to the average OD_{600} across a 1-h window of uptake, with the uptake in DM1 induced with 1,000 μM IPTG set to 1 (see Materials and Methods for additional details).

Results and Discussion

In our previously reported SSO (hereafter referred to as DM1), the transporter was expressed from a T7 promoter on a multi-
copy plasmid (pCDF-bb) in E. coli C41(DE3), and its toxicity mandated carefully controlled induction (14). In its native algal cell, PtNTT2’s N-terminal signal sequences direct its subcellular localization and are ultimately removed by proteolysis. However, in E. coli, they are likely retained, and could potentially contribute to the observed toxicity. Using the cellular uptake of [α-32P]-dATP as a measure of functional transporter expression and as a proxy for the uptake of the unnatural triphosphates, we found that removal of amino acids 1 to 65 and expression of the resulting N-terminally truncated variant PtNTT2(65-575) in E. coli C41(DE3) resulted in significantly lower toxicity, but also significantly re-
duced uptake (Fig. S1 A and B), possibly due to reduced expression (15). Expression of PtNTT2(66-575) in E. coli BL21 (DE3) resulted in significant levels of [α-32P]-dATP uptake with little increase in toxicity relative to an empty vector control, but the higher level of T7 RNA polymerase in this strain (16) was itself toxic (Fig. S1 A and C).

We next explored constitutive expression of PtNTT2(66-575) from a low-copy plasmid or a chromosomal locus, which we antici-
pated would not only eliminate the need to use T7 RNA polymerase but would also impart the SSO with greater autonomy (eliminating the need to induce transporter production), and, importantly, would result in more homogeneous transporter expression and triphosphate uptake across a population of cells, which we reasoned might improve UBP retention. We explored expression of PtNTT2 (66-575) in E. coli BL21(DE3) with the promoters P_lac, P_blb, and P_N25 from a pSC plasmid, and with P_blb, P_tac, P_ACS, P_H207, P_ac, and P_N25 from the chromosomal lacZYA locus (Dataset S1). We also explored the use of a codon-optimized variant of the truncated transporter (see Dataset S1). Although increasing expression of PtNTT2(66-575) (as measured by uptake of [α-32P]-dATP) was correlated with increasing doubling time, indicating that expression of PtNTT2(66-575) still exhibited some toxicity (uptake of [α-32P]-dATP is itself not toxic), each strain exhibited an improved ratio of uptake to fitness compared with DM1 (Fig. 1B). Strain YZ3, which expresses the codon-optimized, chromosomally in-
tegrated PtNTT2(66-575) from the P_ACS promoter, exhibited an optimal compromise of robust growth (<20% increased doubling time relative to the isogenic strain without the transporter), and [α-32P]-dATP uptake, and was thus selected for further characterization.

To determine whether the optimized transporter system of YZ3 facilitates high UBP retention, we constructed three plas-
mds that position the UBP within the 75-nt TK1 sequence (14) [with a local sequence context of d(A-NaM-T)]. These include two high-copy pUC19-derived plasmids, pUCX1 [referred to in previous work as pINF (14)] and pUCX2, as well as one low-copy pBR322-derived plasmid, pBRX2 (Fig. S2). In addition to allowing us to examine the effect of copy number on UBP retention, these plasmids position the UBP at proximal (pUCX1) and distal (pUCX2 and pBRX2) positions relative to the origin of repli-
cation, which we previously speculated might be important (14). Strains YZ3 and DM1 were transformed with pUCX1, pUCX2, or pBRX2 and directly cultured in liquid growth media supple-
mented with [α-32P]-dATP and [α-32P]-dATP [and isopropyl β-D-1-galactopyranoside (IPTG)] for DM1 to induce the transporter, and growth and UBP retention were characterized at an OD_{600} of ~1 (Fig. 2 A and B).

Given that no plasmid locus or copy number biases on UBP retention were observed in YZ3, we chose pUCX2 as a repre-
sentative UBP-containing plasmid to explore the effect of local sequence context on UBP retention, and we constructed 16 pUCX2 variants in which the UBP was flanked by each possible combination of natural base pairs within a fragment of the GFP gene (see Dataset S1). Under the same growth conditions as above, we observed a wide range of UBP retentions, with some sequence contexts showing complete loss of the UBP (Fig. 2C). However, since the development of DM1 with the dNaM-d5SICS UBP, we have determined that ring contraction and sulfur derivatization of d5SICS yields the dNaM-dTPPT3 UBP (Fig. 1A), results in more efficient replication in vitro (17). To explore the in vivo use of dNaM-dTPPT3, we repeated the experiments with YZ3 and each of the 16 pUCX2 plasmids, but with growth in media supplemented with dNaMTP and dTPPT3. UBP retentions were clearly higher with dNaM-dTPPT3 than with dNaM-d5SICS (Fig. 2C).

Although dNaM-dTPPT3 is clearly a more optimal UBP for the SSO than dNaM-d5SICS, its retention is still moderate to poor in some sequence contexts (Fig. 2C). Moreover, several sequences that show good retention in YZ3 cultured in liquid media show poor retention when growth includes culturing on solid media (Fig. S3B). To further increase UBP retention with these challenging sequences and/or growth conditions, we sought to selectively eliminate plasmids that lose the UBP. In prokaryotes, the clustered regularly interspaced short palindromic repeats (CRISPR)-CRISPR-associated (Cas) system provides adaptive immunity against viruses and foreign plasmids (18–20). In type II CRISPR-Cas systems, such as that from Streptococcus pyogenes (21, 22), the endonuclease Cas9 uses encoded RNAs [or their artificial mimics known as single-guide RNAs (sgRNAs) (23)] to introduce double-strand breaks in DNA upstream of a 5′-NGG-3′ protospacer adjacent motif (PAM) (24), which then results in DNA degradation by exonucleases (25) (Fig. 3A). In vitro, we found that the presence of a UBP in the target DNA generally reduces Cas9-mediated cleavage relative to sequences that are fully complementary to the provided sgRNA (Fig. S4). We thus
and regrown to the same OD600. UBP retention in control ex- 
nucleotide deletion in its place, which was not observed with the 
UBP when the correct sgRNA was provided contained a single 
the correct sgRNA was 70% after the second outgrowth (Fig. 
3B). Sequencing revealed that the majority of plasmids lacking a 
UBP when the correct sgRNA was provided contained a single 
nucleotide deletion in its place, which was not observed with the 

hypothesized that, within a cell, Cas9 programmed with sgRNA(s) 
complementary to natural sequences that arise from UBP loss would 
force retention in a population of plasmids (by eliminating 
those that lose the UBP), which we refer to as immunity 
to UBP loss. To test this, we used a p15A plasmid to construct 
Cas9-based editing system. (A) Schematic representation of the 
biotin shift assay used to determin e UBP retention. The plasmid DNA to be analyzed is first amplified in a PCR 
supplemented with the unnatural triphosphates, and the resulting 
products are then incubated with streptavidin and subjected to PAGE 
analysis. X = dNaM or, in the PCR, its biotinylated analog dMMO2bioX. Y = 
dS5CS in the PCR, whereas Y = dTPT3 or dS5CS in the plasmid DNA, depending 
on the experimental conditions. Lane 1 is the product from the oligonucle- 
otide analogous to that used to introduce the UBP during plasmid assembly, 
but with the UBP replaced by a natural base pair (negative control). This 
band serves as a marker for DNA that has lost the UBP. Lane 2 is the product 
from the synthetic oligonucleotide containing the UBP that was used for 
plasmid assembly. The shift of this band serves as a marker for the shift of 
DNA containing the UBP. Lane 3 is the product from the in vitro-assembled 
plasmid before SSO transformation (positive control). The unshifted band 
results from DNA that has lost the UBP during in vitro plasmid assembly. 
Lane 4 is the product from an in vivo replication experiment. (B) UBP re- 
ten tions of plasmids pUCX1, pUCX2, and pBRX2 in strains DM1 and YZ3. 
Error bars represent SD of the mean, n = 4 for pUCX1 and pUCX2, n = 3 for DM1 pBRX2, and n = 5 for YZ3 pBRX2. (C) UBP retentions of 
plasmid variants, wherein the UBP is flanked by all possible combinations of 
natural nucleotides (NXN, where N = dG, C, A, or T) and X = dNaM, in strain 
YZ3 grown in media supplemented with either dNaMTP and dS5CS(TP) (gray bars) or dNaMTP and dS5CS(TP) (black bars).

To further simplify and streamline the SSO, we next constructed strain YZ4 by integrating an IPTG-inducible Cas9 gene at the arcsB locus of the YZ3 chromosome, which allows for the use of a single plasmid that both carries a UBP and expresses the sgRNAs that enforce its retention. Sixteen such “all-in-one” plasmids (pAIO) were constructed by replacing the Cas9 gene in each of the pCas9 variants with a UBP sequence from the corresponding pUCX2 variant (Fig. S2 and Table S1). YZ4 and 
YZ3 (included as a no-Cas9 control due to leaky expression of 
the PAM).

To more broadly explore Cas9-mediated immunity to UBP 
loss, we examined retention using 16 pUCX2 variants with sequences that flank the UBP with each possible combination of 
natural base pairs but also vary its position relative to the PAM, and vary which unnatural nucleotide is present in the strand 
recognized by the sgRNAs (Table S1). We also constructed a corresponding set of 16 pCas9 plasmids that express two sgRNAs, 
one targeting a substitution mutation and one targeting the single 
nucleotide deletion mutation, for each pUCX2 variant. Strain YZ2 
carrying a pCas9 plasmid was transformed with its corresponding 
pUCX2 variant and grown in the presence of the unnatural tri- 
phosphates and IPTG (to induce Cas9), and UBP retention was 
assessed after cells reached an OD600 of ~1. As a control, the 16 
pUCX2 plasmids were also propagated in YZ2 carrying a pCas9 
plasmid with a nontarget sgRNA. For 4 of the 16 sequences 
explored, UBP loss was already minimal without immunity 
(nontarget sgRNA), but was undetectable with expression of the 
correct sgRNA (Fig. 4A). The remaining sequences showed 
moderate to no retention without immunity, and significantly 
higher retention with it, including at positions up to 15 nts from the 
PAM.

Error bars represent SD of the mean, n = 4 for pUCX1 and pUCX2, n = 3 for DM1 pBRX2, and n = 5 for YZ3 pBRX2. (C) UBP retentions of 
plasmid variants, wherein the UBP is flanked by all possible combinations of 
natural nucleotides (NXN, where N = dG, C, A, or T) and X = dNaM, in strain 
YZ3 grown in media supplemented with either dNaMTP and dS5CS(TP) (gray bars) or dNaMTP and dS5CS(TP) (black bars).

To further simplify and streamline the SSO, we next constructed strain YZ4 by integrating an IPTG-inducible Cas9 gene at the arcsB locus of the YZ3 chromosome, which allows for the use of a single plasmid that both carries a UBP and expresses the sgRNAs that enforce its retention. Sixteen such “all-in-one” plasmids (pAIO) were constructed by replacing the Cas9 gene in each of the pCas9 variants with a UBP sequence from the corresponding pUCX2 variant (Fig. S2 and Table S1). YZ4 and 
YZ3 (included as a no-Cas9 control due to leaky expression of 
the PAM).

To more broadly explore Cas9-mediated immunity to UBP 
loss, we examined retention using 16 pUCX2 variants with sequences that flank the UBP with each possible combination of 
natural base pairs but also vary its position relative to the PAM, and vary which unnatural nucleotide is present in the strand 
recognized by the sgRNAs (Table S1). We also constructed a corresponding set of 16 pCas9 plasmids that express two sgRNAs, 
one targeting a substitution mutation and one targeting the single 
nucleotide deletion mutation, for each pUCX2 variant. Strain YZ2 
carrying a pCas9 plasmid was transformed with its corresponding 
pUCX2 variant and grown in the presence of the unnatural tri- 
phosphates and IPTG (to induce Cas9), and UBP retention was 
assessed after cells reached an OD600 of ~1. As a control, the 16 
pUCX2 plasmids were also propagated in YZ2 carrying a pCas9 
plasmid with a nontarget sgRNA. For 4 of the 16 sequences 
explored, UBP loss was already minimal without immunity 
(nontarget sgRNA), but was undetectable with expression of the 
correct sgRNA (Fig. 4A). The remaining sequences showed 
moderate to no retention without immunity, and significantly 
higher retention with it, including at positions up to 15 nts from the 
PAM.
OD_{600} of ~1 to 2 (Fig. 4B). Despite variable levels of retention in the absence of Cas9 (YZ3), with induction of Cas9 expression in YZ4, loss was minimal to undetectable in 13 of the 16 sequences. Although retention with the three problematic sequences—d(C-NaM-C), d(C-NaM-A), and d(C-NaM-G)—might be optimized, for example, through alterations in Cas9 or sgRNA expression, the undetectable loss of the UBP with the majority of the sequences after a regimen that included growth both on solid and in liquid media, which was not possible with our previous SSO DM1, attests to the vitality of YZ4.

Finally, we constructed a pAIO plasmid, pAIO2X, containing two UBPs: dNaM paired opposite dTPT3 at position 453 of the sense strand of the GFP gene and dNaM at position 36 of the sense strand of the SerT tRNA gene, as well as encoding the sgRNAs targeting the most common substitution mutation expected in each sequence (Fig. S2). YZ4 and YZ3 (again used as a control) were transformed with pAIO2X and subjected to the challenging growth regime depicted in Fig. 5, which included extensive high-density growth on solid and in liquid media, thus making them an underestimate of actual cell doublings.

**Materials and Methods**

**Strains, Plasmids, and Oligonucleotides.** A complete list of strains and plasmids and the sequences of oligonucleotides used in this work can be found in Dataset S2; for information regarding strain construction and plasmid cloning, as well as additional experimental details, see SI Materials and Methods. Unless otherwise stated, liquid bacterial cultures were grown in 2×YT (casein peptone 16 g/L, yeast extract 10 g/L, NaCl 5 g/L) supplemented with potassium phosphate (50 mM, pH 7), referred to hereafter as “media,” and incubated at 37 °C in a 48-well flat-bottomed plate (CELLSTAR; Greiner Bio-One) with shaking at 200 rpm. Solid growth media was prepared with that is more autonomous (it is naturally competent to import the unnatural triphosphates) and which stores increased information with a fidelity approaching that of natural information. However, unlike any natural organism, the SSO includes an inanimate, man-made component: a UBP that allows it to store information with the virtually unrestricted ability to maintain increased information, the optimized SSO now provides a suitable platform for efforts to retrieve the increased information and create organisms with wholly unnatural attributes and traits not found elsewhere in nature.

**Conclusion**

Since the last common ancestor of all life on Earth, biological information has been stored with the same four-letter, two-base-pair genetic alphabet. By combining chemical optimization with genetic and immunological engineering, we have created an SSO

---

**Fig. 4.** Variation in Cas9-mediated immunity with sequence context. (A) UBP retentions of pUCX2 variants in strain YZ2 with a pCas9 plasmid that expresses a nontarget sgRNA (gray) or an on-target sgRNA (black). Error bars represent SD of the mean, and n = 3 transformations for all sequences except on-target CXA and CXG, where n = 5. (B) UBP retentions of pAIO plasmids in strain YZ3 (gray), which does not express Cas9, or in strain YZ4 (black) with expression of Cas9. In A and B, the nucleotides immediately flanking X = dNaM are indicated, as is distance to the PAM. “(M)” denotes the nucleotide X in the sgRNA that targets a substitution mutation of the UBP; all pCas9 and pAIO plasmids also express an sgRNA that targets a single nucleotide deletion of the UBP. Error bars represent SD of the mean, and n ≥ 3 colonies; see Table S1 for exact values of n, sequences, and IPTG concentrations used to induced Cas9 in YZ4. See Materials and Methods for additional experimental details.

**Fig. 5.** Simultaneous retention of two UBPs during extended growth. Strains YZ3 and YZ4 were transformed with pAIO2X and plated on solid media containing dNaMTP and dTPT3TP, with or without IPTG to induce Cas9. Single colonies were inoculated into liquid media of the same composition, and cultures were grown to an OD_{600} of ~2 (point 1). Cultures were subsequently diluted 30,000-fold and regrown to an OD_{600} of ~2 (point 2), and this dilution–regrowth process was then repeated two more times (points 3 and 4). As a no-immunity control, strain YZ3 was grown in the absence of IPTG, and two representative cultures are indicated in gray. Strain YZ4 was grown in the presence of varying amounts of IPTG, and averages of cultures are indicated in green (0 μM, n = 5), blue (20 μM, n = 5), and orange (40 μM, n = 4). Retentions of the UBP in gfp and serf are indicated with solid or dotted lines, respectively. After the fourth outgrowth, two of the YZ4 cultures grown with 20 μM IPTG were subcultured on solid media of the same composition. Three randomly selected colonies from each plate (n = 6 total) were inoculated into liquid media of the same composition, and each of the six cultures was grown to an OD_{600} of ~1 (point 5), diluted 300,000-fold into media containing 0, 20, and 40 μM IPTG, and regrown to an OD_{600} of ~1 (point 6). This dilution–regrowth process was subsequently repeated (point 7). The pAIO2X plasmids were isolated at each of the numbered points and analyzed for UBP retention (Fig. S6). Cell doublings are estimated from OD_{600} (Materials and Methods) and did not account for growth on solid media, thus making them an underestimate of actual cell doublings. Error bars represent SD of the mean.
2% (wt/vol) agar. Antibiotics were used, as appropriate, at the following concentrations: carbenicillin, 100 μg/mL; streptomycin, 50 μg/mL; kanamycin, 50 μg/mL; and tetracycline, 15 μg/mL, for chromosomal integrants. All selective agents were purchased commercially. Cell growth, indicated as OD_{600}, was measured using a PerkinElmer Envision 2103 Multilabel Reader with a 95020-mm filter.

Natural oligonucleotides were purchased from IDT with standard purification and desalting. Gene synthesis of the codon-optimized PNTT2 and GFP gene sequences was performed by GeneArt Gene Synthesis (Thermo Fisher) and GenScript, respectively, and kindly provided by Synthorx. Sequencing was performed by Eton Biosciences or Genevet. Plasmids were isolated using commercial miniprep kits (QiAprep; Qiagen, or ZR Plasmid Miniprep Classic; Zymo Research).

The {cα}^{[32P]}dATP (3,000 Ci/mmol, 10 mCi/mL) was purchased from PerkinElmer. Triphosphates of dNAM, dSSCs, dtTTP, and dMMOQ^{25} were synthesized as described previously (5, 7, 10, 17) or kindly provided by Synthorx. The dNAM-containing TK1 oligonucleotide was described previously (14). All other unnatural oligonucleotides containing dNAM were synthesized by Biosearch Technologies with purification by reverse phase cartridge and kindly provided by Synthorx.

dATP Uptake Assay. Radioactive uptake assays were conducted as described (26), with the following modifications: C41(DE3) and BL21(DE3) strains carrying plasmid-based transporters and their appropriate empty plasmid controls were grown, prepared, and electroporated as described above for YZ3. Comparable isogenic transporter controls, grown with appropriate isogenic transporter control, were grown overnight with appropriate antibiotics (streptomycin for pCDF plasmids and chloramphenicol for pSCl plasmids and integrants) in 500 μL of media. Cultures were diluted to an OD_{600} of 0.5 in 500 μL of fresh media, grown for 2.5 h, induced with IPTG (0 mM to 1 mM, pCDF strains only) or grown (all other strains) for 1 h. The cultures were spiked with 20 μL of dATP (final concentration of 250 μM [0.5 μC/mL]) for ~1 h. This experimental scheme is analogous to the protocol used to prepare cells for transformation with UBPC-containing plasmids, with the 1 h of dATP incubation simulating the 1 h of recovery in the presence of unnatural triphosphates following electroporation. A duplicate 48-well plate without [cα]^{[32P]}dATP was grown in parallel to monitor growth.

Following incubation with dATP, 200 μL of each culture was collected through a vacuum filter plate (Multiscreen; EMD Millipore) under vacuum and washed with cold potassium phosphate (320 μM, 50 mM, pH 7) and cold ddH_{2}O (1 × 200 μL). Filters were removed from the plate and exposed overnight to a storage phosphor screen (BAS-IP MS; GE Healthcare Life Sciences), which was subsequently imaged using a flatbed laser scanner (Typhoon 9410; GE Healthcare Life Sciences). The resulting image was quantified by densitometric analysis using Image Studio Lite (LI-COR). Raw image intensities of each sample were subtracted against the average background determined from a background image.

In Vivo Plasmid Replication Experiments. Electrocompetent YZ3 cells were prepared by overnight growth in ~5 mL of media supplemented with chloramphenicol, dilution to OD_{600} of 0.02 in the same media (variable volumes, ~10 mL of media per transformation), and growth to OD_{600} of ~0.3 to 0.4. Cells were then rapidly chilled in an ice water bath with shaking, pelleted (2,500 × g, 10 min), and washed twice with an ice cold volume of ice-cold ddH_{2}O. Electrocompetent cells were then resuspended in ice-cold ddH_{2}O (50 μL per transformation), mixed with a Golden Gate assembled plasmid (~1 μL, ~1 ng) containing the UBP, and transferred to a prechilled 0.2-cm-gap electroporation cuvette. Cells were electroporated (Gene Pulse II; Bio-Rad) according to the manufacturer’s recommendations (voltage 25 kV, capacitor 25 μF, resistor 200 Ω), then immediately diluted with 950 μL of prewarmed media supplemented with chloramphenicol. An aliquot (10 μL to 40 μL) of this dilution was then immediately diluted fivefold with the same prewarmed media, but additionally supplemented with dNAMTP (250 μM) and dSSCSdTTP (250 μM). The samples were incubated (37 °C, 1 h), and then ~15% (vol/vol) of the sample was used to inoculate media (final volume 250 μL to 300 μL) supplemented with chloramphenicol, carbenicillin, dNAMTP (250 μM), and dSSCSdTTP (250 μM). Cells were then monitored for growth, collected at the density (OD_{600}) indicated in Results and Discussion, and subjected to plasmid isolation. Dilutions of the recovery mixture were also spread onto solid media with chloramphenicol and carbenicillin to ascertain transformation efficiencies. Experiments with dNAMTP (150 μM) and dtTTPdTTP (75 μM) were performed analogously.

Experiments with DM1 were performed analogously using media supplemented with streptomycin, with the additional step of inducing transporter expression with IPTG (1 mM, 1 h) before pelleting the cells. All media following electrocompetent cell preparation was also supplemented with streptomycin and IPTG (1 mM) to maintain expression of the transporter.

In Vivo Plasmid Replication Experiments with Cas9 (Liquid Culture Only). Electrocompetent Y22 cells were transformed with various pCas9 guide plasmids, and single clones were used to inoculate overnight cultures. Cells were then grown, prepared, and electroporated as described above for Y2Z, with the following modifications: All media was additionally supplemented with zeocin (to select for pCas9) and 0.2% glucose, electrocompetent cells were stored in 10% (vol/vol) ddH_{2}O DMSO at ~80 °C until use, and recovery and growth media were supplemented with dNAMTP (250 μM) and dtTTPdTTP (75 μM). Varying concentrations of IPTG (0 μM to 100 μM) were added to the growth media (but not the recovery media) to induce Cas9 expression. The sgRNA (GCR) sequence corresponding to the dATP and dtTTPdTTP target sites, for all sgRNAs except for the d(GCAT)-containing sequence itself, the nontarget guides for which correspond to the d(GXT) sequence, and all experiments with nontarget sgRNAs were conducted with the addition of IPTG (10 μM) to the growth media. For growth and regrowth experiments, cells were grown to an OD_{600} of 3.5 to 4.0, then diluted 1,250 and regrown to an OD_{600} of 3.5 to 4.0, after which plasmids were isolated.
Biotin Stain Assay. The retention of the UBP(s) in isolated plasmids was determined as previously described and validated (14), with the following modifications: 0.5 μL of Streptavidin-alkaline phosphatase or Golden Gate assembled plasmids (0.5 μL, 0.5 ng/μL to 5 ng/μL), or dNMA-containing oligonucleotides (0.5 fmol), were PCR-amplified with dNTPs (400 μM), 1× SYBR Green, MgSO4 (2.2 mM), primers (500 nM each), dSSICSTP (65 μM), dEMO2AM (65 μM), OneTag DNA polymerase (0.018 U/μL), and DeepVent DNA polymerase (0.007 U/μL) in 1× OneTag standard reaction buffer (final volume 15 μL), under the following thermocycling conditions: [20 °C (95 °C 0:15 / x°C 0:15 / 68 °C 4:00); see Dataset S2 for a list of primers and their corresponding annealing temperatures (x °C) used in this assay. After amplification, 1 μL of each reaction was mixed with streptavidin (2.5 μL, 2 μg/μL, Promega) and briefly incubated at 37 °C. After incubation, samples were mixed with loading buffer and run on a 6% (wt/vol) polyacrylamide (29:1 acrylamide: bis-acrylamide) Tris/borate/EDTA (TBE) gel, at 120 V for ~30 min. Gels were then stained with 1× SYBR Gold dye (Thermo Fisher) and imaged using a Molecular Imager Gel Doc XR+ (Bio-Rad) equipped with a 520DF39 filter (Bio-Rad).

Calculation of UBP Retention. UBP retention was assessed by densitometric analysis of the gels (Image) or Image Studio Lite; LICOR) from the biotin stain assay and calculation of a percent raw shift, which equals the intensity of the streptavidin-shifted band divided by the sum of the intensities of the shifted and unshifted bands for each band. For representative gels. Reported UBP retention is normalized values.

Unless otherwise indicated, for experiments not involving plating on solid media, UBP retention was normalized by dividing the percent raw shift of each propagated plasmid sample by the percent raw shift of the Golden Gate input plasmid. We assume that the starting UBP content of the cellular plasmid population is equivalent to the UBP content of the input plasmid, which is a valid assumption given direct inoculation of the transplasmid population is equivalent to the UBP content of the input plasmid. Plating enables clonal isolation of UBP-containing plasmids from full natural plasmids that arose during plasmid construction [some of which may contain sequences that are not recognized by the sgRNA(s) used]. Because there is no PCR-mediated loss of the UBP in the oligonucleotide template, normalization to the oligonucleotide template is a better indicator of absolute UBP retention than normalization to the input plasmid. Under the conditions used in the biotin shift assay, most oligonucleotide templates and sequence contexts give ~98% raw shift, with <2% shift for a cognate fully natural template (i.e., UBP misincorporation during the biotin shift assay is negligible).

Plating allows for the differentiation between UBP loss that occurs in vivo from that occurs in vitro, with the exception of clonally derived samples that give <2% shift, for which we are unable to differentiate between whether the UBP was completely lost in vivo or if the sample comes from a transformant that originally received a fully natural plasmid. Such samples are excluded from reported average values when other samples from the same transformation give higher shifts.

ACKNOWLEDGMENTS. This material is based upon work supported by the National Institutes of Health Grant GM060005 (to F.R.). The National Science Foundation Graduate Research Fellowship Grant DGE-1346837 (to A.W.F.), the National Natural Science Foundation of China Grant 21472036 (to L.L.), and the National Aeronautics and Space Administration Grant NNX10AB66G (to T.L.). The NanoBio-ICMG platforms (FR 2607) are acknowledged for their support. B.M.L. was supported by a postdoctoral fellowship from the American Cancer Society, Illinois Division.
Supporting Information

Zhang et al. 10.1073/pnas.1616443114

SI Materials and Methods

General Methods for Strain and Plasmid Construction. Unless otherwise stated, all molecular biology reagents were obtained from NEB and were used according to the manufacturer’s protocols. PCRs for cloning and strain construction were performed with Q5 DNA polymerase. Thermocycling was performed using a PTC-200 thermocycler (MJ Research), except for the PCRs used to generate UBP-containing Golden Gate inserts and the PCRs used in the biotin shift assay, which were performed with a CFX-Connect Real-Time Thermal Cycler (Bio-Rad) to monitor product amplification with SYBR Green (Thermo Fisher). Where necessary, primers were phosphorylated using T4 polynucleotide kinase. Plasmids linearized by PCR were treated with DpnI to remove the plasmid template, and ligations were performed with T4 DNA ligase. PCRs and Golden Gate assembled plasmids were purified by spin column (DNA Clean and Concentrator-5; Zymo Research). DNA fragments isolated by agarose gel electrophoresis were purified using the Zymoclean Gel DNA recovery kit (Zymo Research). Colony PCRs were performed with Taq DNA polymerase. Natural DNA fragments and plasmids were quantified by A260/280 using a NanoDrop 2000 (Thermo Fisher) or an Infinite 200 PRO Microplate Reader (Tecan) to 246 bp downstream of the lac operator. For colony PCR, 1 μL of the PCR product was amplified with primers DM052 and YZ612/pSC-PNTT2. A single clone of the resulting plasmid was isolated, and confirmed by sequencing using primers DM002 and YZ580.

Construction of PNTT2 Plasmids. Construction of pCDF-1b-PNTT2 was described previously (14). To create pCDF-1b-PNTT2(66-575), phosphorylated primers YZ552 and pCDF-1b-fwd were used to linearize pCDF-1b-PNTT2 by PCR, and the resulting product was intramolecularly ligated. Plasmids from single clones were isolated and confirmed by sequencing the PNTT2 gene using primers T7 seq and T7 term seq.

To create plasmids pS-C::lacZ expression cassette and Cm1 kb of sequence 3′ of the lac gene was introduced via primers DM053 and YZ700, which are complementary to the lac promoter and the lac UV5 terminator, cloned by homologous recombination in E. coli DH5α, referred to hereafter as 426.trunc, was digested with PvuII and BamHI, and the resulting 5,969-bp plasmid fragment was generated analogously to pCdClacZYA::Pt::lacUV5::PNTT2(66-575)-T0 CmR [a promoterless plasmid generated analogously to 426.lacZYA::Pt::lacUV5::PNTT2(66-575)-T0 CmR using pSC-PNTT2(66-575)]-T0, referred to hereafter as 426.trunc, was digested with PvuII and AvrII, and the resulting 5,969-bp plasmid fragment was isolated by agarose gel electrophoresis and purification, and gap-repairing with PCR and sequencing of the transporter with primers DM053 and YZ50. Plasmid 426.lacZYA::Pt::lacUV5::PNTT2(66-575)-T0 CmR was generated from the linearization template plasmids with the PNTT2(66-575) expression cassette and CmR flanked by ~1 kb of sequence 5′ to lacZ and ~1 kb of sequence 3′ to 246 bp downstream of the lac operator. The lacZYA locus was chosen so that integration of the transporter would also knock out the lactose permease lacY, thus creating a BL21(DE3) strain that allows for uniform cellular entry of IPTG, and thereby homogenous, finely titratable induction of promoters containing lac operators (32).

To create the integration template plasmids, pRS426 (33) was digested with PvuII-HF, and the resulting 3,810-bp plasmid fragment was isolated by agarose gel electrophoresis and purification. This fragment was then gap-repaired in the S. cerevisiae strain BY4741 via lithium acetate-mediated chemical transformation (34) of the plasmid fragment and PCR products of the following primer/template combinations: YZ/YZ12/pBR322, YZ613/YZ850/E. coli genomic DNA, YZ614/615/E. coli genomic DNA, and DM052/YZ612/pSC-Pnt::lacUV5(66-575)-T0. The resulting plasmid, 426.lacZYA::Pt::lacUV5::PNTT2(66-575)-T0 CmR, was isolated (ZymoPrep Yeast Plasmid Miniprep; Zymo Research), digested with PvuII and XbaI (to reduce background during integration, because the pRS426 shuttle plasmid also contains an E. coli PMBI origin), and used as the template to generate a linear integration fragment via PCR with primers YZ616 and YZ617. Integration of this fragment into BL21(DE3) to generate strain YZ2 was performed using pKD46 as described previously (35). Integrants were confirmed by colony PCR of the 5′ and 3′ junctions using primers YZ618 and YZ887 (1,601-bp product) and YZ69 and YZ619 (1,402-bp product), respectively, detection of lacZ deletion via growth on plates containing X-gal (80 μg/mL) and IPTG (100 μM), and PCR and sequencing of the transporter with primers DM053 and YZ50.

Plasmid 426.lacZYA::Pt::lacUV5::PNTT2(66-575)-T0 CmR was generated from the linearization of 426.lacZYA::Pt::lacUV5::PNTT2(66-575)-T0 CmR using phosphorylated primers YZ850 and pCDF-1b-rev, and subsequent intramolecular ligation. The 426.lacZYA::Pt::lacUV5::PNTT2(66-575)-T0 CmR was then integrated into BL21(DE3) to create an isogenic, transporter-less control strain for dATP uptake assays.

To create plasmids 426.lacZYA::Pt::lacUV5::PNTT2(66-575)-T0 CmR, plasmid 426.lacZYA::Pt::lacUV5::PNTT2(66-575)-T0 CmR [a promoterless plasmid generated analogously to 426.lacZYA::Pt::lacUV5::PNTT2(66-575)-T0 CmR using pSC-PNTT2(66-575)]-T0, referred to hereafter as 426.trunc, was digested with PvuII and AvrII, and the resulting 5,969-bp plasmid fragment was isolated by agarose gel electrophoresis and purification, and gap-repaired using PCR products of the following primer/template combinations: I, YZ12/YZ850/426.trunc; 2, DM053/YZ610/pSC-P-lac::lacUV5::PNTT2(66-575)-T0; and 3, YZ851/YZ50/pSC-P-Pnt::PNTT2(66-575)-T0 CmR Plasmids 426.lacZYA::Pt::lacUV5::PNTT2(66-575) Co-T0 CmR were generated analogously, except fragments 2 were replaced with fragments corresponding to the promoters Plac, P255, and P1407, which were generated by annealing and extension of primer pairs YZ703/YZ704, YZ707/YZ708, YZ709/YZ710, and YZ711/YZ712, respectively, with Klengow fragment. Plasmids 426.lacZYA::Pt::lacUV5::PNTT2(66-575)-T0 CmR were then used to integrate the transporter into BL21(DE3) using primers YZ616 and YZ617, and recombineering, as described above. Strain YZ3
denotes BL21(DE3) integrated with lacZYA:PtacUV5PNTT2(66-575 co)-T0, CmR.

To create strain YZ4, the 4,362-bp fragment of SpeI- and AvrII-digested pcAs9-Multi was ligated into SpeI-digested pKIKOarsBKS (30) and the resulting plasmid, pKIKOarsB:Ptac-Cas9-Term (KmR), was used as the template to generate a linear integration fragment via PCR with primers YZ720 and YZ721. The fragment was then integrated into BL21(DE3) as described above, and confirmed by colony PCR with primers YZ720 and YZ721 and sequencing of the product with primers TG1-TG6. pUCuvGGPNTT2(66-575 co)-T0, CmR was subsequently integrated into this strain, as described above, to generate strain YZ4.

Construction of Golden Gate Destination Plasmids for pUCX1, pUCX2, and pBRX2. Although we previously cloned the UBP into plasmids via circular polymerase extension cloning (CPEC) (14, 36), the method results in a doubly nicked plasmid that cannot be treated with T5 exonuclease to degrade unincorporated linear plasmid and inserts, and thus makes it difficult to accurately quantify the yield of the cloning reaction and control the amount of input plasmid used to transform cells during an in vivo replication experiment. Furthermore, the unincorporated linear plasmid and inserts of a CPEC reaction can also template PCR reactions with the primers used in the bidirectional assay, and thus, in vivo shift assays on CPEC products do not truly reflect the UBP content of the plasmids that are actually transformed into cells. To circumvent these complications, the UBP was incorporated into plasmids using Golden Gate Assembly (37).

To create pUCX1 GG and pUCX2 GG, the Golden Gate destination plasmids for pUCX1 and pUCX2, respectively, pUC19 was linearized with phosphorylated primers pUC19-lin-fwd and pUC19-lin-rev, and the resulting product was intramolecularly ligated to delete the natural 75-nt TK1 sequence. The resulting plasmid was then linearized with phosphorylated primers YZ51 and YZ52, and the resulting product was intramolecularly ligated to mutate the BsaI recognition site within the ampicillin resistance promoter and terminator were introduced by PCR (37, 36). This plasmid was then linearized with primers pUC19-lin-fwd and pUC19-lin-rev (for pUCX1), or primers YZ95 and YZ96 (for pUCX2), and ligated to an inserted generator from PCR with phosphorylated primers YZ93 and YZ94 and template pCas9-Multi, to introduce two BsaI recognition sites (for cloning by Golden Gate Assembly) and a zeocin resistance marker (a stuffer cassette used to differentiate between plasmids with or without an insert) into pUC19.

To create pBRX2 GG, the Golden Gate destination plasmid for pBRX2, the 2,934-bp fragment of Avai- and EcoRI-HF-digested pBR322 was end-filled with Klenow fragment and intramolecularly ligated to delete the tetracycline resistance cassette. The BsaI recognition site within AmpR was mutated as described above. The plasmid was then linearized with primers YZ95 and YZ96 and ligated to the BsaI-zeoR-BsaI cassette as described above. Thus, pBRX2 is a lower-copy analog of pUCX2.

Construction of Golden Gate Destination Plasmids for pcAs9 and pAIO. To create pcAs9-Multi, the Golden Gate destination plasmid for cloning sgRNA cassettes alongside Cas9, pPDAZ (38), and a PCR-amplified Cas9 gene (Primers JL126 and JL128, template Addgene plasmid #41815) were digested with KpnI and XbaI, and ligated to create pcAs9(-). This plasmid and a PCR-amplified GFPPT2-sgRNA cassette (template Addgene plasmid #41820, which contains the sgRNA sequence; the prok promoter and terminator were introduced by PCR) were digested with SalI and ligated to create pcAs9-GFPPT2. This plasmid was then linearized with primers BL557 and BL558 (to remove the BsmBI recognition sites within Cas9) and circularized via Gibson Assembly (39). The resulting plasmid was then linearized with primers BL559 and BL560 (to reintroduce two BsmBI sites in the plasmid backbone), and circularized via Gibson Assembly to yield pcAs9-Multi, which was confirmed by sequencing with primers TG1 to TG6. Digestion of pcAs9-Multi with BsmBI results in a linearized plasmid with overhangs that allow for the simultaneous cloning of one or more sgRNAs by Golden Gate Assembly (see SI Materials and Methods, sGRNA Cloning into pcAs9 and pAIO).

To create pAIO-Multi, pcAs9-Multi was linearized with primers BL731 and BL732 (to remove Cas9 and introduce BsaI recognition sites for UBP cloning), phosphorylated, and intramolecularly ligated, and confirmed by sequencing with primer BL540. Digestion of pAIO-Multi with BsaI results in a linearized plasmid with overhangs identical to the ones produced by BsaI digestion of the pUCX2 destination plasmid, and thus PCR-generated inserts for cloning the UBP into pUCX2 can also be used to clone the UBP into pAIO-Multi and its derivatives. After the sgRNA cassettes were cloned into pAIO-Multi (see SI Materials and Methods, sGRNA Cloning into pcAs9 and pAIO), the Golden Gate Assembly protocol for cloning in a UBP was identical to the one described above for pUCX2, except the product of pAIO-Multi (with sgRNAs) amplified with BL731 and BL732 was used in place of the plasmid itself.

sGRNA Cloning into pcAs9 and pAIO. Dual sgRNA cassettes were cloned into pcAs9-Multi or pAIO-Multi via Golden Gate Assembly. To generate the first sgRNA cassette of each pair, pcCas9-GFPPT2 (1 ng) was PCR-amplified with primers first sgRNA GG (200 nM) and BL562 (200 nM), and OneTaq DNA polymerase, under the following thermocycling conditions: [30 × (94 °C 0:30 | 52 °C 0:15 | 68 °C 0:30)]. PCR products were purified by agarose gel electrophoresis and purification. The first sgRNA GG primer is a 70-nt primer that possesses (from 5' to 3') a BsmBI restriction site, 10 nt of homology with the prok promoter, an 18-nt variable guide (spacer) complementary to a UBP mutation, and 25 nt of homology to the nonvariable sgRNA scaffold. To generate the second sgRNA cassette, pcCas9-GFPPT2 was PCR-amplified with primers BL563 and second sgRNA Rev, and primers BL566 and second sgRNA Fwd, and the resulting two products were combined and amplified by overlap extension PCR using primers BL563 and BL566, followed by agarose gel electrophoresis and purification.

To assemble the guide plasmids, pcCas9-Multi (40 ng) or pAIO-Multi (28 ng), purified DNA of the first sgRNA cassette (4:5 ng) and second sgRNA cassette (8 ng), T4 DNA ligase (200 U), BsmBI (5 U), and ATP (1 mM) were combined in 1× NEB CutSmart reaction buffer (final volume 20 μL) and thermocycled under the following conditions: [5 × (37 °C 6:00 | 16 °C 8:00) 15 × (55 °C 6:00 | 16 °C 8:00)]. Assembled plasmids were transformed into electrocompetent cells for subsequent sequencing and testing.

To assemble pcAs9-TK1-A, a plasmid containing only one sgRNA cassette, pcCas9-GFPPT2 was amplified with primers BL566 and BL567, and the resulting product was ligated into pcAs9-Multi by Golden Gate Assembly as described above.

To assemble pcAs9-HEGF, a plasmid containing a nontarget sgRNA cassette for TK1 experiments, primers BL514 and BL515 were annealed and ligated, by Gibson Assembly, into pcCas9-GFPPT2 linearized with primers BL464 and BL465.

Construction of pAIO2X. The pAIO2X GG, the Golden Gate destination plasmid for pAIO2X, is derived from three plasmids, using PCR-generated inserts and multiple steps of cloning by restriction enzyme digest and ligation. Inserts from pSYN36, which contains a codon-optimized superfolder gfp with a Golden Gate entry site for cloning in sequences that correspond to nucleotides 409 to 483 of gfp, and pET-22b-ESerG, which contains an E. coli SerT gene with a Golden Gate entry site for cloning in sequences that correspond to nucleotides 10 to 65 of serT, were cloned into pAIO dual guide BsmBI, a version of pAIO-Multi that contains two sgRNA cassettes, with the targeting guide (spacer) sequences

Zhang et al. www.pnas.org/cgi/content/short/1616443114 2 of 9
replaced by two orthogonal pairs of BsmBI recognition sites that enable guide cloning using annealed primer duplexes.

To create pAIO2X-GFP151/Eser-69 GG, annealed primer duplexes of YZ310/YZ316 and YZ359/YZ360 were ligated into pAIO2X GG using the same Golden Gate Assembly reagents and thermocycling conditions used for UBP cloning, with the exception that BsaI was replaced by BsmBI, each primer duplex was used at a 50:1 insert/plasmid molar ratio with 30 fmol of destination plasmid, and the reaction was scaled by one third to 10 μL. Following assembly, the reaction was not digested with additional enzymes or purified, and was directly transformed into chemically competent E. coli DH5α. Following isolation of single plasmid clones and confirmation of the guides by sequencing using primer BL450, the UBPs were cloned into the plasmid by Golden Gate assembly with BsaI, as described in Golden Gate Assembly of UBP-Containing Plasmids.

Cas9 In Vitro Cleavage Assay. To generate the DNA substrates for in vitro Cas9 cleavage assays, templates BL408, BL409, BL410, BL487, BL488, and BL489 (1 ng per 50-μL reaction) were PCR-amplified with primers BL415 (400 nM) and BL416 (400 nM), and OneTaq DNA polymerase in 1× OneTaq standard reaction buffer supplemented with dNaMTP (100 μM), dTP3TP (100 μM), and MgCl₂ (1.5 mM), under the following thermocycling conditions: [25 × (95 °C 0:15 | 56 °C 0:15 | 68 °C 1:30)]. To generate the DNA templates for in vitro transcription of sgRNAs, templates BL318, BL484, BL485, and BL486 (1 ng per 50-μL reaction), which contain the T7 promoter and a CRISPR RNA (crRNA) spacer sequence, were PCR-amplified with primers BL472 (200 nM) and BL473 (200 nM), and OneTaq DNA polymerase in 1× OneTaq standard reaction buffer supplemented with MgCl₂ (6 mM), under the following thermocycling conditions: [20 × (95 °C 0:15 | 60 °C 0:15 | 68 °C 1:30)]. DNA from this first PCR (0.5 μL) was then transferred into a second PCR (100 μL) containing primers BL472 (400 nM), BL439 (500 nM), and BL440 (600 nM), and thermocycled under the following conditions: [4 × (95 °C 0:15 | 68 °C 0:15 | 68 °C 1:30) 20 × (95 °C 0:15 | 60 °C 0:15 | 68 °C 1:30)]. In vitro transcription of the PCR products with T7 RNA polymerase was performed according to the manufacturer’s protocol, and transcribed sgRNAs were purified by PAGE, band excision, and extraction (37 °C, overnight) into an aqueous solution of NaCl (200 mM) and EDTA (1 mM, pH 7), followed by concentration and purification by ethanol precipitation.

For in vitro cleavage reactions, Cas9 nuclease (125 nM) was incubated with each transcribed sgRNA (125 nM) in 1 × Cas9 nuclease reaction buffer for 5 min, then DNA substrate was added, and the reaction was incubated (37 °C, 10 min). The reaction was quenched with SDS/PAGE loading buffer [62 mM Tris-HCl, 2.5% (wt/vol) SDS, 0.002% bromophenol blue, 0.7 M β-mercaptoethanol, and 10% (vol/vol) glycerol], heat-denatured (95 °C, 10 min), and then loaded onto an SDS/PAGE gel. The resulting cleavage bands were quantified by densitometric analysis using ImageJ (40). For each sgRNA, raw cleavage efficiencies were divided by the maximum cleavage observed for that sgRNA across the set of the six DNA substrates, to account for differences in sgRNA activity and/or minor variations in preparation. Experiments were performed in technical triplicate, and averages represent an average of three in vitro cleavage reactions performed in parallel.

Biotin Shift Depletion and in Vivo Mutation Analysis. To determine the mutational spectrum of the UBP in isolated plasmid samples, biotin shift assays were performed as described in Materials and Methods, Biotin Shift Assay. Nonshifted bands, which correspond to natural mutations of the UBP-containing sequences, were excised and extracted (37 °C, overnight) into a minimal amount of an aqueous solution of NaCl (200 mM) and EDTA (1 mM, pH 7), followed by concentration and purification by ethanol precipitation. A sample of extract (1 μL) was PCR-amplified under standard conditions (natural dNTPs only), with OneTaq DNA polymerase and the same primers used for the biotin shift PCR, and the resulting products were sequenced by Sanger sequencing.
Fig. S1. The dATP uptake and growth of cells expressing PtNTT2 as a function of inducer (IPTG) concentration or promoter strength, strain background, and presence of N-terminal signal sequences. (A) Uptake of [α-32P]-dATP in strains with inducible PtNTT2. Error bars represent SD of the mean, and n = 3 cultures; r.d.u. is relative decay units, which corresponds to the total number of radioactive counts per minute normalized to the average OD$_{600}$ across a 1-h window of uptake, with the uptake of C41(DE3) pCDF-1b PtNTT2(1-575) (i.e., DM1) induced with 1,000 μM IPTG set to 1. Deletion of the N-terminal signal sequences drastically reduces uptake activity in C41(DE3), but activity can be restored with higher levels of expression in BL21(DE3). (B) Growth curves of C41(DE3) strains. Induction of T7 RNAP in BL21(DE3) is toxic (see empty vector traces), which masks the effect of deleting the N-terminal signal sequences of PtNTT2 on cell growth. (C) Growth curves of BL21(DE3) strains. Induction of T7 RNAP in BL21(DE3) is toxic (see empty vector traces), which masks the effect of deleting the N-terminal signal sequences of PtNTT2 on cell growth. (D) Uptake of [α-32P]-dATP in strains that constitutively express PtNTT2(66-575) from the indicated promoters. (E) Growth curves of plasmid-based and chromosomally integrated transporter strains. All PtNTT2 strains are non-codon-optimized for plasmid-based expression and codon-optimized for chromosomal expression, unless otherwise indicated. Strain Y24 also contains a chromosomally integrated Cas9 gene.
Fig. S2. Plasmid maps. Promoters and terminators are denoted by white and gray features, respectively; pMB1* denotes the derivative of the pMB1 origin from pUC19, which contains a mutation that increases its copy number (41). Plasmids that contain a UBP are generally indicated with the TK1 sequence (orange), but, as described in Results and Discussion and indicated above, pUCX2 and pAIO variants with other UBP-containing sequences also position the UBP in the approximate locus shown with TK1 above; sgRNA \( (N) \) denotes the guide RNA that recognizes a natural substitution mutation of the UBP, with \( N \) being the nucleotide present in the guide RNA; sgRNA \( (\Delta) \) denotes the guide RNA that recognizes a single nucleotide deletion of the UBP; this sgRNA and its associated promoter and terminator (indicated by \( \dagger \)) are only present in certain experiments. serT and gfp do not have promoters.
**Fig. S3.** Additional characterization of UBP propagation. (A) Growth curves for the experiments shown in Fig. 2B. YZ3 and DM1 (induced with 1 mM IPTG) were transformed with the indicated UBP-containing plasmids, or their corresponding fully natural controls, and grown in media containing dNaMTP and dSSC3TP. Each line represents one transformation and subsequent growth in liquid culture. The x axis represents time spent in liquid culture, excluding the 1 h of recovery following electroporation (Materials and Methods). Growth curves terminate at the OD$_{600}$ at which cells were collected for plasmid isolation and analysis of UBP retention. Staggering of the curves along the x axis for replicates within a given strain and plasmid combination is likely due to minor variability in transformation frequencies between transformations (and thus differences in the number of cells inoculated into each culture), whereas differences in slope between curves indicate differences in fitness. Growth of YZ3 is comparable between all three UBP-containing plasmids (and between each UBP-containing plasmid and its respective natural control), whereas growth of DM1 is impaired by the UBP-containing plasmids, especially for pUCX1 and pUCX2. (B) Retentions of gfp pUCX2 variants propagated in YZ3 by transformation, plating on solid media, isolation of single colonies, and subsequent inoculation and growth in liquid media, in comparison with retentions from plasmids propagated by transformation and growth of YZ3 in liquid media only. Cells were plated from the same transformations used in the experiments for Fig. 2C. Solid and liquid media both contained dNaMTP and dTPT3TP. Cells were harvested at OD$_{600}$ of ~1. Five colonies were inoculated for each of the pUCX2 variants indicated, but some colonies failed to grow (indicated by a blank space in the table). Retentions for samples isolated from transformants grown solely in liquid media were assayed from the same samples shown in Fig. 2C, but were assayed and normalized to an oligonucleotide control in parallel with the plated transformant samples to facilitate comparisons in retention. See Materials and Methods for additional details regarding UBP retention normalization. For samples with near-zero shift, we cannot determine whether the UBP was completely lost in vivo or if the sample came from a colony that was transformed with a fully natural plasmid (some of which arises during plasmid assembly, specifically during the PCR used to generate the UBP-containing insert).
Fig. S4. Effect of dNaM-dTPT3 on Cas9-mediated cleavage of DNA in vitro. Cas9-mediated in vitro cleavage was assessed for six DNA substrates, wherein the third nucleotide upstream of the PAM is one of the four natural nucleotides, dTPT3, or dNaM. The four sgRNAs that are complementary to each natural template were prepared by in vitro transcription with T7 RNAP. To account for differences in sgRNA activity and/or minor variations in preparation, a relative percent maximal cleavage for each sgRNA vs. all six DNA substrates is shown in parentheses (SI Materials and Methods). Values represent means ± 1 SD (n = 3 technical replicates). In several cases, the presence of an unnatural nucleotide significantly reduced cleavage compared with DNA complementary to the sgRNA. These data suggest that Cas9 programmed with sgRNA(s) complementary to one or more of the natural sequences would preferentially cut DNA that had lost the UBP.

Fig. S5. (A) The sgRNA sequences used to enhance retention of the UBP in Fig. 3B (red denotes guide RNA nucleotides mismatched with the DNA target; the position of dTPT3 is denoted by Y and shown in green); hEGFP is a nontarget sgRNA. (B) Sanger sequencing chromatogram illustrating mutation of dNaM to dT in the absence of an sgRNA to target Cas9 nuclease activity. (C) Sanger sequencing chromatogram illustrating that, in the presence of Cas9 and a targeting sgRNA (TK1-A), sequences containing the dNaM to dT mutation are likely depleted by Cas9 cleavage, thus resulting in the accumulation of other mutations that are either not targeted by the TK1-A sgRNA (Δ, a single nucleotide deletion of dNaM) or targeted by the TK1-A sgRNA, but less efficiently because of a mismatch between the guide and the mutation sequence (dNaM to dG). UBP-containing species were depleted before sequencing (SI Materials and Methods). The position of the mutation in the chromatograms shown in B and C is indicated by an arrow.
Fig. S6. Representative biotin shift assay gels for Fig. 5. Each lane (excluding the oligonucleotide controls) corresponds to a pAIO2X plasmid sample isolated from a clonally derived YZ4 culture, grown with the IPTG concentration indicated, after an estimated 10^8 cell doublings in liquid culture (point 7 of Fig. 5). Each plasmid sample is split and analyzed in parallel biotin shift reactions that assay the UBP content at the gfp and serT loci (red and blue primers, respectively).

Fig. S7. Representative growth curves of YZ4 replicating pAIO2X. Growth curves are for the first dilution–regrowth (point 2) in Fig. 5. Curves terminate at the OD_{600} at which cultures were collected for both plasmid isolation and dilution for the next regrowth. Doubling times are calculated from the timepoints collected between OD_{600} 0.1 and 1.0 for each curve and averaged for each strain and/or IPTG condition.
### Table S1. Cas9 NXN sequences

<table>
<thead>
<tr>
<th>pUCX2 or pAIO UBP context</th>
<th>Sequence</th>
<th>Guides, N/Δ</th>
<th>UBP retentions in YZ3 (-Cas9)</th>
<th>UBP retentions in YZ4 (+Cas9)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>GXG</td>
<td>TCACACATGTAGXGATCAGGG</td>
<td>T/Δ</td>
<td>59</td>
<td>36</td>
</tr>
<tr>
<td>GXA</td>
<td>ACCGAGATGGGAXAACCACCACGGG</td>
<td>T/Δ</td>
<td>111</td>
<td>97</td>
</tr>
<tr>
<td>AXG</td>
<td>ACCGAGATGGGCAACCACCACGGG</td>
<td>T/Δ</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>AXA</td>
<td>TCACAACATGAXAATCACTCAAGGG</td>
<td>T/Δ</td>
<td>43</td>
<td>49</td>
</tr>
<tr>
<td>AXC</td>
<td>ACCGAGATGGGCAACCACCACGGG</td>
<td>T/Δ</td>
<td>92</td>
<td>92</td>
</tr>
<tr>
<td>AXT</td>
<td>TGTTCACACTCTGACCGYGGGG</td>
<td>G/Δ</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>CXG</td>
<td>TGTTCACACTCTGACCGYGGGG</td>
<td>G/Δ</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>CXC</td>
<td>TCACAATGTACXCATCAAGGG</td>
<td>T/Δ</td>
<td>16</td>
<td>21</td>
</tr>
<tr>
<td>CXT</td>
<td>ACCGAGATGGGCAACCACCACGGG</td>
<td>T/Δ</td>
<td>84</td>
<td>73</td>
</tr>
<tr>
<td>TXG</td>
<td>TCACAATGTATXCATCAAGGG</td>
<td>T/Δ</td>
<td>13</td>
<td>20</td>
</tr>
<tr>
<td>TXA</td>
<td>TCACAATGTATYATCAAGGG</td>
<td>A/Δ</td>
<td>54</td>
<td>54</td>
</tr>
<tr>
<td>TXC</td>
<td>ACCGAGATGGGTYCAACCACGGG</td>
<td>G/Δ</td>
<td>43</td>
<td>38</td>
</tr>
<tr>
<td>TXT</td>
<td>ATTTCACATXTCTTATTAGGG</td>
<td>T/Δ</td>
<td>104</td>
<td>111</td>
</tr>
</tbody>
</table>

The 22 nt of each UBP-containing sequence examined in Fig. 4 is shown. X = dNaM; Y = dTPT3. The sequence of the sgRNA targeting the substitution mutation of the UBP (N) is the 18-nt sequence 5′ to the NGG PAM, with X or Y replaced by the natural nucleotide indicated. The sequence of the sgRNA targeting the deletion mutation of the UBP (Δ) is the 19-nt sequence 5′ to the NGG PAM, but without X or Y. In the SSO, the activity of Cas9-sgRNA complexes is tuned by modulating the induction of Cas9 with IPTG; different sequences require different IPTG concentrations, likely because the different sgRNAs have different activities, by virtue of being different sequences. Several IPTG concentrations were explored for each sgRNA and sequence context pair, and the optimal IPTG concentrations (and their associated retentions) are reported. YZ3 experiments were performed without IPTG.Retentions shown in Fig. 4B are averaged from the values and number of colonies indicated here.

**Other Supporting Information Files**

- Dataset S1 (PDF)
- Dataset S2 (XLSX)
In plasmids containing the UBP, its position is denoted by N.

>`pCDF-1b-PtNTT2(66-575) (4959 bp)`
>426.lacZYA::PtNTT2(66-575)-T0 CmR  (9925 bp)
TGATAGATGACTAAGCCCTTACAAGCTTCTGAGATTGTCTGCAAGACGCTCGGTGCGCCGCGGCGTTTTTTATTT
GGTGAGAATTGGTGCGGATGGCATTTGAGAAGCACACGGTCACACTGCTTCCGGTAGTCAATAAACCGGTAAACCA
GCAATGCTGCTCAGACTCACCGCAGCTACTGTTGTAATTCATTAAGCATTCTGGCAATGAAATATGTTGA
CAAGGTTGACATTCTTACGACATCACCTTCTGCTTTTTCATTCTGACAAAGATTTCCGAGTACCATCACGAC
GGCCGGGAGAATGTTGCAATAAAGCCGGCTATTTAACGACCCTGCCCTGAACCGACGACCGGGTCGAATTTGCTT
AATTTCTGCCATTCATCCGCTTATTATCACTTATTCAGGCGTAGCACCAGGCGTTTAAGGGCACCAATACTGCCTT
AAAAAAATTAACCCGGCTGATGTCTATTATTTCTCTACGGCGAATTGGAACCTCTTACGTGCCGATCAACGTCTCATT
GCTCGCTAAGAATCAGCCATCTGCTTGCTTTTTTCATTCTCAGGCTAAATTGCTGAAAATTCGCTCTCTCTCC
GATATACTGAGGTGTAATTACCGAGTATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
AATAAGAAGATGTAAGTGTAATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
AATAAGAAGATGTAAGTGTAATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
AATAAGAAGATGTAAGTGTAATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
AATAAGAAGATGTAAGTGTAATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
AATAAGAAGATGTAAGTGTAATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
AATAAGAAGATGTAAGTGTAATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
AATAAGAAGATGTAAGTGTAATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
AATAAGAAGATGTAAGTGTAATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
AATAAGAAGATGTAAGTGTAATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
AATAAGAAGATGTAAGTGTAATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
AATAAGAAGATGTAAGTGTAATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
AATAAGAAGATGTAAGTGTAATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
AATAAGAAGATGTAAGTGTAATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAA TCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATT
TTCGCCAAAAGTCAACCCGGTGATCCAGCTATGGCTATAAGGCCGAGGAGTTCACTTTTGCGATAATAGCTGGG
AGCACCCCTGTCACACCATGGATAAAGCCCGGAGGAGTACTTTCTTTTGCGATAATAGCTGGG
ATCGGAGATCAGTCTTATTATAATCCAGGCGGATCCACCTGGGAGAATATTGCTGAAATACGGTGAC
>pUCX1 GG destination plasmid (3174 bp)

AAGCCGAGCAACGGGCGCCCTTTTTTACGCTGCCTTGGCTGCTGTTTTTGTGACGACGTTATATGGTATAGCTAGCT

> pUCX2 GG destination plasmid (3174 bp)

AAGCCGAGCAACGGGCGCCCTTTTTTACGCTGCCTTGGCTGCTGTTTTTGTGACGACGTTATATGGTATAGCTAGCT
>pCas9-Multi GG destination plasmid (6238 bp)
TTGGAGCGAACTGCTACCGGAACTGAGTGCAGGCGAATGAGACAAACGCGCCATAACAGCGGAATGACAC CGGTAACCCGAAAGGCCAGGAACAGGAGACGCGCACGAGGGAGCCGCCAGGGGGAAACGCCTGGTATCTTTATAGTCCT GTCGGTTTCGCCACCAGTTTGAGCGTCAGATTTCGATGCTGTCAGGGGGGCGGAGCCTATGGAAA