A methylation-directed, synthetic pap switch based on self-complementary regulatory DNA in an all-*E. coli* cell-free expression system

Emanuel Gregor Worst[1], Ömer Kurt[1], Marc Finkler[1], Marc Schenkelberger[1], Volkhard Helms[2], Vincent Noireaux[3], and Albrecht Ott[1,*]

[1]Saarland University, Department of Experimental Physics, Saarbrücken, 66041, Germany
[2]Saarland University, Center for Bioinformatics, Saarbrücken, 66041, Germany
[3]University of Minnesota, School of Physics and Astronomy, Minneapolis, 55455, USA
[*]albrecht.ott@physik.uni-saarland.de

ABSTRACT

Pyelonephritis-associated pili (pap) enable migration of the uropathogenic *Escherichia coli* strain (UPEC) through the urinary tract. UPEC can switch between a stable 'ON phase' where the corresponding *pap* genes are expressed and a stable 'OFF phase' where their transcription is repressed. Hereditary, alternate DNA methylation of only two GATC motives within the regulatory region stabilizes the respective phase over many generations. The underlying molecular mechanism is only partly understood. Previous investigations suggest that *in vivo* phase-variation stability results from cooperative action of the transcriptional regulators Lrp and PapI. Here, we use an *E. coli* cell-free expression system to study the function of pap regulatory region based on a specially designed, synthetic construct flanked by two reporter genes encoding fluorescent proteins for simple readout. Based on our observations we suggest that Lrp and the conformation of the self-complementary regulatory DNA play a strong role in the regulation of phase-variation. Our work not only contributes to better understand the phase variation mechanism, but it represents a successful start for engineering stable, hereditary and strong expression control based on methylation.

KEYWORDS: *E. coli,* pyelonephritis associated pili, pap switch, Lrp, methylation patterns, cell-free expression

1 INTRODUCTION

Synthetic biology is a growing field that rapidly expands the limits of bioengineering. Developments in synthetic biology are applied in genomics[1,2], transcriptomics[3,4], proteomics[5,6] and metabolomics[7,8]. Possible applications reach from artificial cells to pharmacological robots at the nanoscale[9,10]. Cellular devices have been developed to detect and kill pathogenic bacteria in cellular systems by rational engineering[11]. The incorporation of noncanonical amino acids into model proteins[12,13] conveys new functions[14], likely to open new directions in medicine. Synthetic oscillators[15], boolean logic gates[16,17] and switches[18] constitute realizations of simple electronic elements at the molecular level.

The control of protein synthesis is a major asset that synthetic biology builds upon. Strategies to adjust the amounts of mRNA molecules consist of varying the transcription initiation rate or termination frequency, of modulating RNA stability[19] or active removal of mRNA[20]. At the same time gene expression needs to be switched at the transcriptional level[21,22]. In this context it is important to identify and characterize new regulatory parts that expand the repertoire of DNA functional modules[10,23]. Research in this direction cannot proceed without proper understanding of the molecular interplay that the genetic modules are made of[16]. Conversely, this means that synthetic biology represents a valuable tool to investigate and understand related cooperative molecular actions in unprecedented detail. DNA methylation patterns are well known to convey hereditary information about genetic regulation in eukaryotes, however, epigenetic regulation remains poorly understood.

Recently we investigated the role of the methyl-CpG binding domain (MBD) from the eukaryotic MeCP2 protein, a transcriptional regulator that, in conjunction with the BDNF (Brain-Derived neurotrophic factor) promoter sequence, acts as a methylation dependent transcriptional modulator in humans[24]. The observed gene expression in an *Escherichia coli* (*E. coli*) cell-free expression system reproduced the regulation well in spite of its eukaryotic nature[24]. We found evidence of MBD acting by deforming the DNA grooves on distances beyond its small footprint of only two basepairs[24]. This unexpected way of function (through DNA) may contribute to the well-known fragility of the transcriptional regulator towards point mutations manifesting as the Rett syndrome[25].

Methylation based regulatory mechanisms play a role in bacteria as well[26]. In the *E. coli* genome, 1.5% of adenine bases are methylated and about 0.75% of cytosine bases are methylation targets[27]. DNA methylation influences the time of DNA replication, DNA repair, the distribution of chromosomes to daughter cells, and the time of transposition of plasmids[28–31] in bacteria. For these processes, the target DNA sequence must be hemi-methylated so that the related proteins bind to it[32].

A prokaryotic example of a methylation dependent genetic switch is the transcription-based phase variation mechanism of uropathogenic *E. coli* (UPEC) that we study in this work. UPEC can switch between the ‘ON phase’, in which the genes for pathogenic virulence factors (pyelonephritis-associated pili, Pap) are expressed, and an ‘OFF phase’, in which the expression of these genes is repressed[33] (fig. 1). The system is suspected to be highly cooperative so that the observed, strong genetic switching, stable over generations, is achieved with only two methylation sites in the regulatory unit. We present the details of the pap regulating unit in more detail in what follows.

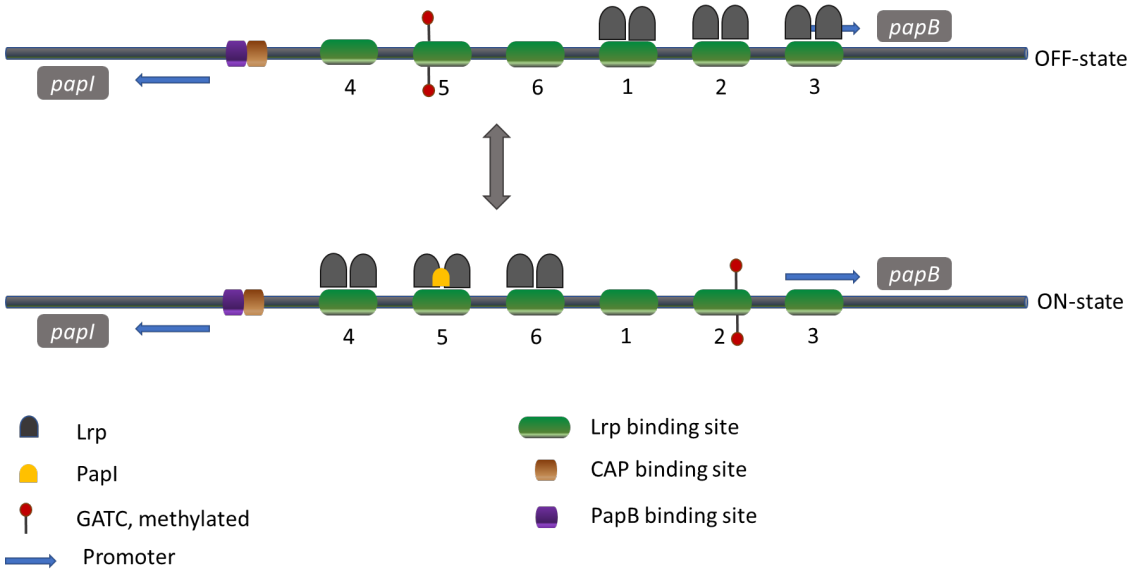


Figure 1. Scheme of the genomic organization of the pap regulatory region. The genes *papI* and *papB* (designated by grey quads) at the flanking sites are oriented in opposite directions. The corresponding promoter regions are highlighted by blue arrows. Six Lrp binding sites (green) are located between the genes *papI* and *papB*. The Lrp binding sites 5 and 2 each contain a single DNA adenine methylase (Dam) recognition sequence (GATC). The PapB binding site 1 (purple) and CAP binding site (brown) are within the regulatory region between the *papI* promoter and the distal Lrp binding site. To simplify, the scheme neither includes further pap regulatory proteins such as CAP or PapB, nor the downstream genes coding for proteins involved in the pili forming process. Moreover, PapB binding sites 2 and 3 are not depicted. The promoter of the downstream genes coding for protein as part of the pili forming process, is called the *papBA* promoter. Scheme is based on Marinus & Løbner-Olesen (2013)[27].**Top:** the putative OFF- state configuration where the *papI* gene is expressed while the *papB* gene is repressed. **Bottom:** the ON-state leading to the reversed expression pattern of both genes.

1.1 Structure of the pap regulatory unit

The pap regulatory region is in between the *papI* gene (upstream) and the *papB* gene (downstream)[34]. Within the pap regulatory region there are three binding sites (BS) for PapB[35]. As illustrated in fig. 1, PapB BS 1 is located between the *papI* promoter and Lrp BS 4[35]. PapB BS 2 overlaps with the RNA polymerase recognition site of the *papBA* promoter and PapB BS 3 is located within the *papB* gene[35] (PapB BS 2 and 3 are not depicted in fig. 1 not to overload the figure). The genes of the co-regulators PapB and PapI are divergently transcribed in opposite directions[36] . The pap regulatory region includes six leucine-responsive regulatory protein (Lrp) binding sites[34]. Lrp has a weight of 18.8 kDa[37,38]. It is a global *E. coli* transcription regulator[37,38]. The six Lrp BS form two functional units[39]. Left we have Lrp BS 4, 5, 6 with the recognition sequence for the DNA adenine methylase (Dam) $GATC^{dist}$ as part of the Lrp BS 5, distal to the *papBA* promoter[39]. On the right we have Lrp BS 1, 2, 3 with the Dam recognition sequence $GATC^{prox}$ as part of the Lrp BS 2, proximal to the *papBA* promoter whose -35 and -10 element enclose Lrp BS 3[39].

1.2 Phases of the pap regulatory unit

Lrp binds to several Lrp BS cooperatively[39,40]. Four Lrp dimers are likely to combine to an octameric structure that binds to three Lrp BS[41–43]. During the transcriptional regulation of the phase variation[33], methylation of Lrp BS 2 and Lrp BS 4 play an essential role[44]. In phase OFF, only $GATC^{dist}$ is methylated whereas in phase ON only the $GATC^{prox}$ position is methylated[44]. *E. coli* strains with knocked out *lrp* genes were unable to express *pap* genes[44]. It was suggested[34] that during phase ON Lrp cooperatively occupies Lrp BS 4-6 in conjunction with papI. This way Dam-mediated methylation can only occur at $GATC^{prox}$ [34]. This inhibits cooperative Lrp binding to Lrp BS 1-3, promoting the expression of *pap* genes[34]. During phase OFF Lrp cooperatively binds to Lrp BS 1-3, sterically inhibiting Dam from methylating $GATC^{prox}$ and, at the same time, preventing RNA polymerase from binding to the corresponding promoter[34]. *Mutual exclusive binding* predicts that cooperatively bound Lrp on Lrp BS 1-3 reduces the binding affinity of Lrp binding to Lrp BS 4-6 by roughly a factor of 10, thereby stabilizing the phase OFF in addition to the Lrp-blocking effect of methylation at $GATC^{dist}$ [34]. Binding of PapB to the PapB BS 1 activates *papI* gene expression[35]. PapI increases the Lrp binding affinity to Lrp BS 2 and Lrp BS 5[45]. The Lrp-PapI complex binds more strongly to the hemimethylated than to the fully methylated $GATC^{dist}$ [45].

1.3 Phase switching

A common model for switching from phase OFF to phase ON involves DNA replication that dissociates Lrp from BS 1-3, leading to hemimethylated $GATC^{dist}$ and $GATC^{prox}$ sites[45–47] , abolishing the *mutual exclusive binding* effect[34]. GATC methylation by Dam within the pap regulatory region proceeds non-processively, so that methylation of $GATC^{prox}$ does not necessarily lead to subsequent methylation of $GATC^{dist}$ or vice versa[47,48]. This means that after DNA replication, Lrp and Dam compete for binding to Lrp BS 2 or Lrp BS 5[45]: In case Dam binds first, this leads to methylation of $GATC^{prox}$ (or $GATC^{dist}$) blocking cooperative Lrp binding to Lrp BS 1-3 (or 4-6). In case Lrp binds first, cooperative Lrp binding to Lrp BS 1-3 (or 4-6) prevents Dam from methylating $GATC^{prox}$ (or $GATC^{dist}$)[45]. Low papI concentrations after DNA replication favor cooperative binding of Lrp to Lrp BS 4-6, preventing full methylation of hemi-methylated $GATC^{dist}$ by Dam[45]. This leads to full methylation of $GATC^{prox}$ hindering cooperative Lrp binding to Lrp BS 1-3, enabling RNA polymerase binding and expression of the *pap* genes[45]. In contrast, in case Lrp cooperatively binds to Lrp BS 1-3, *pap* gene transcription is blocked and the phase OFF is preserved[45]. To stabilize the phase ON, a second round of DNA replication eventually leads to a situation where the $GATC^{dist}$ is fully unmethylated and Lrp cooperatively binds to Lrp BS 4-6 while a fully methylated $GATC^{prox}$ hinders cooperative Lrp binding to Lrp BS 1-3, stabilizing *pap* gene expression[45].

Cooperative binding of Lrp to the Lrp binding sites 1, 2, 3 or 4, 5, 6 as for instance discussed in [34,45,46] reduces the gene expression rate by sterically blocking the RNA polymerase[49]. Hernday *et al.*[34] suggested that the topology of the regulating DNA may well play a role in the phase variation mechanism. A related proposition was made by Peterson[46].

The binding of the Catabolite Activator Protein (CAP) to the CAP BS (fig. 1) next to PapB BS 1 was found a requirement for *pap* gene expression[50,51]. Binding of the Lrp-PapI complex to BS 4-6 may bend the pap regulatory DNA, enabling CAP to interact with the RNA polymerase, however, CAP-dependent *pap* gene transcription does not result from the interaction between CAP and the *papI* promoter or the coding region[51].

In this work we study the transcription-based, epigenetic phase variation mechanism of UPEC as a putative, synthetic expression switch, using an all *E. coli*, cell-free expression system[52,53]. In this system the transcription and translation is performed by the native *E. coli* molecular machinery, making it an ideal tool for the study of the *E. coli* epigenetic phase variation mechanism. We show that in conjunction with the co-regulators Lrp and PapI strong genetic switching can be realized as a response to different methylation patterns. We suggest that Lrp binding to the promoter cooperatively diminishes the likelihood of cruciform conformations and stabilizes the linear conformation of the regulatory DNA that promotes gene expression.

2 METHODS

2.1 Standard molecular-biological procedures

Restriction enzymes, Taq DNA Polymerase, Q5® High Fidelity DNA Polymerase and Gibson Assembly Cloning Kits were from New England Biolabs GmbH (Frankfurt am Main, Germany).
Unless otherwise noted, standard PCRs were performed using either Taq DNA Polymerase or Q5® High Fidelity DNA Polymerase in the corresponding reaction buffer. Plasmid DNA or linear DNA was used as a template. All PCR primers were synthesized and desalinated by Metabion international AG (Martinsried, Germany). PCRs were performed with the MultiGene™ 96-Well Gradient Thermal Cycler (Labnet International Inc., Edison, USA). The hybridization temperature of the primers was predicted using the Tm Calculator (New England Biolabs Inc., Ipswitch, USA) based on work by SantaLucia[54], including corrections for salt concentrations according to Owczarzy *et al.*[55]. After completion of the PCR, a small amount of the PCR reaction (typically 10 µl) was used to check the purity of the PCR products by horizontal agarose gel electrophoresis. If only the DNA fragments of expected size were present in the gel, the remaining sample was purified using the PureLink® PCR Purification kit (Fisher Scientific GmbH, Schwerte, Germany), according to manufacturer's

instructions, and the DNA was dissolved in ultrapure water. In case of additional DNA fragments of unexpected sizes, the entire remaining sample was separated by agarose gel electrophoresis. DNA bands corresponding to DNA strands of the expected length were cut out of the gel. The DNA of interest was extracted and purified from the gel using the PureLink® Quick Gel Extraction Kit (Fisher Scientific GmbH, Schwerte, Germany), according to manufacturer's instructions and dissolved in ultrapure water.

For plasmid amplification *E. coli* strain KL740 (Coli Genetic Stock Center (Yale), CGSG#: 4382[56]), was used. Plasmid preparation was executed using the GenElute™ HP Plasmid Miniprep Kits or GenElute™ HP Plasmid Midiprep Kits (Sigma-Aldrich Chemie GmbH, Taufkirchen, Germany).

Gibson assembly[57,58] of PCR products was undertaken using the 2x Gibson-Assembly®-Master-Mix (New England Biolabs GmbH, Frankfurt am Main). The 20 µl assembly reactions were prepared and incubated according to manufacturer's instructions. For a reaction, the total concentration of all DNA fragments to be assembled was between 1 nM and 25 nM. The total concentration always consisted of equal concentrations of the different DNA strands to be assembled. If not stated otherwise the prepared reactions were incubated between 15 min - 30 min at 50 °C in a MultiGene™ 96-Well Gradient Thermal Cycler (Labnet International Inc., Edison, USA).

2.2 DNA Templates, Plasmids and synthetic pap regulatory construct

For the assembly of the linear synthetic pap regulatory DNA constructs, different DNA templates were used. All plasmids in this study were based on the plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-deGFP-T500 that originates from pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 (Addgene plasmid # 40019), but contains the p15a instead of the ColE1 origin of replication[52]. The vector backbone was pBEST*luc*™ (*Promega Corporation*, Fitchburg, USA), where the PTacI-Promotor[59] was substituted through the strong lambda repressor Cro promoter[22] [GenBank: J02459.1] having a single mutation[52]. The untranslated region of pBEST*luc*™ was substituted by UTR1[52]. UTR1 [GenBank: M35614.1] is an untranslated region that contains the T7 g10 leader sequence that enables a highly efficient translation initiation[60]. T500 is a transcription terminator[61].

To complete the plasmid pBEST-p15a-OR2-OR1-Pr-deGFP-T500, the *luc* gene of pBEST*luc*™ was substituted by the *degfp* gene[52] that codes for a variant of the enhanced green fluorescent protein [GenBank: CAD97424.1] truncated and modified in N- and C-terminal and which is based on [62]. Accordingly, the plasmid pBEST-p15a-OR2-OR1-Pr-mCherry-T500 was created by insertion of the *mcherry* gene [GenBank: AY678264.1, based on work[63]]. The plasmids pBEST-p15a-OR2-OR1-Pr-Lrp-T500 and pBEST-p15a-OR2-OR1-Pr-PapI-T500 were created[64] by replacement of the *degfp* gene by the genes coding for Lrp [GenBank: 949051] or respectively PapI [GenBank: 1039535]. The DNA sequences of regulatory elements and genes of the resulting plasmids pBEST-p15a-OR2-OR1-Pr-

UTR1-Lrp-T500 and pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500 can be found in the supplementary information (fig. S1).

Our synthetic pap regulatory DNA preserves the native DNA sequence (GenBank: X14471.1, based on Blyn *et al.*[33], fig. S3) including all six Lrp BS (Lrp BS according to Hernday et al.[45]) except for the -35 and -10 elements of OR2-OR1-Pr that enclose Lrp BS 3, as well as Lrp BS 4. These elements initiate the transcription of the genes coding for the fluorescent proteins mCherry and deGFP as indicated in fig. 2. At each side UTR1 links the promoter to either gene. While the right -35 and -10 elements are a substitute for the native promoter elements controlling *pap* gene expression, the left -35 and -10 elements are added to the native sequence to enable *mcherry* gene expression.

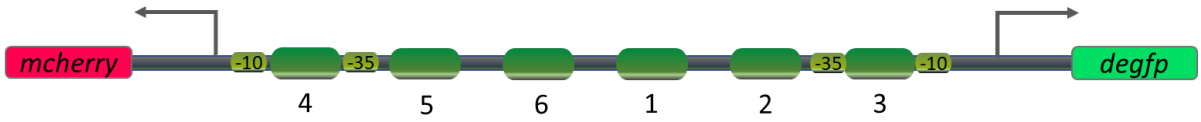


Figure 2. Scheme of synthetic pap regulatory DNA construct. The DNA sequence includes all six Lrp BS (green and numbered, Lrp BS according to Hernday et al.[45]). The synthetic sequence corresponds to the native pap regulatory DNA except for inserted -35 and -10 elements of OR2-OR1-Pr that enclose the Lrp BS 3 and Lrp BS 4, followed by UTR1 and flanked by *mcherry* and *degfp* genes, Sequence (GenBank: X14471.1, based on Blyn et al.[33]). See also fig. S3 for more information.

2.3 Construction of linear synthetic pap regulatory DNA constructs

Braaten *et al.*[44] showed that the single mutation of adenine to cytosine does not influence the binding affinity of Lrp to Lrp BS 5 or Lrp BS 2, however, the mutation prevents Dam methylation of the corresponding DNA motif. Accordingly, three linear synthetic pap regulatory DNA constructs were prepared. They differed in sequence at the Lrp binding sites 5 or 2, the DNA constructs pap(1,0), pap(0,1) and pap(1,1). Here 1 represents the native DNA sequence GATC (possible Dam methylation site), 0 represents the DNA sequence GCTC.

In paragraph 2.3 we explain how the linear three pap regulatory DNA constructs were synthesized. In the paragraph 2.4 the protocol for the subsequent Dam methylation of these products in order to achieve the methylation states ((1,1), (1,0), (0,1)) is illustrated.

2.3.1 Pap(1,1) DNA construct

Two partial fragments, A and B (PCR products), were covalently linked using the Gibson assembly and subsequently amplified by PCR.

For the PCR synthesis of partial fragment A, the plasmid pBEST-p15a-OR2-OR1-Pr-mCherry-T500 was used as PCR template with the primers 1 and 2.

For the PCR synthesis of partial fragment B, a DNA construct consisting of the transcription-translation regulatory and coding elements 5`-UTR1-synthetic-pap-regultaory-region-UTR1-deGFP-T500 (synthesized and cloned in vector pEX-K4 by Eurofins Genomics GmbH, Ebersberg, Germany, sequence in fig. S2) was used with the primers 3 and 4.

As a difference to the standard protocol of the Gibson assembly, for incubation the following scheme was used to assemble the partial fragments A and B: 1 min at 50 °C, 1 min at 45 °C, 1 min at 50 °C, 1 min at 45 °C, 26 min at 50 °C.

Subsequently, a PCR using Taq DNA polymerase (ThermoPol reaction buffer conditions) was executed in order to amplify the assembled linear synthetic pap regulatory DNA using the primers 1 and 4. It was found that this PCR only produced a product of correct length, if 10 µl of the Gibson assembly reaction as PCR template DNA was directly transferred into the PCR reaction of a final volume of 300 µl. The annealing temperature was 50 °C, the elongation time was 1 min and 55 sec, and the PCR had 18 cycles. However, many byproducts of other lengths were produced at the same time. The DNA strands of correct length were purified by agarose gel extraction. Surprisingly, we did not find any condition to efficiently amplify this purified DNA construct via an additional PCR. Furthermore, we found that the utilization of primers 5 and 6 resulted in higher yields of the pap regulatory DNA construct if we executed a PCR after the Gibson assembly of partial fragments A and B as described above (300 µl PCR mix containing 10 µl of the Gibson assembly mix from the previous assembly reaction). Additionally, we increased PCR efficiency using an elongation time of 1 min and 55 sec, an annealing temperature of 56 °C and running the PCR for 15 cycles. The sequence of the resulting synthetic linear pap regulatory DNA construct is given in fig. S3; an alternative way to synthesize this DNA construct, using two different partial fragments, was also successful as described in paragraph 0.4 of the supplementary material.

Table 1. The primers used in this work with their sequences. Mutant nucleotide bases within the Dam recognition sequence are highlighted in red and emphasized by brackets.

Primer No.	**DNA Sequence 5`to 3`direction**	**Forward (F) / Reverse (R) Primer**
1	GAGCTGACTGGGTTGAAGG	F
2	GCTAGCAATAATTTTGTTTAACTTTAAGAAGGAGATATAC	R

3	GTTAAACAAAATTATTGCTAGCTGCAACCATTATGCG	F
4	CACCATCAGCCAGAAAACC	R
5	CGCTATGAATTCCTCGAGTTACTTGTACAGCTC	F
6	CCATGGAAGCTTTTAGATCCCGGCGGCGGTCA	R
7	CATCAGCCAGAAAACCGAGCTGACTGGGTTGAAG	F
8	AGATTTAAACGA[G]CTTTTAACCCACAAAAC	R
9	TAAAAG[C]TCGTTTAAATCTTGACATACAACATAAAAAAC	F
10	CTGACTGGGTTGAAGGCACCATCAGCCAGAAAAC	R
11	CAGCATAAAAGA[G]CGTCTAAATGTTGACATAC	R
12	AGACG[C]TCTTTTATGCTGTAAATTCAATTTGC	F

2.3.2 Pap(1,0) and pap(0,1) DNA construct

We covalently linked in a Gibson assembly reaction and subsequently amplified by PCR: two PCR products of partial fragments C and D to produce the pap regulatory DNA construct pap(1,0); as well as two PCR products of partial fragments E and F to produce the pap regulatory DNA construct pap(0,1).

The PCR-amplified pap regulatory DNA construct from the Gibson assembly (using the primers 1 and 4 as described in paragraph 2.3.1) was the template DNA for the PCR synthesis of partial fragments C and D, or E and F, respectively. For the PCR synthesis of partial fragment C, the primers 7 and 8 were used. For the PCR synthesis of partial fragment D, the primers 9 and 10 were used. For the PCR synthesis of partial fragment E, the primers 7 and 11 were used. For the PCR synthesis of partial fragment F, the primers 12 and 10 were used.

The primers 8 and primer 9 led to complementary ends of the partial fragments C and D, so that these could be covalently linked in a Gibson assembly reaction. The primers hybridizing to the template DNA on the Lrp binding site 2 region, had a single mutation (illustrated in square brackets in table 1) leading to the mutation[44] GCTC described in paragraph 2.3 in both PCR products C and D.

The primer 11 and primer 12 led to complementary ends of partial fragments E and F, so that they could be covalently linked in a Gibson assembly reaction. The primers hybridizing to the template DNA on the Lrp binding site 5 region, had a single mutation (illustrated in square brackets in table 1) leading to the mutation[44] GCTC described in paragraph 2.3 in both PCR products E and F.

Subsequently, partial fragments C and D, as well as E and F were covalently linked following the Gibson assembly incubation protocol described in paragraph 2.3.1. The PCR amplification using the primers 5 and 6 (elongation time of 1 min and 55 sec, an annealing temperature of 56 °C and running the PCR for 15 cycles) and purification by agarose gel extraction of the resulting pap regulatory DNA construct was also executed as described in paragraph 2.3.1.

2.4 Dam methylation of pap regulatory DNA constructs

DNA adenine methylase (Dam), the 10x reaction buffer supplied and S-adenosylmethionine (SAM) (New England Biolabs GmbH, Frankfurt am Main) were used for *in vitro* methylation of the purified, unmethylated pap(1,1), pap(1,0) and pap(0,1) constructs (see 2.3). The Dam methylation protocol used differs from manufacturer's instructions. It was established in[64]. After Dam methylation, the pap regulatory DNA constructs were purified, using the PureLink® PCR Purification kit (Fisher Scientific GmbH, Schwerte) according to manufacturer's instructions, and subsequently dissolved in ultrapure water. Proper methylation was checked by digestion of the pap regulatory DNA constructs with the restriction enzyme DpnI (New England Biolabs GmbH, Frankfurt am Main, Germany) cutting between A and T only at methylated GATC sites. This leads to a characteristic fragmentation pattern in case of successful methylation as observed by agarose gel electrophoresis.

2.5 Cell-free expression and fluorescence measurements

All experiments were performed using an all *E. coli* cell-free system described previously[52,53], known as the kit myTXTL (Arbor Biosciences).

For the experiments presented in fig. 3 and S4, cell-free expression master mix solutions were split into volumes of 9.6 µl. A split reaction volume was completed with 2.4 µl of DNA solution (25 nM) of one of the three pap regulatory DNA constructs pap(1,1), pap(1,0) or pap(0,1) or ultrapure water to a total volume of 12 µl. Accordingly, the concentrations of a pap regulatory DNA construct was 5 nM.

For the experiments presented in fig. 4 and S5, cell-free expression master mix solutions were split into volumes of 9.6 µl. A split reaction volume was completed with 1.2 µl DNA solution of the plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 and 1.2 µl DNA solution of the plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500 at the concentrations 0 nM (ultrapure water), 1 nM and 10 nM. The final concentrations of the plasmids above in the reaction mix were 0 nM, 0.1 nM or 1 nM. The pap regulatory DNA constructs pap(1,1), pap(1,0) or pap(0,1) in the cell-free reaction mix were at a final concentration of 5 nM.

To protect the linear DNA against *E. coli* RecBCD induced degradation, the GamS protein[65] was added to the cell-free reactions (3.3 µM).

Cell-free reactions supplemented with ultrapure water served as a blank in which no (pap regulatory) DNA was present. The fluorescence intensities of the blank as determined with the GFP and mCherry filter sets was determined as outlined below. The blank was subtracted to obtain the intensity due to the presence of fluorescent protein.

For cell-free expression, reactions were incubated at 29 °C. Fluorescence measurements were performed using a microplate reader (POLARstar Optima, BMG Labtech, Ortenberg, Germany observed through bottom optics). For fluorescence measurements, cell-free reactions (10 µl) were transferred into a 384-well plate (Nunc 384-well optical bottom plate, Thermo Fisher Scientific, Waltham, USA) that was sealed (Nunc sealing tape, Thermo Fisher Scientific, Waltham, USA).

During kinetic measurements, cell-free reactions were incubated in a microplate reader for 16 hours. For endpoint measurements, cell-free reactions were first incubated in 1.5 ml reaction tubes (Eppendorf Safe–Lock Tube 1.5 ml, PCR clean, Eppendorf, Hamburg, Germany) for 16 hours and then transferred to the 384-well plate. Cell-free synthesized deGFP was excited at 485 nm and the emission intensity determined at 520 nm. For mCherry, the excitation wavelength was 595 nm while its emission intensity was determined at 620 nm. For all fluorescence measurements the same parameters were used (gains were set to 1100 and 2200 for the deGFP and mCherry channel, respectively). The measured intensities were normalized to the highest deGFP or mCherry fluorescence intensity values (fig. 3 and 4).

2.6 Prediction of pap regulatory DNA secondary structures

Prediction of the secondary structure was performed using the mfold web server[66]. The underlying DNA sequence corresponds to the sequence of the synthetic pap regulatory region including the upstream and downstream UTR1 sequences and the respective beginnings of the *degfp* and *mcherry* genes (base pairs 685-1053 in sequence presented in fig. S3). The structure was predicted for a temperature of 29°C (incubation temperature of cell-free reactions) using the standard settings (DNA type: linear DNA, ionic conditions: 1M Na+, correction type: oligomer, percent suboptimality number: 5, upper bound on the number of computed foldings: 50, maximum distance between paired bases: no limit). These conditions were chosen to reflect the observation by Schoen *et al.* that DNA hybridizes much faster in the cell than *in vitro*[67]. The program Rna Viz[68] generated fig. S6.

3 RESULTS

3.1 Methylation-dependent regulation of the pap regulatory region without cell-free synthesis of Lrp and PapI

We determined the fluorescence emission intensities of deGFP and mCherry as a function of methylation of the regulatory DNA in the case where the cell free reactions were devoid of *lrp*-coding and *papI*-coding plasmid. The methylation states under examination were (0,1): unmethylated $GATC^{dist}$ site and methylated $GATC^{prox}$ site; (1,0): methylated $GATC^{dist}$ and unmethylated $GATC^{prox}$ site; and (1,1): methylated $GATC^{dist}$ site and methylated $GATC^{prox}$ site.

The methylation pattern (0,1) led to the highest deGFP fluorescence intensity while (1,1) produced the lowest deGFP fluorescence intensity. Conversely, the highest mCherry fluorescence intensity appeared for methylation state (1,0). The mean of the mCherry fluorescence intensity in case of methylation state (0,1) was slightly higher than for methylation state (1,1) albeit the errors of the two signals overlap. We understand that a methylated GATC site produced increased fluorescence intensity from its adjacent gene only if the distant GATC site was unmethylated.

In addition to the fluorescence endpoint measurements presented in fig. 3, kinetic fluorescence measurements were performed. The results of the endpoint and the kinetic measurements agreed well (compare fig. S4 and fig. 3). However, the statistical error in the kinetic measurement was increased compared to the endpoint measurements. We could not find a reason for this observation.

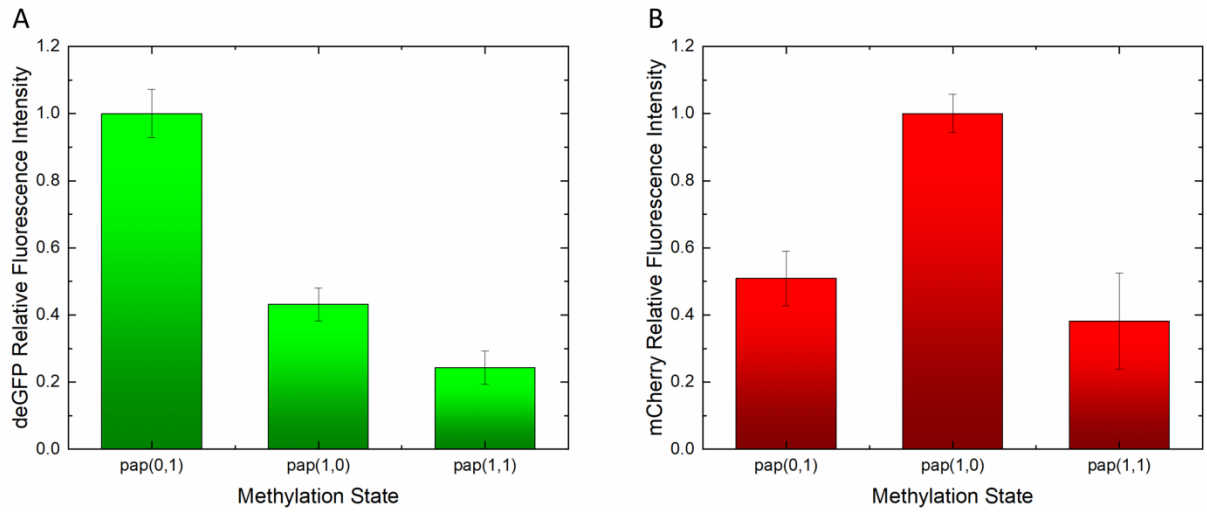


Figure 3. Fluorescence intensities of deGFP (A), mCherry (B) as a function of pap regulatory DNA constructs (5nM) with methylation states (0,1), (1,0) and (1,1). Symbols on the horizontal axis refer to the different methylation states of the DNA construct that were present in the cell-free reactions. (0,1): Lrp BS 5 unmethylated, Lrp BS 2 methylated. (1,0): Lrp BS 5 methylated, Lrp BS 2 unmethylated. (1,1): Lrp BS 2 and 5 are both methylated. Fluorescence intensities were normalized relative to the highest observed intensity of the respective reporter protein. Error bars represent one standard deviation from the mean from five measurements.

3.2 Methylation-dependent regulation of the pap regulatory unit in conjunction with cell-free synthesis of Lrp and PapI

We determined the fluorescence intensities of deGFP and mCherry emission as a function of the methylation pattern of the pap regulatory DNA in the presence of cell-free synthesis of Lrp and PapI. The methylation state (0,1) led to the highest deGFP fluorescence intensities compared to the other

two methylation states (1,0) and (1,1) at the same concentration of the *lrp*-coding plasmid (fig. 4A). At any *lrp*-coding plasmid concentration, we observed the highest mCherry fluorescence intensities for (1,0) compared to the methylation states (0,1) and (1,1) (fig. 4B). The methylation pattern (1,1) resulted in rather weak fluorescence intensity signals of both reporter proteins (compared to the methylation states (1,0) or (0,1)) at different *lrp*-coding plasmid concentrations (fig. 4A, B).

The presence of the *papI*-coding plasmid weakly influenced the reporter fluorescence intensities. Only in case of methylation pattern (0,1) (fig. S5A), in the absence of the *lrp*-coding plasmid we observed that increased concentrations of the *papI*-coding plasmid led to increased deGFP fluorescence intensity while the mCherry fluorescence intensity decreased. In case of methylation state (1,1), again in absence of the *lrp*-coding plasmid, the analog behavior of mCherry fluorescence intensity occurred, but it was less pronounced.

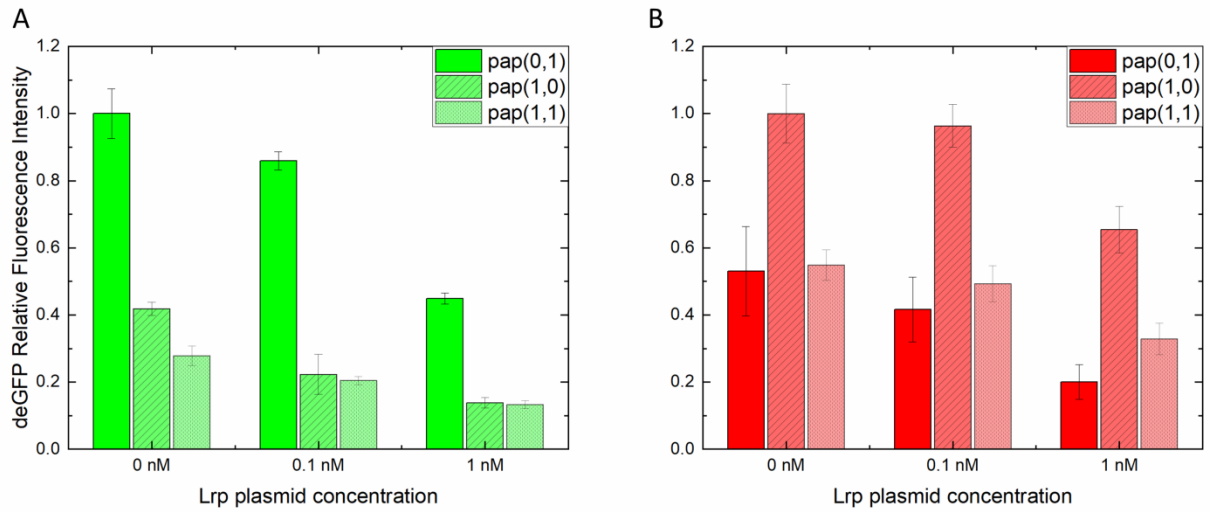


Figure 4. Fluorescence intensities of deGFP (A), mCherry (B) in the presence of pap regulatory DNA constructs with methylation states (0,1), (1,0) and (1,1) as a function of varying *lrp*-coding plasmid concentrations in the absence of *papI*-coding plasmid. The plotted data corresponds to the data presented in fig. S5 where no *papI*-coding plasmid is added. Symbols on the horizontal axis refer to the different methylation states of the DNA construct that were present in the cell-free reactions. (0,1): Lrp BS 5 unmethylated, Lrp BS 2 methylated. (1,0): Lrp BS 5 methylated, Lrp BS 2 unmethylated. (1,1): Lrp BS 2 and 5 both methylated. Fluorescence intensities were normalized to the highest occurring value. Error bars represent one standard deviation from the mean as gained from two measurements.

3.3 Predictions of DNA conformations of the synthetic pap regulatory region

We studied the secondary structures of the pap regulatory DNA using the mfold web server[66]. There were only minor differences among the three predicted secondary structures (fig. 5). In all three structures Lrp BS 2 and 5 as well as Lrp BS 3 and 4 were paired (at least partially). All structures exhibited almost full pairing of $GATC^{prox}$ and $GATC^{dist}$.

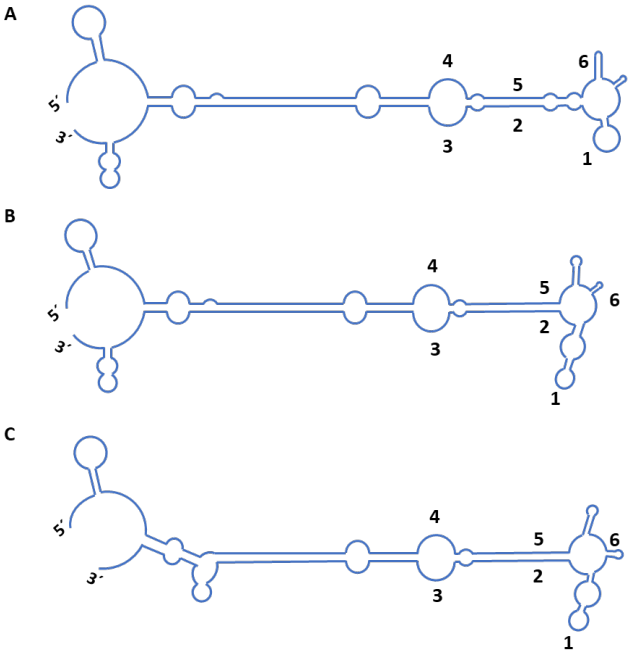


Figure 5. Scheme of the three (A, B, C) predicted secondary structures for the DNA of the synthetic pap regulatory region (more detailed structures are given in fig. S5). In each structure Lrp BS 1 and 6 as well as 3 and 4 are at least partially paired. In all three structures almost full pairing between $GATC^{prox}$ and $GATC^{dist}$ occurs.

4 DISCUSSION

In this work we characterized a synthetic pap regulatory DNA with bidirectional fluorescent reporter gene expression, in an *E. coli*, cell-free expression system[52,53]. In this system gene expression is based on the *E. coli* molecular transcription-translation machinery only. We used the *mcherry* and the *degfp* genes for reporting. In what follows we interpret increased fluorescence intensities due to increased concentrations of the corresponding fluorescent reporter proteins as a higher cell-free expression rate of the corresponding genes. Based on previous data[52] we expect the concentrations of the cell-free synthesized co-regulators Lrp and PapI to scale with the concentrations of their respective plasmids as supplemented to the cell-free reaction.

During preparation of the cell-free system we removed proteins with a molecular weight of less than 10 kDa by dialysis. Since the molecular weight of Lrp is about 19 kDa, this global *E. coli* transcription regulator[37,38] must be expected to be present in our cell free system. We saw that the reporter gene expression varied with the amount of synthesized PapI even without cell-free Lrp synthesis (fig. S5A). This confirms the presence of Lrp since PapI is not able to specifically bind to the pap regulatory DNA without Lrp[69]. Moreover, *E. coli* strains with knocked down *lrp* genes were found unable to express *pap* genes[44]. At the same time, however, native PapI cannot be part of the cell free system since the *papI* gene from UPEC is not present in the bacterial genome of the *E. coli* BL21 Rosetta^{TM} 2 cells used for production of the cell extract. Even if this were not so, with a molecular weight of approximately 8.7 kDa PapI would be removed from the cell extract during dialysis.

We observed that the methylation pattern (0,1) resulted in the highest *degfp* expression rate compared to the other methylation states (fig. 3, 4, S4, S5). This finding suggests that GATC^{prox} must be methylated and GATC^{dist} must be unmethylated for *pap* gene expression, in excellent agreement with previous observations from the native system[44]. Increased PapI concentrations led to decreased *mcherry* expression and increased *degfp* expression (fig. S5A), indicating that PapI sustains the ON state. This again is in excellent agreement with the observed increase of the Lrp binding affinity to Lrp BS 5 due to the presence of PapI[45].

Increased concentrations of *lrp*-coding plasmid resulted in decreased reporter gene expression. This decrease was especially pronounced for the pap regulatory DNA with a single methylated site, either GATC^{prox} or GATC^{dist} (fig. 4). We understand that sites that weaken expression upon Lrp binding were far from saturation at the Lrp concentration given by system preparation. Accordingly, in the presence of the *lrp-coding* plasmid, the additionally synthesized Lrp exerted the role of a suppressor by occupying these sites.

It is clear that our observations cannot be explained by simple, methylation dependent, Lrp blocking that comes with an occupied GATC site close to the gene of interest. Rather, the finding that methylation of GATC^{prox} or GATC^{dist} influenced the reporter gene transcription rate (fig. 3, 4, S4, S5),

although the -35 and -10 elements of both promoters enclose the outer Lrp BS 1 or 6 respectively, sustains the hypothesis[34,39–43] of cooperative Lrp binding on several Lrp BS depending on methylation. We find that in the absence of PapI, contrary to the native system, Lrp alone is sufficient to produce a methylation-dependent interaction between the subunits composed of Lrp BS 1-3 and Lrp BS 4-6 of the pap regulatory region. Sterical blocking of the RNA polymerase[49] through cooperative binding of Lrp [34,45,46] to the unmethylated region (fig. 6) may well be the mechanism that reduced transcription of the reporter genes in our experiments regarding methylation patterns (1,0) and (0,1). Following this mechanism[44] (fig. 6A), however, the methylation pattern (1,1) should result in increased expression levels of both, *degfp* and *mcherry*, but this not what we observe. It is not immediately clear how this methylation pattern translates to molecular binding conformations on the linear DNA regulatory unit, in particular in our simplified synthetic setting.

We observed that the relevant secondary structures of our synthetic pap regulatory sequence (section 3.3, fig. 5) represent three (very similar) DNA cruciform structures (fig. 6B). Cruciform structures are known to play decisive roles in many biological processes, among them gene regulation[70]. The pap regulatory unit is self-complementary except for Lrp BS 1 and 6 as well as 3 and 4. At the same time Lrp BS 2 and 5 exhibit almost full complementarity. In thermodynamic equilibrium, in the absence of binding elements, many cruciform conformations with different length ratios among their arms are likely to coexist (fig. 6C, S7). In cruciform conformations, genes may or may not be expressed as shown in fig. 6C. Due to the entropy arising from all possible cruciform conformations, the linear conformation is not very likely to form spontaneously. Accordingly the role of Lrp in methylation states (1,0) and (0,1) is not only to suppress expression by binding to three binding sites collectively but, equally important, to enable expression by avoiding the cruciform structure of the regulatory portion of the DNA. This could explain the requirement of a cooperative binding complex of Lrp molecules. In such a setting Lrp functions as a suppressor and activator at the same time.

If the pap regulatory DNA of linear conformation in the methylation state (0,1) (or (1,0)) folds according to fig. 6B (left or middle), the newly formed GATC and GATC sites will be hemi-methylated. Hemi-methylated GATC sites have reduced large furrows and are structurally metastable[71], resulting in specific interactions with proteins.

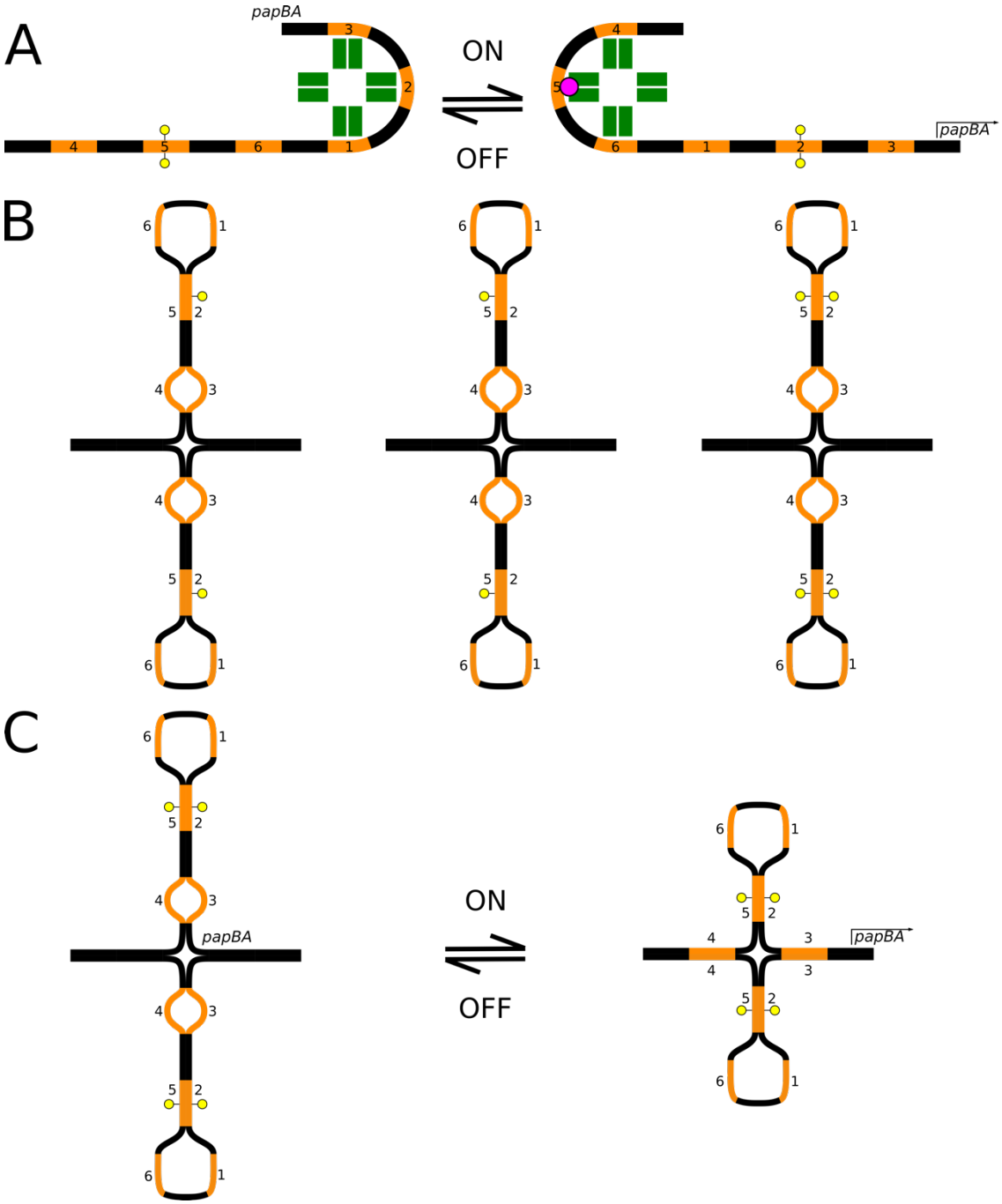


Figure 6. (A) A proposed model of phase variation due to conformational DNA changes based on Peterson (2008)[46]. Phase OFF (left): Lrp cooperatively binds to Lrp BS 1, 2 and 3, forming an octamer (green rectangles), preventing RNA polymerase from binding to the papBA promoter. $GATC^{dist}$ as part of Lrp BS 5 is fully methylated (yellow dots). Phase ON (right): the Lrp-PapI (Lrp: green rectangles, PapI: pink circle) complex binds to Lrp BS 4, 5 and 6. $GATC^{prox}$ as part of Lrp BS 2 is fully methylated (yellow dots). The diagram does not include the contributions of other pap regulatory proteins such as CAP or PapB. Other downstream genes for the pili forming peptides are not shown. (B) Scheme of the different methylation states of the cruciform DNA conformation of the pap regulatory DNA. Lrp BS of the top strand and the bottom strand are highlighted in orange. Methylation of a DNA strand is represented by a single yellow dot. As the methylation states (0,1), (1,0) and (1,1) from the linear DNA conformation transform to the cruciform structure, the newly emerging GATC sites have different methylation states. Left: Hemimethylation resulting from methylation pattern (0,1) (fully methylated $GATC^{prox}$ in the linear conformation). Middle:

Hemimethylation resulting from methylation pattern (1,0) (fully methylated $GATC^{dist}$ in the linear conformation). Right: Full methylation resulting from methylation pattern (1,1) (fully methylated $GATC^{prox}$ and $GATC^{dist}$). (C) In thermodynamic equilibrium, in the absence of binding elements, secondary structures where the cruciform forms to different extend are about equally likely to form. With such conformations, genes may (right) or may not be expressed (left).

In vitro binding experiments by Hernday *et al.*[45] showed that hemi-methylation of the $GATC^{dist}$ site has influence on the binding behavior of the coregulators Lrp and PapI. Moreover, the type of hemi-methylation also affects the binding behavior of these coregulators[45]. These observations apply to the different hemi-methylated GATC sites formed from paired nucleotides of Lrp BS 2 and 5 as part of a cruciform (fig. 6B left and middle). One may speculate that Lrp promotes expression, even in these configurations (fig. S7). In this case Lrp binding will stabilize the cruciform conformation. However, further experiments that control the conformation of the DNA to a much further extent are required to decide if indeed such conformations play significant roles.

5 CONCLUSION

Here, we constructed a synthetic pap DNA regulatory element that controls gene expression upstream and downstream as a function of two methylation sites. We monitored its gene expression and regulation in a cell-free expression system[52,53] based on the endogenous *E. coli* transcription-translation machinery. Our results excellently reproduced previous observations by Hernday and colleagues[45], showing converse regulatory effects of methylation of either $GATC^{prox}$ or $GATC^{dist}$ sites as part of the phase variation mechanism of UPEC. We confirmed an interaction between methylation of both sites. A natural interpretation of our results includes cruciform secondary structures of the pap regulatory DNA as an integral part of the two state phase variation mechanism. To confirm this idea the formation of DNA cruciforms and their role in the phase variation mechanism must be investigated in further experiments. Using magnetic tweezers and Förster Resonance Energy Transfer (FRET), the predicted pap regulatory DNA cruciforms could be checked by placing donor and acceptor pairs within the DNA[72].

Supporting Information Available

This material is available free of charge via the Internet at http://pubs.acs.org/.

Author contributions

AO and EGW designed the research with the help of MS. EGW performed the research with the help of MF. EGW and AO analyzed and interpreted data. MF and ÖK replotted data already published in the PHD thesis[73]. EGW, ÖK, MF and AO wrote the paper. VN, VH and AO supervised the research.

Acknowledgement

We acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG) within the collaborative research center (CRC/SFB 1027) as well as financial support from the Human Frontiers Science Program Organization (HFSP, RGP0037/2015). We thank Julia Flesch and Lucina Kainka for their assistance with the amplification of DNA templates. We thank Po-Hsien for fruitful discussions.

Additional note

Results were previously published as part of the PhD thesis of EGW[73].

6 LITERATURE

(1) Tian, J.; Gong, H.; Sheng, N.; Zhou, X.; Gulari, E.; Gao, X.; Church, G. Accurate Multiplex Gene Synthesis from Programmable DNA Microchips. *Nature* **2004**. https://doi.org/10.1038/nature03151.

(2) Gibson, D. G. Enzymatic Assembly of Overlapping DNA Fragments. In *Methods in Enzymology*; **2011**. https://doi.org/10.1016/B978-0-12-385120-8.00015-2.

(3) Holtz, W. J.; Keasling, J. D. Engineering Static and Dynamic Control of Synthetic Pathways. *Cell* **2010**. https://doi.org/10.1016/j.cell.2009.12.029.

(4) Zhang, F.; Carothers, J. M.; Keasling, J. D. Design of a Dynamic Sensor-Regulator System for Production of Chemicals and Fuels Derived from Fatty Acids. *Nat. Biotechnol.* **2012**. https://doi.org/10.1038/nbt.2149.

(5) Dueber, J. E.; Wu, G. C.; Malmirchegini, G. R.; Moon, T. S.; Petzold, C. J.; Ullal, A. V.; Prather, K. L. J.; Keasling, J. D. Synthetic Protein Scaffolds Provide Modular Control over Metabolic Flux. *Nat. Biotechnol.* **2009**. https://doi.org/10.1038/nbt.1557.

(6) Good, M. C.; Zalatan, J. G.; Lim, W. A. Scaffold Proteins: Hubs for Controlling the Flow of Cellular Information. *Science* **2011**. https://doi.org/10.1126/science.1198701.

(7) Huang, C. C.; Chen, M. W.; Hsieh, J. L.; Lin, W. H.; Chen, P. C.; Chien, L. F. Expression of Mercuric Reductase from *Bacillus Megaterium* MB1 in Eukaryotic Microalga *Chlorella Sp.* DT: An Approach for Mercury Phytoremediation. *Appl. Microbiol. Biotechnol.* **2006**. https://doi.org/10.1007/s00253-005-0250-0.

(8) Burgard, A. P.; Pharkya, P.; Maranas, C. D. OptKnock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization. *Biotechnol. Bioeng.* **2003**. https://doi.org/10.1002/bit.10803.

(9) Sanchez-Garcia, L.; Martín, L.; Mangues, R.; Ferrer-Miralles, N.; Vázquez, E.; Villaverde, A. Recombinant Pharmaceuticals from Microbial Cells: A 2015 Update. *Microb. Cell Fact.* **2016**. https://doi.org/10.1186/s12934-016-0437-3.

(10) Citorik, R. J.; Mimee, M.; Lu, T. K. Bacteriophage-Based Synthetic Biology for the Study of Infectious Diseases. *Curr. Opin. Microbiol.* **2014**. https://doi.org/10.1016/j.mib.2014.05.022.

(11) Saeidi, N.; Wong, C. K.; Lo, T. M.; Nguyen, H. X.; Ling, H.; Leong, S. S. J.; Poh, C. L.; Chang, M. W. Engineering Microbes to Sense and Eradicate *Pseudomonas Aeruginosa*, a Human Pathogen. *Mol. Syst. Biol.* **2011**. https://doi.org/10.1038/msb.2011.55.

(12) Worst, E. G.; Exner, M. P.; De Simone, A.; Schenkelberger, M.; Noireaux, V.; Budisa, N.; Ott, A. Cell-Free Expression with the Toxic Amino Acid Canavanine. *Bioorg. Med. Chem. Lett.* **2015**. https://doi.org/10.1016/j.bmcl.2015.06.045.

(13) Worst, E. G.; Exner, M. P.; de Simone, A.; Schenkelberger, M.; Noireaux, V.; Budisa, N.; Ott, A. Residue-Specific Incorporation of Noncanonical Amino Acids into Model Proteins Using an *Escherichia Coli* Cell-Free Transcription-Translation System. *J. Vis. Exp.* **2016**. https://doi.org/10.3791/54273.

(14) Nash, P. D. Why Modules Matter. *FEBS Lett.* **2012**. https://doi.org/10.1016/j.febslet.2012.04.049.

(15) Elowitz, M. B.; Leibler, S. A Synthetic Oscillatory Network of Transcriptional Regulators. *Nature* **2000**. https://doi.org/10.1038/35002125.

(16) Siuti, P.; Yazbek, J.; Lu, T. K. Synthetic Circuits Integrating Logic and Memory in Living Cells. *Nat. Biotechnol.* **2013**. https://doi.org/10.1038/nbt.2510.

(17) Bonnet, J.; Yin, P.; Ortiz, M. E.; Subsoontorn, P.; Endy, D. Amplifying Genetic Logic Gates. *Science* **2013**. https://doi.org/10.1126/science.1232758.

(18) Gardner, T. S.; Cantor, C. R.; Collins, J. J. Construction of a Genetic Toggle Switch in *Escherichia Coli*. *Nature* **2000**. https://doi.org/10.1038/35002131.

(19) Liu, D.; Hoynes-O'Connor, A.; Zhang, F. Bridging the Gap between Systems Biology and Synthetic Biology. *Front. Microbiol.* **2013**. https://doi.org/10.3389/fmicb.2013.00211.

(20) Finkler, M.; Ott, A. Bead-Based Assay for Spatiotemporal Gene Expression Control in Cell-Free Transcription–translation Systems. *Biotechniques* **2019**. https://doi.org/10.2144/btn-2018-0097.

(21) Romanel, A.; Jensen, L. J.; Cardelli, L.; Csikász-Nagy, A. Transcriptional Regulation Is a Major Controller of Cell Cycle Transition Dynamics. *PLoS One* **2012**. https://doi.org/10.1371/journal.pone.0029716.

(22) Lodish, H.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Baltimore, D.; Darnell, J. *Molecular Cell Biology*, 4th ed.; **2000**.

(23) Cheng, A. A.; Lu, T. K. Synthetic Biology: An Emerging Engineering Discipline. *Annu. Rev. Biomed. Eng.* **2012**. https://doi.org/10.1146/annurev-bioeng-071811-150118.

(24) Schenkelberger, M.; Shanak, S.; Finkler, M.; Worst, E. G.; Noireaux, V.; Helms, V.; Ott, A. Expression Regulation by a Methyl-CpG Binding Domain in an *E. Coli* Based, Cell-Free TX-TL System. *Phys. Biol.* **2017**. https://doi.org/10.1088/1478-3975/aa5d37.

(25) Kucukkal, T. G.; Yang, Y.; Uvarov, O.; Cao, W.; Alexov, E. Impact of Rett Syndrome Mutations on MeCP2 MBD Stability. *Biochemistry* **2015**. https://doi.org/10.1021/acs.biochem.5b00790.

(26) Casadesús, J.; Low, D. Epigenetic Gene Regulation in the Bacterial World. *Microbiol. Mol. Biol. Rev.* **2006**. https://doi.org/10.1128/MMBR.00016-06.

(27) Marinus, M. G.; Løbner-Olesen, A. DNA Methylation. *EcoSal Plus* **2013**. https://doi.org/10.1128/ecosalplus.ESP-0003-2013.

(28) Barras, F.; Marinus, M. G. The Great GATC: DNA Methylation in *E. Coli*. *Trends Genet.* **1989**. https://doi.org/10.1016/0168-9525(89)90054-1.

(29) Modrich, P. Methyl-Directed DNA Mismatch Correction. *J. Biol. Chem.* **1989**.

(30) Palmer, B. R.; Marinus, M. G. The Dam and Dcm Strains of *Escherichia Coli* - a Review. *Gene* **1994**. https://doi.org/10.1016/0378-1119(94)90597-5.

(31) Wion, D.; Casadesús, J. N^6-Methyl-Adenine: An Epigenetic Signal for DNA–protein Interactions. *Nat. Rev. Microbiol.* **2006**. https://doi.org/10.1038/nrmicro1350.

(32) Han, J. S.; Kang, S.; Lee, H.; Kim, H. K.; Hwang, D. S. Sequential Binding of SeqA to Paired Hemi-Methylated GATC Sequences Mediates Formation of Higher Order Complexes. *J. Biol. Chem.* **2003**. https://doi.org/10.1074/jbc.M304923200.

(33) Blyn, L. B.; Braaten, B. A.; White-Ziegler, C. A.; Rolfson, D. H.; Low, D. A. Phase-Variation of Pyelonephritis-Associated Pili in *Escherichia Coli*: Evidence for Transcriptional Regulation. *EMBO J.* **1989**. https://doi.org/10.1002/j.1460-2075.1989.tb03416.x.

(34) Hernday, A.; Krabbe, M.; Braaten, B.; Low, D. Self-Perpetuating Epigenetic Pili Switches in Bacteria. *Proc. Natl. Acad. Sci.* **2002**. https://doi.org/10.1073/pnas.182427199.

(35) Forsman, K.; Göransson, M.; Uhlin, B. E. Autoregulation and Multiple DNA Interactions by a Transcriptional Regulatory Protein in *E. Coli* Pili Biogenesis. *EMBO J.* **1989**. https://doi.org/10.1002/j.1460-2075.1989.tb03501.x.

(36) Båga, M.; Göransson, M.; Normark, S.; Uhlin, B. E. Transcriptional Activation of a Pap Pilus Virulence Operon from Uropathogenic *Escherichia Coli. EMBO J.* **1985**. https://doi.org/10.1002/j.1460-2075.1985.tb04162.x.

(37) Newman, E. B.; D'Ari, R.; Lin, R. T. The Leucine-Lrp Regulon in E. Coli: A Global Response in Search of a Raison d'Être. *Cell* **1992**. https://doi.org/10.1016/0092-8674(92)90135-Y.

(38) Calvo, J. M.; Matthews, R. G. The Leucine-Responsive Regulatory Protein, a Global Regulator of Metabolism in *Escherichia Coli*. *Microbiol. Rev.* **1994**.

(39) Nou, X.; Braaten, B.; Kaltenbach, L.; Low, D. A. Differential Binding of Lrp to Two Sets of Pap DNA Binding Sites Mediated by Pap I Regulates Pap Phase Variation in *Escherichia Coli*. *EMBO J.* **1995**. https://doi.org/10.1002/j.1460-2075.1995.tb00267.x.

(40) Cui, Y.; Midkiff, M. A.; Wang, Q.; Calvo, J. M. The Leucine-Responsive Regulatory Protein (Lrp) from *Escherichia Coli*. Stoichiometry and Minimal Requirements for Binding to DNA. *J. Biol. Chem.* **1996**. https://doi.org/10.1074/jbc.271.12.6611.

(41) Leonard, P. M.; Smits, S. H. J.; Sedelnikova, S. E.; Brinkman, A. B.; Vos, W. M. de; Oost, J. van der; Rice, D. W.; Rafferty, J. B. Crystal Structure of the Lrp-like Transcriptional Regulator from the Archaeon *Pyrococcus Furiosus*. *EMBO J.* **2001**. https://doi.org/10.1093/emboj/20.5.990.

(42) De los Rios, S.; Perona, J. J. Structure of the *Escherichia Coli* Leucine-Responsive Regulatory Protein Lrp Reveals a Novel Octameric Assembly. *J. Mol. Biol.* **2007**. https://doi.org/10.1016/j.jmb.2006.12.032.

(43) Kawamura, T.; Vartanian, A. S.; Zhou, H.; Dahlquist, F. W. The Design Involved in PapI and Lrp Regulation of the Pap Operon. *J. Mol. Biol.* **2011**. https://doi.org/10.1016/j.jmb.2011.01.058.

(44) Braaten, B. A.; Nou, X.; Kaltenbach, L. S.; Low, D. A. Methylation Patterns in Pap Regulatory DNA Control Pyelonephritis-Associated Pili Phase Variation in *E. Coli*. *Cell* **1994**. https://doi.org/10.1016/0092-8674(94)90120-1.

(45) Hernday, A. D.; Braaten, B. A.; Low, D. A. The Mechanism by Which DNA Adenine Methylase and PapI Activate the Pap Epigenetic Switch. *Mol. Cell* **2003**. https://doi.org/10.1016/S1097-2765(03)00383-6.

(46) Peterson, S. N.; Reich, N. O. Competitive Lrp and Dam Assembly at the Pap Regulatory Region: Implications for Mechanisms of Epigenetic Regulation. *J. Mol. Biol.* **2008**. https://doi.org/10.1016/j.jmb.2008.07.086.

(47) Peterson, S. N.; Reich, N. O. GATC Flanking Sequences Regulate Dam Activity: Evidence for How Dam Specificity May Influence Pap Expression. *J. Mol. Biol.* **2006**. https://doi.org/10.1016/j.jmb.2005.11.003.

(48) Mashhoon, N.; Carroll, M.; Pruss, C.; Eberhard, J.; Ishikawa, S.; Estabrook, R. A.; Reich, N. Functional Characterization of *Escherichia Coli* DNA Adenine Methyltransferase, a Novel Target for Antibiotics. *J. Biol. Chem.* **2004**. https://doi.org/10.1074/jbc.M408182200.

(49) Weyand, N. J.; Low, D. A. Regulation of Pap Phase Variation. Lrp Is Sufficient for the Establishment of the Phase OFF Pap DNA Methylation Pattern and Repression of Pap Transcription *in Vitro*. *J. Biol. Chem.* **2000**. https://doi.org/10.1074/jbc.275.5.3192.

(50) Göransson, M.; Forsman, K.; Nilsson, P.; Uhlin, B. E. Upstream Activating Sequences That Are Shared by Two Divergently Transcribed Operons Mediate cAMP-CRP Regulation of Pilus-Adhesin in *Escherichia Coli. Mol. Microbiol.* **1989**. https://doi.org/10.1111/j.1365-2958.1989.tb00141.x.

(51) Weyand, N. J.; Braaten, B. A.; van der Woude, M.; Tucker, J.; Low, D. A. The Essential Role of the Promoter-Proximal Subunit of CAP in Pap Phase Variation: Lrp- and Helical Phase-Dependent Activation of papBA Transcription by CAP from -215. *Mol. Microbiol.* **2008**. https://doi.org/10.1046/j.1365-2958.2001.02338.x.

(52) Shin, J.; Noireaux, V. Efficient Cell-Free Expression with the Endogenous *E. Coli* RNA Polymerase and Sigma Factor 70. *J. Biol. Eng.* **2010**. https://doi.org/10.1186/1754-1611-4-8.

(53) Sun, Z. Z.; Hayes, C. A.; Shin, J.; Caschera, F.; Murray, R. M.; Noireaux, V. Protocols for Implementing an *Escherichia Coli* Based TX-TL Cell-Free Expression System for Synthetic Biology. *J. Vis. Exp.* **2013**. https://doi.org/10.3791/50762.

(54) SantaLucia Jr., J. A Unified View of Polymer, Dumbbell, and Oligonucleotide DNA Nearest-Neighbor Thermodynamics. *Proc. Natl. Acad. Sci.* **1998**. https://doi.org/10.1073/pnas.95.4.1460.

(55) Owczarzy, R.; You, Y.; Moreira, B. G.; Manthey, J. A.; Huang, L.; Behlke, M. A.; Walder, J. A. Effects of Sodium Ions on DNA Duplex Oligomers: Improved Predictions of Melting Temperatures. *Biochemistry* **2004**. https://doi.org/10.1021/bi034621r.

(56) Sussman, R.; Jacob, F. Sur Un Système de Répression Thermosensible Chez Le Bactériophage d'*Escherichia Coli*. *C. R. Hebd. Seances Acad. Sci.* **1962**.

(57) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison III, C. A.; Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **2009**. https://doi.org/10.1038/nmeth.1318.

(58) Gibson, D. G.; Smith, H. O.; Hutchison III, C. A.; Venter, J. C.; Merryman, C. Chemical Synthesis of the Mouse Mitochondrial Genome. *Nat. Methods* **2010**. https://doi.org/10.1038/nmeth.1515.

(59) De Boer, H. A.; Comstock, L. J.; Vasser, M. The Tac Promoter: A Functional Hybrid Derived from the Trp and Lac Promoters. *Proc. Natl. Acad. Sci.* **1983**. https://doi.org/10.1073/pnas.80.1.21.

(60) Olins, P. O.; Devine, C. S.; Rangwala, S. H.; Kavka, K. S. The T7 Phage Gene 10 Leader RNA, a Ribosome-Binding Site That Dramatically Enhances the Expression of Foreign Genes in *Escherichia Coli*. *Gene* **1988**. https://doi.org/10.1016/0378-1119(88)90329-0.

(61) Larson, M. H.; Greenleaf, W. J.; Landick, R.; Block, S. M. Applied Force Reveals Mechanistic and Energetic Details of Transcription Termination. *Cell* **2008**. https://doi.org/10.1016/j.cell.2008.01.027.

(62) Li, X.; Zhang, G.; Ngo, N.; Zhao, X.; Kain, S. R.; Huang, C. C. Deletions of the *Aequorea Victoria* Green Fluorescent Protein Define the Minimal Domain Required for Fluorescence. *J. Biol. Chem.* **1997**. https://doi.org/10.1074/jbc.272.45.28545.

(63) Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y. Improved Monomeric Red, Orange and Yellow Fluorescent Proteins Derived from *Discosoma Sp.* Red Fluorescent Protein. *Nat. Biotechnol.* **2004**. https://doi.org/10.1038/nbt1037.

(64) Schenkelberger, M. Cooperative Biomolecular Binding High Specificity of Competitive Single Stranded DNA, Universität des Saarlandes, **2013**.

(65) Murphy, K. C. The λ Gam Protein Inhibits RecBCD Binding to dsDNA Ends. *J. Mol. Biol.* **2007**. https://doi.org/10.1016/j.jmb.2007.05.085.

(66) Zuker, M. Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction. *Nucleic Acids Res.* **2003**. https://doi.org/10.1093/nar/gkg595.

(67) Schoen, I.; Krammer, H.; Braun, D. Hybridization Kinetics Is Different inside Cells. *Proc. Natl. Acad. Sci.* **2009**. https://doi.org/10.1073/pnas.0901313106.

(68) De Rijk, P.; Wuyts, J.; De Wachter, R. RnaViz 2: An Improved Representation of RNA Secondary Structure. *Bioinformatics* **2003**. https://doi.org/10.1093/bioinformatics/19.2.299.

(69) Kaltenbach, L. S.; Braaten, B. A.; Low, D. A. Specific Binding of PapI to Lrp-Pap DNA Complexes. *J. Bacteriol.* **1995**. https://doi.org/10.1128/JB.177.22.6449-6455.1995.

(70) Brázda, V.; Laister, R. C.; Jagelská, E. B.; Arrowsmith, C. Cruciform Structures Are a Common DNA Feature Important for Regulating Biological Processes. *BMC Mol. Biol.* **2011**. https://doi.org/10.1186/1471-2199-12-33.

(71) Bae, S.-H.; Cheong, H.-K.; Cheong, C.; Kang, S.; Hwang, D. S.; Choi, B.-S. Structure and Dynamics of Hemimethylated GATC Sites. Implications for DNA-SeqA Recognition. *J. Biol. Chem.* **2003**. https://doi.org/10.1074/jbc.M306038200.

(72) Swoboda, M.; Grieb, M. S.; Hahn, S.; Schlierf, M. Measuring Two at the Same Time: Combining Magnetic Tweezers with Single-Molecule FRET. In *Fluorescent Methods for Molecular Motors. Experientia Supplementum*; 2014. https://doi.org/10.1007/978-3-0348-0856-9_12.

(73) Worst, E. G. Evolution Wachsender, Sich Reproduzierender Polymere. Zellfreie Genexpression Zum Einbau Nicht-Kanonischer Aminosäuren in Proteine Sowie Zur Analyse Des Epigenetischen *Escherichia-Coli*-Pili-Phasenvariationsmechanismus., Universität des Saarlandes, **2017**.

SUPPLEMENTARY INFORMATION

DNA sequences

0.1 Relevant DNA sequences of the transcription-translation-regulatory and coding elements of the plasmids pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 and pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500

(A)

```
1   TGAGC TAACA CCGTG CGTGT TGACA ATTTT ACCTC TGGCG GTGAT AATGG TTGCA GCTAG
61  CAATA ATTTT GTTTA ACTTT AAGAA GGAGA TATAC CATGG TAGAT AGCAA GAAGC GCCCT
121 GGCAA AGATC TCGAC CGTAT CGATC GTAAC ATTCT TAATG AGTTG CAAAA GGATG GGCGT
181 ATTTC TAACG TCGAG CTTTC TAAAC GTGTG GGACT TTCCC CAACG CCGTG CCTTG AGCGT
241 GTGCG TCGGC TGGAA AGACA AGGGT TTATT CAGGG CTATA CGGCG CTGCT TAACC CCCAT
301 TATCT GGATG CATCA CTTCT GGTAT TCGTT GAGAT TACTC TGAAT CGTGG CGCAC CGGAT
361 GTGTT TGAAC AATTC AATAC CGCTG TACAA AAACT TGAAG AAATT CAGGA GTGTC ATTTA
421 GTATC CGGTG ATTTC GACTA CCTGT TGAAA ACACG CGTGC CGGAT ATGTC AGCCT ACCGT
481 AAGTT GCTGG GGGAA ACCCT GCTGC GTCTG CCTGG CGTCA ATGAC ACACG GACAT ACGTT
541 GTTAT GGAAG AAGTC AAGCA GAGTA ATCGT CTGGT TATTA AGACG CGCTA ACTCG AGCAA
601 AGCCC GCCGA AAGGC GGGCT TTTCT GT
```

(B)

```
1   TGAGC TAACA CCGTG CGTGT TGACA ATTTT ACCTC TGGCG GTGAT AATGG TTGCA GCTAG
61  CAATA ATTTT GTTTA ACTTT AAGAA GGAGG ATCCA AATGA GTGAA TATAT GAAGA ATGAA
121 ATACT GGAAT TTCTG AACAG GCATG ATGGA GGGAA AACGG CAGAA ATTGC AGAGG CGCTG
181 GCGGT AACGG ATTAC CAGGC CCGCT ATTAC CTGTT ATTAC TGGAA AAAGC AGGTA TGGTT
241 CAGCG CTCAC CATTA AGACG GGGCA TGGCA ACATA CTGGT TTCTG AAAGG AGAGA AGCAG
301 GCGGG GCAGA GCTGT TCTTC CACAA CTTAA CTCGA GAGTC GACCA AAGCC CGCCG AAAGG
361 CGGGC TTTTC TGT
```

Figure S1. Relevant DNA sequences of the transcription-translation-regulatory and coding elements of the plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-Lrp-T500 (A) and plasmid pBEST-p15a-OR2-OR1-Pr-UTR1-PapI-T500 (B) for the cell-free expression of the genes *lrp* or *papI*, respectively. (A, B) Turquoise: OR2-Or1-Pr with a single mutation. Grey: transcription terminator T500. (A) Purple: *lrp* gene. Yellow: UTR1. (B) Purple: *papI* gene. Yellow: modified UTR1. Details related to transcription-translation-regulatory and coding elements can be found in section 2.2.

0.2 Relevant DNA sequences of the transcription-translation-regulatory and coding elements of the DNA template used for the PCR synthesis of partial DNA fragment B

```
1    GGTAT ATCTC CTTCT TAAAG TTAAA CAAAA TTATT GCTAG CTGCA ACCAT TATGC GAAAA
61   GAAAG TCCGT ATGTC AACAT TTAGA CGATC TTTTA TGCTG TAAAT TCAAT TTGCC ATGAT
121  GTTTT TATCT GAGTA CCCTC TTGCT ATTAG TGTTT TGTTC TAGTT TAATT TTGTT TTGTG
181  GGTTA AAAGA TCGTT TAAAT CTTGA CATAC AACAT AAAAA ACTAA ATAAT GGTTG CAGGA
241  TCCAA TAATT TTGTT TAACT TTAAG AAGGA GATAT ACCAT GGAGC TTTTC ACTGG CGTTG
301  TTCCC ATCCT GGTCG AGCTG GACGG CGACG TAAAC GGCCA CAAGT TCAGC GTGTC CGGCG
361  AGGGC GAGGG CGATG CCACC TACGG CAAGC TGACC CTGAA GTTCA TCTGC ACCAC CGGCA
421  AGCTG CCCGT GCCCT GGCCC ACCCT CGTGA CCACC CTGAC CTACG GCGTG CAGTG CTTCA
481  GCCGC TACCC CGACC ACATG AAGCA GCACG ACTTC TTCAA GTCCG CCATG CCCGA AGGCT
541  ACGTC CAGGA GCGCA CCATC TTCTT CAAGG ACGAC GGCAA CTACA AGACC CGCGC CGAGG
601  TGAAG TTCGA GGGCG ACACC CTGGT GAACC GCATC GAGCT GAAGG GCATC GACTT CAAGG
661  AGGAC GGCAA CATCC TGGGG CACAA GCTGG AGTAC AACTA CAACA GCCAC AACGT CTATA
721  TCATG GCCGA CAAGC AGAAG AACGG CATCA AGGTG AACTT CAAGA TCCGC CACAA CATCG
781  AGGAC GGCAG CGTGC AGCTC GCCGA CCACT ACCAG CAGAA CACCC CCATC GGCGA CGGCC
841  CCGTG CTGCT GCCCG ACAAC CACTA CCTGA GCACC CAGTC CGCCC TGAGC AAAGA CCCCA
901  ACGAG AAGCG CGATC ACATG GTCCT GCTGG AGTTC GTGAC CGCCG CCGGG ATCTA ACTCG
961  AGCAA AGCCC GCCGA AAGGC GGGCT TTTCT GTGTC GACCG ATGCC CTTGA GAGGG TTTTC
1021 TGGCT GATGG TG
```

Figure S2. Relevant DNA sequences of the transcription-translation-regulatory and coding elements of the DNA template used for the PCR synthesis of partial DNA fragment B. Yellow: UTR1. Blue: -10 elements. Turquoise: -35 elements. Grey: Lrp binding sites 4 – 6 and 1 – 3 according to Hernday *et al.*[45]. Green: *degfp* gene. Black: transcription terminator T500. Details related to transcription-translation-regulatory and coding elements can be found in paragraph 2.2.

0.3 DNA sequence of the linear synthetic pap regulatory DNA products pap(1,1), pap(0,1) and pap(1,0)

```
1    CGCTA TGAAT TCCTC GAGTT ACTTG TACAG CTCGT CCATG CCGCC GGTGG AGTGG CGGCC
61   CTCGG CGCGT TCGTA CTGTT CCACG ATGGT GTAGT CCTCG TTGTG GGAGG TGATG TCCAA
121  CTTGA TGTTG ACGTT GTAGG CGCCG GGCAG CTGCA CGGGC TTCTT GGCCT TGTAG GTGGT
181  CTTGA CCTCA GCGTC GTAGT GGCCG CCGTC CTTCA GCTTC AGCCT CTGCT TGATC TCGCC
241  CTTCA GGGCG CCGTC CTCGG GGTAC ATCCG CTCGG AGGAG GCCTC CCAGC CCATG GTCTT
301  CTTCT GCATT ACGGG GCCGT CGGAG GGGAA GTTGG TGCCG CGCAG CTTCA CCTTG TAGAT
361  GAACT CGCCG TCCTG CAGGG AGGAG TCCTG GGTCA CGGTC ACCAC GCCGC CGTCC TCGAA
421  GTTCA TCACG CGCTC CCACT TGAAG CCCTC GGGGA AGGAC AGCTT CAAGT AGTCG GGGAT
481  GTCGG CGGGG TGCTT CACGT AGGCC TTGGA GCCGT ACATG AACTG AGGGG ACAGG ATGTC
541  CCAGG CGAAG GGCAG GGGGC CACCC TTGGT CACCT TCAGC TTGGC GGTCT GGGTG CCCTC
601  GTAGG GGCGG CCCTC GCCCT CGCCC TCGAT CTCGA ACTCG TGGCC GTTCA CGGAG CCCTC
661  CATGT GCACC TTGAA GCGCA TGAAC TCCTT GATGA TGGCC ATGTT ATCCT CCTCG CCCTT
721  GCTCA CCATG GTATA TCTCC TTCTT AAAGT TAAAC AAAAT TATTG CTAGC TGCAA CCATT
781  ATGCG AAAAG AAAGT CCGTA TGTCA ACATT TAGAC GATCT TTTAT GCTGT AAATT CAATT
841  TGCCA TGATG TTTTT ATCTG AGTAC CCTCT TGCTA TTAGT GTTTT GTTCT AGTTT AATTT
901  TGTTT TGTGG GTTAA AAGAT CGTTT AAATC TTGAC ATACA ACATA AAAAA CTAAA TAATG
961  GTTGC AGGAT CCAAT AATTT TGTTT AACTT TAAGA AGGAG ATATA CCATG GAGCT TTTCA
1021 CTGGC GTTGT TCCCA TCCTG GTCGA GCTGG ACGGC GACGT AAACG GCCAC AAGTT CAGCG
1081 TGTCC GGCGA GGGCG AGGGC GATGC CACCT ACGGC AAGCT GACCC TGAAG TTCAT CTGCA
1141 CCACC GGCAA GCTGC CCGTG CCCTG GCCCA CCCTC GTGAC CACCC TGACC TACGG CGTGC
1201 AGTGC TTCAG CCGCT ACCCC GACCA CATGA AGCAG CACGA CTTCT TCAAG TCCGC CATGC
1261 CCGAA GGCTA CGTCC AGGAG CGCAC CATCT TCTTC AAGGA CGACG GCAAC TACAA GACCC
1321 GCGCC GAGGT GAAGT TCGAG GGCGA CACCC TGGTG AACCG CATCG AGCTG AAGGG CATCG
1381 ACTTC AAGGA GGACG GCAAC ATCCT GGGGC ACAAG CTGGA GTACA ACTAC AACAG CCACA
1441 ACGTC TATAT CATGG CCGAC AAGCA GAAGA ACGGC ATCAA GGTGA ACTTC AAGAT CCGCC
1501 ACAAC ATCGA GGACG GCAGC GTGCA GCTCG CCGAC CACTA CCAGC AGAAC ACCCC CATCG
1561 GCGAC GGCCC CGTGC TGCTG CCCGA CAACC ACTAC CTGAG CACCC AGTCC GCCCT GAGCA
1621 AAGAC CCCAA CGAGA AGCGC GATCA CATGG TCCTG CTGGA GTTCG TGACC GCCGC CGGGA
1681 TCTAA AAGCT TCCAT GG
```

Figure S3. DNA sequence of the linear synthetic pap regulatory DNA products pap(1,1). The DNA sequence of pap(0,1) and pap(1,0) differ by the single mutation of the base adenine to cytosine at position 817 or 919 respectively according to the idea of Braaten *et al.*[44]. Red: *mcherry* gene. Yellow

(732–764): UTR1. Blue (778–782): -10 element. Grey (783–800): Lrp BS 4. Turquoise (801–806): -35 element. Grey (813–830): Lrp BS 5. Grey (844–861): Lrp BS 6. Grey (876–893): Lrp BS 1. Grey (907–924): Lrp BS 2. Turquoise (931–936): -35 element. Grey (937–954): Lrp BS 3. Blue (955–959): -10 element. Yellow (973–1005): UTR1. Green (1008–1685): *degfp* gene. $GATC^{dist}$: 816–819. $GATC^{prox}$: 918–921. Further GATC sites: 232–235, 628–631, 968–971, 1493–1496, 1641–1644, 1679–1682. The DNA sequence of the synthetic pap regulatory region (778–959) corresponds the native sequence (GenBank: X14471.1, based on Blyn *et al.*[33]. except the insertion of the -35 and -10 elements flanking the Lrp binding sites 4 and 3 (778-782: ATTAT instead of TTTCT. 801-806: TGTCAA instead of AAAATT. 931–936: TTGACA instead of AATATT. 955–959: ATAAT instead of ATTTA). All six Lrp BS according to Hernday *et al.*[45]. Details related to transcription-translation-regulatory and coding elements can be found in paragraph 2.2.

0.4 Production of the partial fragments A1 and B1 for assembly of the pap regulatory region pap(1,1) via Gibson assembly reaction

We used the linear synthetic pap regulatory DNA with the sequence shown in fig. S3 as PCR template in order to produce the two alternative partial fragments A1 and B1. The forward primer with the DNA sequence 5‘-CGCTATGAATTCCTCGAGTT-3‘ and reverse primers with the DNA sequence 5‘-AGATAAAAACATCATGGCAAATTG-3‘ were used for the synthesis of A1 in a standard PCR. The forward primer with the DNA sequence 5‘-TTTGCCATGATGTTTTTATCTGA -3‘ and reverse primers with the DNA sequence 5‘-CCATGGAAGCTTTTAGATCC -3‘ were used for the synthesis of B1 in a standard PCR. Assembling these two partial fragments via Gibson assembly reaction as described in paragraph 2.3.1 and subsequent PCR amplification, led again to the regulatory pap DNA construct pap (1,1).

0.5 Kinetic fluorescence measurements. Methylation-dependent regulation of the pap regulatory DNA without cell-free synthesis of Lrp and PapI

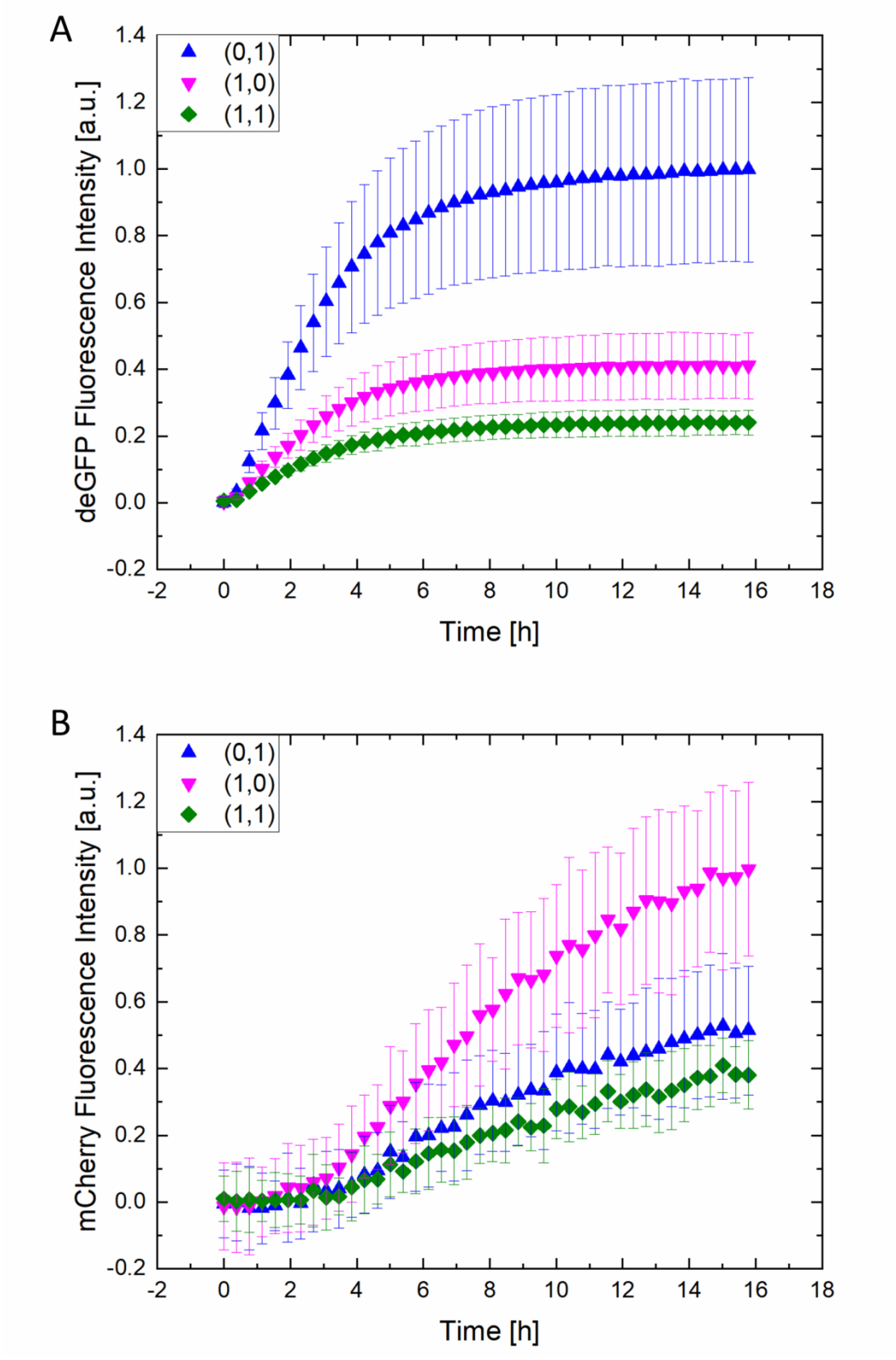


Figure S4. Kinetic measurements of the fluorescence intensities of deGFP (A) and mCherry (B) as a function of time for the different methylation states of the pap regulatory DNA construct (5nM). Symbols refer to the different methylation states (see insert). (0,1): Lrp BS 5 unmethylated, Lrp BS 2 methylated. (1,0): Lrp BS 5 methylated, Lrp BS 2 unmethylated. (1,1): Lrp BS 2 and 5 are both

methylated. Fluorescence intensities were normalized to the highest occurring value. Error bars represent one standard deviation from the mean as gained from five measurements. To simplify reading, only every seventh data point of each kinetic measurement is visualized.

0.6 Results from endpoint fluorescence measurements in the presence of Lrp and PapI plasmids.

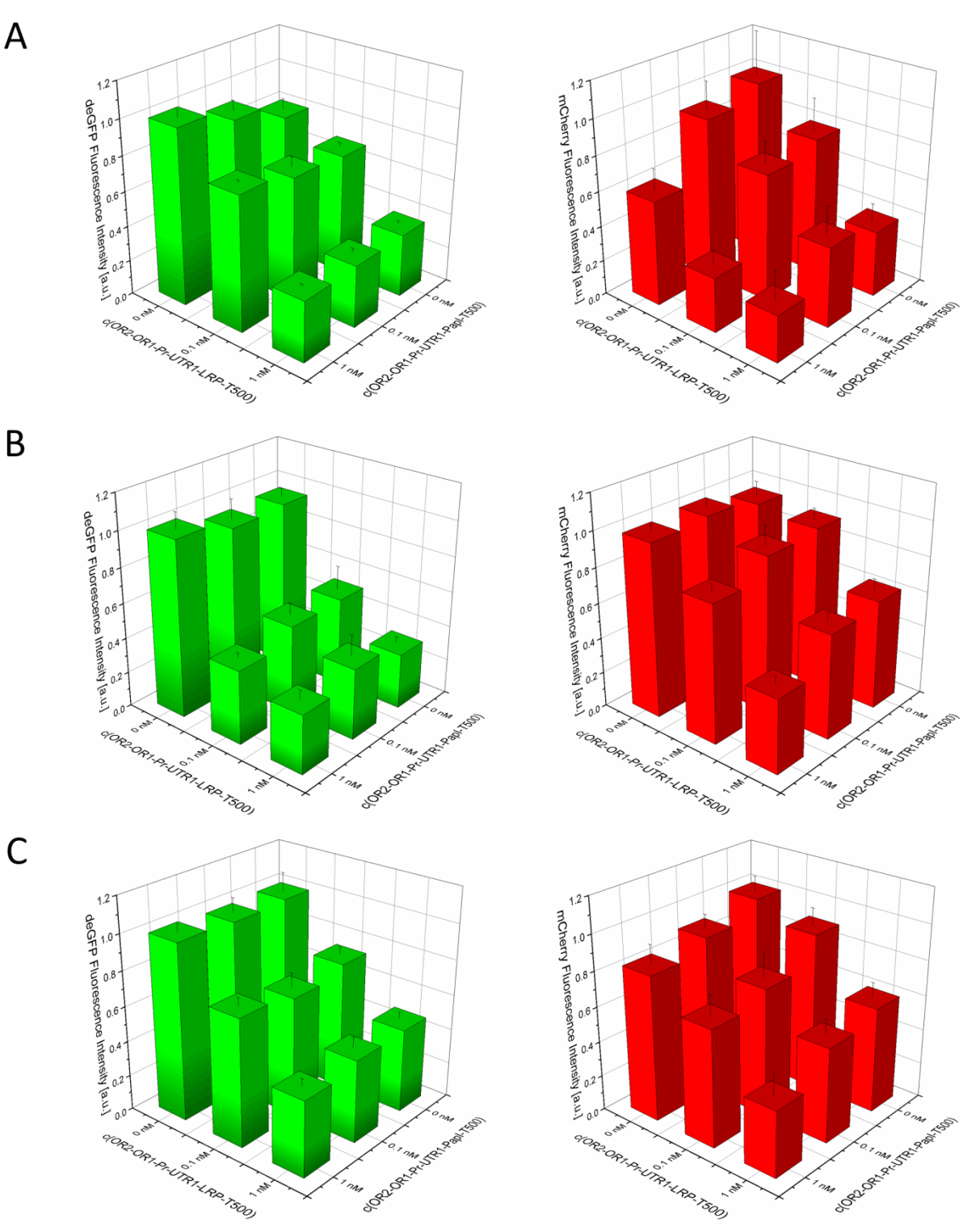
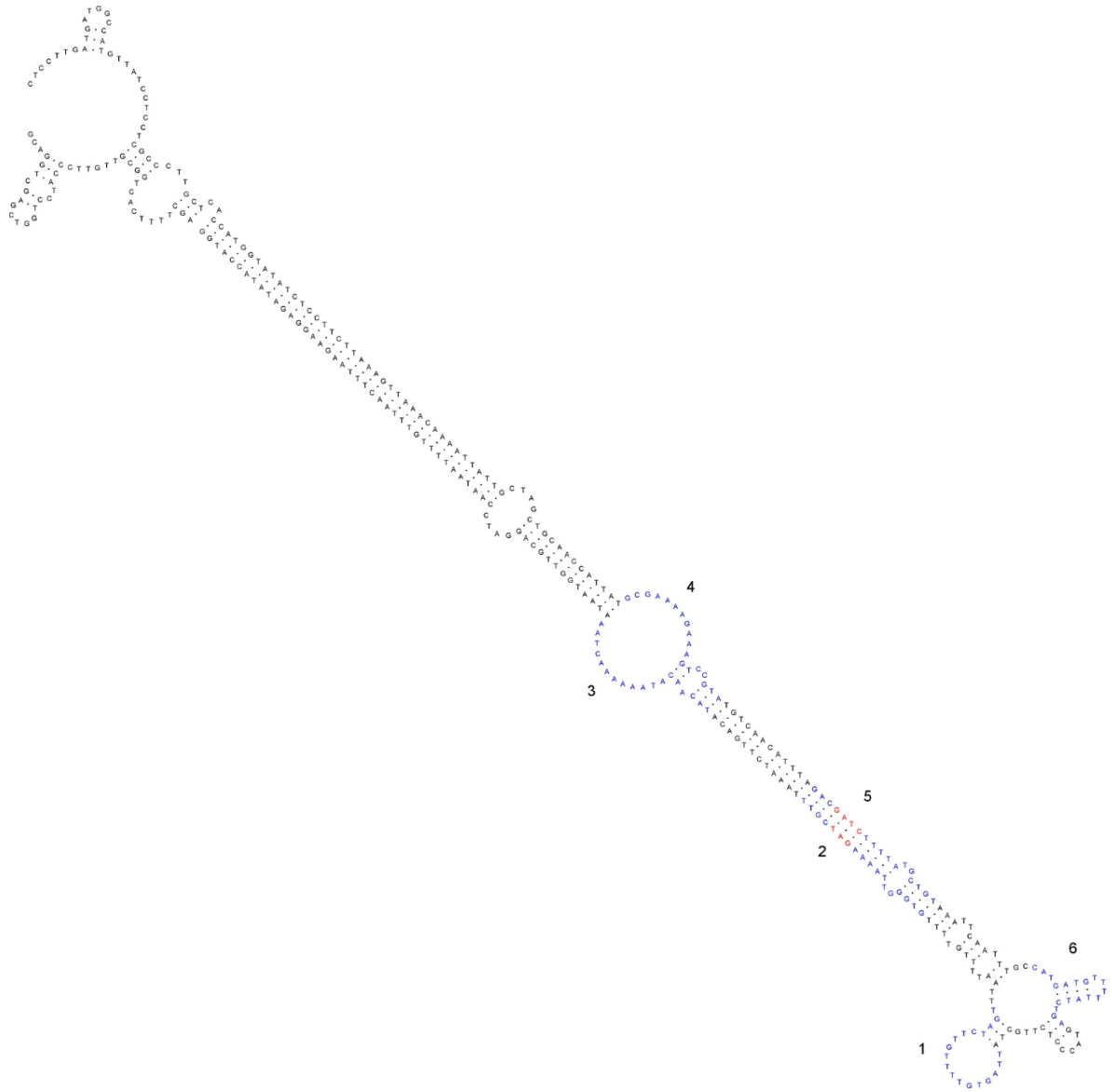


Figure S5. Fluorescence intensities (vertical axis) of deGFP (green bars), mCherry (red bars) in the presence of pap regulatory DNA constructs with methylation states (0,1) in (A), (1,0) in (B) and (1,1)

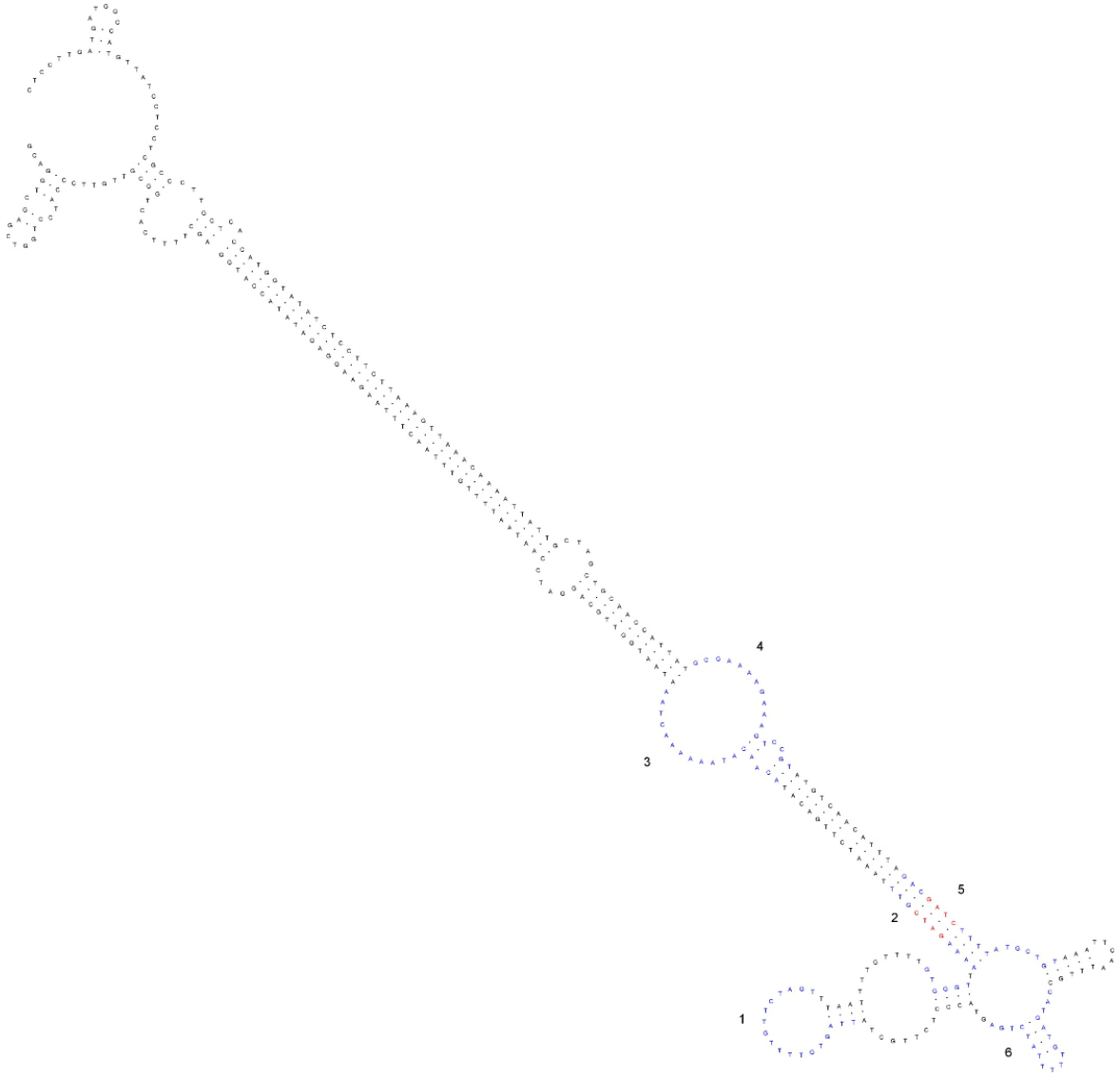
in (C) as a function of *lrp*- and *papI*-coding plasmid concentrations on the horizontal axes. (0,1), (1,0) and (1,1) refer to the different methylation states of the DNA construct that were present in the cell-free reactions. (0,1): Lrp BS 5 unmethylated, Lrp BS 2 methylated. (1,0): Lrp BS 5 methylated, Lrp BS 2 unmethylated. (1,1): Lrp BS 2 and 5 are both methylated. Fluorescence intensities were normalized to the highest occurring value. Error bars represent one standard deviation from the mean gained from two measurements.

0.7 Secondary structures of the synthetic pap regulatory DNA

A



B



C

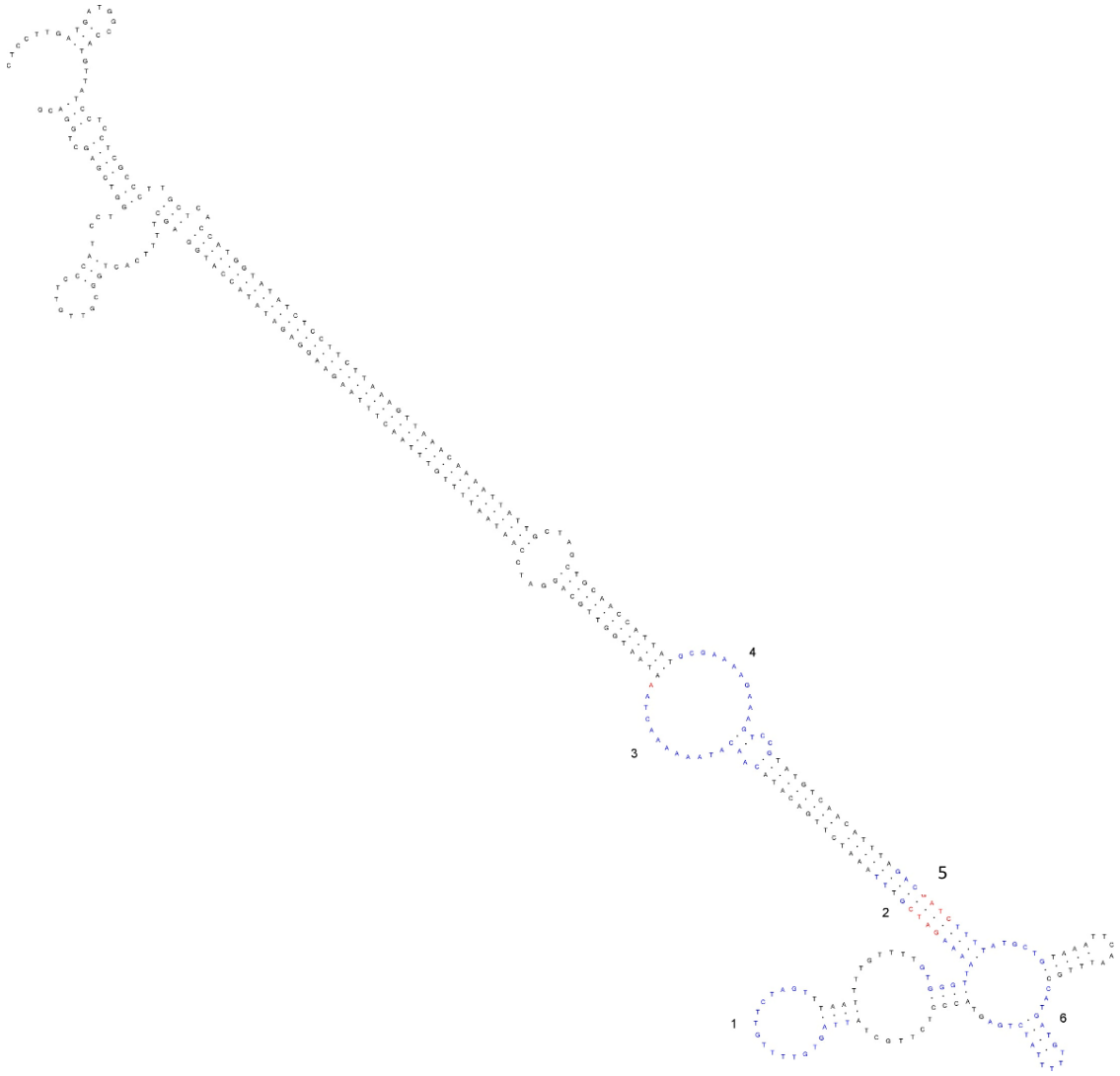


Figure S6. Three most important secondary structures of synthetic pap regulatory DNA as predicted by the mfold web server[66] (default settings, 29 °C, see paragraph 2.6). The six Lrp BS are highlighted in colour. The figure was generated with the program Rna Viz[68].

0.8 Hypothetical conformational changes of the pap DNA cruciform structure due to cooperative Lrp binding

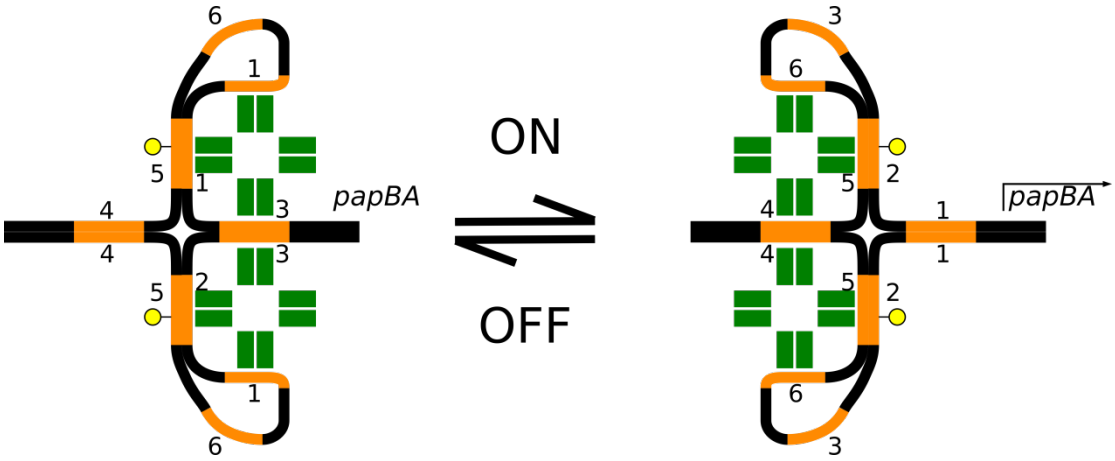


Figure S7. Hypothetical conformational changes of the pap DNA cruciform structure with cooperative binding of the Lrp as an octamer (green rectangles) as a function of the methylation state of the pap DNA. Methylation of a single stranded DNA motif is represented by a yellow dot. Methylation state (1,0) leads to hemi-methylated cruciform structure that prevents cooperative Lrp binding on the left sub-unit of the pap DNA, while Lrp can cooperatively bind to the right sub-unit of the pap DNA, blocking RNA polymerase binding to the *papBA* promotor and leading to "OFF" phase. Methylation state (1,0) leads to hemi-methylated cruciform structure that prevents cooperative Lrp binding on the right sub-unit of the pap DNA, while Lrp can cooperatively bind to the left sub-unit of the pap DNA, enabling RNA polymerase binding to the *papBA* promotor and resulting in the "ON" phase.