

# Assessing the biocompatibility of click-linked DNA in *Escherichia coli*

A. Pia Sanzone<sup>1</sup>, Afaf H. El-Sagheer<sup>1,2</sup>, Tom Brown<sup>1</sup> and Ali Tavassoli<sup>1,3,\*</sup>

<sup>1</sup>Chemistry, University of Southampton, Southampton, SO17 1BJ, UK, <sup>2</sup>Department of Science and Mathematics, Chemistry Branch, Faculty of Petroleum and Mining Engineering, Suez Canal University, Suez, 43721, Egypt and <sup>3</sup>Cancer Research UK Centre, Faculty of Medicine, University of Southampton, Southampton, SO16 6YD, UK

Received May 10, 2012; Revised July 14, 2012; Accepted July 16, 2012

## ABSTRACT

The biocompatibility of a triazole mimic of the DNA phosphodiester linkage in *Escherichia coli* has been evaluated. The requirement for selective pressure on the click-containing gene was probed via a plasmid containing click DNA backbone linkages in each strand of the gene encoding the fluorescent protein mCherry. The effect of proximity of the click linkers on their biocompatibility was also probed by placing two click DNA linkers 4-bp apart at the region encoding the fluorophore of the fluorescent protein. The resulting click-containing plasmid was found to encode mCherry in *E. coli* at a similar level to the canonical equivalent. The ability of the cellular machinery to read through click-linked DNA was further probed by using the above click-linked plasmid to express mCherry using an *in vitro* transcription/translation system, and found to also be similar to that from canonical DNA. The yield and fluorescence of recombinant mCherry expressed from the click-linked plasmid was also compared to that from the canonical equivalent, and found to be the same. The biocompatibility of click DNA ligation sites at close proximity in a non-essential gene demonstrated in *E. coli* suggests the possibility of using click DNA ligation for the enzyme-free assembly of chemically modified genes and genomes.

## INTRODUCTION

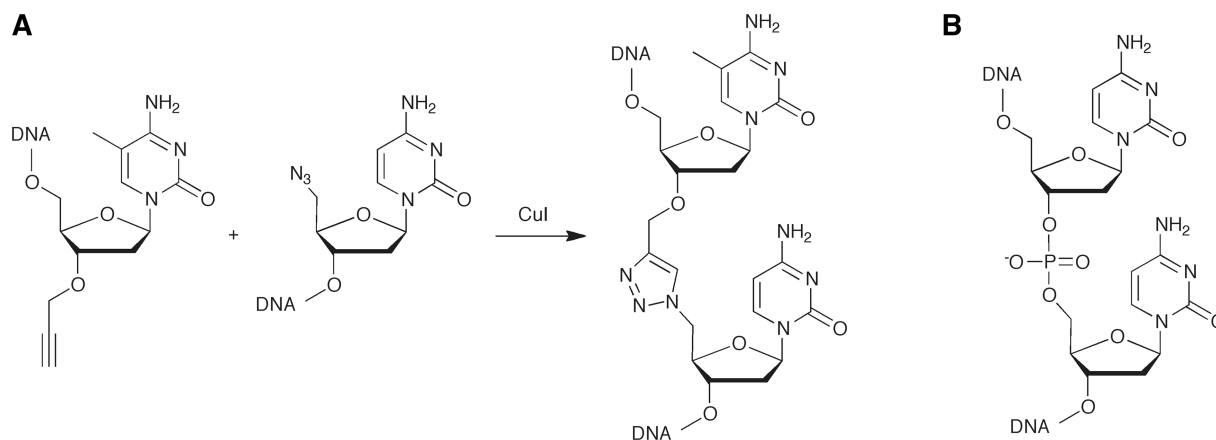
Current methods for the assembly of large DNA molecules (e.g. genes) utilize a mixture of oligonucleotide synthesis, PCR amplification and enzymatic ligation.(1,2) Although extremely powerful (3,4), this approach has several weaknesses, including the additional time and resources required for the enzymatic steps, and the

inability to incorporate epigenetic information, or modified bases into the whole genes and genomes. An alternative approach to DNA assembly may be envisaged that instead of enzymes, uses highly efficient chemical reactions for the ligation of oligonucleotides (5–7). Such an approach would not only eliminate the need for enzymatic ligation and cloning during gene synthesis to enable the full automation of large-scale gene synthesis, but also readily allow the incorporation of modified bases into large DNA fragments. The resulting click-linked DNA will however, contain an unnatural triazole linkage on its backbone at the sites of ligation (in place of the canonical phosphodiester linker). This enzyme-free approach to DNA ligation will therefore only be of use if the unnatural backbone linkage does not adversely affect the rate or fidelity of DNA replication and transcription, thus functioning normally in biological systems.

To this end, we recently reported (8) the rapid, clean and efficient chemical ligation of DNA oligonucleotides using the copper-catalysed alkyne-azide cycloaddition (CuAAC) reaction (Figure 1) (9,10). The DNA sequence around the linker was amplified by PCR with high fidelity, and a TEM-1  $\beta$ -lactamase gene (BLA) encoding ampicillin resistance containing two click linkers (83-bp apart, one on each strand), was found to be functional in *E. coli* at similar levels to the equivalent plasmid with a canonical DNA backbone. The click DNA backbone linker was also shown to be functional in a UvrB deficient strain of *E. coli*, therefore suggesting that nucleotide excision repair does not play a key role in the observed biocompatibility (UvrB is the central component of nucleotide excision repair) (8). Further work has shown that T7 RNA polymerase also reads through this linker *in vitro*, although the effect of the click-linker on the fidelity and rate of this process has not been thoroughly evaluated (11).

The selective pressure exerted on the organism by ampicillin in our previously reported experiment (8) may have pre-selected for colonies that correctly read through the click-linker, possibly masking the true fidelity of the

\*To whom correspondence should be addressed. Tel: +44 2380 592395; Fax: +44 2380 593131; Email: a.tavassoli@soton.ac.uk



**Figure 1.** (A) Click DNA ligation. The CuAAC reaction is used in combination with alkyne and azide-modified DNA bases to join two oligonucleotides together. (B) The structure of the equivalent canonical DNA for comparison of the triazole backbone with the natural phosphodiester.

process; errors in replication or transcription of the click-modified BLA gene that result in a non-functional  $\beta$ -lactamase would subsequently prevent the survival of the host organism on ampicillin containing media, and therefore not be observed. This might account for the high rates of biocompatibility previously observed with the click-containing BLA gene. With a view to using click DNA ligation for the synthesis of a wide variety of genes, we sought to re-assess the biocompatibility of click-linked DNA when incorporated into the non-essential gene encoding the red-fluorescent protein mCherry (12). As the encoded fluorescent protein does not exert selective pressure on the bacterial host, the true fidelity of click-DNA replication and transcription will be determined. We also sought to assess the effect of click DNA linker proximity on biocompatibility; the click linkers in our previous study were 83-bp apart, potentially minimizing any distortive effects on the double helix. Recent structural data has shown that when joining two thymine residues, the click DNA backbone linker causes thermodynamic destabilization of the duplex, spread over the 4–5 bp surrounding the incorporated triazole (13). Here, we utilize site-directed mutagenesis (SDM) for the construction of a plasmid containing two click DNA linkers, 4-bp apart, in the gene that encodes the fluorescent protein mCherry. This click-linked plasmid is used to demonstrate the full biocompatibility of click ligated DNA in a non-essential gene *in vitro* and in *E. coli*.

## MATERIALS AND METHODS

### General method of oligonucleotide synthesis and purification

Standard DNA phosphoramidites, solid supports and additional reagents were purchased from Link Technologies and Applied Biosystems Ltd. All oligonucleotides were synthesized on an Applied Biosystems 394 automated DNA/RNA synthesizer using a standard 0.2 or 1.0  $\mu$ mol phosphoramidite cycle of acid-catalysed detritylation, coupling, capping and iodine oxidation. Stepwise coupling efficiencies and overall yields were determined

by the automated trityl cation conductivity monitoring facility and in all cases were >98.0%. All  $\beta$ -cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C and T monomers was 25 s. Cleavage of the oligonucleotides from the solid support and deprotection was achieved by exposure to concentrated aqueous ammonia solution for 60 min at room temperature followed by heating in a sealed tube for 5 h at 55°C. The oligonucleotides were purified by reversed-phase HPLC on a Gilson system using an XBridge<sup>TM</sup> BEH300 Prep C18 10  $\mu$ M 10  $\times$  250 mm column (Waters) with a gradient of acetonitrile in ammonium acetate (0% to 50% buffer B over 30 min, flow rate 4 ml/min), buffer A: 0.1 M ammonium acetate, pH 7.0, buffer B: 0.1 M ammonium acetate, pH 7.0, with 50% acetonitrile. Elution was monitored by UV absorption at 295 nm. After HPLC purification, oligonucleotides were desalted using NAP-10 columns and analysed by gel electrophoresis.

The sequences of the SDM primers were as shown in Figure 2A; the unmodified primers were synthesized by Eurofins MWG Operon (Ebersberg, Germany) and the click-linked primers were synthesized as outlined below. The template plasmid for SDM was pRSET-mCherry, kindly provided by the Tsien laboratory. Plasmids were isolated and purified using QIAprep Spin Minipreps Kit (QIAGEN, Cat No. 27106) following the manufacturer's instructions. All DNA sequencing was carried out by Eurofins MWG Operon (Ebersberg, Germany). Antibodies were purchased from GE Healthcare.

### Synthesis of 3'-alkyne and 5'-azide oligonucleotides

#### (i) Synthesis of the 3'-alkyne oligonucleotides ODN-1 and ODN-2

3'-Alkyne oligos were synthesized by the attachment of the 5'-O-(4,4'-dimethoxytrityl) –3'-O-propargyl-5-methyldeoxycytidine on solid support according to the previously published method (8) and packing the resin into a twist column (Glen Research) then assembling the

required sequence in the 3'- to 5'-direction by normal standard phosphoramidite synthesis.

#### (ii) Synthesis of the 5'-azide oligonucleotides ODN-3 and ODN-4

Oligonucleotides were assembled on the 1.0 or 0.2  $\mu\text{mol}$  scale (trityl-off) as described above in the general method with a normal 5'-hydroxy group on dC. To convert 5'-hydroxyl to 5'-iodo on the resin, the protected oligonucleotide attached to the synthesis column was treated with a 0.5 M solution of methyltriphenoxyphosphonium iodide in DMF (1.0 ml) which was periodically passed through the column *via* two 1 ml syringes over 15 min at room temperature. The column was then washed several times with dry DMF (14). To convert the 5'-iodo dC to 5'-azido dC, sodium azide (50 mg) was suspended in dry DMF (1 ml), heated for 10 min at 70°C then cooled down and the supernatant was taken into a 1 ml syringe and passed back and forth through the column then left at room temperature overnight. The column was then washed with DMF and acetonitrile and dried by passing a stream of argon gas through it (15). Resultant 5'-azide oligonucleotides were cleaved from the solid support, deprotected and purified by HPLC as explained above in the general method of oligonucleotide synthesis and purification.

#### Templated click ligation

The alkyne, azide and splint oligonucleotides (30.0 nmol of each) in 0.2 M NaCl (400  $\mu\text{l}$ ) were annealed by heating at 80°C for 5 min and cooled down slowly to room temperature. A solution of Cu<sup>I</sup> click catalyst was prepared from *tris*-hydroxypropyltriazole ligand (22) (2.1  $\mu\text{mol}$  in 0.2 M NaCl, 91.0  $\mu\text{l}$ ), sodium ascorbate (3.0  $\mu\text{mol}$  in 0.2 M NaCl, 6.0  $\mu\text{l}$ ) and CuSO<sub>4</sub>·5H<sub>2</sub>O (0.3  $\mu\text{mol}$  in 0.2 M NaCl, 3.0  $\mu\text{l}$ ). The Cu<sup>I</sup> click catalyst solution was added to the annealed oligonucleotides and the reaction mixture was kept at room temperature for 2 h. Reagents were removed using NAP-10 gel-filtration column and the ligated products were analysed and purified by HPLC and denaturing 20% polyacrylamide gel electrophoresis. Yield after HPLC purification was 49–50%.

#### Site-directed mutagenesis

PCR amplification was carried out as follows: a typical reaction was performed in a total volume of 50  $\mu\text{l}$ , containing 10 ng of plasmid DNA, 0.3  $\mu\text{M}$  of mutagenic primers, 5  $\mu\text{l}$  10 $\times$  KOD Hot Start reaction buffer, 4  $\mu\text{l}$  1.5 mM MgSO<sub>4</sub>, 5  $\mu\text{l}$  0.2 mM dNTPs, 0.02 U/ $\mu\text{l}$  KOD Hot Start Polymerase (Novagen, Cat. No. 71 086-3). The reactions were hot-started by heating to 95°C for 2 min. Each reaction was subjected to eight cycles of 95°C for 20 s, 54°C for 10 s and 70°C for 2 min. The solution was allowed to cool to room temperature over 1 h.

#### DpnI digestion and purification of SDM products

*DpnI* restriction endonuclease (NEB, Cat. No. R0176L) was directly added to the product of the above amplification reaction, and incubated at room temperature for 6 h. The SDM product was separated from any remaining

template by gel purification from 0.8% agarose using the QIAquick Gel Extraction Kit (QIAGEN) following the manufacturer's instructions.

#### Transformation of SDM product

A 5  $\mu\text{l}$  aliquot of purified SDM product was added to a 50  $\mu\text{l}$  aliquot of KRX chemical competent cells (Cat. No. L3002, Promega), and incubated for 30 min on ice. The mixture was heat-shocked at 42°C for 30 s and then placed on ice for 2 min. An amount of 450  $\mu\text{l}$  of SOC was added to the mixture and the cells were incubated with shaking at 37°C for 1 h. An amount of 100  $\mu\text{l}$  aliquots of the cell suspension were spread on LB agar plates containing ampicillin (100  $\mu\text{g}/\text{ml}$ ) and 0.1% rhamnose. The plates were incubated overnight at 37°C.

#### Assessing the biocompatibility of click-linked DNA

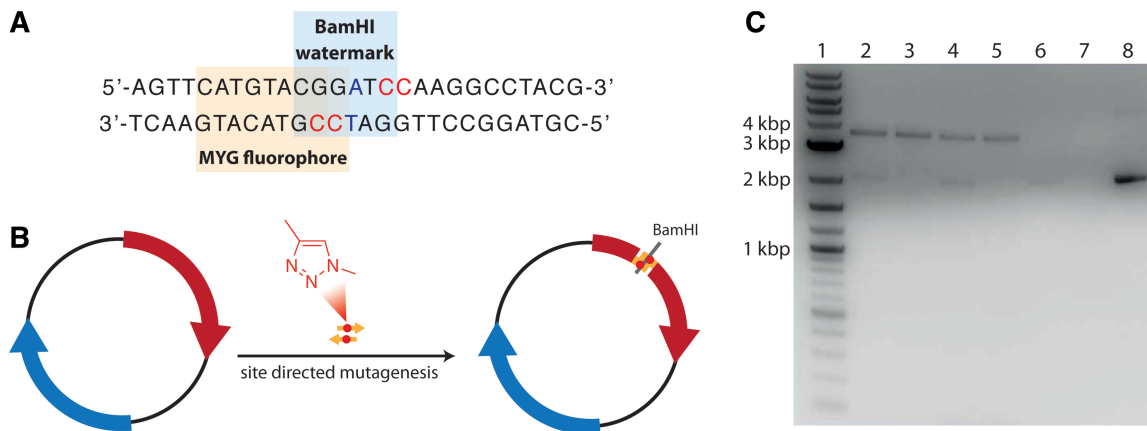
The number of visible colonies on each plate was counted and compared for the click-SDM plates with that from the positive and negative control plates (these experiments were repeated 10 times to generate statistically significant data). Twenty colonies from the click-SDM plates were selected at random and grow overnight in 5 ml LB supplemented with 100  $\mu\text{g}/\text{ml}$  ampicillin, at 37°C. The plasmid was isolated and purified and the mCherry gene sequenced. Ten more random colonies were picked and assessed for the presence of the *BamHI* watermark by restriction digestion using *BamHI*-HF restriction endonuclease (New England Biolabs) according to the manufacturer's instructions.

#### In vitro mCherry expression

The S30 T7 High Yield Protein Expression System (Promega, L1110) was used for *in vitro* protein expression. Reactions were set up following the manufacturer's instructions, except that 1  $\mu\text{g}$  of purified SDM product was used instead of plasmid DNA as template. The template was dialysed against distilled water for 2 h using Millipore filter (Merk Millipore 'V' Series Membranes, 0.025  $\mu\text{m}$ ) prior to use. The reaction mixtures were incubated at 37°C for 1 h, with shaking at 300 rpm, followed by the addition of a 30  $\mu\text{l}$  slurry of Ni-NTA resin (Ni Sepharose™ 6 Fast Flow, GE Healthcare) pre-washed with Ni wash buffer [20 mM Tris base, 20 mM Imidazole, 500 mM NaCl, 10% (v/v) Glycerol, 1% (v/v) Triton X-100]. The mixtures were incubated at 4°C with shaking for 1 h. The Ni-NTA resin was washed three times with 1 ml Ni wash buffer. The bound protein was eluted with 25  $\mu\text{l}$  of elution buffer [Tris base 20 mM, Imidazole 500 mM, NaCl 200 mM, 5% (v/v) Glycerol, 0.1% (v/v) Triton-X 100]. The product was analysed by western blot.

#### Western blot analysis

Samples were resolved by SDS-PAGE (12% polyacrylamide), and the protein transferred to a PROTRAN nitrocellulose membrane (Bio-Rad). Membranes were blocked in 5% non-fat dry milk in PBS-Tween for 1 h, followed by incubation with 5% milk in PBS-Tween with 1:1000



**Figure 2.** (A) The click-oligonucleotides used for site directed mutagenesis contained a silent C to A mutation (shown in blue), that introduces a *BamHI* restriction site not present in the native mCherry gene. The click-linked bases are shown in red. (B) Assembly of the click-linked pRSET-mCherry plasmid by site directed mutagenesis, introducing a *BamHI* watermark. (C) Gel electrophoresis (0.8% agarose gel) of SDM products (expected size 3577bp); Lane 1, 2-log DNA ladder (New England Biolabs); lane 2, pRSET-mCherry SDM with normal primers; lane 3, pRSET-mCherry SDM with normal primers followed by *DpnI* digestion; lane 4, pRSET-mCherry SDM using click primers; lane 5, pRSET-mCherry SDM using click primers followed by *DpnI* digestion; lane 6, negative control (pRSET-mCherry SDM using water instead of primers); lane 7, negative control followed by *DpnI* digestion; lane 8, pRSET-mCherry template plasmid.

dilution of anti-His antibody overnight at 4°C. This was followed by incubation with 5% milk in PBS-Tween containing a 1:1000 dilution of the horseradish peroxidase-conjugated secondary antibody for 1 h at room temperature. ImmunStar Western C (Bio-Rad) was used to detect the horseradish peroxidase-labelled secondary antibody, followed by imaging and analysis on a ChemiDoc MP imaging system (Bio-Rad), and quantification with Image Lab software (Bio-Rad).

### mCherry expression in *E. coli* experiments

The product of SDM with click-linked primers was transformed by heat-shock into KRX competent cells as detailed above. The volume of the recovery mixture was increased to 5 ml with LB media, and grown until OD<sub>600</sub> of 0.6 was reached. Recombinant protein expression was induced at 18°C with 0.1 mM IPTG, and shaken at 18°C overnight. The cells were pelleted and frozen at -80°C. The cell pellet was re-suspended in 2 ml of MHW buffer [20 mM Tris base, 10 mM imidazole, 400 mM NaCl, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 10 mM NaF, 1 mM DTT, Protease inhibitor cocktail tablets (Roche), 20% (v/v) Glycerol, 1% (v/v) Triton X-100] and lysed by sonication. The lysate was centrifuged at 8000 rpm for 20 min and the supernatant applied to 100 μl of a slurry of Ni-NTA resin (Ni Sepharose™ 6 Fast Flow, GE Healthcare) pre-washed with Ni wash buffer [20 mM Tris base, 20 mM Imidazole, 500 mM NaCl, 10% (v/v) Glycerol, 1% (v/v) Triton X-100]. After 1 h incubation at 4°C shaking, the beads with bounded protein were washed three times with 1 ml Ni wash buffer and eluted with 25 μl of elution buffer [Tris base 20 mM, Imidazole 500 mM, NaCl 200 mM, 5% (v/v) Glycerol, 0.1% (v/v) Triton-X 100]. The purified protein was analysed using 12% polyacrylamide SDS-PAGE and a Tecan M200pro fluorescent plate reader.

## RESULTS AND DISCUSSION

### Construction of click-pRSET-mCherry plasmid by SDM

An optimized SDM protocol (16) was used to incorporate click linkers into each strand of the gene encoding the fluorescent protein mCherry (12). Forward and reverse mutagenic primers, each containing a pair of click-linked cytosine nucleosides were designed to overlap in the region encoding the fluorophore of mCherry (Figure 2A), which is formed by a tripeptide MYG motif (methionine 71, tyrosine 72 and glycine 73) in the protein. The primers introduced two click linkers, 4-bp apart, into a region critical for the fluorescence of mCherry. Once transformed into *E. coli*, any deletions or mutations arising from the presence of the click linkers would be expected to result in an easily identifiable, non-fluorescent phenotype. The click SDM primers were also designed to introduce a silent *BamHI* restriction site, to allow the progeny of click-linked plasmids to be readily distinguishable from those of the parent pRSET-mCherry plasmid (Figure 2A). Introduction of this watermark by SDM (Figure 2B) changes the codon used for the glycine of the MYG fluorophore from GGC to GGT, but as both codons have similar relative abundance in *E. coli* (3.0% and 2.8%, respectively), this change was not expected to significantly affect protein expression levels.

The click-linked forward and reverse mutagenic primers were synthesized by the CuAAC reaction between an oligonucleotide with a 3'-alkyne modification and another with a 5'-azide group (ODN-1 with ODN-3 or ODN-2 with ODN-4, Table 1) in the presence of a splint (ODN-5 or ODN-6 respectively, Table 1) and Cu<sup>I</sup> catalyst. In the presence of the splint, the conversion of the CuAAC reaction was found to be quantitative. The click-linked oligonucleotides were purified by HPLC (to remove the splint) and used in subsequent experiments.

The plasmid used in this study was pRSET-mCherry,(12) a high copy number plasmid with a pUC origin that encodes mCherry under the control of bacteriophage T7 promoter. The high copy number of the plasmid maximizes the number of times the click-linked DNA is replicated, therefore propagating any arising errors in replication. The click-containing primers or canonical equivalents (as positive control) were incorporated into pRSET mCherry using either *pfu* polymerase, or KOD polymerase. The procedure was also repeated with water instead of the mutagenic primers as a negative control. The yield of SDM product was found to be significantly higher with KOD polymerase (Supplementary Data), in line with previous reports.(17) KOD polymerase was therefore used in all following experiments. The above products were subsequently digested with *DpnI* restriction endonuclease to remove the methylated template plasmid; note the absence of plasmid template (band at ~2kb, reference lane 8) in lanes 3, 5 and 7 compared to its presence in lanes 2, 4 and 6 in Figure 2C. As expected, there was no SDM product present in the negative control reaction (lanes 6 and 7, Figure 2C). The incubation time with *DpnI* was optimized (Supplementary Data), with the plasmid template fully degraded after 6 h. The SDM product was

further purified by gel electrophoresis to ensure absence of the template in the subsequent reactions.

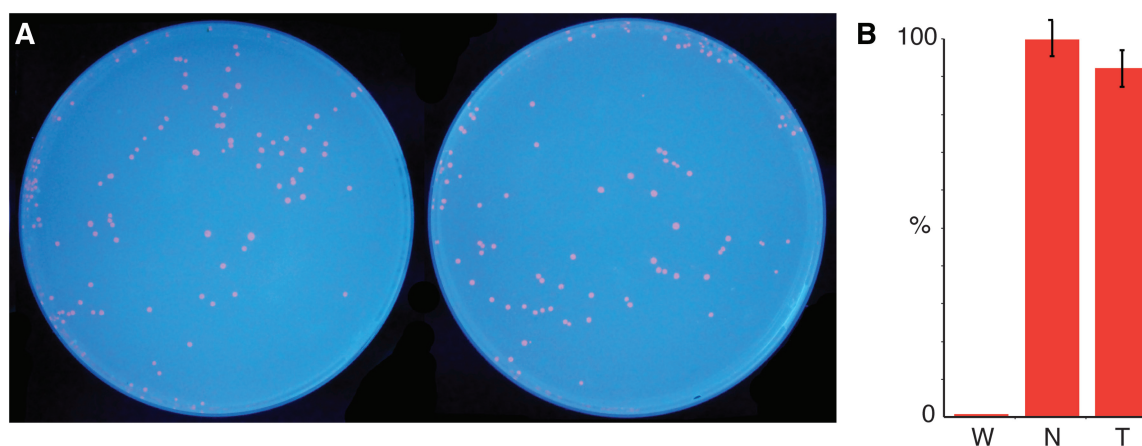
### Biocompatibility of pRSET-click-mCherry in *E. coli*

The purified products of the above SDM reactions (click, positive control and negative control) were initially transformed into high-efficiency *E. coli* DH5 $\alpha$  to assess the yield of the SDM protocol with each primer type. As the strain of DH5 $\alpha$  used lacked a chromosomal copy of T7 RNA polymerase, the red-fluorescent phenotype expected from mCherry (whose expression is under control of a bacteriophage T7 promoter) was not observed. The SDM product using the click-linked primers with KOD polymerase resulted in an average of 179 colonies on LB agar plate supplemented with 100  $\mu$ g/ml ampicillin, the positive control SDM with canonical DNA resulted in 185 colonies, and the negative control SDM with water in place of mutagenic primers resulted in three colonies (Supplementary Table 1). In order to utilize the red-fluorescent phenotype to rapidly identify functional copies of the mCherry gene, the above SDM products were transformed into the KRX strain of *E. coli* (Promega), a high transformation efficiency derivative that express T7 RNA polymerase under control of a rhamnose promoter. Transformation of the SDM product with click-linked DNA primers resulted in 78 colonies, compared to 84 colonies with the positive control using native mutagenic primers. All colonies were visibly pink in colour and displayed red fluorescence when viewed under UV light (Figure 3A). A single colony was observed on the plate containing transformants of the negative control SDM product, indicating the number of colonies arising from background levels of the parent plasmid. These experiment were repeated 10 times (Supplementary Table 2) to generate statistically significant numbers; the colonies containing triazole-containing plasmids were red in all cases, and contained  $92.7 \pm 8.3\%$  of the colonies on the native plates, with the negative

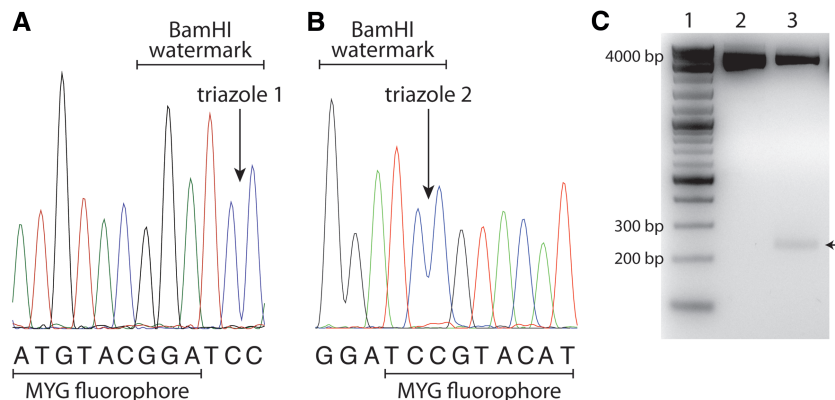
**Table 1.** Oligonucleotides used for the synthesis of the click-linked mutagenic primers

Code	Oligonucleotide sequences (5'-3')
ODN-1	AGTTCATGTACGGAT <sup>MeC<sup>k</sup></sup>
ODN-2	CGTAGGCCTTGGAT <sup>MeC<sup>k</sup></sup>
ODN-3	<sup>z</sup> CAAGGCCTACG
ODN-4	<sup>z</sup> CGTACATGAACT
ODN-5	CATGTACGGATCCAAG
ODN-6	AGGCCTTGGATCCGTA

C<sup>k</sup>, alkyne-modified cytidine, C<sup>z</sup>, azide-modified cytidine.



**Figure 3.** Biocompatibility of click DNA in *E. coli*. (A) The plate on the left contains transformants that have undergone SDM with native primers, and the plate on the right shows transformants of SDM with the triazole-containing primers. Ten replicates of the experiment were performed, with all colonies showing the expected mCherry fluorescence when viewed under a UV light source. (B) Comparison of the number of colonies in the water control (W), native (N) and triazole (T) plates. The triazole plates contained  $92.7 \pm 8.3\%$  of the colonies on the native plates, whereas the negative control plate contained <1%.



**Figure 4.** (A) and (B) Representative sequencing data of the mCherry gene from colonies on the triazole DNA plates. The forward strand is shown in (A), and the complementary strand is shown in (B), with the location of the triazole backbone linkers shown on each strand. (C) Lane 1 is 2-Log DNA ladder (New England Biolabs), lane 2 is *Bam*HI restriction digestion of native pRSET mCherry plasmid, and lane 3 is the progeny of the click-modified variant. The presence of the additional *Bam*HI watermark incorporated by the click-SDM primers results in the highlighted 229-bp fragment in lane 3.

control plates containing <1% of the colonies (Figure 3B).

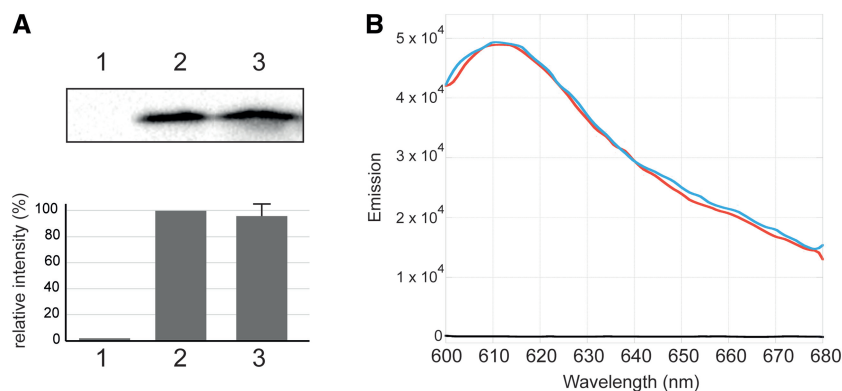
To confirm the origin of the plasmid within the colonies transformed with the product of SDM with click-linked primers, 20 colonies were picked and their pRSET-mCherry plasmid over-expressed, isolated and sequenced. In all cases examined, both the forward and reverse strands contained the *Bam*HI watermark, with no errors observed anywhere in the sequence of the mCherry gene (Figure 4A and B). The parental origin of an additional 10 colonies was probed by restriction digestion with *Bam*HI. The native pRSET-mCherry plasmid contains a *Bam*HI restriction site in the multiple cloning site prior to start of mCherry; the progeny of click-containing plasmids will contain the additional *Bam*HI watermark introduced by the mutagenic click-linked primers in the region encoding the MYG fluorophore. Restriction digestion of this plasmids with *Bam*HI will therefore result in two bands, one at 229 bp and another at 3348 bp, whereas digestion of the native plasmid will lead to a single band at 3577 bp. Two bands were observed in all colonies examined from the click-SDM plate, demonstrating the presence of the *Bam*HI watermark (Figure 4C), and directly linking their heritage to the click-containing SDM primers.

The fidelity of click-DNA replication may also be probed by analysing the progeny of the click-pRSET mCherry plasmids within a single colony. The transformed click-containing plasmid is replicated multiple times in each bacterium, and errors in this process may be observed by isolating, re-transforming and assessing plasmids from a single colony. Plasmids from 10 click-transformed colonies were isolated and re-transformed into KRX cells. All resulting colonies displayed the red-fluorescent phenotype associated with correct replication of the click-containing mCherry gene (Supplementary Figure S4).

#### Protein expression from click-pRSET-mCherry

To probe the compatibility of the click-DNA linker with the transcription machinery, an *in vitro* coupled

transcription/translation system was used. As the *in vitro* system only contains T7 RNA polymerase S30 Extract and the components required for translation (18–20), information may be inferred on the effect of the click-linker on the fidelity and rate of transcription through the linker by T7 RNA polymerase, as reflected by the quantity of fluorescent mCherry produced by the system; protein quantity would be lower than that from canonical DNA if T7 RNA polymerase is slowed by the click linker, while any errors in translation such as skipping through the click-linked bases by the polymerase would result in loss of the fluorescent phenotype of the protein product. It should be noted that such systems are optimized for use with plasmid DNA, therefore products of SDM are typically sub-cloned and over-expressed in *E. coli* prior to use as templates for *in vitro* protein production. In the case of click-pRSET-mCherry however, over-expression of this plasmid by sub-cloning would result in the loss of the click DNA-linker, as DNA replication in *E. coli* results in a phosphodiester backbone. We therefore sought to directly use the SDM product as the template for *in vitro* transcription/translation. As the quantity of purified DNA isolated from each reaction was typically ~0.1 µg, (10-fold lower than the 1 µg of plasmid recommended by the manufacturer), 10 SDM reactions (of each) were carried out in parallel with click-linked primers, canonical primers (positive control) and water (negative control). The products of each set was pooled and used as the template for *in vitro* transcription/translation of mCherry. The SDM product was used directly as template for *in vitro* transcription/translation, as treatment of the SDM product with T4 DNA ligase did not affect the yield of the reaction (Supplementary Figure S5). The mCherry protein produced in the reaction was readily visualized and quantified by western blot; the yield of *in vitro* mCherry production from the click-SDM product was comparable to that from canonical-SDM product (Figure 5A), as was the fluorescence emission associated with mCherry at 610 nm (Figure 5B), with no product or fluorescence observed with the negative control template. These results indicate that transcription from



**Figure 5.** (A) Western blot of the product of *in vitro* transcription/translation using the SDM products with water (lane 1), canonical primers (lane 2) and click-linked primers (lane 3), and as template. The intensity of each band is quantified below the blot (average of three repeats). (B) Fluorescent emission of mCherry (at 610 nm) from *in vitro* transcription/translation of pRSET-mCherry constructed by SDM from click-linked primers (red line), normal primers (blue line); no fluorescence is observed from the negative control (black line).

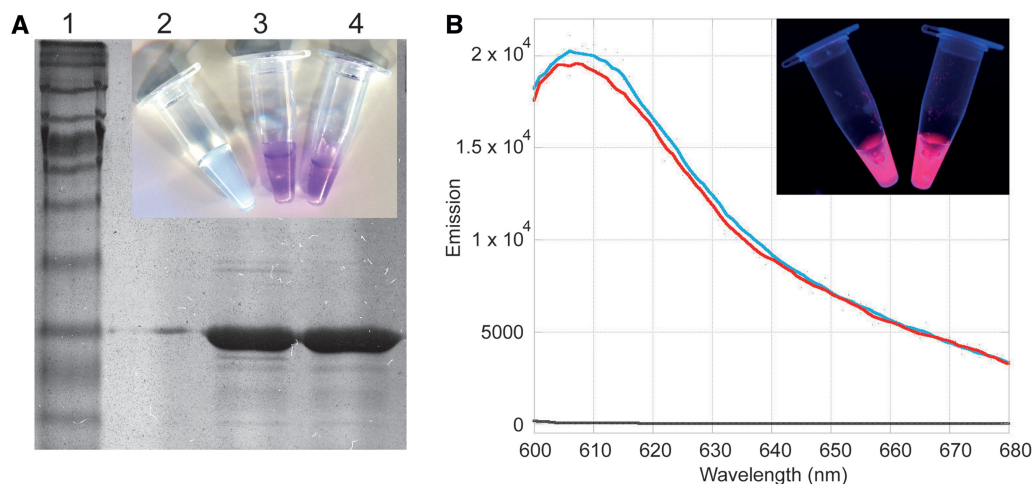
the click linker *in vitro* results in functional mCherry protein, at similar quantity to that from the equivalent canonical DNA.

The effect of the click DNA linker on DNA replication and transcription in *E. coli* was further probed by comparing the fluorescence of the mCherry protein produced from click-linked DNA to that from the canonical equivalent. Rather than determine the fluorescence of mCherry within individual colonies, which will be expected to show stochastic variation (21), we sought to assess total mCherry levels in the population of transformants. Typically a starter culture from a single colony is used for protein expression (to ensure genomic uniformity); in this case however, the whole transformation recovery mixture was used as the starter culture for protein expression. This allowed total mCherry expression in the population of transformants to be assessed, averaging variations in mCherry levels between single cells. The product of SDM with click primers, native primers or water was transformed into KRX competent cells, and after recovery for 1 h, the volume of the solution was increased to 5 ml with LB media, and cells were grown until  $OD_{600}$  0.6. IPTG was added to induce mCherry expression, and the resulting protein was affinity purified. As the transformants were not placed under selective pressure to maintain the plasmid encoding mCherry, the amount of fluorescent protein produced will be affected by the replication rate of the transformed plasmid, informing on the effect of the click DNA linker on this process. The isolated protein product from the click-SDM and native plasmids was visibly pink in colour, whereas the negative control solution was colourless (Figure 6A, insert). Analysis of the purified protein by SDS-PAGE showed a band at 30 kDa, in line with the mass expected for mCherry (Figure 6A). The functionality of the fluorescent protein produced from each plasmid was assessed by comparing the fluorescence of 2  $\mu$ l of each mCherry solution using a fluorescence plate reader. When excited at 587 nm, the protein expressed from click-pRSET-mCherry showed similar levels of fluorescent emission at 610 nm to that from the pRSET-mCherry constructed by SDM with normal

primers (Figure 6B). As expected the solution from the negative control plasmid was not fluorescent. The rate of mCherry expression from the normal and click-linked SDM products was assessed by monitoring protein levels at 1, 2, 3, 4 and 12 h after IPTG induction, and found to be comparable (Supplementary Figure S6). The similarity in mCherry quantity and quality produced from the click-linked plasmid and the equivalent canonical DNA, in both the *in vitro* and *in vivo* experiments (Figures 5 and 6) demonstrates the full biocompatibility of the product of click DNA-backbone ligation and confirms its potential as an alternative to enzymatic DNA ligation.

## CONCLUSIONS

An essential prerequisite for the use of chemically modified DNA in molecular biology is that the modification should be benign, with the modified DNA being a functional mimic of its natural counterpart. The triazole linkage in Figure 1a fulfils this criterion surprisingly well, despite its apparent dissimilarity to a normal phosphodiester bridge; the triazole linkage is accurately and efficiently read through by DNA and RNA polymerases (8,11), making click DNA ligation a promising alternative to enzymatic ligation. Here we have demonstrated the full biocompatibility of a click DNA-backbone linker in a non-essential gene, determining that the presence of selective pressure is not required for its correct function in *E. coli*. That the click-linked mCherry gene is replicated and transcribed with high fidelity despite the very close proximity of the triazole backbone linkers (4-bp apart), demonstrates the robustness with which our click DNA backbone linker mimics canonical DNA. These findings suggest the possibility of an alternative, fully synthetic approach to gene synthesis that will be of potential significance to synthetic biology and industrial biotechnology. This approach is also particularly relevant to the synthesis of large DNA strands containing site-specific epigenetic modifications, as such DNA constructs are difficult to make by PCR-based methods.



**Figure 6.** (A) SDS-PAGE of  $\text{Ni}^{2+}$  affinity purified mCherry (30 kDa) from colonies transformed with the SDM product with water (negative control, lane 2), canonical (positive control, lane 3), and click-linked primers (lane 4). Lane 1 is NEB ColorPlus Prestained Protein Ladder and insert shows the readily visible mCherry protein product at 30 kDa when using normal (left) or click-linked primers (right) in white light. (B) Fluorescent emission of recombinant mCherry (610 nm) expressed from pRSET-mCherry constructed by SDM from click-linked primers (red line), normal primers (blue line) and water control (black line). Insert shows purified mCherry from pRSET-mCherry constructed by SDM from normal (left) and click-linked (right) primers, under UV light at 365 nm.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–3, Supplementary Figures 1–6 and Supplementary Methods.

## ACKNOWLEDGEMENTS

The authors thank Professor Roger Tsien for providing pRSET-mCherry, Dr Patrick Duriez and Dr Fedor Forafonov for helpful discussions, and Dr Paul Duckmanton for taking the photograph in Figure 6B.

## FUNDING

Cancer Research UK [Career Establishment Award 10 263 to A.T.]; Engineering and Physical Sciences Research Council [EP/H04986X/1 to A.T.]; Southampton University Life Sciences Interface [studentship for A.P.S. to A.T.]; the Biotechnology and Biological Sciences Research Council [sLOLA grant BB/J001694/1 to A.H.E.-S., T.B. and A.T.]. Funding for open access charge: Engineering and Physical Sciences Research Council.

*Conflict of interest statement.* None declared.

## REFERENCES

- Gibson,D.G., Smith,H.O., Hutchison,C.A. 3rd, Venter,J.C. and Merryman,C. (2010) Chemical synthesis of the mouse mitochondrial genome. *Nat. Methods*, **7**, 901–903.
- Gibson,D.G., Young,L., Chuang,R.Y., Venter,J.C., Hutchison,C.A. 3rd and Smith,H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, **6**, 343–345.
- Gibson,D.G., Benders,G.A., Andrews-Pfannkoch,C., Denisova,E.A., Baden-Tillson,H., Zaveri,J., Stockwell,T.B., Brownley,A., Thomas,D.W., Algire,M.A. *et al.* (2008) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science*, **319**, 1215–1220.
- Gibson,D.G., Glass,J.I., Lartigue,C., Noskov,V.N., Chuang,R.Y., Algire,M.A., Benders,G.A., Montague,M.G., Ma,L., Moodie,M.M. *et al.* (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, **329**, 52–56.
- El-Sagheer,A.H. and Brown,T. (2009) Synthesis and polymerase chain reaction amplification of DNA strands containing an unnatural triazole linkage. *J. Am. Chem. Soc.*, **131**, 3958–3964.
- Herrlein,M.K., Nelson,J.S. and Letsinger,R.L. (1995) A covalent lock for self-assembled oligonucleotide conjugates. *J. Am. Chem. Soc.*, **117**, 10151–10152.
- Xu,Y.Z. and Kool,E.T. (1998) Chemical and enzymatic properties of bridging 5'-S-phosphorothioester linkages in DNA. *Nucleic Acids Res.*, **26**, 3159–3164.
- El-Sagheer,A.H., Sanzone,A.P., Gao,R., Tavassoli,A. and Brown,T. (2011) Biocompatible artificial DNA linker that is read through by DNA polymerases and is functional in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **108**, 11338–11343.
- Rostovtsev,V.V., Green,L.G., Fokin,V.V. and Sharpless,K.B. (2002) A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem. Int. Ed. Engl.*, **41**, 2596–2599.
- Tornøe,C.W., Christensen,C. and Meldal,M. (2002) Peptidotriazoles on solid phase: [1,2,3]-triazoles by regioselective copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.*, **67**, 3057–3064.
- El-Sagheer,A.H. and Brown,T. (2011) Efficient RNA synthesis by *in vitro* transcription of a triazole-modified DNA template. *Chem. Commun.*, **47**, 12057–12058.
- Shaner,N.C., Campbell,R.E., Steinbach,P.A., Giepmans,B.N., Palmer,A.E. and Tsien,R.Y. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma sp.* *red fluorescent protein*. *Nat. Biotechnol.*, **22**, 1567–1572.
- Dallmann,A., El-Sagheer,A.H., Dehmel,L., Mugge,C., Griesinger,C., Ernsting,N.P. and Brown,T. (2011) Structure and dynamics of triazole-linked DNA: biocompatibility explained. *Chemistry*, **17**, 14714–14717.
- Miller,G.P. and Kool,E.T. (2002) A simple method for electrophilic functionalization of DNA. *Org. Lett.*, **4**, 3599–3601.
- Miller,G.P. and Kool,E.T. (2004) Versatile 5'-functionalization of oligonucleotides on solid support: amines, azides, thiols, and thioethers via phosphorus chemistry. *J. Org. Chem.*, **69**, 2404–2410.



16. Kunkel, T.A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. USA*, **82**, 488–492.
17. Fujii, S., Akiyama, M., Aoki, K., Sugaya, Y., Higuchi, K., Hiraoka, M., Miki, Y., Saitoh, N., Yoshiyama, K., Ihara, K. *et al.* (1999) DNA replication errors produced by the replicative apparatus of *Escherichia coli*. *J. Mol. Biol.*, **289**, 835–850.
18. Nevin, D.E. and Pratt, J.M. (1991) A coupled in vitro transcription-translation system for the exclusive synthesis of polypeptides expressed from the T7 promoter. *FEBS Lett.*, **291**, 259–263.
19. Zubay, G. (1973) In vitro synthesis of protein in microbial systems. *Annu. Rev. Genet.*, **7**, 267–287.
20. Zubay, G. (1980) The isolation and properties of CAP, the catabolite gene activator. *Methods Enzymol.*, **65**, 856–877.
21. Elowitz, M.B., Levine, A.J., Siggia, E.D. and Swain, P.S. (2002) Stochastic gene expression in a single cell. *Science*, **297**, 1183–1186.
22. Chan, T.R., Hilgraf, R., Sharpless, K.B. and Fokin, V.V. (2004) Polytriazoles as copper(I)-stabilizing ligands in catalysis. *Org. Lett.*, **6**, 2853–2855.