Supplementary Data
for

Force-activated DNA substrates for probing individual proteins interacting with single-stranded DNA

Stephen R. Okoniewski[1,2], Lyle Uyetake[1] and Thomas T. Perkins[1,3,*]

[1]JILA, National Institute of Standards and Technology and University of Colorado, Boulder, CO, 80309-0440, USA
[2]Department of Physics, University of Colorado, Boulder, CO, 80309-0440, USA
[3]Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO, 80309-0440, USA

*To whom correspondence should be addressed. Tel: 1-303-492-5291; Email: tperkins@jila.colorado.edu

Contents:

METHODS:

Preparation of 50%-GC DNA substrate

The left handle, right handle, and 50%-GC sequence were each PCR amplified using the primers listed in Table S2 and the conditions listed in Table S3. All PCR products were purified using the Qiagen Qiaquik PCR purification kit (Qiagen 28106), then loaded onto an 1% agarose/TAE (Tris-HCl, Acetic acid, EDTA) gel where the band was excised. They were then processed by Bio-Rad Freeze 'N Squeeze (cat#732-6165), concentrated with a Millipore Amicon Ultra 0.5 mL 10K column (Millipore UFC501096) with TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA) buffer exchanges, and subjected to another Qiaquik kit to remove residual agarose.

The purified PCR products of the left handle and the 50%-GC sequence were then cut with PflmI endonuclease at 37 °C for 3 h. They were combined together in equimolar ratio and ligated together with T4 DNA ligase (New England Biolabs) overnight at 16 °C for 16 h. All ligations were performed in a 50–100 µL volume with ligase at 5% of the total volume (the maximum amount of ligase suggested @ 400 units/µL) and 1–2 pmol of each DNA. The ligated product was run on a 1% agarose/TAE gel and the ligated band was excised and purified as above. The purified ligated product and the right handle were then cut with BbsI at 37 °C for 4.5 h, ligated together, and purified as above.

This fully ligated construct was then incubated at 72 °C for 10 min with Taq polymerase to generate A overhangs for TA cloning. The product was cloned into a bacterial plasmid using TOPO® XL PCR Cloning Kit (Invitrogen K4700-10). The plasmid was transformed into competent *E. coli*, grown up in Luria Broth, and then isolated (Qiaprep Spin Miniprep Kit, Qiagen 27106). After verification by sequencing and PCR, the plasmid constructs were used as a template for PCR labelling using a biotin and a dibenzocyclooctyne (DBCO) primer, where DBCO is a copper-free click chemistry reagent that reacts with an azide moiety. The product was purified as above.

The resulting purified and 5'-labelled PCR product was nicked sequentially with the restriction endonucleases Nt.BspQI and Nb.BsmI (New England Biolabs). Nt.BspQI nicking was performed at 50 units/µg DNA at 50 °C for 4 h, followed by a deactivation step at 80 °C for 20 min. Nb.BsmI nicking was performed at 2 units/µg DNA at 65 °C for 1 h, followed by a deactivation step at 80 °C for 20 min. The nicked constructs were purified using the Qiagen Qiaquick kit and a small amount was analyzed on a 1% alkaline denaturing agarose gel to ensure successful nicking and that the denatured, nicked products were present in the right sizes and proportions.

Preparation of 20-bp hairpin DNA substrate

The left and right halves of this substrate were PCR amplified, purified as above, then cut with BstXI (New England Biolabs) at 37 °C for 3 h, and finally buffered exchanged with TE using a

Millipore Amicon Ultra 0.5 mL 10K. They were next ligated together, purified, and cloned into a bacterial plasmid as above. This fully ligated construct was then PCR amplified from its plasmid using the DBCO and 4x biotin labelled primers, purified as above, and nicked with Nb.BbvCI at 5 units/μg DNA at 37 °C for 1 h. Unlike with the 50%-GC substrates, the nicking was deactivated with 10 mM EDTA, not heat, since it was found that high heat could thermally denature the nicked section. The constructs were then purified as above.

Calculation of k_{on} and comparison to previous data

To determine k_{on}, we visually measured t_1 values from the extension vs. time records. The start time was defined as when the force clamp brought the molecular extension to within ~0.5 nm of the expected activated extension at 8 pN. The unwinding initiation time was determined by a ~0.5 nm change from this same value. We estimate the uncertainty to be ~0.1 s. We next made histograms of the t_1 values using a 9.5-s bin width for both concentrations. Both histograms were then fit with

$$f(t) = \frac{1}{\tau} e^{-\frac{t}{\tau}}, \quad (1)$$

to determine the average on-rate constant τ. To calculate k_{on} , we then assumed τ was dominated by the time needed for RecQ to bind the ssDNA with a negligible contribution from the time needed for RecQ to translocate on the 33 nt ssDNA to start unwinding ($t_{on} \approx 10$ s vs. $t_{trans} \approx 0.1$ s). We then used

$$k_{on} = \frac{1}{[E]\,\tau_{on}}, \quad (2)$$

where $[E]$ is the enzyme concentration and τ_{on} is time constant determined from fitting the histogram of t_1 at each concentration to Eq. 1. Using the fit uncertainties for τ_{on} and an estimated error of 15% for $[E]$, we calculated k_{on} = 3.6 ± 0.6 × 10^8 $M^{-1}s^{-1}$ at $[E]$ = 200 pM and 3.8 ± 0.6 × 10^8 $M^{-1}s^{-1}$ at 100 pM. We calculated the mean to yield our reported value 3.7 ± 0.6 × 10^8 $M^{-1}s^{-1}$. Uncertainties represent the returned best fit ± std. dev. and associated error propagation.

We next wanted to compare our determination of k_{on} with the one previously reported value for $RecQ^{\Delta H}$ (1). In that study, the authors perform an ensemble stopped-flow assay on $RecQ^{\Delta H}$ using 54-nt ssDNA substrates (dT_{54}) at 5 °C. They report a k_{on} of 1.9 ± 0.2 $\mu M^{-1}s^{-1}$, but this value is in terms of the molar concentration of nucleotides in dT_{54} substrates, not $RecQ^{\Delta H}$ binding sites. For an accurate comparison, we need to account for the assay temperature difference (5 °C vs 25 °C) and convert to binding site concentration. Fortunately, Ref. (1) measured k_{on} for wild type RecQ at both 5 °C and 25 °C, and found a 3.3-fold increase at 25 °C. For simplicity, we apply the same factor to $RecQ^{\Delta H}$. To convert to binding site concentration, we subtract the occluded site size of $RecQ^{\Delta H}$ (in nt) from the 54 nt of their substrate. However, obtaining an exact value for the occluded site size is difficult, since the value ranges from 13–27 nt depending on the method used to measure it (1-3). For maximum consistency, we use the occluded site size calculated in Ref. (1) for wild-type RecQ [a recent study demonstrated that the site sizes of $RecQ^{WT}$ and $RecQ^{\Delta H}$ are nearly the same (for a given method) (3)]. Given that we performed our

measurements at saturating [ATP], we use the site size of RecQ with bound AMPPNP (an ATP homolog). This gives a 14-nt site size, leading to a correction factor of 40. Multiplying Ref. (1)'s reported value by 3.3 and 40 yields $k_{\mathrm{on}} = 2.4 \times 10^8$ $M^{-1}s^{-1}$. This corrected value is only ~1.5-fold smaller than our measured value of k_{on}.

Supporting references:

1. Kocsis, Z.S., Sarlos, K., Harami, G.M., Martina, M. and Kovacs, M. (2014) A nucleotide-dependent and HRDC domain-dependent structural transition in DNA-bound RecQ helicase. *J. Biol. Chem.*, **289**, 5938-5949.
2. Harami, G.M., Nagy, N.T., Martina, M., Neuman, K.C. and Kovacs, M. (2015) The HRDC domain of E. coli RecQ helicase controls single-stranded DNA translocation and double-stranded DNA unwinding rates without affecting mechanoenzymatic coupling. *Sci. Rep.*, **5**, 11091.
3. Harami, G.M., Seol, Y., In, J., Ferencziova, V., Martina, M., Gyimesi, M., Sarlos, K., Kovacs, Z.J., Nagy, N.T., Sun, Y., Vellai, T., Neuman, K.C. and Kovacs, M. (2017) Shuttling along DNA and directed processing of D-loops by RecQ helicase support quality control of homologous recombination. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, E466-E475.
4. Smith, S.B., Cui, Y. and Bustamante, C. (1996) Overstretching of B-DNA: The elastic response of individual double-stranded and single stranded DNA molecules. *Science*, **271**, 795-799.
5. McCauley, M.J. and Williams, M.C. (2007) Mechanisms of DNA binding determined in optical tweezers experiments. *Biopolymers*, **85**, 154-168.

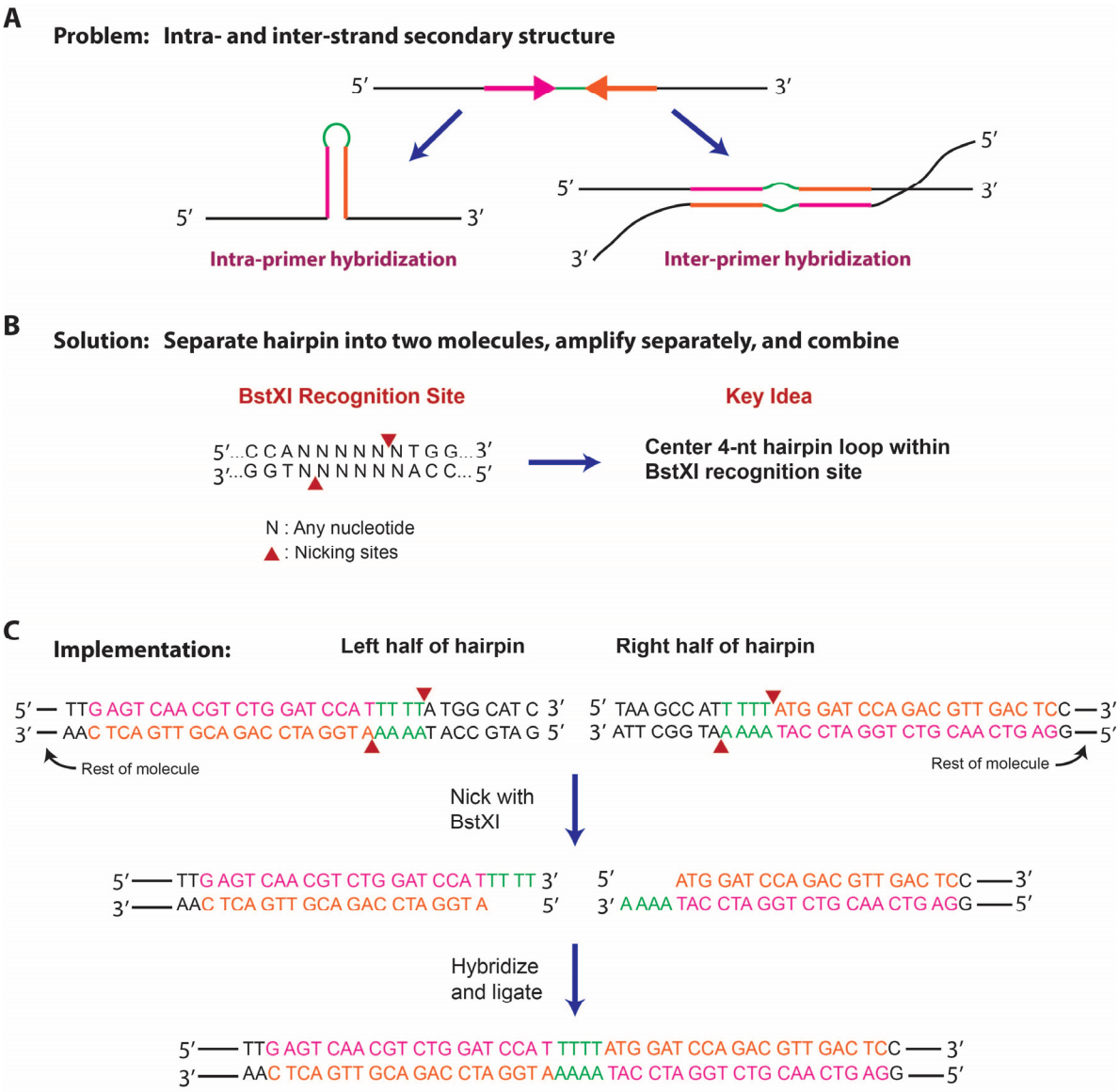

Figure S1. Challenges and solution to the design and construction of the force-activated 20-bp hairpin construct. (**A**) Cartoon depicting the challenges to using a PCR primer that fully encodes a hairpin. Specifically, undesirable intra- and inter-strand hybridization will form. (**B**) The non-palindromic recognition sequence for the endonuclease BstXI contains an internal 6-bp that accommodates any sequence. After cleavage, BstXI leaves a 4-nt overhang. By using a non-palindromic sequence (TTTT) in this region, self-annealing of the same constructs is avoided (*e.g.*, avoiding the two-left halves depicted in panel **C** from annealing). (**C**) Sketch of the steps and sequences used to yield a DNA construct containing an internal hairpin by sequentially nicking the two DNA molecules encoding the two halves of the hairpin and then ligating them together. Note, this hairpin sequence contains three instances of two adjacent GC base pairs, which did not interfere with force activation.

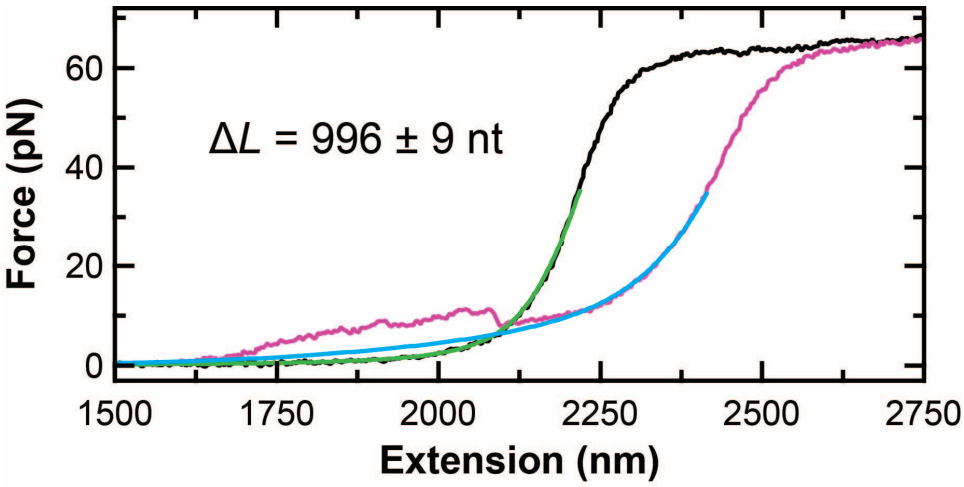

Figure S2. Determining the amount of ssDNA generated after force activation of the 50%-GC construct. Force-extension curves before (black) and after (magenta) force activation. The initial curve was analyzed using an eWLC model as described in Materials and Methods using a fixed dsDNA contour length (2179.8 nm = 0.338 nm/bp × 6,449 bp). The fit (green) to the fully dsDNA construct returned a "ssDNA contour length" value of 32.0 ± 0.3 nm, which accounted for variation in bead size variation and linker lengths (biotin-streptavidin and silane-PEG-azide-DBCO). The force-extension curve of the activated substrate was fit the same way using a fixed dsDNA contour length of 1841.1 nm (0.338 nm/bp × 5447 bp). This second fit (cyan) returned a ssDNA contour length of 589.6 ± 5.1 nm. The difference in contour length was 557.6 ± 5.1 nm. Using a ssDNA contour length of 0.56 nm/nt (4,5), this yields 996 ± 9 nt of ssDNA exposed during activation This finding is excellent agreement with the expected value of 1,002 nt. For consistency, both fits were performed from 15 to 35 pN to avoid the secondary structure unfolding in the activated substrate observed at lower forces. Model eWLC curves were extended to 0 pN for presentation using the determined parameters. Uncertainties represent best fit ± std. dev.

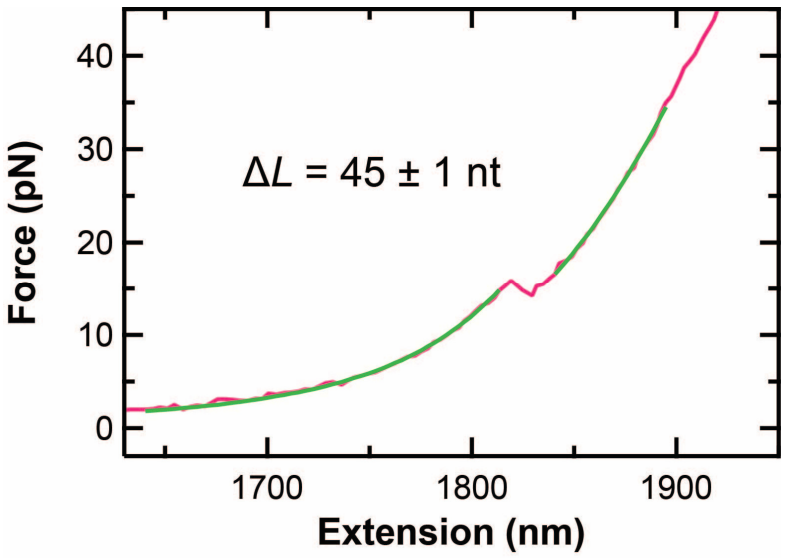

Figure S3. Determining the size of the unfolded hairpin. A force-extension curve of the activated construct (red) shows a rupture of the hairpin at ~15 pN. The lower (2-15 pN) and upper (17-35 pN) sections were both fit as described in methods. Both fits (green) used a fixed dsDNA contour length of 1840.75 nm = 0.338 nm/bp × 5446 bp. In addition, the first fit includes a 2-nm offset to account for the diameter of the DNA hairpin. These fits returned ssDNA contour lengths of 26.2 ± 0.4 nm and 51.6 ± 0.2 nm before and after hairpin unfolding, respectively. The contour length difference was 25.4 ± 0.5 nm, which corresponds to 45 ± 1 nt using 0.56 nm/nt (4,5). This result quantitatively agrees with the expected value of 44 nt. Uncertainties represent best fit ± std. dev.

Table S1. 50%-GC sequence

PflmI and BbsI restriction sequences are colored cyan and grey, respectively. Nt.BspQI and Nb.BsmI nick sequences are colored magenta and green, respectively. GC clamp sequences are colored yellow. The actual nick sites are denoted by "*".

ATACCATGGG*ATGGTATAGGGCGCTCGGCGTGCAGCCGTGCGGGCTCTTCG*TCACAGACACTGTGACA
CTGTCAGTGACTCTGAGAGACTGAGTGAGTGTGAGTGAGTCAGTGACACAGTCACTCTGTCAGAGACACT
CTGACAGTGAGAGTGTGACAGACTGTCTCAGAGTCAGACAGTCAGACTCTCTGAGTCACTGTCACTGACA
CTGACTCAGAGTCAGTCACTCAGTGAGTCTCTGTGTCTGTGTGTCTCTGACTGTGACACTCTCTCTCTGT
GTGAGACACTCAGACACACTCTGTGACAGAGTGTCACACTCACACTGTCTGACAGACACTGAGAGAGTCA
GTGAGAGTCAGTGTGTCAGACTGTGACTCAGTCAGTCACTGAGACTGTCTGACACTCTCTGAGACTGTGA
GTGAGTGTCTCAGAGACTGTGAGACTGTGTCAGTCTGAGTGACACTGACAGAGACTCAGTCAGAGACTCT
CACACTCAGAGTGTCTGAGTGAGAGAGAGTCACTGAGAGTCAGAGACTGACTGTCTCAGTGTGTCTGACA
GTCACACTGACAGTCTCTGACACTGTCTGTCAGAGTCTGTCACTGAGTGTCACAGACAGAGAGAGTGTGT
CACACTGTGACAGTCACAGTGTCAGTGTGTGACACTCACTCTGACTGACAGAGTGTGAGAGACACACACA
CACTCACAGTGTGTGAGTCTGACTCAGTGAGAGTGACAGTCTGTCTCTGTGTGTCTGACACTGACACACT
GTGTGTGTGTGACTCAGACAGAGAGTGTGAGTCACTCAGACAGTGAGTGTGACTCACAGTCAGACACAGA
GTCACTGTGACTCTGACTGAGAGTCTCACTGTCTCTCTGACTCAGACTGTGTGACTGACACACAGAGACT
GTCTCACACTGAGACTCACACTGACTGAGAGACACTCTCACAGTGACAGTCAGTGACTGTGACTGAGAGA
GACTCTGTGAGTCTGTGAGTGAGAGACTGTCTGTGTCACACAGTGACACTGTGAGAGACTCAGAGAGAGT
CTGAG*CATTCGCGATCCGTGCGGGCACGGCCAGGCTCCCGCATAGAAGAC*TTGCAATCAC

Table S2: PCR primers used

All primers were synthesized by Integrated DNA Technologies Inc. The BstXI restriction sequence is colored grey, the Nb.BbvCI nick sequence is colored cyan, and the GC clamps are colored yellow.

50% GC substrate primers:	**Sequence**
Left forward	ATT ACG GTC AAT CCG CCG
Left reverse	CGA GAT CCA TCC CAT GGT TGA GTC ACT GCT AAG TCA CTG CAT ATT CAC TGC CTT GAC AAG AAC CGG ATA TTC
50%GC forward	ACC GCA CAG ATG CGT AAG G
50%GC reverse	GAA ACA GCT ATG ACC ATG ATT A
Right forward	CGC ATA GAA GAC TTG CAA TGA GAG AGT TGC AGC AAG C
Right reverse	AGT TGT TCC TTT CTA TTC TCA CT
DBCO-Left forward	5'DBCO-TEG-ATT ACG GTC AAT CCG CCG
Biotin-Right reverse	5'Biotin-TEG-GCC TCA ACC TCC TGT CAA TGC T
Hairpin substrate primers	**Sequence**
Left hp forward	TCA ATA ATC GGC TGT CTT TCC TTA TCA TTC GGG CGC TCG GCG TGC AGC CGT GCG GGC TTG TTC AGG ACT TAT CTA TTG TTG ATA AAC AG
Left hp reverse	GAT GCC ATA AAA ATG GAT CCA GAC GTT GAC TCA AAT GCT GAG GGC GGG AGC CTG GCC GTG CCC GCT GAG AGA GTT GCA GCA AGC G
Right hp forward	TAA GCC ATT TTT ATG GAT CCA GAC GTT GAC TCC CCG CTG TCT AGA GGA TCC GAC TAT CGA CCT CAG CGC TTT AAT GCG GTA GTT TAT CAC AGT TAA ATT GC
Right hp reverse	CGA AGG TAA CTG GCT TCA GCA GAG CGC AGA TAC CAA ATA CTG TCC
DBCO Right hp forward	5'DBCO-TEG-CGA AGG TAA CTG GCT TCA GCA GAG CGC AGA TAC CAA ATA CTG TCC
4xBiotin-Left hp reverse	5'Biotin-TEG-TCA ATA A-BiodT-C GGC TGT CT-BiodT TCC TTA TCA BiodT-TC

Table S3: PCR conditions used

PCRs were performed using KOD Hot start DNA polymerase (Novagen 71086-4) or Taq polymerase (Life Technologies 10342-053).

PCR Conditions:

PCR product:	Primer 1	Primer 2	template	template amt (ng)	polymerase	units	# of cycles	heat activation (min)	denaturation	anneal	extend	final ext @ 72°C (min)
Left handle	Left forward	Left reverse	M13MP18	2.5	KOD	2	35	2	95°C/20s	70°C/10s	70°C/40s	1
50%GC sequence	50%GC forward	50%GC reverse	pUC57	3	KOD	2	35	2	95°C/20s	60°C/10s	70°C/40s	1
Right handle	Right forward	Right reverse	M13MP18	2.5	KOD	2	35	2	95°C/20s	65°C/10s	70°C/120s	1
Full 50%GC	DBCO-Left forward	Biotin-Right reverse	50%GC plasmid	5	KOD	2	30	2	95°C/20s	60°C/10s	70°C/170s	1
Left half hp	Left hp forward	Left hp reverse	M13MP18	3	TAQ	2	30	3	95°C/45s	60°C/30s	70°C/225s	10
Right half hp	Right hp forward	Right hp reverse	pBR322	3	TAQ	2	30	3	95°C/45s	60°C/30s	70°C/270s	10
Full hp labelling	DBCO-Right hp forward	4xBio-Left hp reverse	hp plasmid	5	KOD	2	30	2	95°C/20s	60°C/10s	70°C/140s	1

10x PCR buffer	1x
MgSO4(KOD) or MgCl2(Taq)	150 nmol
dNTP	20 nmol
Primer forward	30 pmol
Primer reverse	30 pmol
M13MP18 or pBR322 plasmid	2.5-5 ng
KOD or Taq polymerase	2 units
Water	to 100 µl