

DNA cloning for protein expression using Gateway[®] technology

Artem Evdokimov
artem@xtals.org

1. Reagents & their properties. Storage & propagation of Gateway components.

Helpful things to be aware of regarding storage and propagation of vectors containing the *ccdB*/CAT cassette:

- Upon repeated thawing, or after prolonged storage in liquid form these vectors can develop an increased background. This is presumed to be due to breaks in the *ccdB* area which, when repaired by the cellular machinery, result in defunct *ccdB* gene. This problem is not very common, and can be completely avoided by storing these vectors in aliquots and minimizing the freeze/thaw cycle numbers.
- *ccdB*-bearing vectors are designed to kill normal *E. coli*. Propagation of these vectors should be done in specialized cells, such as *ccdB survival* or DB3.1.
- Growth of plasmid-bearing DB3.1 cells is best done at 30°C rather than normal 37°C, in media that is generously supplemented with glucose (0.5-1.0%). At 37°C we have observed (very infrequent) spontaneous change of Gateway vectors. Growth at lower temperature abolishes this problem. We have not observed this issue with *ccdB survival* cells, so far and therefore we grow them at 37°C.
- It is advisable to grow cells bearing Gateway Donor and Destination vectors on media that contains both the main selection antibiotic (Kan-50mg/L, Amp-120mg/L, Carb-100mg/L, Strep-50mg/L etc.) and *chloramphenicol* (Cln-34mg/L). Cells will grow slower.
- Some F'-bearing *E. coli* have increased resistance to *ccdB* and are not recommended for Gateway cloning. We have repeatedly tested the following cells and found them to be 100% killed by *ccdB*-bearing vectors (therefore these are suitable for direct transformation of recombination reactions): DH5 α , TOP10, Mach1-T1, BL-AI, BL21, BL(21)DE3, BL21(DE3)*, Origami[®], Rosetta-gami[®], and several others. Depending on the phenotype and cultivation conditions XL1-blue[®] cells and related strains can generate increased backgrounds if used in Gateway cloning.

BP clonase – this reagent is quite robust and should be stable for at least 6 months (we've used it after a year with only marginal loss of activity) at -80°C. It is a good idea to avoid freeze-thawing, so we routinely aliquot BP clonase into convenient fractions (e.g. 4 or 8 ul) and never re-freeze. Thaw BP clonase on ice.

LR clonase and **LR clonase II** – these reagents are considerably less robust than the BP clonase. It is a very good practice to aliquot LR clonase into 4ul (or whatever is convenient) fractions, and to *never re-freeze more than 1-3 times*. Aliquots normally

remain active after 6 month storage at -80°C. In our hands, LR clonase II gives better results than the original LR clonase and is considerably more convenient to use. Thaw BP clonase on ice, immediately before use.

Proteinase K – normal precautions for storage of proteases apply. We store this enzyme in 10ul aliquots, which get used up after 4-5 freeze/thaws.

Buffers – normal precautions apply. We store buffers in 10ul aliquots that get used up after 4-5 freeze/thaws. Buffers seem to contain DTT (by smell) and should not be left unfrozen for prolonged periods of time to avoid oxidation by air.

Donor vectors (e.g. pDONR221) – just like any other plasmid, storage at -20°C or -80°C guarantees long-term stability. It is always a good idea to have a glycerol stock handy as well. We usually store pDONR221 at its final concentration (150 ng/ul) in 1ul aliquots, frozen. This makes it very convenient to run BP reactions. This vector has a medium/high copy number origin and is KanR and ClmR. It is toxic to normal *E. coli* and must be propagated in either *ccdB survival* or DB3.1 cells.

Entry vectors – these are regular plasmids, which require no special handling precautions. Most of the normal cloning-grade cells are acceptable. We routinely use TOP10, Mach1-T1, and DH5α cells with these vectors.

Destination vectors – same precautions as for Donor vectors. These are conveniently stored at -80°C in 1ul aliquots of 100 ng/ul, ready for LR reaction.

Expression vectors – these are regular plasmids. While transcriptional activity of certain heterologous promoters in *E. coli* has been observed, in our hands this is a fairly infrequent occasion. Therefore yeast, baculovirus-insect cell, Gram-positive bacteria, etc. expression vectors can be propagated in cloning-grade *E. coli* without any special precautions. For vectors known to produce expression in *E. coli* one should always consider toxicity, see parts (4) and (5) for more details.

2. Construction of the AttB amplicon.

a) Primer design and synthesis

Primer design scheme described below corresponds to our normal way of designing Gateway constructs (namely with a protease site at the N-terminus of the protein, included *after* the AttB site).

We begin by identifying the desired 5' end of the DNA of interest. The codon for the initiator methionine can be included or avoided as needed – in this example it is avoided. We usually choose the annealing temperature for the gene-specific portion (GSP) to be 50-65°C, for reasons explained in the next section. There are several methods by which theoretical DNA melting temperature is calculated - it is almost irrelevant which method

is used, as long as it is used consistently[♥]. On the 5' end of the GSP we add the codons for whatever other elements that may be necessary (in this case it is a TVMV protease cleavage sequence), preceded by the sequence of the *AttB1* site.

In order to achieve the right reading frame after the Entry clone is recombined into an Expression clone, the two Lysine codons in the *AttB1* site must be in-frame with the protein-coding portion.

<i>AttB1</i>										<i>TVMV protease site</i>				
G	GGG	ACA	AGT	TTG	TAC	AAA	AAA	GCA	GGC	TTC	GAA	ACC	GTT	CGT
.	K	K	A	G	F	E	T	V	R
TTC	CAG	TCC	CTG	CGG	CCC	AAC	GGG	CAG						
F	Q	S	L	R	P	N	G	Q						
Gene-specific portion														
5' (forward) primer design														

At the 3' side we have similar choices to make – the gene specific portion should have roughly the same annealing temperature as the GSP of the forward primer, and the stop codon can be included (some people prefer to include double stop codons) or omitted – in this example it is included. Other DNA elements can be included as needed (none in this example) followed by the sequence of the *AttB2* site. Naturally, all the sequences included in the 3' primer correspond to the reverse-complement of the actual DNA sequence to be amplified (as shown in the example below).

<i>AttB2</i>															
GGGG	AC	CAC	TTT	GTA	CAA	GAA	AGC	TGG	GTT	CTA	GAA	GAC	GTC	TTT	
											*	F	V	D	K
ACC	GAC	CGA													
G	V	S													
Gene-specific portion (reverse complement)															
3' (reverse) primer design															

These primers tend to be fairly long (45-70 bp typically) – therefore it is convenient to split them into two, and use either a consecutive or nested PCR. Doing so allows us to avoid the need to use longer primers – the latter usually cost more than the two shorter ones[♦], are synthesized with lower yield, and require additional purification steps. An example of a split primer pair (designed for two consecutive PCR reactions) is shown below. Please note that for routine Gateway cloning of N-terminal fusions with TVMV

[♥] This convenient site <http://www.basic.northwestern.edu/biotools/oligocalc.html> provides high-quality calculated DNA parameters.

[♦] The oligo synthesis company that we use (IDT-DNA) provides same-day synthesis for primers less than 45bp in length, which considerably cuts on DNA synthesis time

protease linker sequences it is convenient to order a ‘universal TVMV-overlapped primer’, as shown in the example below. A universal 3’ primer can be designed as well, however it is usually more convenient to use a unique primer pair (this way it is easier to match melting temperatures).

F-I (forward primer part I, 45 nt).

GAA ACC GTT CGT TTC CAG TCC CTG CGG CCC AAC GGG CAG

GSP T_M = 59-65°C

Overall T_M = 74-84°C

Overlap T_M = 54-61°C

F-II (forward primer part II, 52-nt). This ‘universal TVMV-overlapped primer’ has to be ordered only once for all subsequent experiments).

G GGG ACA AGT TTG TAC AAA AAA GCA GGC TTC GAA ACC GTT CGT
TTC CAG TCC

Overall T_M = 72-83°C

Overlap T_M = 54-61°C

R-I (reverse primer part I, 36 nt).

GAA AGC TGG GTT CTA GAA GAC GTC TTT ACC GAC CGA

GSP T_M = 57-65°C

Overall T_M = 67-77°C

Overlap T_M = 52-60°C

R-II (reverse primer part, 39-nt).

GGGG AC CAC TTT GTA CAA GAA AGC TGG GTT CTA GAA GAC

Overall T_M = 68-78°C

Overlap T_M = 52-60°C

Normally the single-step desalted (crude) oligonucleotides are of sufficient quality for this PCR. We store our primers reconstituted in HQ-water (*TE can be quite bad for PCR*) at 50uM final concentration, frozen at either -20°C or -80°C.

b) PCR

In section (a) we designed the primer pairs to produce the amplicon of interest. Very similar PCR conditions can be used in amplification regardless of whether a single-step, dual, or nested PCR strategy is chosen.

We use PfuUltra[®] hot-start polymerase master mix (Stratagene) because of its superior fidelity and productivity. In order to minimize DNA secondary structure, primer-dimers, GC-clamping, etc. we use the PCR mix supplemented with 3-6% DMSO by default. Typically we run two reactions in parallel – one with 3% DMSO and one with 6%, then compare the results and pick the better one. So far we have encountered nearly 100%

fidelity, even after two consecutive PCR runs[^]. The following is a typical PCR set-up using primers described in section (a):

First PCR run

2 x PCR master mix	26ul
Primer F-I	0.2-0.4ul
Primer R-I	0.2-0.4ul
DMSO	1.5ul or 3.0ul (3% or 6%)
<i>Template</i>	1-3ul (50-150 ng)
HQ-Water	to 52ul

Second PCR run

2 x PCR master mix	26ul
Primer F-II	0.2-0.4ul
Primer R-II	0.2-0.4ul
DMSO	1.5ul or 3.0ul (3% or 6%)
<i>First PCR run</i>	1ul
HQ-Water	to 52ul

Both PCR amplifications are cycled as follows (note that the annealing temperature is lowered by DMSO)

initial melting – 2 minutes at 95°C

25 cycles of

30 seconds	at	95°C
30 seconds	at	56°C
1 minute/kB	at	72°C

final extension - 10 minutes at 72°C

If yield is lower than desired, 30 cycles can be used with no detrimental effect. If you choose to follow the nested PCR strategy, the amounts of 1st stage primers should be reduced (0.05-0.1 ul) and the first PCR should be run for 10-12 cycles, following which the second primer pair is added (0.2-0.4 ul) and PCR is continued for another 20-30 cycles. We prefer sequential PCR because if something is not working well this allows us to have more control over the separate stages of the process.

c) Purification of the PCR product

There are reports of PCR reactions used directly for BP recombination, and it is certainly a possibility. However, it is ultimately faster and more convenient to purify the

[^] So far we've encountered a total of two PCR artefacts in more than 100 PCR amplifications varying in size from 300 to 3000 bp. Naturally individual results will vary. Water quality is paramount for high-fidelity performance of this mix – we use double glass distilled, certified water (TekNova).

PCR amplicon prior to BP reaction. Several methods can be used, but we prefer gel-purification because it also allows us to monitor what happened during PCR. Usually we get a clear single band of the right size, which is excised and extracted from the gel via a standard Qiagen gel-purification kit. Alternatively, the PEG precipitation as described in the Gateway manual yields adequate results as well – provided that the DNA pellet is handled with extreme caution as it is invisible and can be easily lost.

If gel-purified DNA is used for BP reaction, the number of false positives is essentially zero. In contrast, the use of unpurified or partially purified amplicon can result in significant proportion of false positives due to recombination events between the Donor vector and primer-dimers*.

3. BP reaction.

In a typical BP reaction we use 100-150 ng of the Donor vector (most commonly pDONR-221) and 100-200 ng of the PCR amplicon. This reaction is fairly forgiving, and as long as the amounts are at least approximately correct it will yield abundant true-positive clones. For the reasons of economy, we have reduced the amounts of Gateway reagents, with no ill effects. Typical reaction set-up is listed below (BP clonase added last):

PCR amplicon (~100 ng total)	1-3ul
pDONR-221 (150ng/ul)	1ul
BP buffer (5x)	2ul
BP clonase	2ul
TE	to 10 ul

It is important to mix the components well, as the BP clonase and the 5x buffer are quite viscous and dense. After mixing, the reaction is allowed to proceed for 1-3 hours at room temperature. The reaction is terminated by addition of 0.5-1.0ul of proteinase K solution (included in the kit) followed by incubation for at least 10 minutes at 37°C.

To obtain numerous colonies 1-3ul of the reaction mixture is transformed into TOP10, DH5 α , or other suitable cloning strain (see section 1 for strain choices). If everything is well, numerous colonies should appear after plating overnight on LB Agar + 50 ug/ml Kanamycin.

4. LR reaction.

The LR reaction is performed in very much the same way as the BP reaction, however it is a little less efficient. Just as for the BP reaction, the amounts of essential components can be reduced for reasons of economy. For a typical LR reaction the following ingredients are brought together, adding the clonase last:

Entry vector (~100 ng total)	1-3ul
Destination vector (100 ng/ul)	1ul

* Compared to the desired amplicon, primer-dimers can be present in the PCR reaction in large molar excess (especially with nested PCR) due to the fact that they are relatively small.

LR clonase II
TE

1ul
to 5ul

The components must be mixed well since the clonase reagent is quite viscous. The reaction is allowed to proceed for several hours at room temperature, or for 2 hours at 35-37°C. The reaction is terminated by addition of 0.5-1.0ul of proteinase K solution (included in the kit) followed by incubation for at least 10 minutes at 37°C.

Typically, to obtain numerous colonies 1-3ul of the reaction mixture is transformed into an appropriate *E. coli* strain. If everything went well, numerous colonies should appear after plating overnight on appropriately selective media (typically LB Agar + 100 ug/ml Ampicillin).

For most LR reactions, the setup above is sufficient to generate numerous colonies. However, in select cases (e.g. very large or very small plasmids) the setup above would be sub-optimal and variations of Entry/Destination vector ratio should be attempted to improve recombination efficiency.

For expression, it is often prudent to first transform the recombination reaction into generic cloning-grade (preferably expression-silent*, *Rec-*, *End-*) cells such as TOP10 or DH5a cells, then purify the expression plasmid. After this, the expression clone is in no way different from any other plasmid and can be propagated and used in virtually any host cell strain. As a shortcut, we have performed numerous transformations of Gateway LR reactions directly into BL *E. coli* and its relatives, with no ill effects provided that the gene is not toxic to the host. Toxic genes can and will generate unwanted background. To minimize toxic effects LR reactions that are transformed into non-expression-silent cells are often plated on LB agar (+ appropriate antibiotic) supplemented with 1% glucose.

5. Verification of Gateway products.

BP products are normally verified by sequencing (although the methods below are also applicable). It is convenient to verify Gateway LR products using colony PCR, since we do not expect any sequence modifications beyond the recombination occurring at the *Att* sites. For colony PCR it is convenient to use one primer from the Destination vector (outside the recombined region) and one primer from the first primer-pair designed to create the Entry vector. For example, a pair of T7-promoter primer and the R-I primer described in section 2a is useful for identification of LR recombinants with any T7-driven Destination vector. A typical procedure for colony-PCR verification is presented below. For convenience, we use PuReTaq Ready-To-Go™ PCR Beads which contain the polymerase, buffer, nucleotides, and magnesium and can be reconstituted in 25ul of water to give a complete PCR cocktail.

* *Caveat emptor* - if expression is driven by P_{Lac}, P_{TAC}, araBAD, or other native-like *E. coli* promoters, some background expression is almost impossible to avoid, even in cloning-grade cells. Therefore one should always watch out for toxicity, which is often manifested as e.g. uneven colony morphology, cells that spontaneously lyse as the colonies grow older, etc. The usual bag of tricks for expression of toxic genes can be applied to minimize these issues.

For colony PCR, a couple of individual colonies are lightly touched with either a sterile toothpick or a sterile inoculation rod, so that a small amount of colony material adheres to the pick. It is important to avoid touching the surface of agar with the pick since agar tends to inhibit PCR reactions. The colony material is then transferred (by inserting the pick into the PCR tube with reaction mix and twirling the pick between fingers) into a PCR tube containing the following mixture:

T7-forward primer	10 pmol
R-I primer	10 pmol (0.2ul of 50uM solution)
PuReTaq bead	1
HQ water	25ul

It is important not to transfer too much of the colony material since the presence of massive amounts of bacteria is usually detrimental to PCR.

PCR is performed according to a standard protocol:

<u>initial melting</u> –	5 minutes at 95°C
25 – 35 cycles of	
30 seconds	at 95°C
30 seconds	at 55°C (temperature depends on primers used)
1 minute/kB	at 72°C
<u>final extension</u> -	5 minutes at 72°C

The resulting products are analyzed on an agarose gel, the presence of a band at the right size indicates correct recombination of the Entry and Destination vectors. Both picked colonies will typically contain the correct product.

6. Conversion of plasmids into Gateway Destination vectors.

Gateway conversion kit (Invitrogen) provides a convenient and easy method of converting nearly any expression plasmid into Gateway Destination vector. In brief the procedure involves the following steps:

- Identification of a convenient restriction site within the right context of the promoter, Shine-Dalgarno/Kozak consensus, fusion proteins, etc. Ideally the cutter should be blunt, however other cut types can be easily blunted or filled in.
- Restriction of the vector followed if necessary by blunting/filling in of the staggered-end cut products with e.g. Klenow fragment or mung bean nuclease.
- Ligation of one of the three reading frame Gateway cassettes into the cut.
- Transformation of the ligation products into *DB3.1* or *ccdB survival E. coli* followed by selection of the transformants on LB-Agar plates containing the

original antibiotic relevant to the expression vector *and* 34 mg/L chloramphenicol.

- e) Preparation of DNA from 6-8 individual colonies.
- f) Verification of the direction of insertion of the Gateway cassette (conveniently done using restriction enzyme(s) cutting once inside the cassette and once outside, so that two orientations give rise to two distinctly different restriction product pairs.

The following considerations apply:

- Expression vectors that confer resistance to Kanamycin *cannot* be used for direct conversion into Gateway if you intend to use Donor vectors that also confer resistance to Kanamycin. *The antibiotic resistances of the Entry and Destination vectors should be orthogonal.* Chloramphenicol-resistant expression vectors should not be used either, if at all possible. We have successfully converted numerous vectors that conferred resistance to Ampicillin, Tetracycline, and Streptomycin into corresponding Gateway Destination vectors.
- It is important to select the right reading frame for conversion. For instance, for N-terminal fusions if the blunted DNA ends in a complete triplet, use reading frame A (RfA) cassette. Therefore in the N-terminal fusion Destination vectors, the lysine codons of ...GTAC AAA AAA... sequence should be in frame with the rest of the relevant 5' sequences. The kit provides cassettes for all three reading frames - more explicit directions can be found in the product manual.
- Both *DB3.1* and *ccdB Survival* cells are somewhat less plasmid-rich than normal *E. coli* – larger amounts of cell mass may be needed to prepare adequate quantities of plasmid DNA.