Sticky-end cloning of PCR products

NOTE 1: There are few things more boring than screening through huge numbers of empty plasmids to find a desired clone. This protocol suggests a number of controls that will immediately suggest where it's all going wrong – take advantage of them.

NOTE 2: The thermostable enzyme in general lab use is *Denville TaqPro* (CB4050-8). While high yield, *TaqPro* is classed as low fidelity; according to Tim Sikorski the error rate is ≈ $1x10^{-6}$/cycle. The higher-fidelity (≈ 47-fold) *Denville SurePol* (CB4095) is preferred for cloning, although the yield tends to be lower. Note that the reaction conditions for these two enzymes are NOT the same, so use the appropriate buffer.

NOTE 3: *Denville TaqPro* (CB4050) leaves a 3'A overhang suitable for direct integration into TA cloning vectors (see below). For blunt-end cloning you will need to polish the ends. *Denville SurePol* (CB4095) leaves no overhang and is suitable for direct blunt-end cloning. General info for both enzymes is available online (www.denvillescientific.com).

There are at least three strategies for cloning PCR products –

1) Direct T/A cloning. Takes advantage of the terminal transferase activity of *Taq* polymerase and other non-proofreading DNA polymerases which adds a single 3'-A overhang to each end of the PCR product (see **NOTE 3**). The resulting PCR product can be directly ligated into a linear vector with a 3´ terminal 'T' or 'U' at both ends.

Advantages: easy, highly-efficient
Disadvantages: only works if your enzyme leaves an overhang (mostly low-fidelity enzymes), T/A vector cloning kits are expensive

2) Direct Blunt-end cloning. Blunt-end PCR products as generated by a proof-reading polymerase (eg. *Denville Surepol*) can be directly cloned into a blunt-end vector

Advantages: straightforward, most MCS's contain EcoRV or SmaI to blunt cut
Disadvantages: Very high backgrounds if you didn't prep the vector properly

3) Restriction Enzyme (RE) digested sticky-end cloning. RE target site(s) are introduced into each of the PCR primers. The resulting PCR product and cloning vector are digested with desired RE(s) to generate complementary ends, which are then ligated.

Advantages: straightforward, gets your product into an interesting vector in one step
Disadvantages: Requires you to think about the experiment in advance and design appropriate overhang primers, you have to further prep the product so higher yields in the initial amplification are desired, high backgrounds if you didn't prep the vector properly

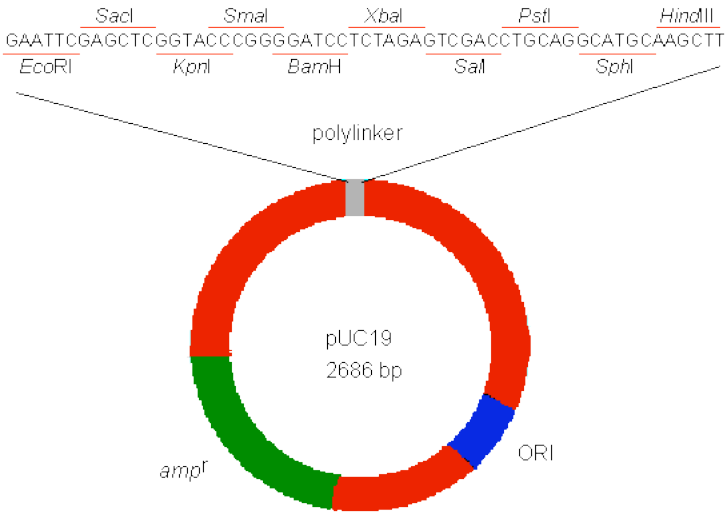
This protocol is particularly concerned with **Option #3** (RE sticky-end cloning): if you can do this then #1 and #2 are a breeze.

Primer design / RE choice

1.1 Look carefully at the region to be amplified. Determine the RE sites that will be present in your final product (obvious, but you'd be surprised how many forget to do this)

1.2 Determine the RE sites available in the vector. Usually, but not always a polylinker (or Multiple Cloning Site, aka MCS) is the intended recipient region.

1.3 Pick REs for cloning. I prefer to employ a limited set of REs (BamHI, EcoR1, EcoRV, HindIII or XbaI) since these are very efficient cutters and all work well in the same buffer (*NEB b^o 2*). However any enzyme in an MCS is usually fairly efficient. For your chosen REs determine: (i) the nature of their overhangs (blunt, 5', 3'), (ii) buffer compatibility (see **1.4**), and (iii) proximity in the vector: DO NOT pick REs right beside each other. In example below [Kpn1 and BamH1 or EcoR1] are fine but not [Kpn1 and Sma1 or Sac1] (see **NOTE 6**).



1.4 Consider the following:

- Will you be doing a blunt or sticky-end ligation ?
- If sticky-end, is it to be a single or double cut ?
- If a double-cut, can you digest with both enzymes simultaneously or are their buffers incompatible ? If the latter, how incompatible are they ? Will adding more salt (eg. *NEB b^o 2 -> b^o 3*) be enough or do you have to purify the DNA between cuts ?

1.5 If it's to be a double-cut choose your two REs**.** Double-check their site preferences: do they have any *isoschizomers* (two different enzymes that cut the same recognitions sequence but not necessarily giving the same ends) ? Are the cut ends compatible? (eg. BamH1 and BglII, Xho1 and Sal1) Are they methylation sensitive? (see **NOTE 4**).

> **NOTE 4:** Answers to all these questions are at: http://rebase.neb.com/rebase/rebase.html. Re. DNA methylation, the only RE you usually have to worry about is Xba1, although (i) it's highly unlikely any MCS would ever contain a methylable Xba1 site (ii) your intended insert in these experiments is a PCR product and thus not methylated, and (iii) the *E.coli* strain in general use in the lab, *DH5α*, is *dam^- dcm^-* anyway.

1.6 Design the Forward and Reverse amplification primers as normal: you want 20-22nt oligos with an annealing temperature (AT^o) >55°C ($AT^o \approx [4x(G+C)] + [2x(A+T)]-3$). Write your primers as if ready to order (*ie.* in the 5' -> 3' direction).

1.7 Decide the cloning direction of your PCR product: eg. I will add an *EcoR1* site (GAATTC) to the Forward primer and a *HindIII* site (AAGCTT) to the Reverse.

Forward Primer + EcoR1: easy to design in 4 easy steps (see **NOTE 6**)

(i) primer sequence = gGA CTg CAg ggT TTC TTC ACC ($AT^o \approx 63^oC$)
(ii) Add the R1 site = GAA TTC gGA CTg CAg ggT TTC TTC ACC
(iii) Add an extension = **GATC** GAA TTC gGA CTg CAg ggT TTC TTC ACC
(iv) Order this primer

Reverse Primer + HindIII: a little more complicated. Let's assume that you'd already decided on a Reverse primer as normal and have it written as a Reverse and Complement (R&C) of the top strand, ready to order in the 5' -> 3' direction.

(i) primer sequence = ggA CCC TTA gAC TgC AAC TgC ($AT^o \approx 63^oC$)
(ii) R&C it so it reads as the top strand: gCA gTT gCA gTC TAA ggg TCC
(iii) Add the HindIII site = gCA gTT gCA gTC TAA ggg TCC AAg CTT
(iv) Add an extension = gCA gTT gCA gTC TAA ggg TCC AAg CTT **gATC**
(v) R&C it = **gATC** AAg CTT ggA CCC TTA gAC TgC AAC TgC
(iv) Now you're back in the 5' -> 3' direction again. Order this primer (see **NOTE 5**)

NOTE 5: For Reverse primer tailing, the 6-step approach is a roundabout way of doing it, but forces you to think about the process (minimizing the possibility of screw ups) and allows you to visualize what the final product will look like.

NOTE 6: We add the 4bp extensions (as above) because otherwise the PCR product has RE sites at the extreme 5' and 3 ends, and REs cut VERY inefficiently under these conditions. The 4bp extensions give them something to hang on to. Re. the sequence, I just jumble a combination of GATC.

Perparing for the ligation

2.1 Perform the PCR (50µl reaction) with these tailed primers as normal. The presence of the tail does not mean you can increase the annealing temperature.

2.2 Run a sample of the product (5µl) If the product is of the expected size but you have reason to check further, perform an internal RE digestion directly on the PCR product (see the specific protocol on http://mckeogh.googlepages.com).

2.3 Depending on how clean the PCR product is you will either want to do a direct PCR-clean up (eg. Qiagen column) or Agarose-gel separation, followed by gel-extraction of the expected band and clean up (eg. QiaexII gel-extraction kit). In either case elute the product in the minimum volume possible (eg. 30µl). Determine the concentration by $OD^{260/280}$ or by running some on a gel and guesstimating the concentration by comparison with the MW marker (eg. Bioline Hyperladder 200bp-2kb; Cat # BIO-33026).

2.4 *Prep the insert:* If at all possible, at **step 1.5** you chose compatible REs to tail your insert. Set up the reaction to digest the PCR product and prepare it for cloning:

50µl reaction: 10x Rb^o 5µl

PCR product	20-30µl (from **Step 2.3**)
RE # 1	2U/µg DNA (guesstimate 0.5µl) (see **NOTE 7**)
RE # 2	2U/µg DNA (guesstimate 0.5µl)
10mg/ml BSA	0.5µl
ddH_2O	to 50µl

Digest at appropriate temp (usually 37°C, but not always – eg. SmaI is 25°C) for 1-2hrs.

Option 1: Column clean the insert, and elute in the smallest volume possible: eg. 10-20µl. Store at -20°C until ready to ligate.

Option 2: Precipitate the DNA by adding 1/10 Vol (5µl) 3M NaOAc (pH5.2) and 2.5 Vol (125µl) 100% EtOH (from -20°C). Leave at -20°C for an hour, collect by micro-centrifugation (14K, 10min, 4°C) and remove the supernatant (you should see a small white pellet). Vortex wash the pellet with 250µl 70% EtOH and re-pellet by centrifugation. Carefully remove the wash, air-dry the pellet for 10-15 mins on the bench and resuspend in 10µl ddH_2O. Store at -20°C until ready to ligate.

NOTE 7: REs are usually sold in tubes of 100 – 5000U, concentrated from 5-15U/µl. 1U is officially the amount required to cut a single site in 1µg of supercoiled DNA in 60mins at optimal temperature and buffer conditions.

2.5 *Prep the vector:* As in **2.4**, I hope you chose compatible REs. Set up the reaction to digest your vector and prepare it for cloning:

50µl reaction:

10x Rb^o	5µl
Vector	≈ 2-5µg (usually cut 10-15µl of a Qiagen mini-prep)
RE # 1	5U/µg DNA (guesstimate 2µl) (see **NOTE 8**)
RE # 2	5U/µg DNA (guesstimate 2µl)
10mg/ml BSA	0.5µl
ddH_2O	to 50µl

Digest at appropriate temp for 2-4hrs. Heat to 65°C for 10mins. In an ideal world the two sticky-ends would be incompatible and you would now immediately proceed to cleaning up the reaction. However if one of your REs cut incompletely you have a mixed population of single- and double-cuts and the singles will snap back together in the ligation increasing your background. Prevent this by Calf-Intestinal Alkaline Phosphatase (CIAP, eg. NEB M0290) treatment to remove the 5'phosphate on the vector backbone. If you did a single cut this step is obligatory.

100µl reaction:

Reaction from above	50µl
10x CiAP b^o	5µl
CiAP	1µl
ddH_2O	to 100µl

Incubate 65°C for 1hr. Add 1µl 0.5M EDTA pH8.0. Heat to 65°C for 10mins. If you did an MCS cut and removed <100bp, column clean the insert, and elute in 50µl. If you removed >100bp run it on a gel, extract the correct sized band, gel-purify and elute in 50µl. In either case, store at -20°C until ready to ligate. You should now have enough cut vector to last you 50 ligations (which is very handy if you use that specific vector a lot for subcloning). I've used it years later with no apparent problems.

NOTE 8: Why so much enzyme ? If the enzyme is old the U calculation is out of date. If the cuts are incomplete you'll spend forever screening empty colonies. Don't worry about *-activity.

Ligation / Transformation

3.1 Set up the ligations as below. In each column the number refers to µl. If you use all the suggested controls and it goes horribly wrong, you'll know exactly why.

Contents	Vector	10 x Lig b^{o}	rATP †	T4 Ligase	Insert	ddH_2O
1. Vector, No ligase	1	1	1	0	0	To 10µl
2. Vector + ligase	1	1	1	1	0	To 10µl
3. Vector + Insert #1	1	1	1	1	1	To 10µl
4. Vector + Insert #2	1	1	1	1	4	To 10µl
5. Insert only	0	1	1	1	4	To 10µl

Some specifics: Enzyme of choice is T4 DNA ligase (eg. NEB M0202). This is supplied with a 10x Reaction buffer and additional ATP (†) is unnecessary. However ATP doesn't like repeated freeze-thaw cycles: it is recommended to aliquot the buffer (20µl per) and store at -20^{o}C.

Each control:

1. Vector, No Ligase: Colonies = both of your RE's are cutting inefficiently
2. Vector + Ligase: Colonies = one of your RE's is cutting inefficiently

3-4. Vector + Insert #1, #2: The optimal ratio Vector : Insert is 1:3 – helps to find it

5. Insert only: Only used if the insert is being sub-cloned from another vector – warns you of backbone carryover.

3.2 Preferred ligation conditions: sticky-end RT^{o}, 15 mins. Inactivate 65^{o}C, 15 mins before transformation. Remove an appropriate number of competent cell aliquots from -80^{o}C and thaw on ice. Our lab uses chemically competent E.coli (DH5α, see **NOTE 9**) from Bioline (BIO-85025; Silver efficiency 10^8 transformants / µg supercoiled DNA). These are in 200µl aliquots and you want 40-50µl per transformation.

3.3 On ice: transfer to an ependorff tube 4µl of ligation mix and 40-50µl of competent cells. Gently mix by flicking. Incubate on ice 30 mins. Place tubes at 42^{o}C for ≈ 45 seconds, then back on ice for 2 mins. Add 350µl SOC buffer (NO selection agent), transfer to a larger culture tube, and incubate 37^{o}C with shaking for 45 mins.

3.4 For color selection (if possible; see **NOTE 9**): spread 20µl X-GAL (Stock 20mg/ml in DMF at -20^{o}C) onto each selection plate (eg. LB^{Amp}, but check the resistance marker in the plasmid) and add 5µl IPTG (1M at -20^{o}C) to each transformation. Plate all of the control ligation (#1, 2, and 5 from **Step 3.1**) transformations, and 50 to 300µl of the #3-4 ligations: for the inexperienced it is advised to plate both volumes on individual plates so you get a feel for the numbers involved (it's difficult to pick colonies if density >300/plate). Incubate overnight 37^{o}C (see **NOTE 10**).

NOTE 9 - *Blue-white color selection*: DH5α contains *lacZΔM15* to facilitate blue-

white color selection (http://en.wikipedia.org/wiki/X-Gal). Plasmids with an MCS in the middle of the lacZ gene (codes for β-galactosidase (β-Gal), the first gene in the lac operon) allow you to use this approach to screen positive integrants. This includes all of the pRS series of yeast plasmids. If you clone a piece of DNA into the MCS the lacZ gene is interrupted, and thus non-functional (**NB**. sometimes, but very rarely, it still works). If we then offer cells containing this plasmid an inducer (IPTG) to turn on the gene, and a substrate (X-gal) that β-Gal can turn into an insoluble blue dye, we can see where activity resides. A blue color means β-Gal is active, and thus no insert.

NOTE 10: Different *E.coli* strains have hugely variable growth rates. Eg. the S28 strain (OK for color-selection) grows very quickly, such that if you plate at 5PM on day 1 you will have very large colonies by 10AM the next AM (so large in fact, that if using Amp as a selection agent and it is limiting you will see satellites appear by noon). DH5α (very good for color selection) grows much more slowly: plate at 3-4PM on day 1 and remove the plates by 11AM on day 2. It will improve color intensity to place the plates at 4°C for an hour or two.

3.5 If all worked well you should have very few colonies on plates #1 and #2 (and all of those should be blue), none on #5, and a fair number on #3 and #4 (with a large number of whites). Screen the white colonies by PCR with primers flanking the insert (eg. for the pRS series use T3 / T7, but note that these have an $AT^{o} = 50^{o}C$). To confirm the inserts, mini-prep representative clones and test by RE analysis (**NOTE 11**).

NOTE 11: When confirming clones by RE digestion keep some things in mind –

(i) Uncut = waste of time: tells you nothing

(ii) Choose RE combinations that confirm insert orientation and identity. Ensure you will definitively know this by the band pattern: if there could be ANY uncertainty pick another enzyme combination. Two or three individual digests should do it.