Cloning, Version 2.0
by Stacey E. Rutledge

I. General Considerations

Cloning is mysterious. What works well for one person will not work well for another person. You have to find things that work well for you, and stick with them. That said, this protocol is meant to provide you with some general guidelines.

When you prepare your insert, you inevitably lose quite a bit of DNA at each step. Therefore, do everything on a fairly large scale (especially when working with libraries). It may take an extra half hour to set up ten extra PCR reactions or even an extra day to gel purify more oligos, but if you play it safe, you won't ever get to the end of your insert preparation and realize you do not have enough DNA to do ligations, thus necessitating weeks of more work to prepare more insert.

II. General preparation strategies

There are a few main strategies that can be used to prepare inserts to be cloned into pCANTAB_5E (or any other vector). In both strategies, a double-stranded DNA is created and then digested on either end with appropriate restriction enzymes. The length of the sequence you wish to insert will dictate what strategy you use to create the double-stranded insert: if the size of the insert is less than 100 bp (including restriction sites and random ends, as discussed below), the double-stranded insert can be created by primer extension of a single synthetic oligo. A longer DNA insert (100-200 bp) can be prepared by mutually primed synthesis.

In addition, short (less than 100 bp) inserts (but not libraries!) can be prepared by annealing of two complementary synthetic oligos. Single sequence inserts can also be created by PCR, using two short primers (usually with overhangs which introduce restriction sites) to amplify a sequence from another vector.

III. General design considerations

Protein-coding inserts must either contain a methionine initiation codon (ATG) or be ligated in-frame to a vector-encoded initiation codon. Your insert must also contain a stop codon or be ligated in-frame to a vector-encoded stop codon.

Ideally, the codons in your insert should be optimized for expression in E.Coli (see any phage display book or manual). Note that you will not have control over this if you are creating your insert by PCR from another vector. You can, however, use site-directed mutagenesis (Stratagene's Quikchange kit is particularly easy and effective) to make small insertions, deletions, or mutations in your vector, which may be used to optimize codons for expression.

It is useful to design your insert so that ligation of your insert into your vector creates a restriction site which can be used as a positive screen for the presence of insert. If you are lucky, the sequence you are inserting will contain a single restriction site which is not contained in your vector. If not, you can use silent mutagenesis to create a unique restriction site. I have found the program Webcutter (available from our links page) to be useful in this regard.

Because most DNA inserts are very small (50-200 bp), it is often difficult to distinguish fully cut insert from uncut or singly cut insert. For this reason, it is useful to design your predigested insert such that restriction digest on either end will cut off 15-20 base pairs or more.

A. Preparing an Insert from a Single Synthetic Oligo via Primer Extension

Oligos can be ordered on an 0.2 or 1.0 µmole scale. They can be ordered from the Keck center at Yale Medical School (which has always worked well for me) or from Operon (see Kevin), or other commercial vendors. If you are ordering long oligos for library construction, Lori (the resident oligo synthesis expert) has the particulars of how they should be ordered to ensure efficient synthesis.

1. Purification (Time estimate: One day per gel + overnight for elution + a few hours for drying)

Primers do not need to be gel purified, and can be used directly after desalting. Longer oligos can be ordered purified (Beware! They are not always very pure!) or they can be easily purified by denaturing PAGE followed by desalting. I generally resuspend oligos on a 1.0 µmole scale in 500 µL dH_20, add 500 µL formamide loading buffer (For 1 mL: 980 µL deionized formamide, 20 µL 0.5 M EDTA, spatula tip of xylene cyanol and bromophenol blue). The oligos are heated to >95 °C for 10 minutes, then quick cooled on dry ice before being loading on an appropriate percentage denaturing acrylamide gel (see table).

I load 200-300 µL of each oligo in each well (3 wells/comb). The oligos are excised from the gel (take care to avoid n–1 contaminants!), crushed through a 5 mL syringe into a 15 mL orange-cap tube, and eluted in 3 volumes TE (10 mM Tris, pH 8.0; 1 mM EDTA) overnight with shaking. The acrylamide is pelleted by centrifugation and the supernatant carefully transferred to eppendorf tubes. The oligos can be dried in the speed-vac and resuspended in 1 mL dH_2O for desalting.

Dye Migration in Polyacrylamide Denaturing Gels

Gel percentage	Bromophenol blue	Xylene cyanol
5%	35 nt	140 nt
6%	26 nt	106 nt
8%	19 nt	75 nt
10%	12 nt	55 nt
20%	8 nt	28 nt

Recommended Polyacrylamide Gel Percentages for Resolution of DNA

Gel percentage	DNA size range
3.5%	100-1000 bp
5%	75-500 bp
8%	50-400 bp
12%	35-250 bp
15%	20-150 bp
20%	5-100 bp

2. Desalting (Time estimate: 1 hour + drying time)

NAP10 columns (Pharmacia) are used for desalting. The columns are equilibrated with three column volumes of water or TE. The load volume for the columns is 1 mL. Oligos can be eluted with 1.5 mL dH$_2$0. The concentration of desalted oligos can be determined by measuring UV absorbance (A260) and converting this to concentration via the Schepartz Lab biopolymer calculator. The oligos are dried in the speed-vac and resuspended to a concentration of 50 μM.

3. Annealing (Time estimate: 1.5 hours + 1.5 hours for agarose gel)

To anneal, equimolar amounts of each oligo (either two long oligos, or a long oligo and a primer) are mixed, heated to >95 °C for 10 minutes, then slowly cooled to room temperature. To monitor the success of the annealing reaction, run a 3% agarose gel with your single-stranded DNA #1 in one lane, DNA #2 in another lane, and then the annealed DNA in a third lane. Run a DNA ladder (100 bp (NEB) or ↗X174 HinfI (Promega)) in another lane for comparison. Note that the DNA ladders are double-stranded, so your single-stranded oligos will not have the same mobility as markers of the same length.

B. Preparation of double-stranded inserts via primer extension

(Time estimate: 3 hours + 1.5 hours for agarose gel + 0.5–2 hours for optional cleanup)

The primer for primer extension reactions should be 20-30 bp in sequence, complementary to either the 3′ or 5′ end of the template oligo, and GC rich. Both ends of the primer should be a G or C base (preferably two in a row). Primer extension reactions are performed as follows:

Step 1. Annealing

        400 pmol long DNA
        400 pmol primer
        40 μL 5x sequenase buffer (USB)
        200 μL total volume

The reaction should be heated to >95 °C for 10 minutes, and then slow cooled to room temperature by removing the hot block from the heating apparatus.

Step 2. Extension

To each annealing reaction, add:

        2 μL 25 mM dNTP's
        2 μL 10 μg/μL BSA
        2 μL 100 mM DTT
        4 μL 13U/μL sequenase (USB)

The reaction is incubated at 37 °C for 30 minutes, and then incubated at 65 °C for 1 hour to heat inactivate the sequenase.

One test reaction should be attempted first, and the success of the primer extension reaction determined by running 5 μL of the reaction on a 3% agarose gel (single-stranded DNA should be run also, as a comparison.) If the reaction is successful, multiple primer extensions can be performed.

After primer extension, digests can be performed immediately on the primer extension mixture (this is the way I have always done it). However, the high salt and/or high protein concentration in the primer extension reaction may interfere with restriction digests. If you find this is a problem, you can clean up your DNA in a number of ways:

EtOH precipitation – add 2.5 vol ice-cold EtOH, 1/10 vol 3 M NaOAc, incubate on dry ice for 30 minutes, spin for 15 minutes, remove supernatant, wash pellet carefully with 1 vol 70% EtOH (room temperature), remove supernatant, dry pellet, resuspend in appropriate volume of dH_2O or TE.
Phenol/Chloroform extraction – add 1 vol 25:24:1 Tris-buffered phenol:chloroform:isoamyl alcohol, vortex, centrifuge for 10 minutes, transfer aqueous layer (top, contains DNA!) to different eppendorf tube.
Nucleotide removal kit (Qiagen) or PCR Purification Kit (Qiagen).

C. Preparing an Insert from Multiple Synthetic Oligos using Mutually Primed Synthesis (MPS)
(Time estimate: 3 hours + 1.5 hours for agarose gel + 0.5–2 hours for optional cleanup)

MPS is just glorified primer extension! The oligos to be used in the MPS reaction should contain a 20-25 bp overlap, should be fairly GC-rich in the overlap region and should contain G or C bases at either end of the overlap region.

The oligos can be annealed and extended under the same conditions as those described above for primer extension (substitute 400 pmol long oligo #2 for primer). The success of the MPS reaction should be monitored by running 5 µL of the MPS reaction and each ssDNA on a 3% agarose gel. Multiple reactions should then be performed. If the MPS reaction is not clean, reaction conditions can be varied. Some things to try are changing the $MgSO_4$ concentration, eliminating BSA, changing the annealing temperature, or changing the extension time.

As with primer extension, these reactions can be cleaned up by one of the methods discussed previously, or can be used directly.

D. Preparing an Insert from a Different Vector using PCR
(Time estimate: 3 hours + 1.5 hours for agarose gel + 0.5–2 hours for cleanup.)

To PCR an insert out of a different vector, you need to order two primers (0.2 µmole scale), one for each end. These PCR primers must contain a 20-25 bp region identical to the 5´ region of the gene to be amplified (5´ primer) or a 20-25 bp region antiparallel to the 3´ region of the gene to be amplified (3´ primer). The hybridization portions of the primer should be GC rich and should contain GC clamps at the ends. In addition, you may want your primers to contain an overhang region coding for restriction sites, an initiation codon (5´ primer) and/or a stop codon (3´ primer).

The success of the PCR reaction can be affected by a number of things (especially primer concentration and Mg_2SO_4 concentration), and thus a number of PCR conditions should be screened. Note that you need only a very small amount of template DNA (1 µL of a 1:100 dilution of miniprepped plasmid is sufficient). A good place to start in testing PCR conditions:

	Volume (µL)					
Primer 1 (10 µM)	4	4	6	6	8	8
Primer 2 (10 µM)	4	4	6	6	8	8
100 µM Mg_2SO_4	2	4	2	4	2	4
25 mM dNTP's	1	1	1	1	1	1
Template DNA	1	1	1	1	1	1
Thermo Pol Buffer	5	5	5	5	5	5
Vent (exo-) Polymerase	2	2	2	2	2	2

50 µl total reaction volume

PCR Program: 94 °C, 1 min, 40 sec
20 x (94 °C, 30 sec, 55 °C, 1 min, 72 °C, 1 min)
72 °C, 6 min

5 to 10 µL of each PCR reaction should be run on an agarose gel to determine the success of the reactions. If the reactions look clean (only your desired product is visible) the reactions can be cleaned up as discussed above. I find the PCR purification kit (Qiagen) works very well for this (use 1 column/ PCR reaction). If the PCR does not look very good, some other things to try are varying the concentrations of primer or Mg_2SO_4 further, adding BSA, using a different polymerase or changing the annealing temperature.

PCR can also be used to amplify small amounts of any double-stranded DNA made by any of the previous methods. To do this, you need primers identical to the 5´ region of the insert to be amplified (5´ primer) and antiparallel to the 3´ region of the insert to be amplified (3´ primer).

E. Restriction Digests of Double-Stranded Inserts
(Time estimate: As long as it takes! (1–3 days, generally) including gels to monitor reactions + cleanup of digested oligos (2 hours – 2 days)

1. Digestion
The NEB catalog should be consulted for buffer requirements and enzyme compatibility. Some enzyme combinations are compatible, others are not. Note that many restriction enzymes cut poorly close to the ends of oligos. Thus, for library inserts, it is essential to prepare A LOT of double-stranded insert, because you will most likely need to purify doubly cut insert away from uncut or singly cut insert.

If the required enzymes are compatible (BglII/NotI, for example), the DNA can be cut with both enzymes at the same time. (For library inserts, I generally do MULTIPLE (5-7) digests in a volume of 50 µL, cutting 10 µg of DNA in each reaction, with an appropriate volume of enzyme). In general, you want to keep the volume of the reaction as small as possible while still cutting as close to completion, so that you can load more DNA across a smaller number of lanes when you purify it, whether on acrylamide or agarose.

If the enzymes are not compatible (SfiI/NotI, for example), the DNA should be cut sequentially. I do this by cutting first with one enzyme (again, 10 mg DNA in 50 µL reaction), EtOH precipitating the DNA after the first digest is complete, resuspending the DNA in an appropriate volume dH$_2$O (generally 30 µL), and digesting with the second enzyme (again in 50 µL reactions). Alternatively, you could cut with the enzyme requiring less salt first, then add salt to the required concentration and cut with the second enzyme.

Digestion should be monitored along the way. If a digest has not progressed sufficiently, it is a good idea to add more enzyme to the reaction and continue the incubation. For sequential digests, I always run an agarose gel to check that the first digest has gone to completion BEFORE I go on to EtOH precipitate the DNA and cut with the second enzyme.

In the end, if the digest appears to have gone to completion, it may not be necessary to gel purify. In this case, the digested DNA could be cleaned up using the PCR purification kit, the nucleotide removal kit, or just by EtOH precipitation, or phenol/chloroform extraction. If this is not the case, gel purification is necessary to purify your digested insert.

2. Agarose gel purification
After the second digest, I gel purify my inserts on 3% agarose gel. I load as much of the digests as I can in each well, excise the appropriate doubly cut bands, and use the QIAquick gel extraction kit to purify the DNA. Even at its best, the kit only recovers 20% of the DNA. Thus, you need LOTS of doubly digested material to recover enough to do ligations, especially in the case of your libraries. As far as the kit goes, a couple of things help recovery: First, the columns can each accept 400 mg of agarose. I find that I get more DNA out if I do not overload the columns with agarose, but similarly do not UNDERLOAD the columns (in other words, use almost exactly 400 mg of agarose/column). Second, always do the optional isopropanol step. Third, elute with 30 µL elution buffer (comes with kit) and wait at least 1 minute before centrifuging. The DNA that comes out of this kit is VERY salty. I find a single EtOH precipitation is enough to clean up the DNA.

3. Acrylamide gel purification
One drawback to using the agarose gel purification columns is that if your insert is less than 100 base pairs you lose even more DNA. For this reason, other people have found other methods for purifying their doubly digested DNA to be more successful. For example, the DNA can be gel purified on **native** acrylamide (15%, 29:1 acryl:bisacryl) (use the fat wells). The DNA needs to be very very clean (i.e., free from contaminating proteins) before loading on the gel or the lanes will be smeary, so it is advisable to doubly phenol/chloroform extract the DNA before loading. First, a test gel should be run, in which various amounts of the DNA (in glycerol loading dye) are loaded in each lane, to determine the best amount of DNA to load (generally, about 1/15 to 1/10 of your 400 pmol scale primer extension reaction). To gauge how far to run the gel, consult the Pharmacia catalog (useful info in the back). Stain the gels in TBE/EtBr. Once you have determined the appropriate amount of DNA to load, run enough gels to purify all of your digests. The appropriate bands should be excised, and eluted in TE as with denaturing purification. The DNA should then be cleaned up by EtOH precipitation (multiple times if necessary).

4. Quantification

The DNA should be quantified by UV prior to ligation. The final concentration of the DNA insert should be around 10 ng/µL for it to be useful in ligations. For cloning of libraries, you need at least 1 mg of insert, but more is better (I usually aim for 3 µg total). For single sequences, 300 ng is probably enough (I usually aim for 1 µg total).

F. Preparing a Vector for Ligation

(Time estimate*:* Overnight + 4-5 hours for maxiprep of vector, 1-2 days for digests and cleanup)

As starting material, a small maxiprep (100-250 mL) of your vector should be way more than enough to digest for use in ligations. It's useful to have extra uncut vector to use as controls for later restriction digest screens.

Digests are performed on the vector in a manner analogous to the insert (sequentially if necessary, etc.). However, the restriction enzymes will cut your vector much more efficiently than your insert, so vector digestion is generally much more quickly than insert digestion and generally does not require as much cleanup. For libraries, I usually cut 5 µg of the vector in a 50 µL reaction (do four or so reactions). For single sequence clones, one digest should be sufficient. Care should be taken to ensure that the vector is indeed doubly digested. If it is only singly digested, the vector will re-ligate to itself with fairly high efficiency, and you will see a lot of background. The double digest can be monitored by agarose gel, but this has two disadvantages: First, you often cannot clearly gauge the success of the second cut without performing additional restriction digests to reduce your vector to two or more pieces of smaller size. Second, very small amounts of uncut vector will not be detected on agarose gels, yet will re-ligate very efficiently.

A better way to ensure that your vector is doubly digested is to digest your insert with a third enzyme that cuts at a site between those of the two enzymes of interest. This will cut the DNA you are attempting to remove in half, thus making it significantly less likely that the "old" insert will be re-ligated to the vector. It will also decrease the chance of any vector remaining singly cut (and thus able to easily re-ligate to itself).

After the digests are complete (monitor on a 1% agarose gel), I combine the digests into 100 µL aliquots, and use the Chromaspin 1000 columns to clean up my doubly digested vector. The protocol is simple: spin the column for 5 minutes to remove the buffer, load your sample (100 µL) and spin column 5 minutes more. The vector is in the eluent. (FYI, I generally lose about 1/3 of my DNA on the column.) The DNA may also be cleaned up by EtOH precipitation or phenol/chloroform extraction.

The final concentration of the doubly digested vector should be around 70ng/µL (for libraries) or 25 ng/µL (for single sequences) for it to be useful in ligations.

G. Ligation*,* the old-fashioned way

(Time estimate: 1 hour to set up + 16 hour incubation)

A number of vector:insert **molar** ratios should be used to investigate ligation efficiency. A good starting point is 1:1, 1:5 and 1:10. You may also do a vector-only ligation reaction with no insert added. Note however that the number of colonies you get

after transforming a vector-only ligation does not necessarily accurately represent the background in your insert-containing reactions.

<div align="center">

10x ligation buffer (NEB)
5 µL 1 M MgCl$_2$
5 µL 2.5 mg/mL BSA
5 µL 100 mM ATP
25 µL 1 M Tris, pH 8
10 µL 0.5 M DTT
50 µL total volume

</div>

Ligations are generally performed on a 150-500 ng scale, although some people have successfully scaled ligation reactions up (for libraries). The reactions should be performed in a total volume of 10 to 15 µL (for 150-500 ng scale), with 1 µL of T4 DNA ligase (NEB). The reactions are incubated at 16 °C (in the hot block in the deli case) for 16 hours. The reactions can be spiked with ligase and ATP, if necessary. These reactions can be transformed directly into cells as described below.

H. FastLink ligation kit
(Time estimate: 1 hour)
Ligations can be performed with the Epicentre Fast-Link ligation kit as described in the manufacturer's protocol. If cloning libraries, a number of vector:insert molar ratios should be tried, to determine which ratio gives the highest ligation efficiency, and hence, the largest number of transformants. The ligation reaction should be assembled as described in a total reaction volume of 15 µL, with 1 µL of Fast-Link DNA Ligase. The reaction is incubated at room temperature for 5 to 15 minutes and then incubated at 70 °C for 15 minutes to heat-inactivate the ligase. The reactions can be transformed directly into cells as described below.

I. Transformation and Plating
(Electroporation, Time estimate: 2–2.5 hours including plating + overnight for plates)
Generally, 1 µL of each ligation reaction is transformed into 40 µL store-bought XL1Blue cells (or any other electrocompetent cells). The cells are thawed on ice, the DNA is added, and the mixture is stirred gently with a pipette tip. Cells and DNA are incubated together for ~15 minutes before transformation. Electroporation cuvettes should be chilled on ice. The cells are added to the electroporation cuvette, and the cuvette is placed in the electroporation safety stand, with the shield down. XL1 Blue cells should be pulsed at 1.38 kV. Push the charge button, and wait until the green light stops flashing. Proceed immediately to recovery (below). Be sure to press the reset button between transformations. The display should show a value around 4 while charging. However, sometimes our electroporator shows a negative number and cannot be reset…when this happens, the transformation should still work (it's a display problem, not a pulsing problem). If the electroporator sparks while pulsing, this means that your mixture of DNA and cells is too salty. The DNA may need to be cleaned up more before you attempt another transformation. Note that any healthy cells will be killed when it sparks, so those cells should be discarded. It is always a good idea to do control transformations along with transformations of your DNA. Water can be transformed in place of DNA for a negative control, and another plasmid (such as pUC) can be transformed as a positive control.

Cells are recovered by immediately adding 960 μL of any rich media (+ glucose for phagemids) lacking antibiotics to the cuvette. The cells can be transferred to 5 mL falcon tubes and are incubated for 1 hour with shaking at 37 °C. Plate 50-100 μL of the recovered cells (neat) and multiple dilutions made in media without antibiotics (10^{-1} to 10^{-5}) on appropriate agar plates (2xYT-AG for phagemid vectors, LB Amp for most other plasmids) to ensure that that you will be able to pick single colonies and/or determine the number of transformants. Plates are incubated at 37 °C for 12-16 hours (not longer!). The remaining cells are mixed with 500 μL of 50% glycerol and frozen on dry ice.

J. Heat Shock
(Time estimate: 1.5-2 hours including plating + overnight for plates)
Heat shock competent cells, including BL21(DE3) and many other strains, are first thawed on ice. 1 μL of DNA is added to 20 μL cells and mixed by gentle stirring with a pipette tip. The cells and DNA are incubated for 5 minutes on ice, and then heated to 42 °C for 30 seconds **exactly**. The cells are incubated for 2 additional minutes (**exactly**) and 80 μL of any rich media (without antibiotics) is added to the tube. The cells are recovered at 37 °C for one hour, with shaking. Plating and controls are performed as described above for electroporation.

K. Screening and sequencing
(Time estimate: Overnight + 1 hour for minipreps + 1 hour for digests + 1.5 hours for agarose gel)
Screen around 10 colonies for single sequences, or at least 20 for libraries, off of the plate with the highest ligation efficiency. Grow 5 mL overnight cultures in appropriate media, each inoculated with a single colony, for 15-16 hours. Minipreps of 2 mL of each overnight culture should be sufficient for screening and sequencing. Two miniprep kits popular in the lab are the Promega Wizard kit and the Qiagen spin kit. Both are easy to use, and the manufacturer's directions can be followed exactly. Two digests of each clone should be performed: one with an enzyme whose site is found in the new insert but not in the vector (this enzyme should cut your clones!), and another with an enzyme whose site is found in the region of the vector cut out, but not in the new insert (this enzyme should not cut your clones!). The digests are performed in a 10 μL volume, with 1 μL enzyme and 5 μL miniprepped DNA for one hour, and then loaded on a 1% agarose gel. Clones that show the correct pattern of digestion can be sent to Keck for sequencing. For some people, sequencing is more successful when they follow the Keck guidelines for sample preparation exactly. Other people find that, in their hands, other amounts of DNA are more likely to be sequenced cleanly. I use 15 μL miniprepped DNA, 7 μL dH$_2$O, 2 μL μM primer.

L. Cloned Libraries
The number of transformants needs to exceed the theoretical diversity of your library. To calculate number of transformants, determine the best plate for counting colonies and multiply:
(# of colonies) ✧ (dilution) ✧ (1000/how much you plated).

For example, if you have 46 colonies on the 10^{-4} dilution plate (and you plated 100 μl), you have 4.6 x 10^6 transformants / 1 μl of ligation. If your diversity is 3.2 x 10^7 (5 residues randomized), you will have to transform multiple aliquots of your ligations and possibly do multiple ligations. To guarantee 90% completion of your library, you need a number of transformants that is at least 2.3 times the theoretical diversity.

M. Preparation of cells for phage display experiments
(Time estimate: Overnight + ~5 hours for maxiprep + 0.5 hours for aliquoting)

The recovered cells from the required number of transformations are combined and grown overnight in a large volume of 2XYT-AG (the volume required will depend on how many transformations you need to combine). You want to have a healthy culture of cells in the morning. From this culture, you can maxiprep your library. In addition, you should make multiple (~30) glycerol stocks (1 mL cells + 500 μL 50% glycerol; freeze on dry ice). These aliquots should each contain multiple copies of every transformant and can be used directly in phage display experiments.