Solutions for Chapter 2

1. 1 mol of 1 bp weighs 660 g. So a 2 kb plasmid is 1320 µg/nmole, or say approximately 1 mg is 1 nmole of plasmid, which is 6 x 10^{14} molecules. Say 1000 plasmids grow in each cell, so it's 6 x 10^{11} cells needed, or ~ 10^{12} cells for a 50% yield. For optimal plasmid recovery, cells are grown to just past log phase, or 3-4 x 10^{9} cells/mL. Should get 1 mg from 250-300 mL of culture.

2. The conversion factors are:
Average dsDNA MW = 660 x N (where N = number of base pairs)

C (µM or pmol/µl) = A260 / (0.01 x N)
C (ng/ml) = (A260 x MW) / (0.01 x N)

For RNA, A260 reading of 1.0 is equivalent to about 40 µg/ml

3. 1 in 4096 (so 0.5-1 for an average plasmid); 1 in 65536

4. Use some uncut plasmid as a control that is transformed and plated next to your ligation. When plating, make sure to dilute the uncut plasmid enough to be able to count the colonies.

5. The smallest NcoI fragment is barely visible, so we can call this the limit. Assume it's a typical screening gel from a miniprep, so you've cut ~200 ng of DNA total. It's about a 250 bp fragment out of a 4 kb fragment, so the amount of DNA in that band is ~12.5 ng. This is a reasonable value for a limit.

6. (1) Some uncut plasmid could be left in and transform with very high efficiency. The control with no insert and no ligase controls for this. If anything grows on this one, throw the ligation away and try again. (2) The plasmid could close back in on itself. This can even happen with incompatible sticky ends, especially if they have short overhangs. The no-insert control controls for this.

7. 3:1 is a good compromise to maximize the chance of the insert finding a vector while still making the probability of getting two inserts per vector very small. **(1)** Rarely, there may be more than one insert lined up in the final product. This is hard to control for; gel screening may show it. The frequency should be roughly the frequency of expected successful ligations^(number of inserts), so extremely rare if successful ligations are 1 in 10^{6}! (2) The plasmid may recombine to give a result much smaller than expected, without

the insert. Gel screening will reveal a vector that's too short and/or no insert. The likelihood of this depends upon the presence of homologous sequences in the plasmid. Sometimes the frequency of this type of recombination is close to 1, and E. coli defective in recombination may need to be used. (3) Part of the insert could be edited out. Gel screening will show an insert that's wrong. The frequency depends upon the properties of the insert; it's more likely with very long or repetitive inserts.

8. Blue/white selection works when an insert is cloned into a polylinker that is deliberately placed inside the lacZ coding sequence. If an insert is cloned, the coding sequence is disrupted and no β-galactosidase will be produced. If there is no insert, the β-galactosidase is made. β-galactosidase turns the colorless substrate X-gal from clear to blue. So no insert = blue colonies. Insert = white colonies.

9. (A) Pick a few colonies and grow in 3-5 mL of LB-Amp, then miniprep. (B) Restriction digest screening of minipreps using XbaI/XhoI. (C) There may be no promoter in the vector, or at least no mammalian promoter. Alternatively, the gene might not express well in mammalian cells and may need to be codon optimized. (D) First lane: the digest just cut out CFP, so you're good to go. Second lane: the digest cut a piece out of the vector. As long as you can separate those last two bands and cut out the one corresponding to CFP, you're good to go. You probably want to run the gel longer than was done in the example to better separate the bands. Third lane: The vector appears to be linearized, but the gene didn't cut out. You're in trouble unless one of your enzymes has gone bad and you can replace it. Fourth lane: XbaI and XhoI cut in a lot of places, but you should be able to isolate the gene. For all except lane #3, your cloning strategy should be to put the isolated gene into a vector that you understand, which has a mammalian promoter. (E) If the gene doesn't cut out with XbaI and XhoI, you have no real choice but to PCR it out. The sequence should be in GenBank and you can sequence the ends of the gene to figure out what lies on either side, then develop PCR primers to flank the gene; or you can use the known sequence to develop primers for precisely the gene ends with no flanking sequence.

10. The simple answer is to knock in a growth gene using a tooth-specific promoter. More advanced students will provide ideas of how to deliver the gene (e.g. stem-cell recombination, transfection of spermatozoa, etc).

11. Get the gene sequence from the most closely related possible cat (perhaps a lynx or lion). Extract mRNA from the smilodon muscle sample and use the lynx cDNA as probe on either Southern or Northern blot to verify it will bind. If successful, use it to probe a Smilodon cDNA library, pull out ryanodine receptor cDNA and sequence. Verify that you have the right gene by comparing to the sequence from the lynx.

12. The sequence may be found at GenBank: GQ354859.1

1 cgtcctcctt attattcaga agagatgttt gaacaagatt ttgaaaaacc caaagaaaaa 61 gatcctccta aagcagagag tccagccact gaaccctcgg gtaattcaac aggtcctgag 121 actaattcta ctcaaccaaa tgcacctgga gggaatcagg gcggaaatga caccagccca 181 acaggaaaca gtggccatgg cccaaacact gtgagtaatc ctacagctca aaacggggtt 241 atctcacccc ctacagttaa tgtttcaggc cagggagtac caagaactca aatctcatgg 301 ggaccaagtc agccaaatat ttatgaaaac tatccaaatc ctaacatccg aaattttcct 361 gcaggaagac aatggcgtcc cactggtact atcatggggc gtagacagaa tgggcctttt 421 tacagaaatc aacaggtcca aaggggtcct cggtggaact cctttgcttt ggaaggcaaa 481 caagcatttc gttcaggaaa tccaatttat cgcaaagctt atgcttctac tgcaagaggc 541 aattctccca ctcatgcagg aaatccagga aatgtcagaa gaaagcctca gggaccaaat 601 aaacacccta tgggaaccaa cgttgctcct ctgggtccca aacatggtac tgttgtccgc 661 aatgaaaaaa tccaaaatcc agcagagaaa ccagtaggtc caaaagaaag aatagtcatt 721 cctacaagag atccatctgg cccctggaga aactctcaag attatggagt taataaatca 781 aactataaac tgcctcatcc tgagagtaac atgctagtcc caaattttaa ttctgttggt 841 caacatgaaa actcttatta cccaagagga gagtccaaaa gagccccaaa ttctgatgga 901 caaacccaaa gccagaattt gcccaaaggg attattttgg agccaagaag aatcccttat 961 gaatcagaaa ctaatcgacc agaattaaag cacagtacat atcaacgtgt ataccctgag 1021 gaaatccctt ccactgcaag agaacatttt cctgctggaa gaaatatttg gaatcatcaa 1081 ggaatctctc cacctattaa ggaagatccc gggatgcagg aagaacactt acttcatcct 1141 tcccatggct caaggggagg tgtttactac cctgactata actcttataa tcccagggaa 1201 aactcaccat accttagaag caatacatgg gatgagagag atgattctcc caatactatg 1261 gggcaaccca aaaattcact gtaccccatg aatactccag acctgaaaga aacagtccct 1321 tataatgaag aggacccaat tgatgcaact ggagatgaaa cttttccagg acaaagtaga 1381 tggggtatgg aggagccaag ctttaaagaa gggccaacag ttaggcacta tgaaggcgag 1441 caatatacct caaatcaacg gaaagaatac cttccctatt ccttagataa tccatcgaaa 1501 cccagggagg atttccctta tggtgaattt tatccctgga accaagatga gaattttcca 1561 tcatataata cagctcccag tgcaccacca cctgtggaga gcaggggcta ttatgctaat 1621 aatgctgttg gacaggaaga aagcactctg ttcccctctt ggaactcctg ggaccatgag 1681 attcaagccc aagggcagaa agaaagaaga ccatatttta acagaaactt ctgggatcga 1741 ccaacaactt tacacaaagc tccccctagt ccaccacatg agaaagagaa ccagccttat 1801 cctagcaatt ccccagctgg gcttcagaaa agtccaacat ggcatgaaag tgagaatttg 1861 aattatggca tgcacattac taggataaat tcaccagaga gagaacagtt ggctttccca 1921 gacttaattc ctccaagtta cccatcaggt caaaaagaag cgcatgtatt tcatctaagc 1981 cagagaggcc cttgctgtgc tggcagctcc acaggacaca aagaagatcc acttgctcta 2041 caggactaca ctccatcctt tggtcttgca ccaggggaga accaagacac cagtcctccg 2101 tatacagaag atagtcatac

taaacatgca agacacacca tctcccctcc aagcaaccta 2161 cctggccaaa gaaacagctc agagaaaaga ctgcctggag aaagtgaaaa cccaagtcct 2221 tttagagatg atgtgtccac cctgaggagg aacacaccat gttccatgaa gaatcagctg 2281 agccaaaggg gaattatgcc ctttcctgaa gccagttccc ttcaatcaaa gaatatgcct 2341 tgcctcaaaa gtgatcttgg aggagatggg gacaatgttt tagaacaaat atttgaaggc 2401 aaccagctca gtgaaagaac tgttgacctt actcctgagc agcttgttat gggtacacct 2461 gaagaagccc ctaatccaga aggaatccaa agtgaagtac aaggaaatga gggtgaaagg 2521 cagcaacaaa gaccatcaag catcctgcag ttaccatgct ttggctccaa attggcaaag 2581 tatcattcct ccagcactgg aactccatct ggcattggaa ggcaaggccc atttgatgaa 2641 gatccgatta cgcctactga aagtcctaac tcattgtcta ggttagctac tggggcacag 2701 tttcagagta taaatgttga cccacttaat gcagatgaac acatttcatt tgattccctt 2761 caaataga

The primers should be similar in Tm and have the appropriate BamHI overhangs.

13. Use either *in vivo* transfection or a viral vector with a strong promoter such as CMV. Inject into your skin in the desired pattern. Don't inject the GFP protein; it won't be stable.

14. There are two parts to the pET expression system. First, the plasmid contains YFP-DNA cloned downstream of a T7 promoter site. This plasmid also contains amp resistance and origin of replication. The host cell has a modified genome such that the T7 RNA pol gene is inserted into its chromosome and can be activated with the addition of IPTG. When IPTG is added, T7 RNA pol is made, binds to the promoter upstream of YFDNA and transcribes it. The cell then translates it and you have tons of YFP.

15. Get answers from the NEB catalog section on "compatible cohesive ends." Neither Sal nor Xho will recleave the first ligation. Xho will recleave the second one.

16. Restriction enzymes are found in all bacteria, and sequence specificities cover just about every possible 4-6 bp sequence. These enzymes protect bacteria against foreign DNA.

17. May be obtained from manufacturers of plasmid-prep kits.

18. Steps as in Fig. 2.10. Before the discovery of the heat-stable polymerase from *T. aquaticus*, polymerases would be destroyed at each denaturation cycle. This meant that someone trying to run a reaction would have to add fresh polymerase each cycle.

19. Called the Ames test. Bacteria that require histidine are used, and mutations make them revert to a state where they do not require histidine.

20. Find sequence on GenBank. Create mutagenic primers. PCR.

21. Answers will vary.

22. Mechanosensitive channel from Mycobacterium tuberculosis. Protects cells from osmotic shock. Restriction sites: BamHI, SspI, FspI.

Use an online codon optimizer to optimize for mammalian expression. Human optimization:

ATG CTG AAA GGC TTC AAA GAA TTT CTG GCC CGA GGA AAT ATT GTG GAC CTG GCG GTA GCC GTT GTA ATA GGA ACC GCG TTT ACT GCG CTG GTA ACA AAA TTC ACG GAT AGC ATC ATC ACC CCT TTG ATA AAT AGG ATA GGA GTG AAC GCG CAA AGC GAT GTG GGG ATC CTC CGC ATA GGC ATA GGA GGT GGT CAG ACT ATT GAC CTT AAC GTA CTT TTG TCA GCG GCG ATC AAT TTC TTC CTC ATA GCC TTT GCC GTT TAT TTT TTG GTA GTT CTC CCC TAT AAC ACC TTG CGC AAG AAA GGA GAG GTC GAG CAG CCG GGA GAC ACA CAA GTT GTC TTG TTG ACT GAG ATC AGA GAT TTG CTG GCA CAA ACG AAT GGG GAT TCA CCA GGG AGG CAT GGA GGA CGA GGA ACA CCG TCA CCT ACA GAC GGA CCC AGA GCT AGT ACA GAG AGC CAA TAG

23. TOPO cloning is a ligase-free method of cloning that takes advantage of the property of Taq polymerase that it adds a single deoxyadenosine (A) to the 3'-end of the PCR products. TOPO vectors are special plasmid vectors that are designed to carry the sequence 5´-(C/T)CCTT-3' at the two linear ends. The linear vector DNA already has the topoisomerase enzyme covalently attached to both 3' ends. This is then mixed with PCR products. When the free 5' ends of the PCR product strands attach to the topoisomerase/3' end of each vector strand, the strands are covalently linked by the topoisomerase.

24. N terminal fusions are usually the safest in terms of protein expression. If protein is desired, then the fusion partner with secretion signals must be on the N terminus. N terminal fusions avoid putting foreign sequences next to translation signals, which can inhibit translation. C terminal fusions can work fine in many cases as well. Trying to put the GFP protein on an extracellular part of an ion channel will almost certainly not work because it will have to pass through the membrane. I know of only one group that claims to have got this to work (Verkman, UCSF). Most N-and C-terminal fusions can be done with simple restriction enzyme cloning, but make sure that the two parts are in frame. Getting them out of frame leads to truncation or just bizarre results. Trying to put the GFP anywhere but the N or C terminus will take PCR cloning. Even if the sequence is perfect, things that can go wrong include the protein getting hung up in the endoplasmic

reticulum or forming inclusion bodies, or simply failing to fluoresce. This is usually obvious from looking at the cells.

25. Either increase or decrease annealing temperature. Increase can prevent formation of nonspecific products; decrease can reduce issues such as primer dimer.

26. Disadvantages: Short product lengths (< 1 kb). Low product amounts. Errors. Advantages: Simple, fast, and robust.

27. The main disadvantage is that it requires the extraterrestrial life to have some universal sequence, such as ribosomal RNA. For a close by planet such as Mars, that is probably a safe assumption. Life cannot have been extinct for more than thousands of years in order to get a signal. The possibility of contamination must be very carefully controlled for, and the experiment must have the ability to reject known Earth sequences. Within these limits, the experiment can be low-cost, robust, and very sensitive.

28. If not inactivated, the enzyme will cut the ligation product, thereby undoing the ligation (unless of course the resulting product doesn't cut with the same enzyme, in which case it could be good to have the enzyme still in there!) Enzymes may be removed by running the ligation on an agarose gel or using a clean-up kit. A list of temperatures is given here: https://www.neb.com/tools-and-resources/usage-guidelines/heat-inactivation

29. The accelerating force on the molecule is ZeE. The retarding force is given by the Stokes formula and is related to the medium viscosity and particle shape; it can be expressed as fv where f is the frictional coefficient and v is the speed. Setting the two forces to balance gives Eq. A1.

30. "Dummy" nucleotides are used to stop synthesis at a specific type of base pair (ACGT). Then the fragments are separated by size, using very long polyacrylamide gels to obtain single base pair resolution.