PROBLEM SET 4: BALANCE AND STABILITY

1. A single substitution to a polar or charged residue at any of 30 sites in the core of protein X produces a non-functional protein. The amino acid replacement rate is 10^{-6} replacements per site per generation, and 40% of all possible replacements at these sites are to polar residues. If the allele is recessive and the selection coefficient against the double null homozygote is 0.1, what is the expected frequency of the null allele in the population? (Assume there are no back-mutations from polar to hydrophobic.)
2. Familiarly, a bacterium's instantaneous growth rate r depends upon the expression of protein P, an enzyme which degrades an abundant environmental toxin. For this bacterium, fitness $w = e^r$. The equilibrium cellular concentration of P is $[P_0]$, but only the folded population of the enzyme is active, so that $r = \alpha[P]_0 \Pr_{\text{fold}}(\Delta G_u) k_{\text{cat}}$ where α summarizes proportionality between enzymatic activity and growth rate. For convenience, we (again) collapse all the constants into $\beta = \alpha[P]_0 k_{\text{cat}} = 1h^{-1}$.

   At the site encoding the surface residue 15 in the gene encoding P, a still-unusual and substantially more promiscuous mutational process turns each amino acid's codons into those of any other, such that at the protein level, all amino acids mutate into each other without bias. The wild-type enzyme has A at site 15, and this allele has stability $\Delta G_u = 5$ kcal/mol.

   We want to determine the steady-state frequency of amino acids at site 15.

   (a) (3pt) Suppose the 19 possible substitutions have $\Delta\Delta G_u$ values ranging from -0.5 to -9.5 kcal/mol in alphabetical order of the one-letter code, excluding A, equally spaced in units of 0.5. Plot the expected frequency of each amino acid for a population of $N = 10$, $N = 1000$, and $N = 10^6$, with lines. Suggestion:

```
# split a string into characters
charlist <- function(x,sep='') {
  t(sapply(x, function(m) {unlist(strsplit(m,sep))})) }
#
pfold <- function(dGu) {
  # you can crib this from earlier...
}
# Equilibrium fixed levels, given many competing alleles.
eq.fix <- function(ddGu.i, dGu, N){
  r.i <- # What is growth rate?

  # equilbrium proportions, given the log fitnesses r.i
  eq.props <- sapply(r.i, function(r) {
      # r will take each of the r.i values in turn.
      # the current code is wrong; you need to
      # replace it with the actual calculation of
```

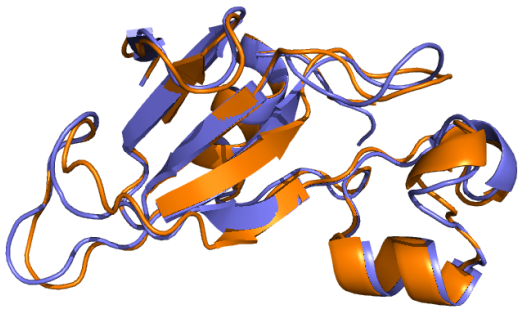
```
      # equilibrium proportions, based on what was
      # covered in class.
      r - mean(r.i)  # THROWAWAY
      })
  return(eq.props)
}

dgu.wt <- 5.0
ddgus <- seq(0,-9.5,-0.5)
# Calculate results
n10 <- eq.fix(ddgus, dgu.wt, 10)     #  N=10
n1000 <- eq.fix(ddgus, dgu.wt, 1000) #  N=1000
n1e6 <- eq.fix(ddgus, dgu.wt, 1e6)   #  N=1M

aas <- charlist('ACDEFGHIKLMNPQRSTVWY')
# Plot results
# 1:20 puts each amino acid at the numbers 1 through 20
plot(1:20, n10, col='blue', type='l', ylim=c(0,1),
     las=1, log='', xaxt='n', xlab='',
     ylab='Frequency of allele')
lines(1:20, n1000, col='red')
lines(1:20, n1e6, col='black')
mtext(aas, 1, at=1:20, line=1)
```

(b) (1pt) Comment on the effect of changing effective population size on the steady-state amino-acid distribution. What effect does increased N have on conservation of the wild-type allele?

3. Barnase and binase have quite similar structures. Download their coordinates (`http://pdb.rcsb.org` and search for `1BUJ` and `1A2P`).
   (a) (1pt) Align the backbones of these two proteins in PyMol. What is the RMSD? Compare this RMSD to the size of a single hydrogen atom.



(b) (1pt) What is the percentage of solvent-exposed residue surface area in barnase for the site of the most, and the least, destabilizing substitution reported by Serrano *et al.* in their Table 3? Note that this does not involve PyMol.

(c) (2pts) Most amino-acid substitutions are destablizing, or so we often hear. According to Serrano *et al.*, roughly half of the barnase/binase mutations are stabilizing. Indicate and explain your level of surprise.