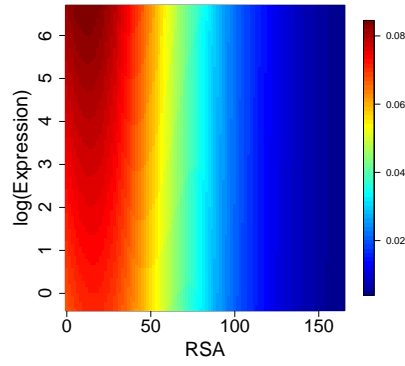


Supplementary File: Roles of solvent accessibility and gene expression in modeling protein sequence evolution

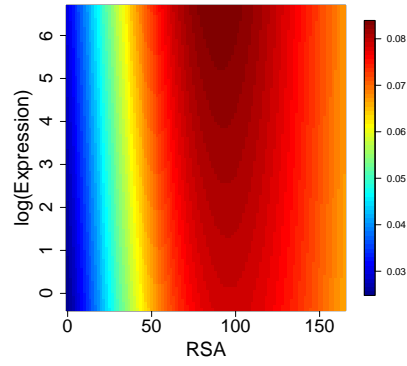
February 4, 2015

Supplementary Figure 1: Predicted probabilities of 16 amino acids across possible ranges of RSA and gene expression

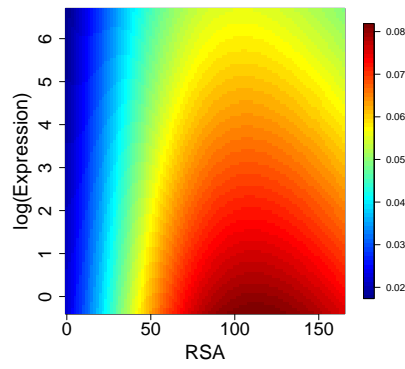
Probabilities for each amino acid were calculated from the estimated RSA and gene expression coefficients (see Table 5). The horizontal axis covers the possible range of RSA and the vertical axis is for gene expression. Red regions indicate relatively high probabilities, while blue regions show low probabilities. The 16 amino acids depicted here are those that were not depicted in Figure 3.



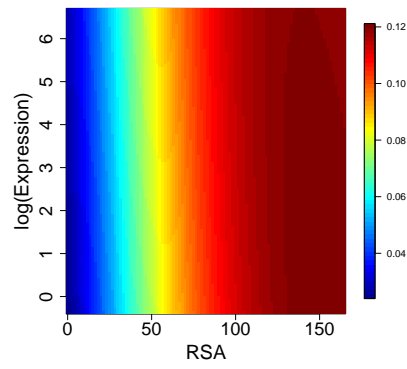
(a) Alanine (A)



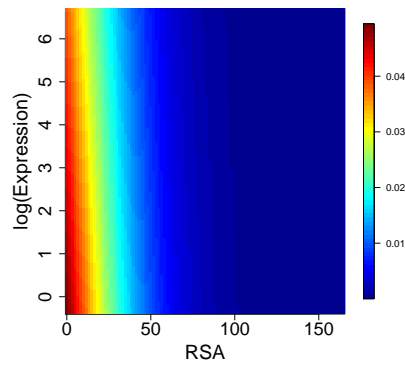
(b) Arginine (R)



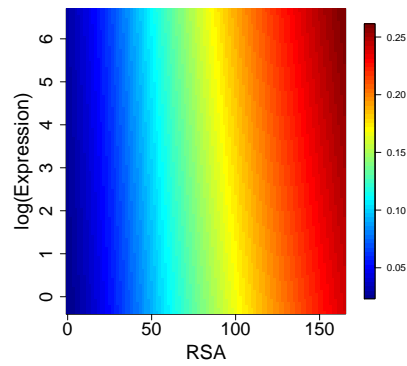
(c) Asparagine (N)



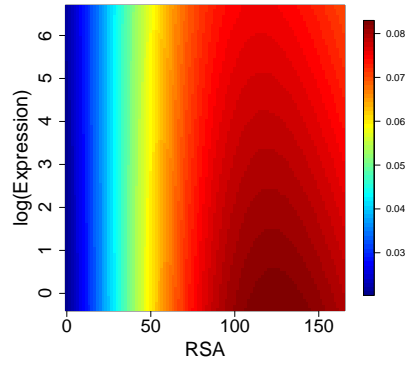
(d) Aspartic acid (D)



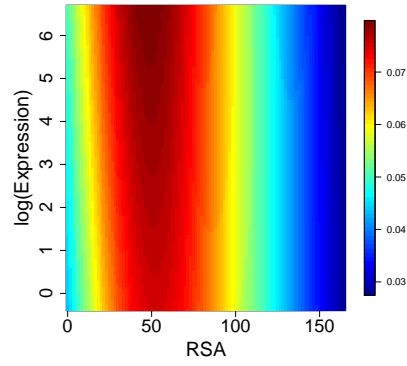
(e) Cysteine (C)



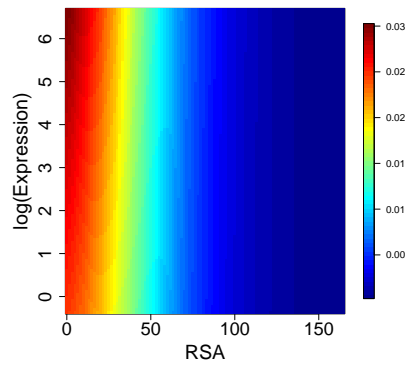
(f) Glutamic acid (E)



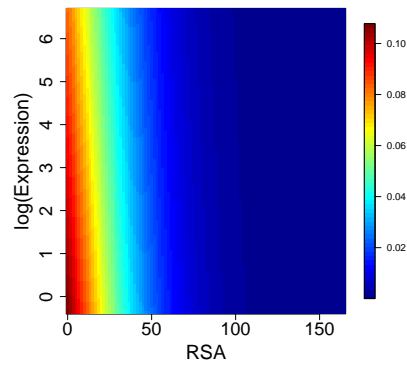
(g) Glutamine (Q)



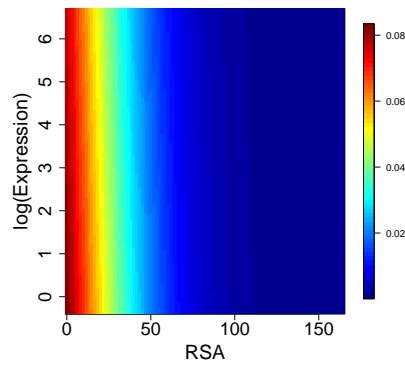
(h) Glycine (G)



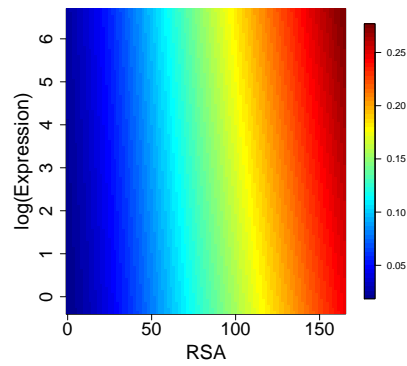
(i) Methionine (M)



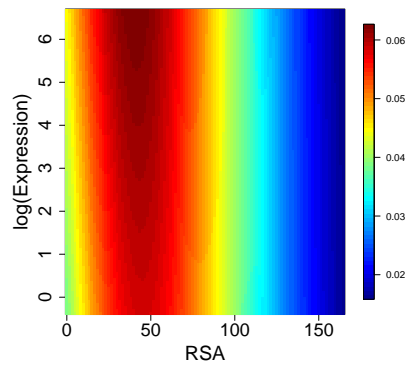
(j) Isoleucine (I)



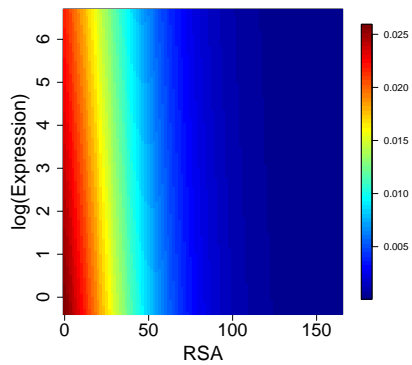
(k) Phenylalanine (F)



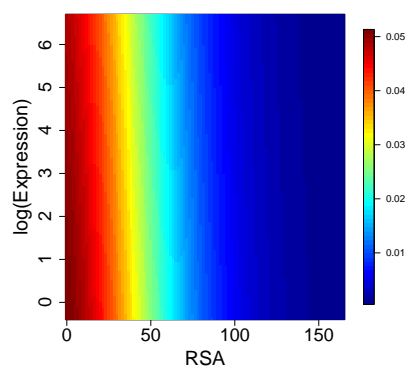
(l) Lysine (K)



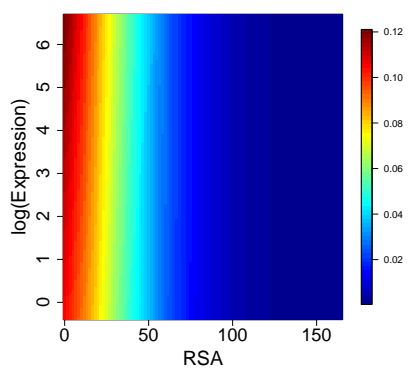
(m) Threonine (T)



(n) Tryptophan (W)



(o) Tyrosine (Y)



(p) Valine (V)