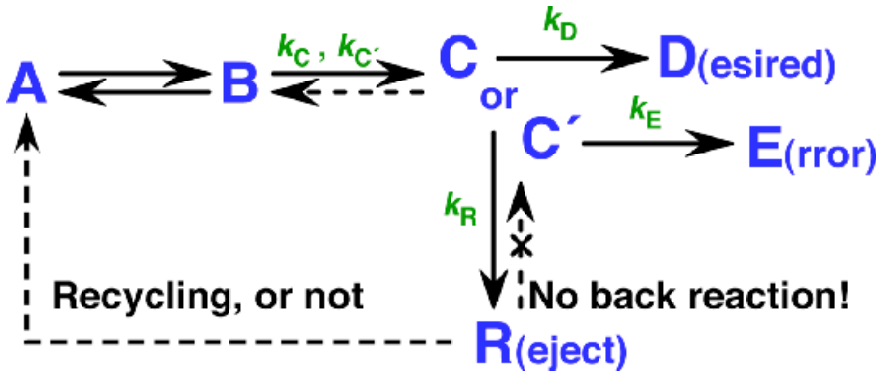
Proofreading with Kinetic Partitioning

I. General Concept

Biochemical systems have evolved to **spend free energy to increase fidelity**, typically by hydrolyzing ATP or GTP over and above what is required for the chemical transformation in question. **Consider the general reaction scheme below:**



The idea is that the system seeks to convert the C intermediate to desired product D, while converting the C´ intermediate to reject product R instead of converting C´ to error product E. There is probably some discrimination at the step of converting B to C or C´, which will make [C] > [C´]; the goal of proofreading is to improve upon this level of intrinsic discrimination. To the extent that B → C or C´ is reversible, the rate of production of D and E in the absence of any rejection depends on the overall net rate constant from B to D vs. E (we will skip the math, for which see Cleland, 1975, but the net rate constants do depend on k_D and k_E). This is **kinetic partitioning**, the idea that for irreversible transformations the product distribution depends on the kinetics of product formation, not the thermodynamics. **Proofreading** is a particular example of kinetic partitioning, in which free energy is dissipated in the process of increasing fidelity over what is achievable through discrimination in a non-branched pathway. Proofreading requires a rejection path: if the B → C or C´ step is irreversible and there is no rejection path, then the ratio of yields of D and E is independent of the rate constants k_D and k_E.

The main idea here is that **the existence of a rejection pathway allows the system to take full advantage of the different in rate constants k_D and k_E.** There may or may not be discrimination built in to the rejection step; here we assume there is none. **The rejection step must be irreversible**, otherwise the error rate is increased by conversion of R to C´, short-circuiting any discrimination from the A ⇔ B → C or C´ steps. **Free energy is dissipated in ensuring this irreversibility of the discrimination step, but the actual chemical step that generates the needed free energy need not be the rejection step**: C and C´ may be high-energy intermediates because of previous steps. This is very general in ATP-consuming reaction cycles: the actual hydrolysis may or may not be the irreversible step in the cycle.

II. Application to Biological Information Processing

The general reaction scheme above can be applied to replication, transcription, and translation.

For DNA or RNA polymerization we have:

C = correct incorporation at the primer terminus (length *n*)
C´ = a mismatch incorporated at the primer terminus (error frequency ~ $1/10^5$)
D = extended correct product (length *n+1*)
E = extended mismatch (i.e. keeping the mistake)
R = shortened DNA primer (length *n-1*)
k_C or $k_{C'}$ = nucleotide incorporation; all the finger closure etc. is here
k_D = extension of complementary terminus
k_E = extension of a mismatch
k_R = exonuclease reaction to remove 3´ terminus
Free energy is dissipated in the rejection step, because the hydrolysis of a phosphodiester to give a dNMP product is irreversible. The exo is not very selective because its rate is presumably determined by fraying of the ends of the primer-template, which is not a high-fidelity process.

For translation, we have:

C = cognate tRNA bound at the A site of the ribosome
C´ = near-cognate tRNA partially bound at the A site
D = correct amino acid incorporated at C-terminus
E = incorrect amino acid incorporated at C-terminus
R = ribosome with empty A site ready to bind a new tRNA.
k_C or $k_{C'}$ = GTP hydrolysis by EF-Tu during insertion of tRNA
k_D = accommodation of cognate tRNA, GDP release, eventual correct peptidyl transfer
k_E = accommodation, GDP release, and eventual incorporation of near-cognate tRNA
k_R = GDP release and ejection of charged tRNA to allow for binding a new one
Free energy is dissipated by GTP hydrolysis to GDP + P_i in the B $\rightarrow$ C or C´ step. Even though no covalent bonds are broken in the rejection step it is irreversible because the C or C´ intermediate is produced extremely slowly in the absence of free energy input from GTP hydrolysis. The ribosome burns a GTP at every step to increase fidelity.
Once the wrong amino acid is added to the C-terminus, there is no way to remove it.
We are ignoring all of the proofreading that goes into charging of tRNA with the correct amino acid.

III. Mathematical Analysis

The idea of kinetic partitioning is that the product distribution between D and E depends on the ratio of rate constants (not the thermodynamics) corresponding to the different possible outcomes. For example, C can be converted to D or to R. The yield of D is then given by the rate constant for the pathway leading to D divided by the sum of rate constants for all processes that deplete C. Assuming the rate constant for the rejection step is the same for C or C´, we have:

$$\text{Yield of D} = \frac{k_D}{k_D + k_R}[C] \quad \text{and} \quad \text{Yield of E} = \frac{k_E}{k_E + k_R}[C']$$

So

$$\frac{\text{Yield of D}}{\text{Yield of E}} = \frac{1}{\text{Error frequency}} = \frac{k_D}{k_E} \cdot \frac{k_E + k_R}{k_D + k_R} \cdot \frac{[C]}{[C']}$$

We can see that the error frequency depends on both the relative rates of conversion of C or C´ and also on the rate of the rejection reaction.

Let's look at three limiting cases:

If $k_D \gg k_R$ and $k_R \gg k_E$, the optimal situation, we have

$$\text{Error frequency} = \frac{k_E}{k_D} \cdot \frac{k_D + k_R}{k_E + k_R} \cdot \frac{[C']}{[C]} \approx \frac{k_E}{k_D} \cdot \frac{k_D}{k_R} \cdot \frac{[C']}{[C]} = \frac{k_E}{k_R} \cdot \frac{[C']}{[C]}$$

Kinetic partitioning of the C´ intermediate between acceptance and rejection has improved fidelity at modest cost.

If $k_R \gg k_D$ and $k_R \gg k_E$, i.e. a very fast rejection path, we have

$$\text{Error frequency} = \frac{k_E}{k_D} \cdot \frac{k_D + k_R}{k_E + k_R} \cdot \frac{[C']}{[C]} \approx \frac{k_E}{k_D} \cdot \frac{k_R}{k_R} \cdot \frac{[C']}{[C]} = \frac{k_E}{k_D} \cdot \frac{[C']}{[C]}$$

The error frequency depends only on relative rates of conversion of C and C´. This may have very high fidelity, but at a high cost in speed and/or ATP.

If $k_D \gg k_R$ and $k_E \gg k_R$, i.e. if there is no rejection path, we have

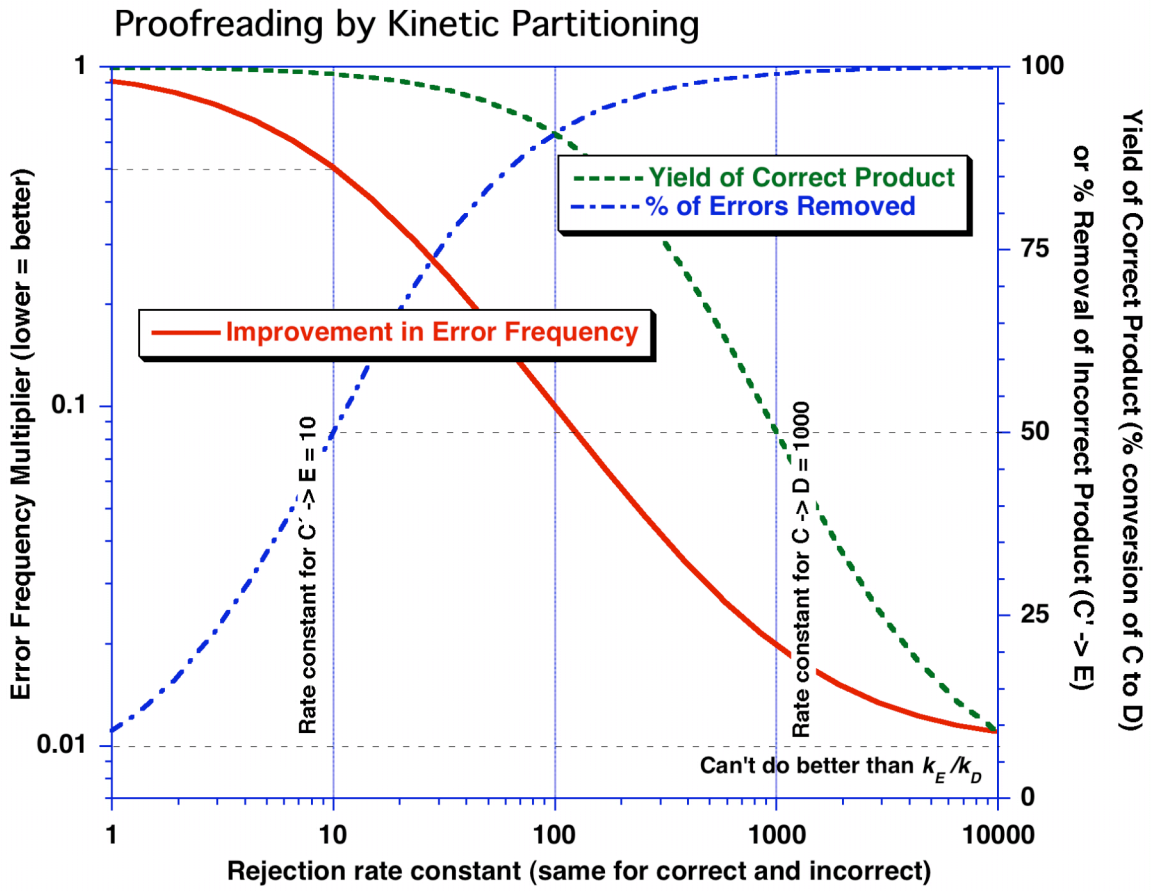
$$\text{Error frequency} = \frac{k_E}{k_D} \cdot \frac{k_D + k_R}{k_E + k_R} \cdot \frac{[C']}{[C]} \approx \frac{k_E}{k_D} \cdot \frac{k_D}{k_E} \cdot \frac{[C']}{[C]} = \frac{[C']}{[C]}$$

There has been no improvement in fidelity.

The other way to think of this process is that **the error rejection pathway sets a molecular clock**: the half-life of the rejection reaction is $\tau_R = (\ln 2)/k_R$, and if this is much less than the time required for the forward reaction $\tau_E = (\ln 2)/k_E$, then the intermediate will be rejected. The ATPase- or GTPase-dependent molecular clock is also very general, viz. G-proteins.

IV. Graphical Analysis

The plot below shows the improvement in error frequency (the factor by which the error frequency is multiplied) due to proofreading, as a function of the rate constant for the error rejection pathway. For the purpose of this illustration we assume that $(k_E/k_D) = 100$ and that the rejection pathway is non-selective. You can see that for very slow rejection, there is no improvement in the error frequency. This is a sensible limiting case. When the rate constant of the error rejection pathway equals the rate constant of the incorrect incorporation pathway, we have a 2-fold improvement in error frequency as expected; a factor of 2 in this business is rather underwhelming. The optimum situation is when the error rejection rate constant is roughly the geometric mean of the correct and incorrect incorporation rate constants: we see that the error frequency is 10-fold lower at a cost of removing only 10% of correct incorporations. When the rejection rate constant equals the correct incorporation rate constant, half of the correct product is destroyed, and we can reach the maximum possible fidelity improvement (k_E/k_D) only with a maximally wasteful rejection pathway.



The bottom line:

There ain't no such thing as a free lunch!