Eur Biophys J (1985) 12:115–119

European Biophysics Journal

The tRNA cycle and its relation to the rate of protein synthesis

H. Liljenström*, G. von Heijne, C. Blomberg, and J. Johansson

Research Group for Theoretical Biophysics, Department of Theoretical Physics, Royal Institute of Technology, S-10044 Stockholm, Sweden

Received September 20, 1984/Accepted in revised form February 28, 1985

Abstract. With the aid of a kinetic model, we have investigated how the adaptation between the various components of the tRNA cycle and the codon frequencies affects the rate of protein synthesis. Depending on the relative amounts of total tRNA, synthetase and ribosomes, the optimal correlations vary between a situation where all tRNA species are either present in equal amounts or are present in amounts proportional to the square-root of the corresponding codon frequencies, and a situation where the amounts of the different tRNA species present are linearly proportional to the codon frequencies.

Key words: tRNA cycle, optimization, elongation rate, codon frequency, tRNA synthetase

Introduction

It seems quite obvious that one of the main aims of prokaryote cells is maximum growth rate. The efficiency of the translation apparatus is one key to accomplishing this aim. Recent work (Blomberg 1983; Kurland and Ehrenberg 1984; Ehrenberg and Kurland 1984) has emphasized the importance of applying theoretical optimization principles to cellular systems with respect to speed, accuracy and free energy dissipation. For instance, it has been shown that the cellular growth rate is critically dependent upon the rate of translation (Ehrenberg and Kurland 1984).

In this work we focus on the tRNA cycle and optimize the concentrations of tRNA and amino-acyl synthetase pools with respect to translation rate. By setting up and solving the kinetic equations for all major steps in the tRNA cycle, we calculate the optimal correlations between the codon frequencies and the corresponding total tRNA and amino-acyl synthetase levels.

Previous theoretical work (von Heijne and Blomberg 1979; Chavancy and Garel 1981; Ehrenberg and Kurland 1984) has shown that the optimal correlation between the concentrations of the individual Tu-bound tRNA species (ternary complexes) and the corresponding codon frequencies in the ribosomal *A*-sites is a square root dependence. With the more complete analysis presented here, however, where *all* the tRNA pools are taken into consideration and where only the *total* amounts of tRNA, ribosomes and synthetases are directly varied, a much more complicated correlation picture becomes evident. Experimentally, the observed correlation between the total tRNA levels and codon frequencies is close to linear (Garel 1974; Ikemura 1981), an observation that can be explained by our model.

Theory

Although the translation rate is directly determined by the size and composition of the Tu-bound tRNA pool, the cell can only effect changes in this pool through variations in the *total* amounts of tRNAs synthesized. Thus, in a more realistic optimization, the whole tRNA cycle must be taken into account.

In an extensive treatment of the tRNA cycle, one should include all the different steps, as shown in Fig. 1. This leaves us with (at least) six different tRNA pools: 1) free, unacylated tRNA, 2) tRNA bound to amino-acyl synthetase, tRNA-*E*, 3) free, acylated tRNA, *aa*-tRNA, 4) tRNA in the ternary complex, *aa*-tRNA-EFTu-GTP, 5) tRNA at the *A*-site of the ribosome, *aa*-tRNA-EFTu-GTP-R^A, and 6) tRNA at the *P*-site of the ribosome, *aa*-tRNA-R^P. For simplicity, we will neglect the *aa*-tRNA pool, which should be comparatively

* To whom offprint requests should be sent

Abbreviations: EFTu, Elongation factor Tu

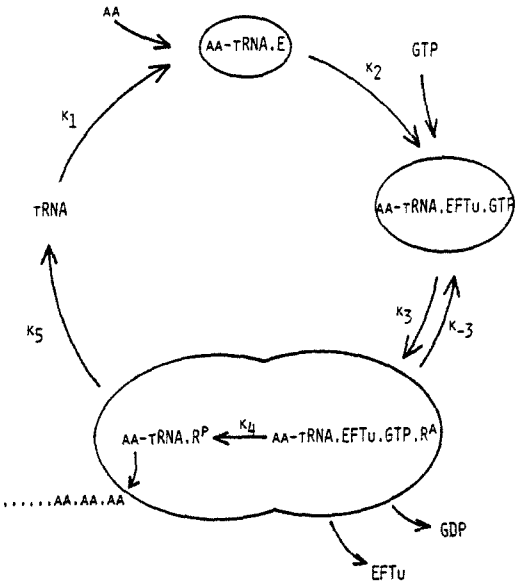

Fig. 1. Model of the tRNA cycle

small (assuming that the binding to EFTu is rapid and, at least in a kinetic sense, essentially irreversible). For similar reasons we will neglect the pool of tRNAs incorrectly bound to non-cognate codons at the *A*-site. All the backward reactions are omitted, except for the one from the *A*-site. These simplifications seem justified at the moment and should not, with an appropriate choice of forward rate constants, significantly affect the relative pool sizes.

More importantly, in order not to complicate the model too much at this stage, we do not include a full description of the amino-acyl synthetase reactions, but assume that the initial amino acid charging reaction is fast (i.e. all free synthetases have a charged amino acid bound), and that all the tRNA-charging reactions are characterized by the same rate constant. Furthermore, we do not take into account any proof-reading steps that may be present at the synthetases and the ribosomes, since proof-reading will have no major effect on the elongation rate (it may decrease it by some 10% to 30%, Ehrenberg and Kurland 1984). Finally, it is assumed that the amino acid pools, the EFTu pools (as well as other elongation factor pools), the GTP and ATP pools are constant.

We let the total amounts of tRNA, synthetase and ribosomes be designated by T, E and R, respectively. The amounts of each species, i, are represented by T_i and E_i (by assumption, only the T_i's, E_i's and R are externally controlled). We now define the concentration of the various pools of tRNA type i as T_i^f (free tRNA), T_i^E (synthetase-bound), T_i^T (Tu-bound), T_i^A (at the *A*-site) and T_i^P (at the *P*-site).

Similarly, we define the concentration of free synthetase by E_i^f and the concentration of "open" *A*-sites with codons matching the tRNA's of type i by R_i. The mean codon frequencies in the mRNA population are denoted by z_i.

With the rate constants k_1, k_2, k_3, k_{-3}, k_4 and k_5, which we here assume to be equal for all tRNA species, we can set up the following kinetic equations for the complete tRNA cycle (see Fig. 1, and note that k_5 may be expressed as $k_4 \cdot \sum T_i^A/R$, see below, leaving us with five independent rate constants):

$$dT_i^f/dt = -k_1 E_i^f T_i^f + k_5 T_i^P , \quad (1)$$

$$dT_i^E/dt = -k_2 T_i^E + k_1 E_i^f T_i^f , \quad (2)$$

$$dT_i^T/dt = -k_3 R_i T_i^T + k_{-3} T_i^A + k_2 T_i^E , \quad (3)$$

$$dT_i^A/dt = -(k_{-3} + k_4) T_i^A + k_3 R_i T_i^T , \quad (4)$$

$$dT_i^P/dt = -k_5 T_i^P + k_4 T_i^A , \quad (5)$$

$$dR_i/dt = -k_3 T_i^T R_i + k_{-3} T_i^A + k_4 z_i \sum T_i^A . \quad (6)$$

The conservation conditions are

$$T_i = T_i^f + T_i^E + T_i^T + T_i^A + T_i^P , \quad E_i = E_i^f + T_i^E , \quad (7)$$
$$T = \sum T_i , \quad E = \sum E_i , \quad R = \sum T_i^P = \sum R_i + \sum T_i^A .$$

To simplify our calculations, we let the index i run from 1 to 20 for the 20 amino acids, letting the corresponding iso-acceptor tRNA's, synthetases and codons represent one species for each amino acid (the results do not change significantly if we use more than 20 species).

The first five equations give the kinetics for the five separate tRNA pools in the cycle, whereas the sixth equation gives the kinetics for the open ribosomes, waiting for tRNA's to arrive at the *A*-site. This equation may need some clarification. The pool of ribosomes with tRNA-free codons decreases as a ternary complex enters the *A*-site. It increases as a ternary complex leaves this site, either by going back to the ternary complex pool or by irreversibly translocating to the *P*-site. The last term in Eq. (6) corresponds to the latter case, where the R_i pool increases proportionally to the codon frequency z_i (i.e. the new codon entering the *A*-site is uncorrelated with the preceeding one) and to the total amount of tRNA's leaving the *A*-sites (or *P*-sites). Note that $k_4 z_i \sum T_i^A$ is the number of tRNAs of type i translocating from *A*-site to *P*-site per time unit and this number must be equal to the number of tRNA's of type i leaving the *P*-site, which is described by $k_5 z_i \sum T_i^P$.

We also have that $\sum T_i^P = R$, since we assume that all active ribosomes have occupied *P*-sites. Thus, $k_4 z_i \sum T_i^A = k_5 z_i \sum T_i^P = k_5 z_i R$, or $k_5 = k_4 \sum T_i^A/R$. This result may also be obtained directly from Eq. (5) at steady state. There is no need to have a separate equation for the synthetase pool, since it is described by the negative of Eq. (2).

The twenty times six kinetic equations have been solved numerically for the steady-state case. By varying the five independent rate constants and using experimental data for the total amounts of tRNA, synthetase and ribosomes, as well as known values for amino acid frequencies in *E. coli* (Lehninger 1970), we have been able to get sizes of the different pools that correspond to the values given by Gouy and Grantham (1980) for *E. coli* cells growing at 2 doublings/hour (although some of these values are no more than informed guesses, our conclusions are not critically dependent upon an exact knowledge of the pool sizes). In this way we have arrived at a set of rate constants ($k_1 = 10$, $k_2 = 1$, $k_3 = 10$, $k_{-3} = 1$, $k_4 = 1.2$) that we have used for optimizing the system.

Results

By inspection of the six kinetic equations at steady-state, one immediately sees that some of the tRNA pools depend linearly on the codon frequency, z_i, and that some have a more complicated dependence. We may write the dependence thus:

$$\begin{aligned} &T_i^f \cdot E_i^f \propto z_i\,, \\ &T_i^E \propto z_i\,, \\ &T_i^T \cdot R_i \propto z_i\,, \\ &T_i^A \propto z_i\,, \\ &T_i^P \propto z_i\,. \end{aligned} \tag{8}$$

Summing the different tRNA pools yields (9)

$$T_i = T_i^f + T_i^E + T_i^T + T_i^A + T_i^P = (a + b/E_i + c/R_i) \cdot z_i\,.$$

For simplicity, we use a power-law approximation, $T_i \propto z_i^{\alpha}$ (Savageau 1976) and optimize the correlation exponent, α, with respect to the elongation rate.

Similarly, a correlation exponent, β, may be introduced to give an approximate expression for the correlation between the synthetases and the codon frequencies:

$$E_i = E_i^f + T_i^E \propto z_i^{\beta}\,. \tag{10}$$

Clearly, by varying the values of α and β for fixed values of the total amounts of tRNA, synthetase and ribosomes, the maximum elongation rate, given by $k_4 \sum T_i^A/R$, can be found.

Depending on the relative degree of saturation of the ribosomes and the synthetases, two different situations can be distinguished at high levels of tRNA. If (i) the ribosomes are more saturated (i.e. rate-limiting), most of the tRNA will pile up in the Tu-bound pool (T_i^T will dominate and α will tend to 0.5, c.f. von Heijne and Blomberg (1979)). If (ii) the synthetases are more saturated, most of the tRNA will instead accumulate in the free, unacylated tRNA pool, T_i^f (with all T_i^f's being equal at the optimum, i.e. $\alpha = 0$, see Fig. 4).

Thus, as the total tRNA level increases, for fixed amounts of ribosomes and synthetases, the optimal correlation between the tRNA species and the codon frequencies will tend towards a square root dependence ($\alpha_{opt} = 0.5$) in case (i), and a situation with equal amounts of all tRNA species ($\alpha_{opt} = 0$), in case (ii).

At the other extreme, when the tRNA/ribosome ratio becomes low, most of the tRNA will, in both cases, be found on the ribosomes and on the synthetases (i.e. T_i^E and T_i^P will dominate); both these pools have a linear correlation with the codon frequency ($\alpha = \alpha_{opt} = 1$, see Eq. (8)).

Between these two extremes, the correlation exponent α will vary between 0.5 or 0, respectively, and 1. The synthetase correlation exponent, β, is in most cases 1 or close to 1, and falls towards 0 only when the tRNA level becomes extremely low (see Fig. 4).

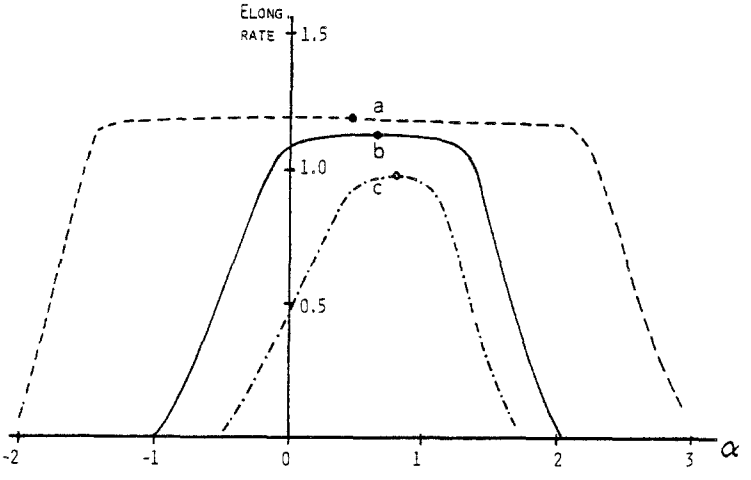

Fig. 2. Elongation rate (arbitrary units) dependence on the tRNA correlation exponent, α, for various tRNA/ribosome ratios: 100 (curve *a*), 10 (curve *b*), and 5 (curve *c*). α is defined by $T_i \propto z_i^{\alpha}$, where T_i denotes the total amount of tRNA of type i and z_i denotes the corresponding codon frequency. (Amount of ribosomes, $R = 10$, amount of synthetases, $E = 15$, and the rate constants $k_1 = k_3 = 10$, $k_2 = k_{-3} = 1$, $k_4 = 1.2$, and $k_5 = k_4 \sum T_i^A/R$, where T_i^A denotes the amount of tRNAs of type i at the ribosomal *A*-sites). The optimal correlations are marked

In Figs. 2 and 3, we have plotted the elongation rate versus α and β for three different tRNA levels corresponding to case (i). From these figures it is clear that the elongation rate becomes less dependent on the tRNA correlation exponent, the larger the tRNA/ribosome ratio is. The reverse is true for the synthetase correlation. Overall, the elongation rate is more critically dependent on the synthetase correlation with codon frequency than on the tRNA correlation. These figures also show that, for "normal" tRNA levels ($T/R \approx 10$), there is very little gain in elongation rate even for a large increase in the amount of tRNA. It is also obvious that the optimal values of the correlation exponents change as the tRNA level is altered. These optimal values, for both cases (i) and (ii), are plotted against the tRNA/ribosome ratio in Fig. 4.

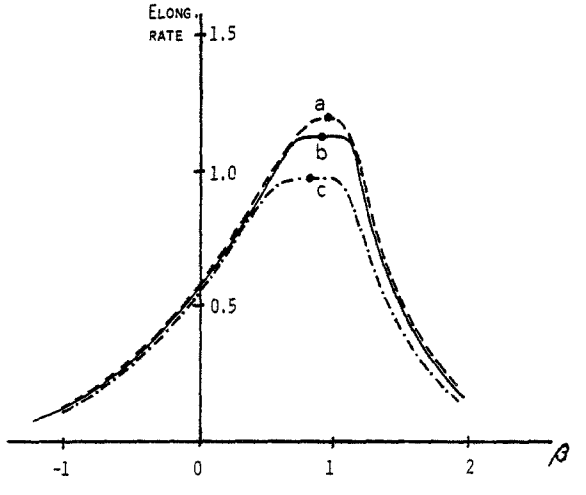

Fig. 3. Elongation rate dependence on the synthetase correlation exponent, β, for various tRNA/ribosome ratios: 100 (curve *a*), 10 (curve *b*), and 5 (curve *c*). β is defined by $E_i \propto z_i^{\beta}$, where E_i denotes the total amount of synthetase of type i and z_i denotes the corresponding codon frequency. (All parameters are the same as in Fig. 2). The optimal correlations are marked

Discussion

By solving the kinetic equations for the different steps in the complete tRNA cycle, we have been able to optimize the correlation between the tRNA and synthetase levels and their corresponding codon frequencies, with respect to translation rate. It is found that the optimal correlations are dependent on the total amounts of tRNA, ribosomes and synthetases. These amounts determine the relative degree of saturation of the ribosomes and synthetases and thereby determine the relative sizes of the various tRNA pools. Since each pool separately has its own optimal correlation with codon frequency, the optimal correlation for the total tRNA and synthetase levels will be different, depending on which pools are dominating. Hence, it will be a result of blending different proportions of constant, square root and linear correlations.

At high tRNA levels, two different situations are distinguishable, depending on whether the ribosomes or the synthetases are more saturated. In the former case the "excess" tRNA tends to be in the Tu-bound pool, which yields a square root correlation as an optimum. When the tRNA/ribosome ratio becomes extremely low most of the tRNA will, in all cases, be attached to the ribosomes and the synthetases. These tRNA pools have a linear dependence on codon frequency, as given by the kinetic equations for the system. The optimal correlation for the total tRNA levels will thus drift between a square root and a linear correlation, depending on the relative sizes of all these pools. The free, unacylated tRNA pool is comparatively small in this case. In the other case, however, where the synthetases are more saturated and become rate limiting,

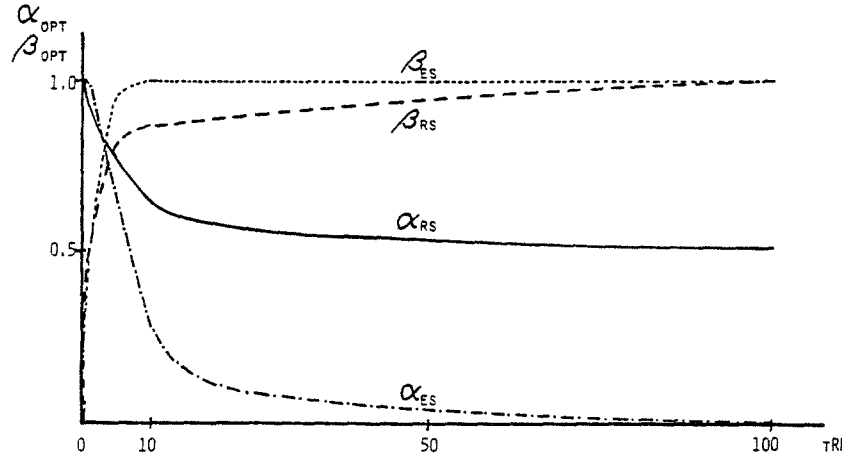

Fig. 4. The optimal correlation exponents as a function of the tRNA/ribosome ratio. RS = ribosomes more saturated (case (i), $R = 10$, $E = 15$), ES = synthetases more saturated (case (ii), $R = 10$, $E = 8$)

the unacylated tRNA pool becomes predominant. This pool has a correlation exponent equal to zero at the optimum, and hence, when this pool becomes larger the optimal correlation exponent for the total tRNA tends towards zero. This means that almost equal amounts of all tRNA species is optimal in this case, providing the tRNA/ribosome ratio is large (larger than, say, 10). If it is small, we get the same result as in the previous case, i.e. a linear correlation being optimal.

The optimal correlation between the synthetases and the codon frequencies will be similar in both cases, going from a constant to a linear correlation as the tRNA/ribosome ratio increases. (Here, we have treated all the synthetases as if they had the same properties and rate constants. The actual situation is more complex, however, and will be the subject of future studies.)

In rich media, where conditions allow for faster proliferation, the number of ribosomes in the cell increases more than does the number of tRNA molecules and thus the tRNA/ribosome ratio decreases for higher growth rates (in vivo, this ratio varies between 15 and 5, depending on growth rate) (Ehrenberg and Kurland 1984). At higher growth rates, then, the tRNA correlation becomes more critical. Therefore, it would be a good strategy for the cell to optimize the system for high growth rate, where the tRNA correlation is important. In less favourable growth conditions, where the cellular tRNA/ribosome ratio increases, the linear correlation optimal for the higher growth rates will be almost as good as the optimal correlation for lower growth rates, where the exact value of the correlation exponent is not so important. This may explain why the experimentally found correlation is approximately linear.

Acknowledgement. This work was supported by a grant from the Swedish Natural Sciences Research Council.

References

Blomberg C (1983) Free energy cost and accuracy in branched selection processes of biosynthesis. Q Rev Biophys 16: 415–519

Chavancy G, Garel JP (1981) Does quantitative tRNA adaptation to codon content in mRNA optimize the ribosomal translation efficiency? Proposal for a translation system model. Biochimie 63:187–195

Ehrenberg M, Kurland CG (1984) Costs of accuracy determined by a maximal growth rate constraint. Q Rev Biophys 17:45–82

Garel JP (1974) Functional adaptation of tRNA population. J Theor Biol 43:211–225

Gouy M, Grantham R (1980) Polypeptide elongation and tRNA cycling in *Escherichia coli:* a dynamic approach. FEBS Lett 115:151–155

von Heijne G, Blomberg C (1979) The concentration dependence of the error frequencies and some related quantities in protein synthesis. J Theor Biol 78:113–120

Ikemura T (1981) Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurence of the respective codons in its protein genes. J Mol Biol 146:1–21

Kurland CG, Ehrenberg M (1984) Optimization of translational accuracy. Prog Mol Biol Nucl Acids Res 31:191–219

Lehninger AL (1970) Biochemistry. Worth Publishers, New York

Savageau MA (1976) Biochemical systems analysis. Addison-Wesley, Reading, Mass