A model for mis-sense error in protein synthesis: 
mis-charged cognate tRNA versus mis-reading of codon

Annwesha Dutta and Debashish Chowdhury

Department of Physics, Indian Institute of Technology, Kanpur 208016, India.

The sequence of amino acid monomers in the primary structure of protein is decided by the corresponding sequence of codons (triplets of nucleic acid monomers) on the template messenger RNA (mRNA). The polymerization of a protein, by incorporation of the successive amino acid monomers, is carried out by a molecular machine called ribosome. Transfer RNA (tRNA) molecules, each species of which is "charged" with a specific amino acid, enters the ribosome and participates in the reading of the codon by the ribosome. Both mis-reading of mRNA codon and prior mis-charging of a tRNA can lead to "mis-sense" error, i.e., erroneous substitution of a correct amino acid monomer by an incorrect one during the synthesis of a protein. We develop a theoretical model of protein synthesis that allows for both types of contributions to the "mis-sense" error. We report exact analytical formulae for several quantities that characterize the interplay of mis-charging of tRNA and mis-reading of mRNA. The average rate of elongation of a protein is given by a generalized Michaelis-Menten-like formula. We discuss the main implications of these results. These formulae will be very useful in future in analyzing the data collected during experimental investigations of this phenomenon.

Corresponding author (E-mail: debch@iitk.ac.in)
I. INTRODUCTION

Proteins are linear polymers whose monomeric subunits are amino acids. Twenty different species of amino acids have been used by nature for synthesizing proteins. The specific sequence of the amino acids in a given protein is directed by the corresponding sequence of codons (triplets of nucleotide monomers) on the corresponding template messenger RNA (mRNA). The template-directed polymerization of a protein, called translation, is carried out by a molecular machine called ribosome [1-6] that consists of two loosely connected subunits designated as large and small. Transfer RNA (tRNA) molecules play crucial roles in translation [7]. When “charged” (amino-acylated) by a specific enzyme, called amino-acyl tRNA synthetase (aa-tRNA synthetase) [8-10], one end of each species of these “adapter” molecules carries a specific amino acid. The amino acid brought in by a correctly charged cognate tRNA is also the correct one, as directed by the corresponding template; the other end of the same cognate tRNA molecule, referred to as anti-codon, matches perfectly, by complementary base pairing, with the codon on the template mRNA. In contrast, increasing degree of mismatch makes the tRNA near-cognate or non-cognate.

However, because of the intrinsic stochasticity of aminoacylation, and occasional failure of the editing mechanism of aa-tRNA synthetase, a mis-charged tRNA may be produced [8, 10]. Therefore, even when it turns out to be a cognate tRNA for a given codon, such a mis-charged tRNA compromises the translational fidelity by contributing an amino acid which is different from that dictated by the mRNA template. Erroneous substitution of one amino acid by another is called mis-sense error. Pre-translational mis-charging of tRNA is not the only possible cause of mis-sense error. Erroneous selection of a correctly charged near-cognate or non-cognate tRNA, i.e., a co-translational mis-reading of a codon, also contributes to mis-sense error [11-15]. The more stringent is the mechanism of selection of incoming tRNA the lower is the mis-reading error. But, increasing the probability of rejecting near-cognate and non-cognate tRNAs would increase the likelihood of accepting not only correctly charged cognate tRNA but also that of a mis-charged cognate tRNA.

Most of the experimental works on mis-reading error have been carried out for bacteria. So far as the eukaryotes are concerned, a comprehensive analysis of translational mis-reading error in budding yeast has been reported by Kramer et al. [16]. Mis-charging of tRNA and the failure of the editing mechanisms have been investigated separately for a long time [8, 10]. However, to our knowledge, the relative contribution of mis-charging error to the overall mis-sense error has not been measured quantitatively in any experiment on translational kinetics. It is worth pointing out that a mis-charging error is not always detrimental for biological function of a cell and are believed to play some regulatory roles under special conditions [17-20].

The dependence of the frequencies of mis-reading error on the codon usage and tRNA concentration have been investigated extensively in the past [21-20]. But, to our knowledge, mis-charging has not been incorporated so far in any mathematical model of kinetics of translation because under normal circumstances mis-charging error is as low as 1 in $10^4$ (or even lower). But, when subjected to various types of stress, at least ten fold increase in mis-charging has been observed [27]. The main aim of this paper is to develop a theoretical model for the stochastic kinetics of the elongation state of translation incorporating both these possibilities within a single mathematical framework. Our model also distinguishes between the concentrations of four distinct types of tRNA molecules, namely, correctly charged cognate tRNA, incorrectly charged cognate tRNA, correctly charged near-cognate tRNA and correctly charged non-cognate tRNA. The exact analytical results and a few illustrative plots display the interplay of these two sources of error in the overall mis-sense error in translation. The exact analytical expression that we derive for the average rate of protein elongation is a generalized Michaelis-menten-like equation. The model and the analytical formulae derived here will be useful in analyzing the data collected in future experiments that might be performed for investigation of the same phenomenon.

II. MODEL

Sharma and Chowdhury [25] developed a 5-state stochastic kinetic model (from now onwards referred to as SC model) for the elongation cycle of translation (see Fig.1). The arrival of a aa-tRNA molecule, bound to GTP and EF-Tu, and its rejection because of the codon-anticodon mismatch are captured by the forward and reverse transitions 1 $\rightleftharpoons$ 2. The second stage of quality control (kinetic proofreading) involves hydrolysis of GTP by EF-Tu (2 $\rightarrow$ 3) followed by either rejection (3 $\rightarrow$ 1) or incorporation (3 $\rightarrow$ 4) in the growing protein by formation of a peptide bond. The first of the two-step translocation process consists of the Brownian rotation of large subunit of the ribosome with respect to the small subunit and simultaneous reversible transitions of the tRNAs between the so-called “classical” and “hybrid” states. The second, and final, step of translocation, driven by hydrolysis of another molecule of GTP by EF-G completes the cycle irreversibly. More detailed stochastic models of mecano-chemical kinetics of each elongation cycle have been developed (see, for example, [23]). However, in order to capture some other aspects of translational kinetics, we describe the kinetics of elongation cycle by the simpler SC model. Nevertheless, the strategy of modeling
followed here can be implemented also using the more detailed descriptions as the basic models of elongation cycle.

The SC model was used further to account for the stochastic alternating pause-and-translocation kinetics of a single ribosome [30] as well as for analyzing collective spatio-temporal organization of ribosomes in a polysome [31]. Because of the extreme simplicity of the SC model model, no clear distinction could be made, in terms of different rate constants, between processes involving near-cognate and non-cognate tRNAs. More importantly, the SC model captured the possibility of mis-sense error arising from only mis-reading of the codons; it was not possible to incorporate the contributions from both mis-reading and mis-charging errors explicitly. The non-trivial extension of the SC model that we present here does not suffer from any of the above mentioned limitations of the original SC model.

We begin formulation of the model with the four alternative elongation cycles shown in Fig.2 which correspond to the four different mutually exclusive pathways that open up with the arrival of (a) correctly charged cognate tRNA, (b) incorrectly charged cognate tRNA, (c) correctly charged near-cognate tRNA, and (d) correctly charged non-cognate tRNA. Note that each of these cycles is formally identical to the only cycle that appeared in the original SC model. However, by opening up the possibility of four distinct pathways, each associated with a distinct identity of aa-tRNA, this model not only allows for a clear distinction between non-cognate, near-cognate and cognate tRNAs but also that between correctly and incorrectly charged cognate tRNAs.

Next we simplify the model by exploiting some well known facts from the existing literature [1,3]. First, we note that \( \omega_a \) is proportional to the concentration of the corresponding aa-tRNA species; therefore, we assume:

\[
\begin{align*}
\omega_a &= \omega_a^{0,1}[\text{tRNA}]_{c1} \quad \text{(correctly charged cognate tRNA)} \\
\omega_a' &= \omega_a^{0,2}[\text{tRNA}]_{c2} \quad \text{(incorrectly charged cognate tRNA)} \\
\omega_a'' &= \omega_a^{0,n}[\text{tRNA}]_n \quad \text{(correctly charged near-cognate tRNA)} \\
\omega_a''' &= \omega_a^{0,N}[\text{tRNA}]_N \quad \text{(correctly charged non-cognate tRNA)}
\end{align*}
\]

where the symbol \([ \cdot ]\) denotes the concentration of the corresponding tRNA species and the prefactors are measures of the intrinsic rates of the reactions for unit concentration of the tRNA species. Thus, as stated in the introduction, concentrations of all the four types of tRNA molecules are incorporated explicitly.

Since the well orchestrated movement of the tRNA through the inter-subunit space in the translocation step is not expected to depend on which species of amino acid they bring in, we do not distinguish 4' from 4 and 5' from 5 because both the pathways 3 \( \rightarrow \) 4 \( \rightarrow \) 5 and 3' \( \rightarrow \) 4' \( \rightarrow \) 5' involve cognate tRNAs (see Fig.3). Similarly, assuming the rates of translocation of near-cognate and noncognate tRNA molecules to be comparable, but discriminating these from the corresponding cognate tRNAs, we assume 4'' \( \equiv \) 4''' \( \equiv \) 4* \( \neq \) 4 and 5'' \( \equiv \) 5''' \( \equiv \) 5* \( \neq \) 5 (see Fig.3). These assumptions help in combining the four pathways shown in Fig.2 within the single and simpler kinetic scheme depicted in Fig.3 thereby also reducing the number of parameters (rate constants). From now onwards, unless stated otherwise, all our discussions will be based on the model kinetic scheme shown in Fig.3.

Since most often the prefactors \( \omega_a^{0,1}, \omega_a^{0,2}, \omega_a^{0,n} \) and \( \omega_a^{0,N} \) in eq.(1) are not available, for simplicity, we’ll assume that \( \omega_a^{0,1} = \omega_a^{0,2} = \omega_a^{0,n} = \omega_a^{0,N} = \omega_a^0 \). However, for our graphical analysis we’ll treat \( \omega_a, \omega_a', \omega_a'', \) and \( \omega_a''' \) directly as the measures of the corresponding concentrations. Thus, we define

\[
\begin{align*}
C_{c1} &= \left( \frac{\omega_a}{\omega_a + \omega_a' + \omega_a'' + \omega_a'''} \right) \\
C_{c2} &= \left( \frac{\omega_a'}{\omega_a + \omega_a' + \omega_a'' + \omega_a'''} \right) \\
C_n &= \left( \frac{\omega_a''}{\omega_a + \omega_a' + \omega_a'' + \omega_a'''} \right) \\
C_N &= \left( \frac{\omega_a'''}{\omega_a + \omega_a' + \omega_a'' + \omega_a'''} \right)
\end{align*}
\]

as the relative concentrations of the four species of charged tRNA.

We use the symbol \( P_\mu(j,t) \) to denote the probability at time \( t \) that the ribosome is in the “chemical” state \( \mu \) and is decoding the \( j^{th} \) codon. In the steady state, all the probabilities \( P_\mu(j,t) \) become independent of time. We define translational fidelity by the fraction

\[
\phi = \frac{\omega_p P_3}{\omega_p P_3 + \omega_p P_3' + \Omega_p P_3'' + \Omega_p' P_3'''} = \frac{\omega_p P_3}{\omega_p P_3 + \omega_p P_3' + \Omega_p P_3'' + \Omega_p' P_3'''}
\]

where we have used the relation \( \Omega_p P_3'' + \Omega_p' P_3''' = \Omega_{p2} P_5' \).
The total mis-sense error \( E = 1 - \phi \) is defined by the relation

\[
E = \frac{\omega_r P_3 + \Omega h_2 P_5^*}{\omega_P P_3 + \omega_r P_3 + \Omega h_2 P_5^*}
\]

which is the sum of the total mis-charged mis-sense error (i.e., mis-sense error arising solely from mis-charged tRNAs) is

\[
E_{mc} = \frac{\omega_r P_3 + \Omega h_2 P_5^*}{\omega_P P_3 + \omega_r P_3 + \Omega h_2 P_5^*}
\]

and the total mis-reading mis-sense error (i.e., mis-sense error arising only from mis-reading of codons) is

\[
E_{mr} = \frac{\Omega h_2 P_5^*}{\omega_P P_3 + \omega_r P_3 + \Omega h_2 P_5^*}
\]

Similarly, the fraction

\[
\epsilon_{mc} = \frac{\omega_r P_3 + \Omega h_2 P_5^*}{\omega_P P_3 + \omega_r P_3 + \Omega h_2 P_5^*}
\]

is the fraction of mis-sense error caused by mis-charged cognate tRNAs, while the corresponding fraction of mis-sense error caused by misreading is defined by

\[
\epsilon_{mr} = \frac{\Omega h_2 P_5^*}{\omega_P P_3 + \omega_r P_3 + \Omega h_2 P_5^*}
\]

Obviously, the average velocity of a ribosome in the steady-state can be obtained by substituting the expressions of \( P_5 \) and \( P_5^* \) into the defining relation

\[
V = \ell_c (\omega h_2 P_5 + \Omega h_2 P_5^*)
\]

where \( \ell_c \) is the length of a codon. We also note that the average velocity \( V \) of a ribosome is same as the average rate of elongation of the protein that it polymerizes.

We define the total rejection factor \( \mathcal{R}_{tot} \) by the relation

\[
\mathcal{R}_{tot} = C_{c1} \mathcal{R} + C_{c2} \mathcal{R}' + C_n \mathcal{R}'' + C_N \mathcal{R}'''
\]

where

\[
\mathcal{R} = \left( \frac{\omega_r}{\omega_r + \omega_h} \right) \left( \frac{\omega_r}{\omega_r + \omega} \right)
\]

\[
\mathcal{R}' = \left( \frac{\omega_r}{\omega_r + \omega_h} \right) \left( \frac{\omega_r}{\omega_r + \omega} \right)
\]

\[
\mathcal{R}'' = \left( \frac{\omega_r'}{\omega_r' + \omega_h'} \right) \left( \frac{\omega_r'}{\omega_r' + \omega} \right)
\]

\[
\mathcal{R}''' = \left( \frac{\omega_r''}{\omega_r'' + \omega_h''} \right) \left( \frac{\omega_r''}{\omega_r'' + \omega} \right)
\]
The master equations governing the time evolution of the probabilities can be written as:

\[
\begin{align*}
\frac{dP_1(t)}{dt} &= -(\omega_a + \omega_a' + \omega_a'' + \omega_a''')P_1(t) \\
&\quad + \omega_{r1}P_2(t) + \omega_{r1}'P_2'(t) + \omega_{r1}''P_2''(t) \\
&\quad + \omega_{r1}'''P_2'''(t) + \omega_{r2}P_3(t) + \omega_{r2}'P_3'(t) \\
&\quad + \omega_{r2}''P_3''(t) + \omega_{r3}P_4(t) + \omega_{r3}'P_4'(t) \\
&\quad + \omega_{h2}P_5(t) + \Omega_{h2}P_5'(t) \\
\frac{dP_2(t)}{dt} &= \omega_aP_1(t) - (\omega_{r1} + \omega_{h1})P_2(t) \\
\frac{dP_3(t)}{dt} &= \omega_{h1}P_2(t) - (\omega_p + \omega_r)P_3(t) \\
\frac{dP_4(t)}{dt} &= \omega_pP_3(t) + \omega_p'P_3'(t) + \omega_{br}P_5(t) - \omega_{bf}P_4(t) \\
\frac{dP_5(t)}{dt} &= \omega_{bf}P_4(t) - (\omega_{h2} + \omega_{br})P_5(t) \\
\frac{dP_2'(t)}{dt} &= \omega_a'P_1(t) - (\omega_{r1}' + \omega_{h1}')P_2'(t) \\
\frac{dP_3'(t)}{dt} &= \omega_{h1}'P_2'(t) - (\omega_{r2}' + \omega_p')P_3'(t) \\
\frac{dP_2''(t)}{dt} &= \omega_a''P_1(t) - (\omega_{r1}'' + \omega_{h1}'')P_2''(t) \\
\frac{dP_3''(t)}{dt} &= \omega_{h1}''P_2''(t) - (\Omega_a + \omega_{r2}'')P_3''(t) \\
\frac{dP_2'''(t)}{dt} &= \omega_a'''P_1(t) - (\omega_{r1}''' + \omega_{h1}''')P_2'''(t) \\
\frac{dP_3'''(t)}{dt} &= \omega_{h1}'''P_2'''(t) - (\Omega_a' + \omega_{r2}''')P_3'''(t) \\
\frac{dP_4'(t)}{dt} &= \Omega_{p}P_3'(t) + \Omega_{br}P_5'(t) + \Omega_{bp}P_3''(t) - \omega_{bf}P_4'(t) \\
\frac{dP_5'(t)}{dt} &= \Omega_{bf}P_4'(t) - (\Omega_{br} + \Omega_{h2})P_5'(t)
\end{align*}
\]

(12)

The analytical results for this model that we report here are exact, i.e., these are derived without making any mathematical approximations. The derivations of these analytical expressions do not require imposition of any conditions on the numerical values of the rate constants. However, we now list some biologically motivated constraints on the relative magnitudes of the rate constants that we will use later in this paper only for presenting the results graphically for biologically relevant situations. Based on the levels of base-pair complementarity between the codon and the anticodon of the incoming tRNA, we expect that under normal physiological conditions the following conditions would be satisfied: \(\omega_{r1}'' > \omega_{r1}' > \omega_{r1}'' = \omega_{r1}'). Motivated by similar considerations, for graphical plots, we also assume \(\omega_{r2}' > \omega_{r2}' > \omega_{r2}'' = \omega_{r2}'). Continuing similar justification for the reduction in the number of model parameters, we assume \(\Omega_a' \approx \Omega_a'' \approx \omega_p' < \omega_p''.

III. RESULTS

We begin our theoretical analysis by first solving the master equations (12) under steady-state conditions to get the corresponding expressions for \(P_n\); the full analytical expressions are given in appendix A. Then using those expressions for \(P_n\) we calculate the quantities of our interest namely, \(\phi, E_{mc}, E_{mr}, \epsilon_{mc}, \epsilon_{mr}\) and \(V\). The results are listed below.

\[
\phi = \frac{A}{A + B + C + D}
\]

(13)

and, hence,

\[
E = 1 - \phi = \frac{B + C + D}{A + B + C + D}
\]

(14)
which is sum of the the two contributions

\[ E_{mc} = \frac{B}{A + B + C + D} \]  \hspace{1cm} (15) \\

and

\[ E_{mr} = \frac{C + D}{A + B + C + D}. \]  \hspace{1cm} (16)

Similarly, we get

\[ \epsilon_{mc} = \frac{B}{B + C + D} \]  \hspace{1cm} (17) \\

and

\[ \epsilon_{mr} = 1 - \epsilon_{mc} = \frac{C + D}{B + C + D}. \]  \hspace{1cm} (18)

In all the expressions (13)-(18) \( A, B, C \) and \( D \) are given by

\begin{align*}
A &= \frac{\omega_a}{[1 + (\omega_{11}/\omega_{1})][1 + (\omega_{22}/\omega_{p})]} \\
B &= \frac{\omega'_a}{[1 + (\omega'_{11}/\omega_{1})][1 + (\omega'_{22}/\omega_{p})]} \\
C &= \frac{\omega''_a}{[1 + (\omega''_{11}/\omega_{1})][1 + (\omega''_{22}/\omega_{p})]} \\
D &= \frac{\omega'''_a}{[1 + (\omega'''_{11}/\omega_{1})][1 + (\omega'''_{22}/\omega_{p})]} \hspace{1cm} (19)
\end{align*}

The expressions (13)-(18) can be easily justified by intuitive physical arguments. Let us first consider the special case where \( \omega_{11} = \omega'_{11} = \omega''_{11} = 0 = \omega_{22} = \omega'_{22} = \omega''_{22} = 0 \). In this case the expressions for \( A, B, C \) and \( D \) reduce to \( A = \omega_a, B = \omega'_a, C = \omega''_a \) and \( D = \omega'''_a \), respectively. Consequently, \( \phi = \omega_a/\omega_a + \omega'_a + \omega''_a + \omega'''_a \) is the probability of following the path 1 \( \rightarrow 2' \), instead of the other three alternatives, namely, 1 \( \rightarrow 2' \), 1 \( \rightarrow 2'' \) and 1 \( \rightarrow 2''' \). Similarly, in this special case, the expression \( \epsilon_{mc} = \omega'_a/\omega_a + \omega''_a + \omega'''_a \) is expected because \( \epsilon_{mc} \) is the probability of following the path 1 \( \rightarrow 2' \), instead of the two alternatives 1 \( \rightarrow 2'' \) and 1 \( \rightarrow 2''' \).

In the general case, the rate of transition 1 \( \rightarrow 2 \) \( \rightarrow 3 \) \( \rightarrow 4 \) is given by

\[ \frac{\omega_a}{\omega_{h1} + \omega_{r1}} \frac{\omega_p}{\omega_{h2} + \omega_{r2}} = A \]  \hspace{1cm} (20)

The quantities \( B, C \) and \( D \) have similar interpretations as rates for the transitions 1 \( \rightarrow 2' \) \( \rightarrow 3' \) \( \rightarrow 4 \), 1 \( \rightarrow 2'' \) \( \rightarrow 3'' \) \( \rightarrow 4^* \), and 1 \( \rightarrow 2''' \) \( \rightarrow 3''' \) \( \rightarrow 4^* \), respectively. Once the system reaches the state 4 it cannot return to state 1 without completing the full cycle. Therefore, fidelity \( \phi \) is the ratio \( A/(A + B + C + D) \). The expressions (15)-(18) for \( E_{mc}, E_{mr} \) and \( \epsilon_{mc}, \epsilon_{mr} \) also follow from the same interpretations of \( A, B, C \) and \( D \).

Ribosome is an enzyme; interestingly, at any given instant of time its substrate-specificity depends on the codon that it is engaged in translating. In recent years, the average rate of translation has been shown to be a generalization of the rate of enzymatic reactions. Recall that for the Michaelis-Menten (MM) enzymatic reaction

\[ E + S \xrightarrow{k_{+}} [ES] \xrightarrow{k_{-}} E + P \]  \hspace{1cm} (21)

the rate of the reaction under steady-state condition is given by the MM equation

\[ \frac{1}{V} = \frac{1}{V_{max}} + \frac{K_M}{V_{max} [S]} \]  \hspace{1cm} (22)

where the Michaelis constant \( K_M = (k_{-1} + k_2)/k_1 \) and \( V_{max} = k_2[E]_0 \), with \([E]_0\) being the initial (total) concentration of the enzyme. In the past the average rate of translation by a ribosome have been shown to follow a generalized MM-like equation where the concentration of aa-tRNA is interpreted as the substrate concentration. For simpler
models of translation reported earlier, the average rate of translation has been expressed as generalized MM equation \[24, 32\].

For the full kinetic model shown in Fig.3 the average velocity \(V\) is given by

\[
\frac{1}{V} = \frac{1}{A + B + C + D} + \frac{A}{A + B + C + D} \left( \frac{1}{V_A} \right) + \frac{B}{A + B + C + D} \left( \frac{1}{V_B} \right) + \frac{C}{A + B + C + D} \left( \frac{1}{V_C} \right) + \frac{D}{A + B + C + D} \left( \frac{1}{V_D} \right) \quad (23)
\]

where

\[
\frac{1}{V_A} = \left[ \frac{1}{\omega_{h1}} \left( 1 + \omega_{r2} \right) + \frac{1}{\omega_p} \left( 1 + \omega_{br} \right) + \frac{1}{\omega_{h2}} \right]
\]

\[
\frac{1}{V_B} = \left[ \frac{1}{\omega_{h1}} \left( 1 + \omega_{r2} \right) + \frac{1}{\omega_p} \left( 1 + \omega_{br} \right) + \frac{1}{\omega_{h2}} \right]
\]

\[
\frac{1}{V_C} = \left[ \frac{1}{\omega_{h1}} \left( 1 + \omega_{r2} \right) + \frac{1}{\omega_p} \left( 1 + \omega_{br} \right) + \frac{1}{\omega_{h2}} \right]
\]

\[
\frac{1}{V_D} = \left[ \frac{1}{\omega_{h1}} \left( 1 + \omega_{r2} \right) + \frac{1}{\omega_p} \left( 1 + \omega_{br} \right) + \frac{1}{\omega_{h2}} \right] \quad (24)
\]

Eqn.\((23)\) is a generalized version of the MM equation \((22)\) for our model. An intuitive derivation of the expression \((23)\), that provides a deeper physical interpretation of this formula, is given in appendix B.

The connection between the two can be elucidated by considering a special case of our model. In the limit \([tRNA]_{c2} = [tRNA]_n = [tRNA]_N = \omega_{r2} = \omega_{br} = 0\) and \(\omega_p \to \infty, \omega_f \to \infty, \omega_{h2} \to \infty\), the model reduces to

\[
\text{Ribosome} + tRNA_{c1} \xrightarrow{\omega_{v1}} 2 \text{charged tRNA} \quad (25)
\]

which is, formally, identical to the MM reaction \((21)\). In this limit the expression \((23)\) reduces to

\[
\frac{1}{V} = \frac{1}{\omega_{h1}} + \left( \frac{(\omega_{h1} + \omega_{r1})/\omega_{0,c1}}{\omega_{h1}} \right) \frac{1}{[tRNA]_{c1}} \quad (26)
\]

which is identical to MM equation \((22)\) because of the correspondence \(\omega_{h1} \leftrightarrow k_2 = V_{max}\) and \((\omega_{h1} + \omega_{r1})/\omega_{0,c1} \leftrightarrow (k_2 + k_3)/k_1 = K_M\). Thus, on the right hand side of the equation \((23)\) the sum of the last four terms is the generalized counterpart of \(1/V_{max}\) while the last term is the generalized analog of \(K_M/V_{max}[S]\).

For plotting the graphs, we have used the parameter values as given in the table I:

| for correctly charged cognate tRNA | for incorrectly charged cognate tRNA | for near-cognate tRNA | for non-cognate tRNA |
|------------------------------------|-------------------------------------|----------------------|---------------------|
| \(\omega_0 = 25 \text{ s}^{-1}\)   | \(\omega_0 = 10 \text{ s}^{-1}\)  | \(\omega_0 = 5 \text{ s}^{-1}\)  | \(\omega_0 = 5 \text{ s}^{-1}\)  |
| \(\omega_{r1} = 5 \text{ s}^{-1}\) | \(\omega_{r1} = 5 \text{ s}^{-1}\) | \(\omega_{r1} = 10 \text{ s}^{-1}\) | \(\omega_{r1} = 25 \text{ s}^{-1}\) |
| \(\omega_{r2} = 5 \text{ s}^{-1}\) | \(\omega_{r2} = 5 \text{ s}^{-1}\) | \(\omega_{r2} = 20 \text{ s}^{-1}\) | \(\omega_{r2} = 25 \text{ s}^{-1}\) |
| \(\omega_{h1} = 25 \text{ s}^{-1}\) | \(\omega_{h1} = 25 \text{ s}^{-1}\) | \(\omega_{h1} = 10 \text{ s}^{-1}\)  | \(\omega_{h1} = 5 \text{ s}^{-1}\)  |
| \(\omega_p = 25 \text{ s}^{-1}\)   | \(\omega_p = 10 \text{ s}^{-1}\)  | \(\omega_p = 10 \text{ s}^{-1}\)  | \(\omega_p = 10 \text{ s}^{-1}\)  |
| \(\omega_f = 25 \text{ s}^{-1}\)   | \(\omega_f = 25 \text{ s}^{-1}\)  | \(\omega_f = 10 \text{ s}^{-1}\)  | \(\omega_f = 10 \text{ s}^{-1}\)  |
| \(\omega_{br} = 25 \text{ s}^{-1}\) | \(\omega_{br} = 25 \text{ s}^{-1}\) | \(\omega_{br} = 10 \text{ s}^{-1}\) | \(\omega_{br} = 10 \text{ s}^{-1}\) |

The two-dimensional plots of the error fractions \(\epsilon_{mc}\) and \(\epsilon_{mr}\) against the rejection factors \(R''\) and \(R'''\) are shown in Figs.4 and 5 respectively. Both show how the error fraction \(\epsilon_{mr}\) decreases, while the fraction \(\epsilon_{mc}\) increases with increasing \(R''\) and \(R'''\). The total mis-sense error can also decrease because, under favorable conditions, the increase of \(E_{mc}\) with \(R''\) is more than compensated by the simultaneous decrease of \(E_{mr}\), as shown in fig.6.
IV. SUMMARY AND CONCLUSION

In this paper we have developed a theoretical model that includes both the effects of mis-charging of tRNA and mis-reading of mRNA during the elongation cycle of gene translation. It also allows explicit distinction between (i) correctly charged cognate tRNA, (ii) incorrectly charged cognate tRNA, (iii) correctly charged near-cognate tRNA, and (iv) correctly charged non-cognate tRNA. For each of the four species, the master equations capture only the essential steps of the elongation cycle. From these equations we derive exact analytical formulae (13)-(18) that characterize various aspects of the erroneous protein polymerization process, particularly the average speed and fidelity of polymerization. The average speed of the ribosome, i.e., the average rate \( \langle 23 \rangle \) of elongation of the protein, is an interesting generalization of the Michaelis-Menten equation that governs the average rate of a very simple enzymatic reaction. Some important implications of the analytical results reported here have been emphasized by plotting the results graphically. In particular, the plots show how with increasing rates of rejection of the near-cognate and non-cognate tRNAs the relative contribution of the mis-charged cognate tRNAs to the overall mis-sense error increases.

Next we point out some features of the model that should be reflected in the experimental set up which may, in near future, be analyzed with the analytical formulae reported in this paper. For a natural mRNA molecule, because of its sequence inhomogeneity, the identity of the cognate tRNA keep changing from one codon to another. On the other hand, the rate constants in our mathematical derivation is based on the assumption that the rates are independent of the position of the ribosome, i.e., independent of the identity of the codons. Thus, the sequence heterogeneity of natural mRNA molecules make those unsuitable for direct comparison with the analytical formulae reported here. Nevertheless, the model can be simply extended by assigning codon-dependent rate constants; but, in that case the results cannot be derived analytically (with closed form mathematical expressions) although all the quantities can be evaluated numerically. Since no experimental data for direct comparison is available at present, we have not carried out numerical study of the sequence-dependent model.

As an alternative to sequence inhomogeneous real mRNA, a synthetic mRNA with homogeneous sequence can be used to directly test the validity of our analytical formulae reported here. For example, poly-U, along with the necessary start-, stop codons and untranslated region (UTR) upstream from the start codon \( \text{[30]} \) could be a good candidate for this purpose. For any study of mis-sense error, the tRNA species which contribute the successive amino acids of the growing protein have to be identified. Fluorescence-based optical techniques \( \text{[33, 34]} \) seem to be ideally suited for this purpose. We hope that the relative quantitative contributions of mis-charging of tRNA and mis-reading of mRNA will be measured experimentally in near future and the analytical formulae derived here will find use in analyzing the experimental data.

ACKNOWLEDGEMENTS

DC thanks Joachim Frank, Ruben Gonzalez Jr. and Michael Ibba for useful correspondence. This work is supported by “Prof. S. Sampath Chair” professorship (DC) and a J.C. Bose Fellowship (DC).
Appendix A: Stead-state probabilities

The steady-state solutions $P_\mu$ of the equations (12) are given by

$$P_1 = \frac{1}{X1} \quad \text{(A1)}$$

where

$$X1 = A \left[ \frac{1}{\omega_{h1}} \left( 1 + \frac{\omega_r}{\omega_p} \right) + \frac{1}{\omega_p} + \frac{1}{\omega_{bf}} \left( 1 + \frac{\omega_{br}}{\omega_{h2}} \right) \right] +$$

$$B \left[ \frac{1}{\omega_{h1}'} \left( 1 + \frac{\omega_{r}'}{\omega_p} \right) + \frac{1}{\omega_p'} + \frac{1}{\omega_{bf}'} \left( 1 + \frac{\omega_{br}'}{\omega_{h2}} \right) \right] +$$

$$C \left[ \frac{1}{\omega_{h1}''} \left( 1 + \frac{\omega_{r}''}{\Omega_p} \right) + \frac{1}{\Omega_p} + \frac{1}{\Omega_{bf}} \left( 1 + \frac{\Omega_{br}}{\Omega_{h2}} \right) \right] +$$

$$D \left[ \frac{1}{\omega_{h1}'''} \left( 1 + \frac{\omega_{r}'''}{\Omega_p'} \right) + \frac{1}{\Omega_p'} + \frac{1}{\Omega_{bf}'} \left( 1 + \frac{\Omega_{br}'}{\Omega_{h2}} \right) \right]$$

$$P_2 = A \left[ \frac{1}{\omega_{h1}'} \left( 1 + \frac{\omega_{r}'}{\omega_p} \right) \right] P_1 \quad \text{(A2)}$$

$$P_3 = A \left[ \frac{1}{\omega_p r} \right] P_1 \quad \text{(A3)}$$

$$P_4 = (A + B) \left[ \frac{1}{\omega_{h2}'} \left( 1 + \frac{\omega_{br}'}{\omega_{h2}} \right) \right] P_1 \quad \text{(A4)}$$

$$P_5 = (A + B) \left[ \frac{1}{\omega_{h2}''} \right] P_1 \quad \text{(A5)}$$

$$P'_2 = B \left[ \frac{1}{\omega_{h1}'} \left( 1 + \frac{\omega_{r}'}{\omega_p} \right) \right] P_1 \quad \text{(A6)}$$

$$P'_3 = B \left[ \frac{1}{\omega_p'} \right] P_1 \quad \text{(A7)}$$

$$P''_2 = C \left[ \frac{1}{\omega_{h1}'} \left( 1 + \frac{\omega_{r}''}{\Omega_p} \right) \right] P_1 \quad \text{(A8)}$$

$$P''_3 = C \left[ \frac{1}{\Omega_p} \right] P_1 \quad \text{(A9)}$$

$$P'''_2 = D \left[ \frac{1}{\omega_{h1}'''} \left( 1 + \frac{\omega_{r}'''}{\Omega_p'} \right) \right] P_1 \quad \text{(A10)}$$

$$P'''_3 = D \left[ \frac{1}{\Omega_p'} \right] P_1 \quad \text{(A11)}$$

$$P'_4 = (C + D) \left[ \frac{1}{\Omega_{bf}} \left( 1 + \frac{\Omega_{br}}{\Omega_{h2}} \right) \right] P_1 \quad \text{(A12)}$$

$$P'_5 = (C + D) \left[ \frac{1}{\Omega_{h2}} \right] P_1 \quad \text{(A13)}$$
Appendix B: Intuitive derivation of the expression for average rate of translation

Following Cleland’s approach \[35\] for replacing complex network of biochemical pathways by a simpler equivalent network and deriving the effective rates of the transitions of that network, we derive the expression for the average velocity of the ribosome in our model. To illustrate the method, we consider a simpler reaction

\[ X \xrightleftharpoons[k_{-1}]{k_1} Y \xrightarrow{k_2} Z \]

The effective rate constant, \( k'_1 \), for \( X \rightarrow k'_1 Y \), is given by

\[ k'_1 = \frac{k_1 k_2}{k_{-1} + k_2} \]

The same treatment can be applied to our model.

Let us first assume that only correctly charged cognate tRNAs are present (i.e. assuming that mischarged cognate, near cognate and non cognate tRNAs are absent in the surrounding). For the five consecutive steps of the cycle, we denote the transit times by \( T_1, T_2, T_3, T_4 \) and \( T_5 \), respectively. It is straightforward to see that

\[ T_1 = \frac{1}{\omega_a} \left( 1 + \frac{\omega_r}{\omega_h} \right) \left( 1 + \frac{\omega_r}{\omega_p} \right) = \frac{1}{A} \]  

\[ T_2 = \frac{1}{\omega_{h1}} \left( 1 + \frac{\omega_r}{\omega_p} \right) \]  

\[ T_3 = \frac{1}{\omega_p} \]  

\[ T_4 = \frac{1}{\omega_{bf}} \left( 1 + \frac{\omega_r}{\omega_{h2}} \right) \]  

\[ T_5 = \frac{1}{\omega_{h2}} \]

Therefore, for the above cycle, i.e., when only correctly charged cognate tRNAs present in the surrounding, the average velocity of the ribosome would be

\[ \frac{1}{V_{c1}} = T_{c1} = T_1 + T_2 + T_3 + T_4 + T_5 \]  

(B7)

Similarly, for the other three cycles we can specify the transit times in a similar manner. For mischarged cognate tRNA, the corresponding symbols are \( T'_2, T'_3, T'_4, T'_5 \) respectively, while for near cognate tRNA, the symbols are \( T''_2, T''_3, T''_4, T''_5 \) respectively. For non cognate tRNA, we have \( T'''_2, T'''_3, T'''_4, T'''_5 \) respectively.

Next let us consider the general case when all the four different substrates are present in the surrounding; the kinetics of the system is described by full model shown in fig.3. The transit times are analogous to resistances in electrical circuits, which means that for a series of reaction, the transit times are additive and for parallel reaction pathways, the reciprocals of the transit times are additive. Hence, the average velocity for the complete model is

\[ \frac{1}{V_{tot}} = T_{tot} = \frac{A}{A + B + C + D} \left( T_{1,eff} + T_2 + T_3 + T_4 + T_5 \right) \]  

\[ + \frac{B}{A + B + C + D} \left( T_{1,eff} + T''_2 + T''_3 + T''_4 + T''_5 \right) \]  

\[ + \frac{C}{A + B + C + D} \left( T_{1,eff} + T'''_2 + T'''_3 + T'''_4 + T'''_5 \right) \]  

\[ + \frac{D}{A + B + C + D} \left( T_{1,eff} + T''''_2 + T''''_3 + T''''_4 + T''''_5 \right) \]

(B8)

(B9)

(B10)

(B11)

where

\[ \frac{1}{T_{1,eff}} = \frac{1}{T_1} + \frac{1}{T''_1} + \frac{1}{T'''_1} + \frac{1}{T''''_1} \equiv A + B + C + D \]

(B12)
III. REFERENCES

[1] A.S. Spirin, Ribosomes (Springer, 2002).
[2] M. V. Rodnina, W. Wintermeyer and R. Green (eds.) Ribosomes: Structure, Function, and Dynamics (Springer, 2011).
[3] J. Frank (ed.) Molecular machines in biology (Cambridge University Press, 2011).
[4] J. Frank and R. L. Gonzalez, Structure and Dynamics of a Processive Brownian Motor: The Translating Ribosome, Annu. Rev. Biochem. 79 (2010) 381-412.
[5] D. Chowdhury, Stochastic mechano-chemical kinetics of molecular motors: A multidisciplinary enterprise from a physicist’s perspective, Phys. Rep. 529, 1 (2013).
[6] D. Chowdhury, Modeling Stochastic Kinetics of Molecular Machines at Multiple Levels: From Molecules to Modules, Biophys. J. 104, 2331 (2013).
[7] S. Kim, Aminoacyl-tRNA synthetases in biology and medicine, Springer (2014).
[8] J. Ling, N. Reynolds and M. Ibba, Aminoacyl-tRNA Synthesis and Translational Quality Control, Annu. Rev. Microbiol. 63 (2009) 61-78.
[9] N. M. Reynolds, B. A. Lazazzera and M. Ibba, Cellular mechanisms that control mistranslation, Nat. Rev. Microbiol. 8 (2010) 849-856.
[10] S. S. Yadavalli and M. Ibba, Quality control in aminoacyl-tRNA synthesis, Adv. Protein Chem. and Struct. Biol. 86 (2012) 1-43.
[11] J. Parker, Errors and alternatives in reading the universal genetic code, Microbiol. Rev. 53 (1989) 273-298.
[12] L. Cochella and R. Green, Fidelity in protein synthesis, Curr. Biol. 136 (2005) R536-R540.
[13] H. S. Zaher and R. Green, Fidelity at the molecular level: lessons from protein synthesis, Cell 136 (2009) 746-762.
[14] M. Johansson, M. Lovmar and M. Ehrenberg, Rate and accuracy of bacterial protein synthesis, Curr. Opin. Microbiol. 11 (2008) 141-147.
[15] M. V. Rodnina, Quality control of mRNA decoding on the bacterial ribosome, Adv. Protein Chem. and Struct. Biol. 86 (2012) 95-128.
[16] E.B. Kramer, H. Vallabhaneni, L. M. Mayer and P.J. Farabaugh, A comprehensive analysis of translational missense errors in the yeast Saccharomyces cerevisiae, RNA 16, 1797 (2010).
[17] T. Pan, Adaptive translation as a mechanism of stress response and adaptation, Annu. Rev. Genet. 47 (2013) 121-137.
[18] L.R. de Pouplana, M.A.S. Santos, J.H. Zhu, P.J. Farabaugh and B. Javid, Trends Biochem. Sci. 39, 355-362 (2014).
[19] B. Ruan, S. Paloura, J. Sabina, L. Marvin-Guy, S. Kochhar, Robert A. LaRossa and D. Söll, Quality control despite mistranslation caused by an ambiguous genetic code, Proc. Natl. Acad. Sci. USA 105 (2008) 16502-16507.
[20] Y. Fan and J. Wu and M. H. Ung and N. De Lay and C. Cheng and J. Ling, Protein mistranslation protects bacteria against oxidative stress, Nucleic Acids Res. 43, 1740 (2015).
[21] A. Fluitt, E. Pienaar and H. Viljoen, Ribosome kinetics and aa-tRNA competition determine rate and fidelity of peptide synthesis, Comp. Biol. and Chem. 31, 335 (2007).
[22] H. Zouridis and V. Hatzimanikatis, Effects of Codon Distributions and tRNA Competition on Protein Translation, Biophys. J. 95, 1018 (2008).
[23] A. Basu and D. Chowdhury, Phys. Rev. E 75, 021902 (2007).
[24] A. Garai, D. Chowdhury, D. Chowdhury and T.V. Ramakrishnan, Stochastic kinetics of ribosomes: single motor properties and collective behavior, Phys. Rev. E 80 011908 (2009).
[25] P. Shah and M. A. Gilchrist, Effect of Correlated tRNA Abundances on Translation Errors and Evolution of Codon Usage Bias, PLoS Genet. 6(9), e1001128 (2010).
[26] S. Rudorf and R. Lipowsky, Protein Synthesis in E. coli: Dependence of Codon-Specific Elongation on tRNA Concentration and Codon Usage, PLOS ONE 10 e0134994 (2015).
[27] N. Netzer et al., Innate immune and chemically triggered oxidative stress modifies translational fidelity, Nature 462 522 (2010).
[28] A. K. Sharma and D. Chowdhury, Quality control by a mobile molecular workshop: Quality versus quantity, Phys. Rev. E 82 (2010) 031912.
[29] P. Xie, Dynamics of tRNA occupancy and dissociation during translation by the ribosome, J. Theor. Biol. 316 (2013) 49-60.
[30] A.K. Sharma and D. Chowdhury, Distribution of dwell times of a ribosome: effects of infidelity, kinetic proofreading and ribosome crowding, Phys. Biol. 8 (2011) 026005
[31] A.K. Sharma and D. Chowdhury, Stochastic theory of protein synthesis and polysome: Ribosome profile on a single mRNA transcript, J. Theor. Biol. 289 (2011) 36-46
[32] D. Chowdhury, Michaelis-Menten at 100 and allostery at 50: driving molecular motors in a hailstorm with noisy ATPase engines and allosteric transmission, FEBS J. 281, 601 (2014).
[33] J. Chen, A. Tsai, S. E. OLeary, A. Petrov and J.D. Puglisi, Unraveling the dynamics of ribosome translation, Curr. Opin. Struct. Biol. 22, 804 (2012).
[34] S. Uemura, C.E. Aitken, J. Korlach, B.A. Flusberg, S.W. Turner and J.D. Puglisi, Real-time tRNA transit on single translating ribosomes at codon resolution, Nature 464, 1012 (2010).
[35] W. W. Cleland, *Partition analysis and the concept of net rate constant as tools in enzyme kinetics*, Biochemistry 14, 3220 (1975).
FIG. 1. Pictorial depiction of the elongation cycle in the SC model. (see the text for details).

FIG. 2. Pictorial depiction of the four possible alternative mutually exclusive pathways that open up, in each chemo-mechanical cycle of a single ribosome, upon the arrival of (a) correctly charged cognate aa-tRNA, (b) incorrectly charged cognate aa-tRNA, (c) correctly charged near-cognate tRNA, and (d) correctly charged non-cognate tRNA (see the text for details).
FIG. 3. Pictorial depiction of the full chemo-mechanical kinetics in the elongation cycle of a single ribosome, along with the corresponding rate constants. It is obtained from Fig. 2 by combining the four cycles (see the text for details).

FIG. 4. The fraction $\epsilon_{mc}$ of error caused by mis-charged tRNA is plotted in 3D against the rejection factors $R''$ and $R'''$ of the near-cognate and non-cognate tRNAs, respectively. The relative concentrations of the four species of tRNA were kept fixed at $C_{c1} = 0.25$, $C_{c2} = 0.25$, $C_n = 0.25$, $C_N = 0.25$. The values of all the other parameters are listed in table I.
FIG. 5. Same as in Fig.4 except that the fraction $\epsilon_{mr}$ of error caused by misreading of mRNA is plotted against $R''$ and $R'''$. 
FIG. 6. The errors $E_{mc}$, $E_{mr}$ and $E$ are plotted against the rejection factor $R''$ for a fixed value of $R'''$ (see text for the definitions of the symbols). The relative concentrations of the four species of tRNA and the values of all the other model parameters are same as those used in Figs 4 and 5.