Distribution of dwell times of a ribosome: effects of infidelity, kinetic proofreading and ribosome crowding

Ajeet K Sharma and Debashish Chowdhury

Department of Physics, Indian Institute of Technology, Kanpur 208016, India
E-mail: debch@iitk.ac.in

Received 16 August 2010
Accepted for publication 4 January 2011
Published 24 January 2011
Online at stacks.iop.org/PhysBio/8/026005

Abstract
Ribosome is a molecular machine that polymerizes a protein where the sequence of the amino acid residues, the monomers of the protein, is dictated by the sequence of codons (triplets of nucleotides) on a messenger RNA (mRNA) that serves as the template. The ribosome is a molecular motor that utilizes the template mRNA strand also as the track. Thus, in each step the ribosome moves forward by one codon and, simultaneously, elongates the protein by one amino acid. We present a theoretical model that captures most of the main steps in the mechanochemical cycle of a ribosome. The stochastic movement of the ribosome consists of an alternating sequence of pause and translocation; the sum of the durations of a pause and the following translocation is the time of dwell of the ribosome at the corresponding codon. We derive the analytical expression for the distribution of the dwell times of a ribosome in our model. Wherever experimental data are available, our theoretical predictions are consistent with those results. We suggest appropriate experiments to test the new predictions of our model, particularly the effects of the quality control mechanism of the ribosome and that of their crowding on the mRNA track.

1. Introduction

The primary structure of a protein consists of a sequence of amino acid residues linked together by peptide bonds. Therefore, a protein is also referred to as a polypeptide which is essentially a linear heteropolymer, the amino acid residues being the corresponding monomers. The sequence of the amino acid residues in a polypeptide is dictated by that of the codons, each of which is a triplet of nucleotides, on a messenger RNA (mRNA) that serves as the template. The polymerization of each protein from the corresponding mRNA template is carried out by a macromolecular machine called ribosome [1–3] and the process is referred to as translation (of genetic code). The polymerization of protein takes places in three stages that are identified as initiation, elongation (of the protein) and termination.

The ribosome also utilizes the template mRNA as a track for its own movement; it steps forward by one codon at a time while, simultaneously, it elongates the growing polypeptide by one amino acid monomer. Therefore, ribosome is also regarded as a motor that, like other molecular motors, takes input (free-)energy from the hydrolysis of a nucleotide triphosphate to move along a filamenous track [4]. In fact, a ribosome hydrolyzes two molecules of guanosine triphosphate (GTP) to move forward by one codon.

Enormous progress has been made in the last decade in the fundamental understanding of the structure, energetics and kinetics of ribosomes [5–14]. Recent single molecule studies of ribosomes [15–21] have thrown light on its operational mechanism.

In single molecule experiments, it has been observed that a ribosome steps forward in a stochastic manner; its stepping is characterized by an alternating sequence of pause and translocation. The sum of the durations of a pause and the following translocation defines the time of a dwell of the...
ribosome at the corresponding codon. The time of dwell of a ribosome varies randomly from one codon to another. This randomness arises from two different sources: (i) intrinsic fluctuations associated with the Brownian forces as well as the low of concentrations of the molecular species involved in the chemical reactions and (ii) extrinsic fluctuations caused by the inhomogeneities of the sequence of nucleotides on the template mRNA [22]. Because of the sequence inhomogeneity of the mRNA templates used by Wen et al [23], the dwell time distribution (DTD) measured in their single-molecule experiment reflects a combined effect of the intrinsic and extrinsic fluctuations on the dwell time. In contrast, in this paper, we focus on situations (to be explained in detail in the following sections) where the randomness of the dwell times arises exclusively from intrinsic fluctuations.

The probability density \( f(t) \) of the dwell times of a ribosome, measured in single-molecule experiments [23], has been compared with the corresponding data obtained from computer simulations [24]. It has been claimed that the data do not fit a single exponential thereby indicating the existence of more than one rate-limiting step in the mechanochemical cycle of each ribosome. In fact, the best fit to the simulation data was achieved with five different rate-determining steps [24].

A systematic analytical derivation of the DTD of the ribosomes was presented recently [25] from a kinetic theory of translation [26] that also involves essentially five steps in the mechanochemical cycle of a ribosome. However, the model developed in [26] ignores some of the key features of the mechanochemical cycle of an individual ribosome during the elongation stage. For example, a ribosome deploys an elaborate proofreading mechanism to select the correct amino acid dictated by the template (and to reject the incorrect ones) to ensure high translational fidelity. Nevertheless, occasionally, translational errors take place. In this paper we extend Basu and Chowdhury’s original model [26] by capturing the processes of proofreading and allowing for the possibility of imperfect fidelity of translation. Moreover, the identification of the mechanochemical states as well as the nature of the transitions among these states is revised in the light of the empirical facts established in the last couple of years. Using this revised and extended kinetic model of translation [27], we analytically calculate the probability density \( f(t) \) of the dwell times of ribosomes.

The DTD derived in this paper is qualitatively similar to that observed by Wen et al in their single molecule experiments [23]. However, because of the sequence inhomogeneity of the template mRNA used in the experiment of Wen et al [23], their data cannot be compared quantitatively with our analytical expression for \( f(t) \). Therefore, we propose a concrete experimental setup required for a quantitative testing of our theoretical predictions. However, till such experiments are actually carried out, the results reported here will continue to provide insight into the mechanistic origin of the qualitative features of the DTD. Moreover, attempts are being made to extend our model to capture sequence inhomogeneities of mRNAs and to calculate the corresponding DTDs of ribosomes analytically.

Interestingly, the inverse of the mean dwell time is the average velocity \( V \) of the ribosome motor. Since, in our model, the ribosome is not allowed to back track on the mRNA template, \( V \) is also the mean rate of elongation of the polypeptide. We show that \( V \) satisfies an equation that resembles the Michaelis–Menten equation for the average rate of simple enzymatic reactions [28].

It is well known that most often a large number of ribosomes simultaneously move on the same mRNA track each polymerizing one copy of the same protein. This phenomenon is usually referred to as the ribosome traffic because of its superficial similarity with vehicular traffic on highways [26, 29–41]. In this paper we also report the effect of steric interactions of the ribosomes in ribosome traffic on their DTD.

2. The model

We develop here a theoretical model for the polymerization of a protein by ribosomes using an artificially synthesized mRNA template that consists of a homogeneous sequence (i.e. all the codons of which are identical). In the surrounding medium, two species of amino acids are available only one of which is the correct one according to the genetic code. The ribosome deploys a quality control system that rejects the amino acid monomer if it is incorrect. If this quality control system never fails, perfect fidelity of translation would result in a homopolymer whose constituent amino acid monomers are all identical. However, occasionally, wrong amino acid monomer escapes the quality control system. Such translational errors, whereby a wrong amino acid monomer is incorporated in the elongating protein, results in a heteropolymer. We will study the effects of the quality control system on the DTD of the ribosomes.

A ribosome consists of two interconnected subunits which are designated as ‘large’ and ‘small’ (see figure 1). The small subunit binds with the mRNA track and decodes the genetic message of the codon whereas the polymerization of the protein takes place in the large subunit. The operations of the two subunits are coordinated by a class of adapter molecules, called tRNA (see figure 1). One end of a tRNA helps in the decoding process by matching its anti-codon with the codon on the mRNA while its other end carries an amino acid; in this form the complex is called an amino-acyl tRNA (aa-tRNA).

The three main stages in each mechanochemical cycle of a ribosome, shown in figure 2, are as follows: (i) selection of the cognate (i.e. correct) aa-tRNA, (ii) formation of the peptide bond between the amino acid brought in by the selected aa-tRNA and the elongating protein and (iii) translocation of the ribosome by one codon on its track. However, some of these steps consist of important sub-steps. Moreover, the aa-tRNA selected (erroneously) by the ribosome may not be the cognate tRNA; this leads to a branching of pathways. Such branching, in turn, gives rise to the possibility of more than one cyclic pathway for a ribosome in a given cycle.

Let us begin with the state labeled by ‘1’; both the E and A sites are empty while the site P is occupied as shown in figure 3. An incoming aa-tRNA molecule, bound to an
elongation factor EF-Tu, occupies the site A; the resulting state is labeled by ‘2’. The transition 1 → 2 takes place at the rate \( \omega_{2} \). However, not all incoming aa-tRNA molecules are automatically selected by the ribosome. In order to ensure translational fidelity [42–45] the ribosome deploys a quality control mechanism to ensure that the aa-tRNA selected is, indeed, cognate (i.e. carries the correct amino acid as dictated by the mRNA template). A non-cognate tRNA is rejected on the basis of codon–anticodon mismatch; the corresponding transition 1 → 2 accounts for the hydrolysis of GTP leave the ribosome. We will later utilize the fact that one amino acid, a fresh molecule of GTP enters bound with an elongation factor EF-G. The arrival of GTP-bound EF-G is not treated explicitly in our model. Alternatively, \( \omega_{p} \) (and \( \Omega_{p} \)) is an effective rate constant that accounts for both the polypeptide elongation and the arrival of the GTP-bound EF-G.

Next, spontaneous Brownian (relative) rotation of the two subunits of the ribosome coincides with the back-and-forth transition between the so-called classical and hybrid configurations of the two tRNA molecules [46–56]. In the classical configuration, both ends of the two tRNA molecules correspond to the locations of P and A sites. In contrast, in the hybrid configuration, the ends of tRNA molecules interacting with the large subunit are found at the locations of E and P sites, respectively. The forward transition 4 → 5 takes place at the rate \( \omega_{bf} \) along the correct branch (and at the rate \( \Omega_{bf} \) along the wrong branch) whereas the reverse transition 5 → 4 takes place at the rate \( \omega_{br} \) along the correct branch (and \( \Omega_{br} \) along the wrong branch). The reversible transition 4 = 5, which is caused by spontaneous Brownian fluctuations, does not need any free energy input from GTP hydrolysis.

Finally the hydrolysis of GTP drives the irreversible transition 5 → 1 which involves the translocation of the ribosome on its track by one codon and, simultaneously, that of the tRNAs inside the ribosome by one binding site following which the deacetylated (i.e. denuded of amino acid) tRNA exits from the E site. The rate of this transition is \( \omega_{h} \) or \( \Omega_{h} \) depending on whether the system is following the correct branch or the wrong branch. In our model, the transition 5 → 1 captures the hydrolysis of GTP, departure of the products of hydrolysis along with EF-G as well as the final exist of the deacetylated tRNA molecule from the E site.

The rate constant \( \omega_{a} \) can be made identical for both species of amino acid monomers by maintaining their concentrations in the medium appropriately. The rate constant \( \omega_{h} \) refers only to the wrong amino acids (i.e. non-cognate tRNA) because, we assume, cognate tRNA is not rejected at all. Since the transition 2 → 3 accounts for the hydrolysis of a GTP molecule by the GTPase EF-Tu, the corresponding rate constant is \( \omega_{h1} \), irrespective of the identity of the attached aa-tRNA. In kinetic proofreading, the rate of rejection of the non-cognate tRNA is much higher than that of a cognate tRNA. For the sake of simplicity, we assume that the cognate tRNA is not rejected at all. Therefore, the rate constant \( \omega_{2} \) refers to the rejection of only non-cognate tRNA. For the remaining steps of the mechanochemical cycle, the rate constants are \( \omega_{p} \), \( \omega_{bf} \), \( \omega_{br} \) and \( \omega_{h2} \) provided a cognate tRNA has been selected finally. On the other hand, the corresponding rate constants are \( \Omega_{p} \), \( \Omega_{bf} \), \( \Omega_{br} \) and \( \Omega_{h2} \), respectively; if a non-cognate tRNA escapes rejection by the quality control system of the ribosome. Since the last step involves not only hydrolysis of GTP by the GTPase EF-G but also translocation of the ribosome and the tRNA molecules, the rate constants \( \omega_{h2} \) and \( \Omega_{h2} \) need not be equal, in general.
Phys. Biol. 8 (2011) 026005
A K Sharma and D Chowdhury

Figure 3. Pictoral depiction of the main steps in the chemomechanical cycle of a single ribosome (see the text for details).

Figure 4. The full mechanochemical kinetic of a ribosome during the elongation state, as modeled in figure 3, can be regarded as a composite of at least four different cycles shown in (a)–(d). Among these, both (a) and (b) are unproductive in the sense that these lead to neither elongation of the polypeptide nor forward stepping of the ribosome. However, (b) is ‘futile’ because it dissipates the free energy input from the hydrolysis of a GTP molecule whereas no fuel is wasted in (a). In contrast, both (c) and (d) lead to elongation of the polypeptide and forward stepping of the ribosome by hydrolyzing two molecules of GTP; however, (c) incorporates the correct amino acid whereas (d) incorporated a wrong amino acid into the growing polypeptide.

Thus, our model is an extension of the generic models for molecular motors based on the stochastic chemical kinetic approach which was pioneered by Fisher and Kolomeisky [57, 58]. Following Hill [59], the mechanochemical kinetics of a ribosome in our model can be regarded as a composite of the four cycles shown in figure 4.

All the numerical data generated from the analytical expressions for the graphical plots have the following set of
values of the rate constants (except in the figures where \( f(t) \) has been plotted for several different values of the parameters \( \omega_{a} \) and \( \omega_{b} \)): \( \omega_{a} = 25 \text{ s}^{-1}, \omega_{b} = 10 \text{ s}^{-1}, \omega_{h} = 25 \text{ s}^{-1}, \omega_{p} = 40 \text{ s}^{-1}, \omega_{b r} = \omega_{b r} = 25 \text{ s}^{-1}, \Omega_{b r} = \Omega_{b r} = 10 \text{ s}^{-1}, \omega_{a 2} = 25 \text{ s}^{-1}, \Omega_{a 2} = 10 \text{ s}^{-1} \). The value of \( \omega_{a} = 25 \text{ s}^{-1} \) is identical to that used in our earlier papers (see [25, 26] and the references therein). For the purpose of plotting our results, the magnitudes of the other parameters have been chosen to be comparable to that of \( \omega_{a} \). The values of some of the parameters, and the ranges over which some of these have been varied, may be unrealistically high or low. But this deliberate choice has been motivated by our intention to demonstrate graphically the interplay of various kinetic processes in translation.

3. Dwell time distribution

For the convenience of mathematical calculations, we assume that, after reaching the state 5 (or 5*) at location \( j \) the system makes a transition to a hypothetical state 0 at location \( j + 1 \) which, then, relaxes to the state 1, at the same location, at the rate \( \delta \). At the end of the calculation our model is recovered by setting \( \delta \to \infty \).

Suppose \( P_{\mu}(j, t) \) denote the probability at time \( t \) that the ribosome is in the 'chemical' state \( \mu \) and is decoding the \( j \)th codon. Let us use the symbol \( \tilde{P}_{1}(j + 1, t) \) to denote the probability of finding the ribosome in the hypothetical state 0 at time \( t \) while decoding the \( (j + 1) \)st codon. The time taken by the ribosome to reach the state 0 at \( j + 1 \), starting from the initial state 1 at \( j \), defines its time of dwell at the \( j \)th codon. Since, in this context, all the 'chemical' states except 0 refer to the \( j \)th codon while \( \tilde{P}_{1} \) corresponds to the \( (j + 1) \)st codon, from now onwards we drop the site index \( j \) (and \( j + 1 \)) to keep the notations simple.

The master equations governing the time evolution of the probabilities \( P_{\mu}(t) \) can be written as

\[
\frac{dP_{1}(t)}{dt} = -\omega_{a}P_{1}(t) + \omega_{a 1}P_{2}(t) + \omega_{a 2}P_{3}(t) \tag{1}
\]

\[
\frac{dP_{2}(t)}{dt} = \omega_{a}P_{1}(t) - (\omega_{a 1} + \omega_{h 1})P_{2}(t) \tag{2}
\]

\[
\frac{dP_{3}(t)}{dt} = \omega_{h 1}P_{2}(t) - (\omega_{p} + \Omega_{p} + \omega_{a 2})P_{3}(t) \tag{3}
\]

\[
\frac{dP_{4}(t)}{dt} = \omega_{p}P_{3}(t) - \omega_{b r}P_{4}(t) + \omega_{b r}P_{5}(t) \tag{4}
\]

\[
\frac{dP_{5}(t)}{dt} = \omega_{b r}P_{4}(t) - (\omega_{h 2} + \omega_{b r})P_{5}(t) \tag{5}
\]

\[
\frac{dP_{4}^{*}(t)}{dt} = \Omega_{p}P_{5}(t) - \Omega_{b r}P_{4}^{*}(t) + \Omega_{b r}P_{5}^{*}(t) \tag{6}
\]

\[
\frac{dP_{5}^{*}(t)}{dt} = \Omega_{b r}P_{4}^{*}(t) - (\omega_{h 2} + \Omega_{b r})P_{5}^{*}(t) \tag{7}
\]

\[
\frac{d\tilde{P}_{1}(t)}{dt} = \omega_{h 2}P_{5}(t) + \Omega_{h 2}P_{5}^{*}(t). \tag{8}
\]

Because of the normalization condition

\[
\sum_{\mu=1}^{5} P_{\mu}(t) + P_{\mu}^{*}(t) + P_{\mu}^{**}(t) + \tilde{P}_{1}(t) = 1 \tag{9}
\]

not all of the equations (1)–(8) above are independent of each other.

For the calculation of the dwell time, we impose the initial conditions

\[
P_{1}(0) = 1 \quad \text{and} \quad P_{2}(0) = P_{3}(0) = P_{4}(0) = P_{5}(0) = P_{4}^{*}(0) = P_{5}^{*}(0) = \tilde{P}_{1}(0) = 0. \tag{10}
\]

Suppose, \( f(t) \) is the probability density of the dwell times. Then, the probability of adding one amino acid to the growing polypeptide in the time interval between \( t \) and \( t + \Delta t \) is \( f(t)\Delta t \) where

\[
f(t) = \frac{\Delta \tilde{P}_{1}(t)}{\Delta t} = \omega_{h 2}P_{5}(t) + \Omega_{h 2}P_{5}^{*}(t). \tag{11}
\]

The calculation of the DTD [60–65] is essentially that of a distribution of first-passage times [66].

The exact probability density of the dwell times is given by

\[
f(t) = e^{-\omega_{h 2}t} \sum_{\mu=1}^{5} \frac{\omega_{a 2} \omega_{b r} \omega_{p} \omega_{h 1} \omega_{a}}{\omega_{a 1} \omega_{h 1} + \omega_{a 2} \omega_{h 2} + \omega_{p} \omega_{h 2} + \Omega_{p} \omega_{h 2} + \omega_{a 2} \omega_{b r}} e^{-\omega_{a 1}t} \tag{12}
\]

where \( \omega_{1, a 2} \) and \( \omega_{3} \) are solution of the cubic equation

\[
\omega^{3} - \omega^{2}(\omega_{1} + \omega_{a 1} + \omega_{a} + \omega_{a 2} + \omega_{p} + \Omega_{p}) + \omega(\omega_{1} \omega_{a 1} + \omega_{a 1} \omega_{a} + \omega_{a 2} \omega_{a} + \omega_{p} \omega_{a} + \omega_{a 2} \omega_{p} + \Omega_{p} \omega_{a} + \Omega_{p} \omega_{a 1} + \Omega_{p} \omega_{a 2} + \Omega_{p} \omega_{h 2} + \Omega_{p} \omega_{b r}) - \omega_{p} \omega_{h 2} \omega_{a} - \omega_{a 2} \omega_{b r} = 0, \tag{13}
\]

\( \omega_{4} \) and \( \omega_{5} \) are the solution of the quadratic equation

\[
\omega^{2} - \omega(\omega_{h 2} + \omega_{b r} + \omega_{b r}) + \omega_{h 2} \omega_{b r} = 0 \tag{14}
\]
and  and  are the solution of the quadratic equation

\[ \Omega^2 - \Omega(\Omega h_2 + \Omega p + \Omega_{bf}) + \Omega_{h2}\Omega_{bf} = 0. \]  

Some of the details of this derivation are given in the appendix. Note that the problem of determining the rates  (i = 1, 2, 3, 4, 5) in terms of the rate constants of the model is similar to that of expressing the normal modes of vibration of a set of coupled harmonic oscillators. It is the ‘backward’ reactions in the mechnochemical cycle of our model which play the role of coupling of the harmonic oscillators. For example, if  and simultaneously  reduce to the simple form  (i = 1, 2, 3, 4, 5) reduce to the simple form  and  are the weight factors associated with the two pathways emanating from the state labeled 3.

The distribution (12) is plotted in figure 5 for a few different values of the parameter  . Note that

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

is a measure of the fidelity of translation. Therefore, increasing  , keeping  fixed, enhances translational fidelity. Moreover, a higher  also corresponds to a faster peptidyl transferase reaction. Therefore, the trend of variation of the most probable dwell time with  is consistent with the intuitive expectation that the slower the peptidyl transferase reaction, the longer the dwell time.

The kinetic parameter  is a measure of the rate of rejection of the aa-tRNA by kinetic proofreading. Consequently, the effect of the variation of the  on the most probable dwell time is opposite to that of  (see figure 6); the higher the frequency of tRNA rejection by kinetic proofreading, the longer the dwell time.

4. Mean rate of polymerization: a Michaelis–Menten-like equation

The average Dwell time can be calculated by substituting (12) into the definition

\[ \langle t \rangle = \int_0^{\infty} tf(t) \, dt. \]  

Hence, the expression for

\[ \langle t \rangle = \frac{1}{\Omega_p} + \frac{1}{\omega_2} + \frac{1}{\omega_3} \left( \frac{1 + \omega_p + \Omega_p}{\omega_p + \Omega_p} \right) \]  

written in terms of  ,  ,  ,  ,  ,  and  has a clear physical meaning. In the special case  ≠ 0, \( \langle t \rangle = \sum_{i=1}^{5} \omega_i^{-1} \). However, if  ≠ 0, then the sum  is replaced by the last two terms of (18) where  and  are the weight factors associated with the two pathways emanating from the state labeled 3.

Finally, in terms of the rate constants for the mechnochemical transitions in the model depicted in figure 3

\[ \langle t \rangle = \frac{1}{\Omega_p} + \frac{1}{\omega_1} + \frac{1}{\omega_2} + \frac{1}{\omega_3} \left( \frac{1 + \omega_p + \Omega_p}{\omega_p + \Omega_p} \right) \]  

Next we will show that equation (19) can be re-expressed in a form that resembles the Michaelis–Menten equation for simple enzymatic reactions [28]. Consider an enzymatic reaction of the type

\[ E + S \overset{\omega_{s1}}{\rightarrow} ES \overset{\omega_{s1}}{\rightarrow} E + P, \]  

where  is the enzyme,  is the substrate and  is the product of the reaction catalyzed by . Given that the total initial concentration of the enzyme is  , the rate  of this reaction is given by

\[ \frac{1}{V} = \frac{1}{V_{\text{max}}} + \frac{K_M}{V_{\text{max}}} \frac{1}{[S]}, \]  

where the maximum possible rate of the reaction is

\[ V_{\text{max}} = \omega_{s2}[E]_0 \equiv \hat{\omega}_2 \]
and the Michaelis constant $K_M$ is given by
\[
K_M = \frac{\omega_{i_1} + \omega_{i_2}}{\omega_{i_1}}.
\] (23)

In the case of our model, we assume that the ‘pseudo’ first-order rate constant $\omega_d$ can be written as $\omega_d = \omega_d^0[tRNA]$ where [tRNA] is the concentration of tRNA molecules in the solution. Treating tRNA molecules as the analogs of the substrates in an enzymatic reaction, equation (19) can be re-expressed as a Michaelis–Menten-like equation [28]
\[
\langle t \rangle = \frac{1}{V_{\text{max}}} + \frac{K_M}{V_{\text{max}}} \frac{1}{[\text{tRNA}]},
\] (24)

where
\[
\frac{1}{V_{\text{max}}} = \frac{1}{\omega_{i_1}} = \frac{1}{\omega_{i_1}} \left( 1 + \frac{\omega_{i_2}}{\omega_p + \omega_{i_2} + \omega_p + \omega_{i_2}} \right) + \frac{1}{\omega_p + \omega_p}
\]
\[
+ \frac{1}{\omega_p + \omega_p} \left( 1 + \frac{\omega_{h_1}}{\omega_{h_1} + \omega_{h_2}} + \frac{1}{\omega_{h_1} + \omega_{h_2}} \right) + \frac{1}{\omega_{h_2} + \omega_{h_2}} + \frac{1}{\omega_{h_2} + \omega_{h_2}}.
\] (25)

and the Michaelis constant
\[
K_M = \frac{\omega_{i_1}^\text{eff} + \omega_{i_2}^\text{eff}}{\omega_i^\text{eff}}.
\] (26)

Thus, the mean dwell time for the ribosomes follows a Michaelis–Menten-like equation. This result is consistent with the experimental observations in recent years [67–72] that, in spite of the fluctuations of an enzymatic reaction catalyzed by a single enzyme molecule, the average rate of the reaction is, most often, given by the Michaelis–Menten equation. What makes the Michaelis–Menten-like equation for the ribosome even more interesting is the fact that a single ribosome is not a single enzyme molecule but it provides a platform for the coordinated operation of several bio-catalysts.

5. Effects of crowding on the DTD

It is well known that most often a large number of ribosomes simultaneously move on the same mRNA track each polymerizing one copy of the same protein. This phenomenon is usually referred to as the ribosome traffic because of its superficial similarity with vehicular traffic on highways [26, 29–41]. Suppose, $\ell$ denotes the number of codons that a ribosome can cover simultaneously. Extending the prescription used in our earlier works on ribosome traffic [26] for capturing the steric interactions of the ribosomes, we replace equations (5), (7) and (8) by
\[
\frac{dP_i(t, i)}{dt} = \omega_{i_1} P_i(t, i) - \omega_{i_2} P_i(t, i) - \omega_{h_2} P_i(t, i) Q(i + \ell)
\] (28)

\[
\frac{dP_{i_1}(t, i)}{dt} = \Omega_{i_2} P_{i_1}(t, i) - \Omega_{i_2} P_{i_1}(t, i) - \Omega_{h_2} P_{i_1}(t, i) Q(i + \ell)
\] (29)

\[
\frac{dP_{i_2}(t, i)}{dt} = \Omega_{i_2} P_{i_2}(t, i) - \Omega_{i_2} P_{i_2}(t, i) - \Omega_{h_2} P_{i_2}(t, i) Q(i + \ell)
\] (30)

where $Q(i, j) = 1 - P(i, j)$ is the conditional probability that, given a ribosome in site $i$, site $j$ is empty. All the other equations for $P_{i_1}(t, i)$ remain unchanged. Note that these equations have been written under the mean-field approximation.

In the limit $L \to \infty$, all the sites are treated on the same footing so that the site dependence of $P_{i_1}(t, i)$ drops out. In this limit, $Q$ takes the simple form [26]
\[
Q(i, j + \ell) = \frac{1 - \rho \ell}{1 + \rho - \rho \ell},
\] (31)

where $\rho$ is the number density of the ribosomes (i.e. number of ribosomes per unit length of the mRNA track).

Therefore, in the ribosome traffic the distribution of the dwell times of the ribosomes is given by expression (12) where $\omega_{i_2}$ and $\Omega_{h_2}$ are replaced by $\omega_{i_2} Q$ and $\Omega_{h_2} Q$, respectively. Using expression (31), we get the DTD for the given number density $\rho$ of the ribosomes. For the purpose of graphical demonstration of the effects of ribosome crowding on the DTD, we use the coverage density $\rho_{\text{cov}} = \rho \ell$. In figure 7 we plot the DTD for a few different values of $\rho_{\text{cov}}$. The higher the density, the stronger the hindrance and longer the dwell time. Moreover, higher coverage density introduces stronger correlations which our mean-field equations ignore. Consequently, our analytical predictions, based on equations written under the mean-field approximation, deviate more and more from the corresponding simulation data as the coverage density increases.

6. Comparison with experimental data

The distribution of dwell times involves essentially five different rate-determining parameters $\omega_i (i = 1, 2, 3, 4, 5)$ if the translational fidelity is perfect. This is consistent with the
The actual coding sequence to be translated consists of \( n \) bulk samples. But the more recent single-molecule FRET of the mRNA) and the other is either near-cognate or non-cognate (corresponding to the codons in the coding sequence and only two species of aa-tRNA one of which is one should use an mRNA template with a homogeneous codon conditions maintained in those experiments. However, for a strongly indicates the possibility that only two of the five well with a difference of just two exponentials [23]. This DTD obtained from the single-molecule experiments was obtained with five rate-determining parameters. However, consider the concrete example shown schematically in figure 8. This example is essentially the protocols used by Uemura et al [20] in their single-molecule studies of translation in real time. The actual coding sequence to be translated consists of \( n_c \) number of identical codons; in figure 8, \( n_c = 6 \) and each codon is UUU which codes for the amino acid phenylalanine (abbreviated Phe or F). The coding sequence is preceded by a start codon AUG and is followed by a stop codon UAA. The start codon itself is preceded by an untranslated region (UTR) at the 5′-end of the mRNA; this is required for assembling the ribosome and for stabilizing the pre-initiation complex. At the 3′-end, the stop codon is followed by a sequence of \( n_{nc} \) non-coding codons UUU (\( n_{nc} = 4 \) in figure 8); this region merely ensures the absence of any ‘edge effect’, i.e. the translation is not affected when the ribosome approaches the 3′-end of the codon sequence. A good choice for the corresponding near-cognate tRNA would be tRNA\(^{\text{Leu}}\) because it is cognate for the codon CUU which codes for leucine (abbreviated L). The two distinct species of aa-tRNA should be labeled by fluorescent dyes molecules of two different colors which could be, for example, Cy5 (red) and Cy3 (green). Equal concentrations of Cy5-labeled ternary complex aa-tRNA\(^{\text{Ph}}\)-EF-Tu-GTP and Cy3-labeled ternary complex aa-tRNA\(^{\text{Leu}}\)-EF-Tu-GTP should be made available in the medium surrounding the ribosome to avoid any bias arising from difference in their concentrations. The color of the fluorescence pulse identifies the amino-acid monomer species that elongates the polypeptide by one unit; monitoring the colors of the fluorescence pulses one would get an estimate of the translational fidelity \( \phi \). Moreover, the time interval between the arrival of the successive aa-tRNA molecules provides an estimate of the dwell times of the ribosome. Thus, using this optical technique one would get not only the DTD but also the fidelity of translation. Usually the coding sequence of such poly-U mRNA strands is quite short. Therefore, for collecting enough data to extract the DTD, the experiment has to be repeated sufficiently large number of times.

The effects of kinetic proofreading on the rate of translational error have been investigated experimentally by several different methods in the last three decades. However, most of those methods (see, for example, [73, 74]) require bulk samples. But the more recent single-molecule FRET technique used in the study of tRNA selection [19] seems to be more suitable for testing our predictions on the effects of kinetic proofreading on the DTD.

The cluster of ribosomes translating the same mRNA simultaneously is usually referred to as a polysome. For studying the effects of ribosome crowding on the DTD, one has to measure simultaneously the DTD and the polysome size [75].

7. Summary and conclusion

In this paper we have presented a theoretical model of translation that captures all the main steps in the mechanochemical cycle of a ribosome during the elongation stage. This model also accounts for translational fidelity, kinetic proofreading and the crowding of the ribosomes on the same mRNA track. In principle, this model can be extended, by increasing the number of ‘chemical’ states, to account for some of the sub-steps which have not been treated explicitly in the version of the model presented here.

In spite of the details already incorporated in this model, we have succeeded in carrying out an analytical calculation of the distribution of the dwell times of the ribosomes at the successive codons. We have compared this theoretical estimate of the DTD with the corresponding numerical data which we have obtained from direct computer simulation of the model. If the motion of the ribosome is not hindered by the presence of any other ribosome on the same mRNA track, our analytical treatment yields the exact expression for the distribution of the times of its dwell at the successive codons. In this case, excellent agreement between theoretical prediction and simulation data is observed provided the data are averaged over sufficiently large number of samples.

However, because of the mean-field approximation made in capturing the effects of crowding of the ribosomes, the corresponding analytical expression is approximate. Therefore, the higher the coverage density, the larger the deviation of the analytical estimate from the simulation data which are averaged over many samples.

We have analyzed the dependence of the DTD on some of the crucially important kinetic parameters of the model to elucidate the physical implications of the result. Our results are in good qualitative agreement with the experimental data reported in the literature [23, 24]. However, for the reasons explained in sections 1 and 2, it is not possible to compare these experimental data quantitatively with the analytical expressions of DTD which we have reported in this paper. We hope our theoretical predictions will stimulate further experimental studies along the lines suggested briefly in section 6 although some technical hurdles may hinder quick progress. A combination of the single-ribosome
experiments and bulk measurements may be required for comparing the theoretically predicted variation of the DTD with the concentrations of tRNA, GTP and other key molecules involved in translation as well as with the increase of futile cycles and crowding.

It would also be desirable to extend our model in future to account for some important features of translation; some of these possible extensions are listed below. (i) Capturing the sequence heterogeneity of real mRNA templates will open up larger number of branched pathways in figure 3, each corresponding to a distinct species (cognate, non-cognate or near-cognate) of amino acid. Moreover, at each step of the ribosome, the rate constants for the correct and incorrect pathways will also depend on the codon under consideration. (ii) The interaction of the growing polypeptide with the exit tunnel in the ribosome and the spontaneous folding of the nascent protein as it comes out of the tunnel may affect the rate of translation. It may be more difficult than on an inhomogeneous sequence. It may be desirable to extend our model allowing the possibility of frameshift errors [76] although, at present, the prescription for this extension is not obvious. The DTD for such a realistic model may be obtained numerically because an analytical derivation of the general expression may not be possible in the foreseeable future.

Acknowledgments

One of the authors (DC) thanks Joachim Frank and Joseph D Puglisi for useful correspondences.

Appendix

Solving equations (1)–(7) under the initial condition (10) we get

\[ P_1(t) = C_1 e^{-\omega t} + C_2 e^{-\omega_2 t} + C_3 e^{-\omega_3 t} \]  
(A.1)

\[ P_2(t) = C_4 e^{-\omega_4 t} \]
(A.2)

\[ P_3(t) = C_5 e^{-\omega_5 t} \]
(A.3)

\[ P_4(t) = C_6 e^{-\omega_6 t} \]
(A.4)

\[ P_5(t) = C_7 e^{-\omega_7 t} \]
(A.5)

\[ P_6(t) = C_8 e^{-\omega_8 t} + C_9 e^{-\omega_9 t} \]
(A.6)

\[ P_7(t) = C_{10} e^{-\omega_{10} t} \]
(A.7)

where \( \omega_4, \omega_5 \) and \( \omega_6 \) are the solutions of the cubic equation (13) while \( \omega_1, \omega_2 \) and \( \omega_3 \) are the solutions of the quadratic equation (14) and \( \omega_4, \omega_5 \) are the solutions of the quadratic equation (15). The coefficients \( C_i \) \( (i = 1, 2, \ldots, 5) \) and \( D_4, D_5 \), which are determined by the initial condition (10), are as follows:

\[ C_1 = \frac{(\omega_1 + \omega_2 - \omega_4)(\omega_2 + \omega_3 + \omega_p - \omega_1)}{(\omega_2 - \omega_1)(\omega_3 - \omega_1)} \]
(A.8)

\[ C_2 = \frac{(\omega_1 + \omega_3 - \omega_2)(\omega_2 + \omega_3 + \omega_p - \omega_2)}{(\omega_1 - \omega_2)(\omega_3 - \omega_2)} \]
(A.9)

\[ C_3 = \frac{(\omega_1 + \omega_2 - \omega_3)(\omega_2 + \omega_3 + \omega_p - \omega_3)}{(\omega_1 - \omega_3)(\omega_2 - \omega_3)} \]
(A.10)

\[ C_4 = \frac{(\omega_2 + \omega_3 - \omega_4)(\omega_2 + \omega_3 + \omega_p - \omega_4)}{(\omega_2 - \omega_3)(\omega_2 - \omega_4)} \]
(A.11)

\[ C_5 = \frac{(\omega_2 + \omega_3 - \omega_5)(\omega_2 + \omega_3 + \omega_p - \omega_5)}{(\omega_2 - \omega_3)(\omega_2 - \omega_5)} \]
(A.12)

\[ D_4 = \frac{(\omega_5 - \omega_1)(\omega_1 - \omega_4)(\omega_2 - \omega_2)(\omega_3 - \omega_4)}{(\omega_5 - \omega_1)(\omega_5 - \omega_4)(\omega_1 - \omega_4)(\omega_2 - \omega_2)} \]
(A.13)

\[ D_5 = \frac{(\omega_1 - \omega_5)(\omega_2 - \omega_3)(\omega_3 - \omega_5)(\omega_4 - \omega_5)}{(\omega_1 - \omega_5)(\omega_2 - \omega_3)(\omega_3 - \omega_5)(\omega_4 - \omega_5)} \]
(A.14)
References

[1] Spirin A S 2000 Ribosomes (Berlin: Springer)
[2] Spirin A S 2002 FEBS Lett. 514 2
[3] Spirin A S 2009 J. Biol. Chem. 284 211203
[4] Cross R A 1997 Nature 385 18
[5] Ramakrishnan V 2010 Angew. Chem. Int. Edn 49 4355
[6] Yonath A 2010 Angew. Chem. Int. Edn 49 4340
[7] Steitz T A 2010 Angew. Chem. Int. Edn 49 4381
[8] Steitz T A 2008 Nat. Rev. Mol. Cell Biol. 9 242
[9] Schmeing T M and Ramakrishnan V 2009 Nature 461 1234
[10] Ramakrishnan V 2002 Cell 108 557
[11] Frank J and Spahn C M T 2006 Rep. Prog. Phys. 69 1383
[12] Mitra K and Frank J 2006 Annu. Rev. Biophys. Biomol. Struct. 35 299
[13] Agirrezabala X and Frank J 2009 Q. Rev. Biophys. 42 159
[14] Frank J and Gonzales R L 2010 Annu. Rev. Biophys. 39 381
[15] Marshall R A, Aitken C E, Dorywalska M and Puglisi J D 2008 Annu. Rev. Biochem. 77 177
[16] Munro J B, Vaiana A, Sanbonmatsu K Y and Blanchard S C 2008 Biopolymers 89 865
[17] Munro J B, Sanbonmatsu K Y, Spahn C M T and Blanchard S C 2009 Trends Biochem. Sci. 34 390
[18] Blanchard S C 2009 Curr. Opin. Struct. Biol. 19 103
[19] Blanchard S, Gonzalez R L Jr, Kim H D, Chu S and Puglisi J D 2004 Nat. Struct. Mol. Biol. 11 1008
[20] Uemura S, Aitken C E, Flusberg B A, Turner S W and Puglisi J D 2010 Nature 464 1012
[21] Aitken C E and Puglisi J D 2007 Nat. Struct. Mol. Biol. 14 793
[22] Buchan J R and Stansfeld I 2007 Biol. Cell 99 475–87
[23] Wen J D, Lancaster L, Hodges C, Zeri A C, Yoshimura S H, Noller H F, Bustamante C and Tinoco I Jr 2008 Nature 452 598
[24] Tinoco I Jr and Wen J D 2009 Phys. Biol. 6 025006
[25] Garai A, Chowdhury D, Chowdhury D and Ramakrishnan T V 2009 Phys. Rev. E 80 011908
[26] Basu A and Chowdhury D 2007 Phys. Rev. E 75 021902
[27] Sharma A K and Chowdhury D 2010 Phys. Rev. E 82 031912
[28] Dixon M and Webb E C 1979 Enzymes (New York: Academic)
[29] MacDonald G, Gibbs J C and Pippin A 1968 Biopolymers 6 1
[30] MacDonald C and Gibbs J 1969 Biopolymers 7 707
[31] Lakatos G and Chou T 2003 J. Phys. A: Math. Gen. 36 2027
[32] Shaw L B, Zia R K P and Lee K H 2003 Phys. Rev. E 68 021910
[33] Shaw L B, Sethna J P and Lee K H 2004 Phys. Rev. E 70 021901
[34] Shaw L B, Kolomeisky A B and Lee K H 2004 J. Phys. A: Math. Gen. 37 2105
[35] Chou T 2003 Biophys. J. 85 755
[36] Chou T and Lakatos G 2004 Phys. Rev. Lett. 93 198101
[37] Dong J J, Schmittmann B and Zia R K P 2006 J. Stat. Phys. 128 21
[38] Dong J J, Schmittmann B and Zia R K P 2007 Phys. Rev. E 76 051113
[39] Cook L J, Zia R K P and Schmittmann B 2009 Phys. Rev. E 81 051904
[40] Ciandrini L, Stansfeld I and Romano M C 2010 Phys. Rev. E 81 051904
[41] Mitarai N, Sneppen K and Pedersen S 2008 J. Mol. Biol. 382 236
[42] Rodnina M V and Wintermeyer W 2001 Annu. Rev. Biochem. 70 415
[43] Daviter T, Gromadski K B and Rodnina M V 2006 Biochimie 88 1001
[44] Ogle J M and Ramakrishnan V 2005 Annu. Rev. Biochem. 74 129
[45] Zaher H S and Green R 2009 Cell 136 746
[46] Moazed D and Noller H F 1989 Nature 342 142
[47] Wilson K S and Noller H F 1998 Cell 92 337
[48] Noller H F 2002 FEBS Lett. 514 11
[49] Horan L H and Noller H F 2007 Proc. Natl Acad. Sci. 104 4881
[50] Korostelev A, Ermolenko D N and Noller H F 2008 Curr. Opin. Chem. Biol. 12 674
[51] Frank J and Agrawal R K 2000 Nature 406 318
[52] Valle M, Zavaliov A, Sengupta J, Rawat U, Ehrenberg M and Frank J 2003 Cell 114 123
[53] Frank J, Gao H, Sengupta J, Gao N and Taylor D J 2007 Proc. Natl Acad. Sci. 104 19671
[54] Pan D, Kirlilov S V and Cooperman B S 2007 Mol. Cell 25 519
[55] Morgan S, J Flanagan J F, Nam Y, Stuart D I, Brieler I and Gilbert R J C 2008 Structure 16 664
[56] Shoja S, Walker S E and Fredrick K 2009 ACS Chem. Biol. 4 93
[57] Fisher M E and Kolomeisky A B 1999 Proc. Natl Acad. Sci. 96 6597
[58] Fisher M E and Kolomeisky A B 2007 Annu. Rev. Phys. Chem. 58 675
[59] Hill T L 2005 Free Energy Transduction and Biochemical Cycle Kinetics (New York: Dover)
[60] Chemla Y R, Molitt J R and Bustamante C 2008 J. Phys. Chem. B 112 6025
[61] Shaevitz J W, Block S M and Schnitzer M J 2005 Biophys. J. 89 2277
[62] Liao J C, Spudich J A, Parker D and Delp S L 2007 Proc. Natl Acad. Sci. 104 3171
[63] Linden M and Wallin M 2007 Biophys. J. 92 3804
[64] Tsygankov D, Linder M and Fisher M E 2007 Phys. Rev. E 75 021909
[65] Garai A and Chowdhury D in preparation
[66] Redner S 2001 A Guide to First-Passage Processes (Cambridge: Cambridge University Press)
[67] Qian H and Elson E L 2002 Biophys. Chem. 101-102 565
[68] English B P, min W, van Oijen A M, Lee K T, Luo G, Sun H, Cherayil B J, Kou S C and Xie X S 2006 Nat. Chem. Biol. 2 87
[69] Kou S C, Cherayil B J, Min W, English B P and Xie X S 2005 J. Phys. Chem. B 109 19068
[70] Min W, English B P, Luo G, Cherayil B J, Kou S C and Xie X S 2005 Acc. Chem. Res. 38 923
[71] Min W, Gopich I V, English B P, Kou S C, Xie X S and Szabo A 2006 J. Phys. Chem. B 110 20093–7
[72] Basu M and Mohanty P K 2009 arXiv:0901.2844
[73] Hopfield J J, Yaman S, Yue V and Coutts S M 1976 Proc. Natl Acad. Sci. USA 73 1164
[74] Yamane T and Hopfield J J 1977 Proc. Natl Acad. Sci. 74 2246
[75] Arava Y, Wang Y, Storey J D, Liu C L, Brown P O and Herschlag D 2003 Proc. Natl Acad. Sci. 100 3889
[76] Farabaugh P J 2000 Prog. Nucl. Acid Res. Mol. Biol. 64 131