PLOS ONE

RESEARCH ARTICLE

Protein Synthesis in *E. coli*: Dependence of Codon-Specific Elongation on tRNA Concentration and Codon Usage

Sophia Rudorf*, Reinhard Lipowsky*

Theory and Bio-Systems, Max Planck Institute of Colloids and Interfaces, Potsdam, Germany

* Sophia.Rudorf@mpikg.mpg.de (SR); Reinhard.Lipowsky@mpikg.mpg.de (RL)

Citation: Rudorf S, Lipowsky R (2015) Protein Synthesis in *E. coli*: Dependence of Codon-Specific Elongation on tRNA Concentration and Codon Usage. PLoS ONE 10(8): e0134994. doi:10.1371/journal.pone.0134994

Editor: Hans-Joachim Wieden, University of Lethbridge, CANADA

Received: March 17, 2015

Accepted: July 15, 2015

Published: August 13, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was funded by the German Science Foundation (Deutsche Forschungsgemeinschaft) via Research Unit FOR 1805. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

To synthesize a protein, a ribosome moves along a messenger RNA (mRNA), reads it codon by codon, and takes up the corresponding ternary complexes which consist of aminoacylated transfer RNAs (aa-tRNAs), elongation factor Tu (EF-Tu), and GTP. During this process of translation elongation, the ribosome proceeds with a codon-specific rate. Here, we present a general theoretical framework to calculate codon-specific elongation rates and error frequencies based on tRNA concentrations and codon usages. Our theory takes three important aspects of *in-vivo* translation elongation into account. First, non-cognate, near-cognate and cognate ternary complexes compete for the binding sites on the ribosomes. Second, the corresponding binding rates are determined by the concentrations of free ternary complexes, which must be distinguished from the total tRNA concentrations as measured *in vivo*. Third, for each tRNA species, the difference between total tRNA and ternary complex concentration depends on the codon usages of the corresponding cognate and near-cognate codons. Furthermore, we apply our theory to two alternative pathways for tRNA release from the ribosomal E site and show how the mechanism of tRNA release influences the concentrations of free ternary complexes and thus the codon-specific elongation rates. Using a recently introduced method to determine kinetic rates of *in-vivo* translation from *in-vitro* data, we compute elongation rates for all codons in *Escherichia coli*. We show that for some tRNA species only a few tRNA molecules are part of ternary complexes and, thus, available for the translating ribosomes. In addition, we find that codon-specific elongation rates strongly depend on the overall codon usage in the cell, which could be altered experimentally by overexpression of individual genes.

Introduction

In the past decades, the function and structure of the ribosome were intensively studied [1, 2]. Each ribosome consists of a small and a large subunit, both of which contain RNA and protein molecules. During the initiation of translation, the two subunits bind to the mRNA upstream of its coding sequence and then form a functional ribosome at the start codon of the mRNA

[3]. The subsequent codon-wise movement of the ribosome along the mRNA is called translation elongation. Many *in-vitro* experiments have been performed to identify the individual steps of the elongation process and to reveal the corresponding kinetics [4–7]. Our current view of this process is summarized in Fig 1.

First, a ternary complex consisting of elongation factor EF-Tu, an aminoacyl tRNA (aa-tRNA), and a GTP molecule binds to the ribosome. If the aa-tRNA is cognate and matches the currently read codon, it its accommodated in the ribosomal A site via GTP hydrolysis. The peptide chain is then extended by forming a peptide bond between the amino acid of the "old" tRNA in the ribosome's P site and the amino acid of the "new" tRNA in the A site. Another elongation factor (EF-G) binds to the ribosome and, fueled by a second GTP hydolysis, causes it to translocate to the next codon. Simultaneously, the A-site tRNA is pushed to the P site and the P-site tRNA is moved to the E site (exit), from where it leaves the ribosome. The ribosomal A and P sites are related to the core tasks of a ribosome, i.e., decoding and peptide bond formation. The function of the E site is not yet fully understood but it is believed that it supports translocation of the ribosome [2]. It is widely assumed that the E-site tRNA is released from the ribosome before a new ternary complex binds to the A site. This was named *2-1-2* pathway of E-site tRNA release and has been corroborated by several *in-vitro* experiments [8–11]. However, for the early cycles of elongation corresponding to a relatively short length of the nascent peptide chain, the alternative *2-3-2* pathway has also been observed [10]. In this pathway, the binding of a new ternary complex to the A site precedes the release of the E-site tRNA [12]. After a tRNA is released from the ribosomal E site, it is recharged with an amino acid by an aminoacyl tRNA synthetase and forms a new ternary complex by binding to an EF-Tu molecule, see Fig 2.

Translation is terminated as soon as the ribosome encounters one of the three stop codons. Release factors unbind the peptide chain from the last P-site tRNA and the two ribosome subunits dissociate from the mRNA [13].

Protein synthesis must comply with high requirements concerning speed and accuracy. It must be fast enough to ensure doubling of protein mass within the time scale of cell division. At the same time, the synthesis of proteins must be very precise, because erroneous proteins are often dysfunctional or even harmful to the cell. Therefore, perturbations that hamper or dysregulate protein synthesis can lead to all kinds of cellular defects up to cell death. Indeed, a well-known strategy to fight bacterial infections is based on antibiotics that block protein synthesis in bacteria, hence killing the microorganisms. On the other hand, many diseases in multicellular organisms arise from genetic mutations that affect protein synthesis. Thus, it is of great general interest to understand what determines the speed and accuracy of protein synthesis.

Although each elongation cycle involves basically the same steps, the speed of translation is not uniform. For two decades, it has been known that *in-vivo* elongation rates are codon-specific, i.e., the local velocity of the ribosome depends on the codon that is translated [14, 15]. Many mechanisms have been proposed to explain this variability in the speed of translation, of which a recent review can be found in [16, 17]. For example, mRNA is often folded into a secondary structure which the ribosome has to open up before it can translate the message [18, 19]. Furthermore, certain codon sequences may interact with the ribosome and slow it down [20]. Also, the peptide chain can interact with the ribosome, especially the chain segment that is still in the ribosomal exit tunnel [21]. In addition, ribosomes can be slowed down by preceding ribosomes translating the same message if they come into close contact [22]. Equivalently, codon- or tRNA-dependent variations in the ribosomal processing, and the presence of co-translational processes, like chaperone interactions or co-transcriptional translation, could affect the efficiency of the elongation process [23, 24].

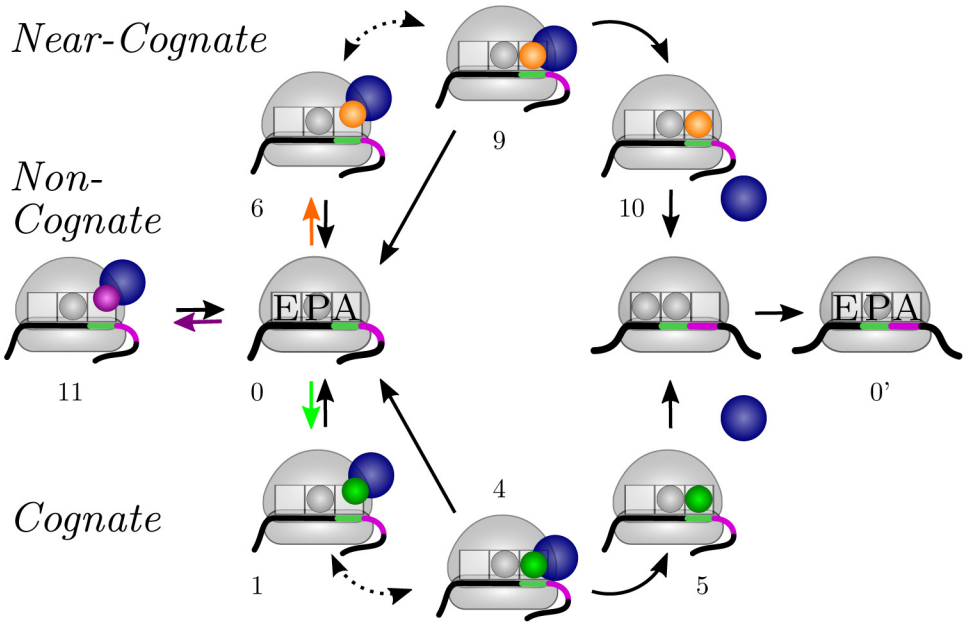

Fig 1. Translation elongation cycle. The ribosome has three tRNA binding sites, the A, P, and E site. A ribosome that has just arrived at a new (green) codon of an mRNA (state "0") has an empty A site, whereas the P site is occupied by a tRNA (here shown as small gray sphere) that is cognate or near-cognate to the preceding codon. Elongation factor EF-Tu (blue spheres), aa-tRNAs (green, orange, and purple small spheres), and GTP molecules (not shown) form ternary complexes. Free cognate, near-cognate and non-cognate ternary complexes bind to the ribosome with rates depending on their respective concentrations (green, orange, and purple arrows from state "0" to states "1", "6", and "11", respectively). Since the initial binding is not codon-specific, all kinds of ternary complexes unbind again from the ribosome with the same dissociation rate. Alternatively, a cognate or near-cognate ternary complex can be recognized by the ribosome (dotted arrows from states "1" and "6" to states "4" and "9", respectively) before the ternary complex is either completely released (arrows from states "4" and "9" to state "0"), brought back to the initial binding state (dotted arrows from states "4" and "9' to states "1" and "6", respectively), or its aa-tRNA is accommodated in the ribosomal A site (arrows from states "4" and "9" to states "5" and "10", respectively). Along with aa-tRNA accommodation, EF-Tu leaves the ribosome. The new A-site tRNA is then further processed and shifted to the P site, while the ribosome translocates to the next (purple) codon. The former P-site tRNA is now in the E site. Depending on the assumed pathway of tRNA release, the E-site tRNA either dissociates very rapidly from the ribosome (2-1-2 pathway), or stays until the next aa-tRNA has been accommodated in the ribosomal A site (2-3-2 pathway). The numerals correspond to the ribosomal states of the codon-specific Markov process introduced below.

doi:10.1371/journal.pone.0134994.g001

One additional factor is believed to determine the local speed of translation, namely the concentration of ternary complexes. In contrast to all of the other influencing factors mentioned above, the concentration of ternary complexes represents a universal control of the speed of translation, independent of the mRNA that is translated or the organism that is studied. Because the ribosome has to wait for a cognate ternary complex to bind before it can proceed, it is plausible that translation is faster when more ternary complexes are present. Based on this idea the tRNA adaptation index (tAI) was introduced and interpreted as a measure for the actual codon-specific elongation rate [25, 26].

From a theoretical point of view, not only the cognate, but also the non- and near-cognate ternary complexes should have an effect on the elongation rate of the corresponding codon, as they can bind to the ribosome as well. Experimentally, their influence has still to be clarified [27–29]. The competition between cognate, near- and non-cognate aa-tRNAs was studied theoretically and was found to strongly affect codon-specific elongation rates [30, 31].

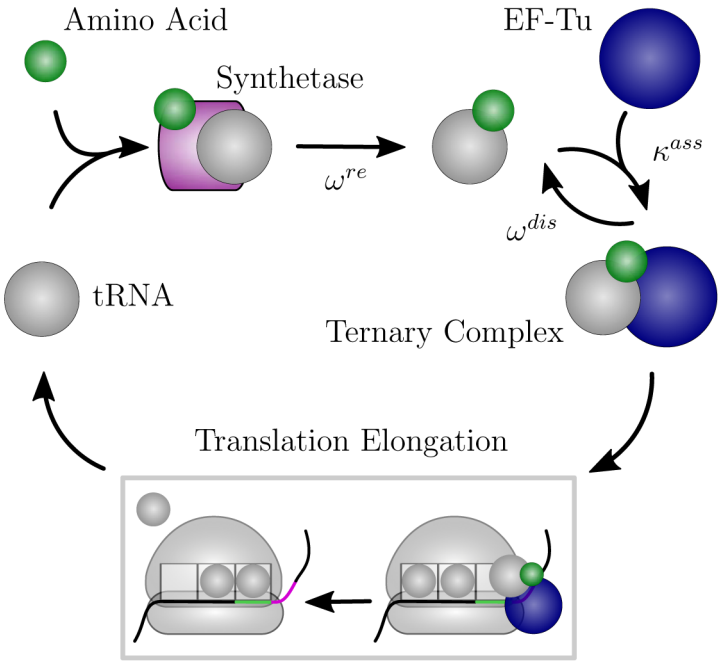

Fig 2. Recharging cycle of tRNAs. During translation elongation an aa-tRNA delivers its amino acid to the elongating peptide chain. After the processed tRNA is released from a ribosomal E site, it binds to an aminoacyl tRNA synthetase. The synthetase recharges the tRNA with a new amino acid, and the recharged aa-tRNA binds to elongation factor EF-Tu to form a new ternary complex.

doi:10.1371/journal.pone.0134994.g002

In addition to the effect of near- and non-cognate aa-tRNAs, the competition between the active ribosomes for ternary complexes also has a strong influence on the elongation speed. Each *Escherichia coli* (*E. coli*) cell contains about twenty to thirty thousand ribosomes actively translating at the same time. As they always keep one tRNA in their P site and sometimes a tRNA in their E site, active ribosomes dramatically reduce the concentrations of tRNAs that can form ternary complexes. Therefore, the simultaneous action of ribosomes can turn abundant into rare ternary complexes. Thus, elongation rates should rather depend on the concentrations of available ternary complexes than on the total amount of tRNAs.

Here, we develop a theoretical framework that allows us to determine the codon-specific speed of translation based on the total, experimentally measured tRNA concentrations. Previously published theories [30–35] on tRNA concentration dependent elongation rates have some limitations because they ignored fundamental aspects of ribosome translation. These aspects include (i) the proper distinction between the concentration of free ternary complexes and the measured tRNA concentrations, a distinction that was not considered in [30, 33]; (ii) the dependence of the free ternary complexes on the recharging of deacetylated tRNA by new amino acids which was not taken into account in [30–32]; (iii) the different *in-vitro* and *in-vivo* values for the rates of aa-tRNA decoding, accommodation, peptide bond formation and translocation [36], a difference that was ignored in [30–32]; and (iv) the competition between cognate, near-, and non-cognate ternary complexes for initial binding to ribosomes, which has a strong influence on the translation process [31] but was ignored in [32, 34, 35]. In the present paper, we include all of these aspects within a single theoretical framework.

In this paper, we will develop a comprehensive theory on translation elongation that takes into account all of the aforementioned mechanisms that influence the concentrations of ternary complexes, as well as the competition of ternary complexes. Furthermore, we use kinetic

rates to describe *in-vivo* translation that were deduced from their corresponding measured *in-vitro* rates by applying a recently published extremum principle, i.e., by minimizing the kinetic distance of the *in-vitro* and *in-vivo* rates [36]. Our theory sheds light on the intricate relationships between the internal dynamics of the ribosome, the availability of ternary complexes, and the codon usages, both for the 2-1-2 and for the 2-3-2 pathway of tRNA release from the E site.

Results

Codons and tRNAs: Some Definitions

We denote by $\mathbf{C}$ the set of sense codons which are labeled by $c = 1, 2, \ldots |\mathbf{C}|$, where in *E. coli* $|\mathbf{C}| = 61$. The different species of tRNA form the set $\mathbf{A}$. These tRNA species will be distinguished by the label $a = 1, 2, \ldots |\mathbf{A}|$. The total number $|\mathbf{A}|$ of different tRNA species depends on the organism. In *E. coli*, there are $|\mathbf{A}| = 43$ different elongator tRNA molecules [37]. For each codon c, there is a set $\mathbf{A}_{\mathrm{co}}(c)$ of cognate tRNAs. All other tRNAs belong either to the set $\mathbf{A}_{\mathrm{nr}}(c)$ of near-cognate or to the set $\mathbf{A}_{\mathrm{no}}(c)$ of non-cognate tRNAs. Therefore, each codon $c \in \mathbf{C}$ leads to a unique partition of the set $\mathbf{A}$ into pairwise disjoint sets $\{\mathbf{A}_{\mathrm{co}}(c), \mathbf{A}_{\mathrm{nr}}(c), \mathbf{A}_{\mathrm{no}}(c)\}$.

In some cases, a certain tRNA species is cognate to more than one codon. The set of codons that are cognate to the tRNA with index a is denoted by $\mathbf{C}_{\mathrm{co}}(a)$, whereas the near- and non-cognate codons are contained in the sets $\mathbf{C}_{\mathrm{nr}}(a)$ and $\mathbf{C}_{\mathrm{no}}(a)$, respectively. Therefore, each tRNA $a \in \mathbf{A}$ leads to a unique partition of the set $\mathbf{C}$ into pairwise disjoint sets $\{\mathbf{C}_{\mathrm{co}}(a), \mathbf{C}_{\mathrm{nr}}(a), \mathbf{C}_{\mathrm{no}}(a)\}$. The complex relationships between cognate, near- and non-cognate tRNAs and codons can be visualized by a large matrix as displayed in Fig 3 for *E. coli*. Note that this matrix has $43 \times 61 = 2623$ elements.

Theoretical Description of the Elongation Cycle

We model translation as a continuous-time Markov process as introduced in [36] to capture the stochastic nature of the ribosomal movement. In contrast to earlier stochastic theories of translation [39, 40], our model of translation elongation as depicted in Fig 1 is based on the detailed chemical kinetics as determined for a certain *in-vitro* assay [7, 41–46]. For each sense codon c, the ribosome can cycle through twelve ribosomal states, numbered from 0 to 11 as shown in Fig 4(A). Each state of the translating ribosome is characterized by the codon in its A site and by the binding of ternary complexes, aa-tRNAs, or tRNAs to the ribosomal A, P, and E sites. A ribosome that has just moved to a codon c has a free A site, but its P site is occupied by a tRNA that is cognate or near-cognate to the preceding codon. The corresponding ribosomal state is denoted by $(c|0)$.

When a ternary complex binds to the ribosome, it can be cognate, near-cognate, or non-cognate to the codon in the ribosomal A site. After binding a cognate ternary complex, the ribosome arrives in state $(c|1)$; the binding of a near-cognate ternary complex leads to state $(c|6)$; finally, if the bound ternary complex is non-cognate to c, the ribosome is in state $(c|11)$. A non-cognate ternary complex always dissociates from the ribosome which then goes back to state $(c|0)$. After initial binding, a cognate ternary complex has to be recognized by the ribosome, which is achieved in state $(c|2)$. The recognition is followed by an activation of GTPase and GTP hydrolysis, leading to ribosomal state $(c|3)$, and release of phosphate and rearrangements of the EF-Tu molecule during the transition from $(c|3)$ to state $(c|4)$. From here, the aa-tRNA gets accommodated and the EF-Tu is released with high probability, leading to state $(c|5)$. However, in rare cases, the ternary complex is released from the ribosome which then returns to state $(c|0)$. We will call the states $(c|1)$ to $(c|5)$ the cognate branch of the Markov process. In addition to the cognate branch, the Markov process also involves an analogous near-cognate branch of another five states: a near-cognate ternary complex is recognized in state

Fig 3. Cognate (blue), near-cognate (purple), and non-cognate tRNAs (white). for all sense codons in *E. coli* following the definitions of "cognate" and "near-cognate" as given in [37] and [38], respectively.

doi:10.1371/journal.pone.0134994.g003

($c|7$) and gets usually rejected during this initial selection step, bringing the ribosome back to state ($c|0$). Alternatively, GTPase is activated, GTP is hydrolyzed and the ribosome with the near-cognate ternary complex proceeds to state ($c|8$). After release of phosphate and EF-Tu rearrangements during the transition to state ($c|9$), the ribosome usually rejects the near-cognate ternary complex and returns to state ($c|0$). The latter step is called proofreading. If the ribosome with the near-cognate ternary complex moves on to state ($c|10$), the near-cognate aa-tRNA becomes fully accommodated in the ribosomal A site and the EF-Tu leaves the ribosome, resulting in a misreading error. After cognate or near-cognate aa-tRNA accommodation, the A-site tRNA is processed which includes peptide bond formation and translocation of the

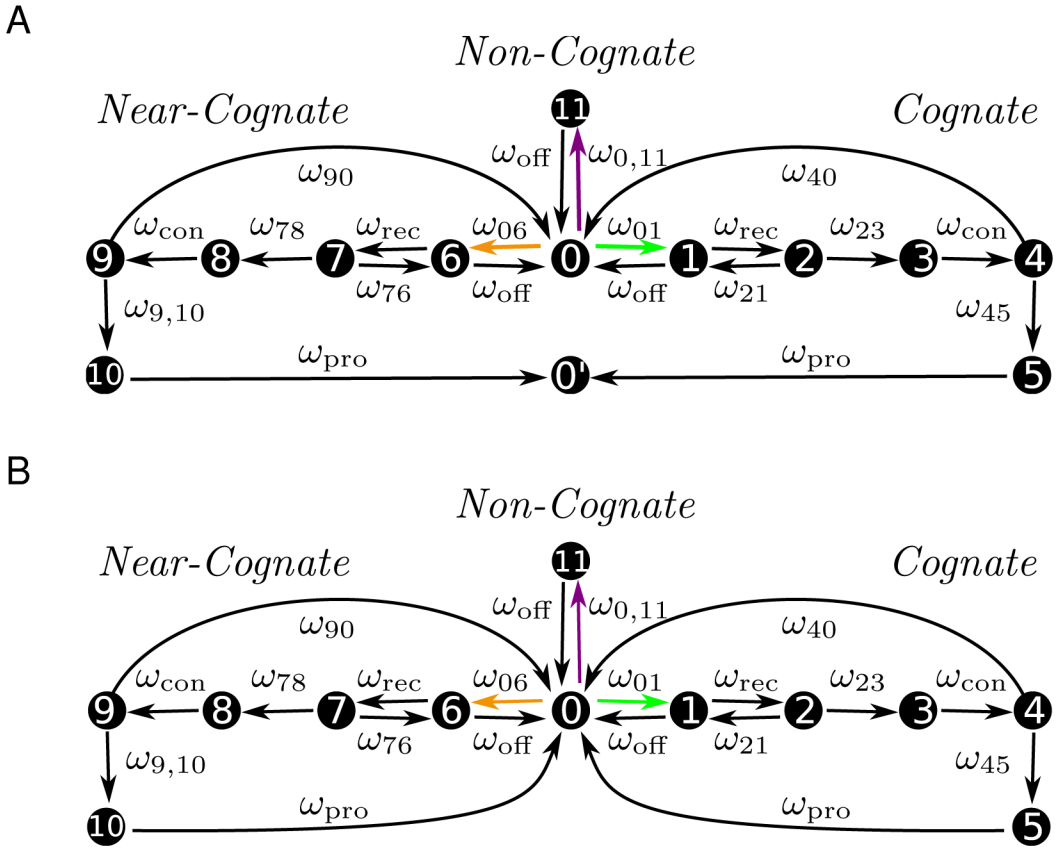

Fig 4. Representation of translation elongation as a Markov process. (A) Each of the substeps of the elongation cycle is represented by one state in the Markov model, which leads to twelve states per sense codon. State ($c|0$) indicates the state assumed by a ribosome reading codon c when it is not bound to a ternary complex. State ($c'|0'$) is attained by a ribosome after translocation to the next codon c'. All rates of transitions between states are taken to be codon-independent, except for the binding rates of cognate, near-, and non-cognate ternary complexes (green, orange and purple arrows). (B) Auxiliary Markov process for the computation of dwell times. The auxiliary Markov process is almost identical to the Markov process of codon-specific elongation depicted in (A), but the absorbing state ($c'|0'$) is omitted. The states ($c|5$) and ($c|10$) are directly connected back to the initial state ($c|0$) with the same transition rate ω_{pro} as in the original process.

doi:10.1371/journal.pone.0134994.g004

ribosome to the next codon along with a transfer of the A-site tRNA to the P site and of the P-site tRNA to the E site. A new elongation cycle on the following codon c' can begin, indicated by the ribosomal state ($c'|0'$) in Fig 4(A). Processed tRNAs leave the ribosome from the E site. Afterwards, they bind to synthetases that recharge them with amino acids, before the recharged aa-tRNAs bind again to EF-Tu to form new ternary complexes, see Fig 2.

Note that some transitions from one state of the Markov process to another state are reversible, which is depicted in Fig 4(A) by pairs of opposing arrows. In contrast, several other transitions are taken to be irreversible as indicated by single arrows. These assumptions of irreversibility are based on experimental findings, i.e., these transitions are irreversible in the sense that the rates of the reverse transitions are too small to be resolved experimentally.

We restrict our theoretical analysis to elongation under low densities of translating ribosomes on mRNAs, i.e., elongation without ribosome-ribosome interactions. In fact, for eukaryotic translation experimental findings suggest that most ribosomes do not interfere under physiological conditions [47]. However, interactions of ribosomes can be easily taken into account if the Markov process is studied by stochastic simulations as for example described in

[36, 40, 48]. In [36], we used such simulations to study the time-dependent incorporation of radioactively labeled amino acids into proteins.

Binding Rates and Internal Transition Rates

The ribosomal transition rate from state i to state j will be denoted by ω_{ij}. Thus, the rates ω_{01}, ω_{06}, and $\omega_{0,11}$ govern the transition from $(c|0)$ to $(c|1)$, $(c|6)$ and $(c|11)$, respectively. The rates ω_{01}, ω_{06}, and $\omega_{0,11}$ depend on the molar concentrations $\hat{\mathrm{X}}_a$ of available, free ternary complexes containing aa-tRNA species a and have the form

$$\omega_{01} = \kappa_{\mathrm{on}} \sum_{a \in \mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a, \tag{1}$$

$$\omega_{06} = \kappa_{\mathrm{on}} \sum_{a \in \mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a, \tag{2}$$

and

$$\omega_{0,11} = \kappa_{\mathrm{on}} \sum_{a \in \mathbf{A}_{\mathrm{no}}(c)} \hat{\mathrm{X}}_a, \tag{3}$$

where κ_{on} is the binding rate constant that we take to be identical for all (cognate, near-cognate, and non-cognate) ternary complexes in agreement with experimental observations [49]. The sets $\mathbf{A}_{\mathrm{co}}(c)$, $\mathbf{A}_{\mathrm{nr}}(c)$, and $\mathbf{A}_{\mathrm{no}}(c)$ contain the tRNA species that are cognate, near-cognate, and non-cognate to codon c as previously defined.

In general, transition rates in the cognate branch of the Markov process differ from their counterparts in the near-cognate branch. As an exception to this rule, based on experimental findings [41, 44] we take the transition rates ω_{12} and ω_{67} to state $(c|2)$ and $(c|7)$ and the transition rates ω_{34} and ω_{89} to state $(c|4)$ and $(c|9)$ to be identical for all cognate and near-cognate ternary complexes. The same assumption is made for the transitions from the states $(c|5)$ and $(c|10)$ to the free state $(c'|0')$ after translocation

$$\omega_{12} = \omega_{67} \equiv \omega_{\mathrm{rec}}, \tag{4}$$

$$\omega_{34} = \omega_{89} \equiv \omega_{\mathrm{con}}, \tag{5}$$

and

$$\omega_{50'} = \omega_{10,0'} \equiv \omega_{\mathrm{pro}}. \tag{6}$$

Furthermore, the dissociation rates from state $(c|1)$, $(c|6)$, or $(c|11)$ back to state $(c|0)$ are taken to be identical for all cognate, near-cognate, and non-cognate ternary complexes [44]

$$\omega_{10} = \omega_{60} = \omega_{11,0} \equiv \omega_{\mathrm{off}}, \tag{7}$$

because dissociation happens before a physical contact between the A-site codon and the anticodon of the aa-tRNA is made and, thus, before a distinction between cognate, near-cognate, and non-cognate ternary complexes is possible. In summary, the 20 transitions of the Markov process in Fig 4(A) are governed by three codon-dependent binding rates ω_{01}, ω_{06}, and $\omega_{0,11}$, and 12 codon-independent internal rates. Numerical values for the latter 12 internal rates are given in S1 Table in the Supporting Information. They were obtained by minimization of the kinetic distance as introduced and described in [36].

Codon-Specific Elongation Rates and Error Frequencies

We define the elongation time $t_{c,elo}$ of codon c as the average time that is needed to finish a complete elongation cycle on codon c. According to the Markov process described in the previous sections, this time is the average time to absorption in state $(c'|0')$ from state $(c|0)$. We define the codon-specific elongation rate $\omega_{c,elo}$ as the inverse of the elongation time $t_{c,elo}$. The codon-specific elongation time can be computed as a mean first passage time using the general theory of Markov processes with absorption [50, 51]. Numerical results for the the codon-specific elongation rates $\omega_{c,elo}$ are given in S2 and S3 Tables in the Supporting Information. Alternatively, we can express the elongation time by the sum of the average dwell times $t_{(c|i)}$ in states $(c|i)$ with $i = 0, 1, \ldots, 11$ per elongation cycle on codon c. This leads to

$$\omega_{c,elo} \equiv t_{c,elo}^{-1} \equiv \left(\sum_{i=0}^{11} t_{(c|i)} \right)^{-1}. \tag{8}$$

To compute the dwell times $t_{(c|i)}$, we introduce an auxiliary Markov process and study its steady state as described by Hill [52]. The auxiliary Markov process is almost identical to the process depicted in Fig 4(A) with one modification: In the auxiliary process, the absorbing state $(c'|0')$ is identified with the state $(c|0)$, see Fig 4(B). Thus, when the ribosome arrives on state $(c'|0')$, it is immediately transferred back to state $(c|0)$. Instead of the transitions $(c|5)$ to $(c'|0')$ and $(c|10)$ to $(c'|0')$, the states $(c|5)$ and $(c|10)$ are connected back to the initial state $(c|0)$ using the same rate ω_{pro} as in the original process. In this way, the states $(c|i)$ are decoupled from all other states that the ribosome visited prior to codon c under consideration. Such a decoupling is possible because the Markov process has no memory. The time dependence of the probabilities $\tilde{P}_{c,i}(t)$ to attain the states $(c|i)$ in the auxiliary process are described by the loss-and-gain (or master) equations

$$\frac{d}{dt} \tilde{P}_{c,i}(t) = \sum_j \left(\tilde{P}_{c,j}(t)\, \omega_{ji} - \tilde{P}_{c,i}(t)\, \omega_{ij} \right), \tag{9}$$

where the transition rates ω_{ij} are defined in the previous section and Fig 4(B). Furthermore, the time dependent probabilities $\tilde{P}_{c,i}(t)$ fulfill the normalization condition

$$\sum_i \tilde{P}_{c,i}(t) = 1. \tag{10}$$

The dwell times $t_{(c|i)}$ can be calculated from the steady state of the auxiliary Markov process. In the steady state, the left side of Eq (9) vanishes and this set of equations together with Eq (10) can be solved for $\tilde{P}^{st}_{c,0}, \tilde{P}^{st}_{c,1}, \ldots, \tilde{P}^{st}_{c,11}$, where $\tilde{P}^{st}_{c,i}$ denotes the steady state value $\tilde{P}^{st}_{c,i} \equiv \tilde{P}_{c,i}(t = \infty)$. The steady state probabilities $\tilde{P}^{st}_{c,i}$ of the auxiliary process are identical to the fractions of time that the original process spends on average in the states $(c|i)$ with initial state $(c|0)$ and absorbing state $(c'|0')$ which implies the relation

$$\tilde{P}^{st}_{c,i} = \frac{t_{(c|i)}}{t_{c,elo}} \tag{11}$$

with the codon-specific elongation time $t_{c,elo}$ defined in Eq (8). Explicit expressions for the resulting dwell times are given in Eqs (S2) to (S13) in the Supporting Information.

It is convenient to express the dwell times in terms of the effective cognate and near-cognate accommodation rates defined by

$$\omega_{c,\mathrm{co}} \equiv \omega_{01}\,\frac{\pi_{12}\pi_{23}\pi_{45}}{1-\pi_{12}\pi_{21}}, \tag{12}$$

$$\omega_{c,\mathrm{nr}} \equiv \omega_{06}\,\frac{\pi_{67}\pi_{78}\pi_{9,10}}{1-\pi_{67}\pi_{76}}. \tag{13}$$

with the probability π_{ij} of transition from state $(c|i)$ to state $(c|j)$ as given by

$$\pi_{ij} = \frac{\omega_{ij}}{\sum_k \omega_{ik}}. \tag{14}$$

Note that the effective accommodation rates $\omega_{c,\mathrm{co}}$ and $\omega_{c,\mathrm{nr}}$ are proportional to the cognate or near-cognate binding rates ω_{01} and ω_{06} multiplied by the respective probabilities that the aa-tRNA of a ternary complex that has bound to a ribosome successfully accommodates in the A site instead of dissociating from the ribosome before accommodation has happened. These probabilities can be computed by first step analysis [50].

Furthermore, we define the probabilities $\mathcal{P}_{c,\mathrm{co}}$ and $\mathcal{P}_{c,\mathrm{nr}}$ for cognate and near-cognate accommodation

$$\mathcal{P}_{c,\mathrm{co}} \equiv \frac{\omega_{c,\mathrm{co}}}{\omega_{c,\mathrm{co}}+\omega_{c,\mathrm{nr}}}, \tag{15}$$

$$\mathcal{P}_{c,\mathrm{nr}} \equiv \frac{\omega_{c,\mathrm{nr}}}{\omega_{c,\mathrm{co}}+\omega_{c,\mathrm{nr}}}. \tag{16}$$

The probabilities of cognate and near-cognate accommodation $\mathcal{P}_{c,\mathrm{co}}$ and $\mathcal{P}_{c,\mathrm{nr}}$ fulfill the normalization condition $\mathcal{P}_{c,\mathrm{co}} + \mathcal{P}_{c,\mathrm{nr}} = 1$. The accommodation probability $\mathcal{P}_{c,\mathrm{co}}$ is a measure of the codon-specific fidelity: if $\mathcal{P}_{c,\mathrm{co}} = 1$, the ribosome translates codon c without any errors and, thus, with maximal fidelity. Likewise, $\mathcal{P}_{c,\mathrm{nr}}$ represents the codon-specific infidelity, which is also named near-cognate missense error frequency.

Codon Usages and Overall Elongation Rate

One quantity that has been determined experimentally for *E. coli* under various growth conditions is the average elongation time per codon, $\langle \mathrm{t}_{c,\mathrm{elo}} \rangle$, [53] which represents the codon-specific elongation times $\mathrm{t}_{c,\mathrm{elo}}$ averaged over all codons in the "coding transcriptome", i.e., the population of all mRNA molecules in the cell. This average involves the codon usages p_c, which represent the normalized frequencies or probabilities that a randomly chosen sense codon in the mRNA population is equal to c. The quantities p_c satisfy the relations

$$0 \le p_c \le 1 \quad \text{and} \quad \sum_{c=1}^{61} p_c = 1. \tag{17}$$

Numerical values of the codon usages in *E. coli* for different specific growth rates are displayed in S7 Table in the Supporting Information. The average elongation time per codon $\langle \mathrm{t}_{c,\mathrm{elo}} \rangle$ and

its inverse, the overall elongation rate ω_{elo}, are then given by

$$\omega_{\mathrm{elo}} \equiv \langle \mathrm{t}_{c,\mathrm{elo}} \rangle^{-1} = \left(\sum_{c=1}^{61} p_c \mathrm{t}_{c,\mathrm{elo}} \right)^{-1} . \tag{18}$$

Probabilities to Attain Ribosomal States

In a cell or in an *in-vitro* translation assay, many mRNAs are usually present at the same time. The total number of all mRNA molecules will be denoted by M. In general, these mRNAs are composed of different codon sequences, although some of them might have the same sequence of codons and, thus, encode the same protein. To identify the state space for translation by non-interacting ribosomes, we first consider the translation of an individual mRNA m of length L_m that is composed of a sequence of codons $c_1, c_2, c_3, \ldots c_{L_m}$, see Fig 5. For the translation process of this mRNA, the ribosomal states described above form the state space

$$\mathbf{S}_m \equiv \bigcup_{x,i} \{(c_x|i)\} \tag{19}$$

where $1 \leq x \leq L_m$ refers to the position of the codon $c_x \in \mathbf{C}$ and $i = 0, 1, \ldots, 11$ indicates the internal state of the ribosome as introduced above. Thus, the state space $\mathbf{S}_m$ consists of $12 \times L_m$ states.

For each state $(c_x|i) \in \mathbf{S}_m$, we can define the time-dependent probability $\mathrm{P}_{c_x,i}(t)$ that a ribosome dwells on codon c_x and attains the ribosomal state i at time t. For the Markov process described above and in Fig 5, the time dependent probabilities fulfill the following master equations

$$\frac{\mathrm{d}}{\mathrm{dt}} \mathrm{P}_{c_{x,0}}(t) = \sum_{j=1,6,11} \left(\mathrm{P}_{c_x,j}(t)\, \omega_{\mathrm{off}} - \mathrm{P}_{c_x,0}(t)\, \omega_{0\mathrm{j}} \right) + \sum_{j=5,10} \mathrm{P}_{c_{x-1},j}(t)\, \omega_{\mathrm{pro}}, \tag{20}$$

for the free states $(c_x|0)$ and $x \geq 2$, and

$$\frac{\mathrm{d}}{\mathrm{dt}} \mathrm{P}_{c_x,i}(t) = \sum_j \left(\mathrm{P}_{c_x,j}(t)\, \omega_{\mathrm{ji}} - \mathrm{P}_{c_x,i}(t)\, \omega_{\mathrm{ij}} \right) \tag{21}$$

for all other states, where the transition rates ω_{ij} are defined in the previous section and Fig 4 (A). Furthermore, the time dependent probabilities $\mathrm{P}_{c_x,i}(t)$ fulfill the normalization condition

$$\sum_{x,i} \mathrm{P}_{c_x,i}(t) = 1. \tag{22}$$

As long as ribosome-ribosome interactions can be ignored, a full Markov model describing translation elongation on all mRNAs in the entire cell or in an *in-vitro* translation assay has a ribosomal state space $\mathbf{S}$ that is composed of all mRNA-specific state spaces $\mathbf{S}_m$

$$\mathbf{S} \equiv \bigcup_m \mathbf{S}_m = \bigcup_m \bigcup_{x,i} \{(c_{x,m}|i)\} \tag{23}$$

where the index m runs over all individual mRNAs $1, \ldots, M$ and $c_{x,m} \in \mathbf{C}$ indicates the codon found at position x in mRNA m. Therefore, the full ribosomal state space of translation elongation consists of $12 \times \Sigma_{m=1}^{M} L_m$ states, where L_m is the length of mRNA m. The average length of mRNAs in bacteria is about 300 codons [54] and there are more than 1000 mRNA molecules

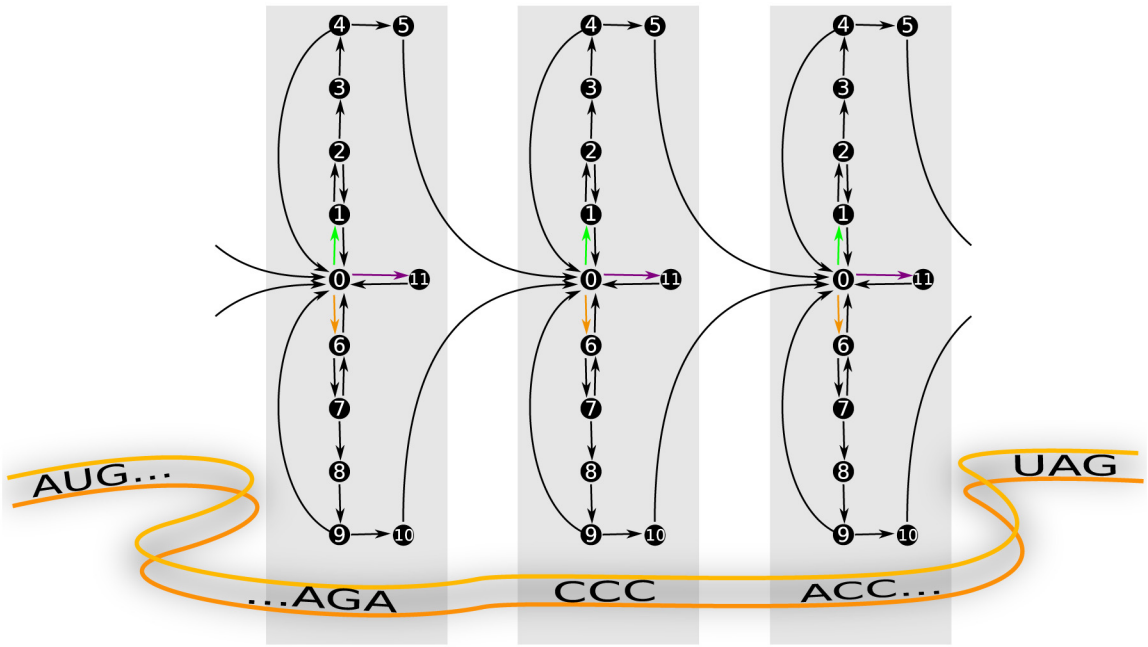

Fig 5. Markov process for the translation of an individual mRNA. The mRNA (orange band) consists of a sequence of codons, where in this example AUG and UAG indicate the start and the stop codon, respectively. The codons are translated sequentially from start to stop. Hence, the Markov process for the translation of an mRNA is given by the concatenation of individual codon-specific Markov processes as defined in the text and Fig 4(A).

doi:10.1371/journal.pone.0134994.g005

in each *E. coli* cell [55]. Thus, for translation in *E. coli* the state space **S** contains more than 3.6×10^6 states that could be attained by each ribosome. Because of this enormous number of states, a detailed time-dependent analysis of translation elongation in the entire cell is hardly feasible.

However, we can analyze the entire Markov process of cell-wide translation elongation under steady state conditions. For a steady state, the flux of ribosomes entering the mRNAs during translation initiation is equal to the flux of ribosomes terminating translation and to the flux of ribosomes elongating on mRNAs. In particular, for every codon in the cell, the flux of arriving ribosomes balances the flux of leaving ribosomes, and this flux is the same on every codon. This means that the steady state probabilities to attain a certain state only depend on the species c of a codon, but not on its position x. In other words, all codons belonging to the same species c can be treated identically. Thus, in a steady state, we can describe translation by a reduced ribosomal state space that consists of $12 \times 61 = 732$ states $(c|i)$, each of which corresponds to a certain codon species c in the ribosomal A site and one of the twelve ribosomal states $i = 0, \ldots, 11$.

Furthermore, in a steady state, the marginal probability $\mathrm{P}_c^{\mathrm{st}}$ to find a ribosome on codon c is simply given by the relative codon-specific elongation time

$$\mathrm{P}_c^{\mathrm{st}} = \frac{p_c \, \mathrm{t}_{c,\mathrm{elo}}}{\langle \mathrm{t}_{c,\mathrm{elo}} \rangle}, \tag{24}$$

with the codon-specific elongation time $\mathrm{t}_{c,\mathrm{elo}}$, the codon usage p_c, and the average elongation time $\langle \mathrm{t}_{c,\mathrm{elo}} \rangle$, as given by the relations Eqs (8), (17), and (18), respectively. Likewise, in a steady state, the joint probability $\mathrm{P}_{c,i}^{\mathrm{st}}$ to find a ribosome on codon c in state $(c|i)$ is equal to

$$\mathrm{P}_{c,i}^{\mathrm{st}} = \mathrm{P}_{i|c}^{\mathrm{st}} \, \mathrm{P}_c^{\mathrm{st}}, \tag{25}$$

where $P^{st}_{i|c}$ is the conditional probability to find a ribosome in state $(c|i)$ under the condition that codon c is in the ribosomal A site. The steady state conditional probabilities $P^{st}_{i|c}$ can be computed by using again the auxiliary Markov process introduced above and in Fig 4(B). In particular, we can identify the steady state probabilities $\tilde{P}^{st}_{c,i}$ of the auxiliary process with the steady state conditional probabilities $P^{st}_{i|c}$

$$\tilde{P}^{st}_{c,i} = P^{st}_{i|c}. \tag{26}$$

Thus, in a steady state, the joint probability $P^{st}_{c,i}$ to find a ribosome in state $(c|i)$ on codon c can be expressed by the dwell times as in Eq (11), and by the marginal probability in Eq (24), which leads to

$$P^{st}_{c,i} = \frac{t_{(c|i)}}{t_{c,elo}} \frac{p_c t_{c,elo}}{\langle t_{c,elo} \rangle} = \frac{p_c t_{(c|i)}}{\langle t_{c,elo} \rangle}. \tag{27}$$

Steady State Concentrations of Free Ternary Complexes

It is important to note that the codon-specific elongation rates defined in Eq (8) and, thus, the expression for the overall elongation rate ω_{elo} as in Eq (18) involve the binding rates ω_{01}, ω_{06}, and $\omega_{0,11}$ for cognate, near-cognate, and non-cognate ternary complexes. The latter rates are given by expressions Eqs (1)–(3), which depend on the concentrations $\hat{X}_a$ of *free* ternary complexes. Therefore, we derive implicit equations to compute these concentrations from measured *total* tRNA concentrations. The details of this derivation are given in the Supporting Information. As a result, we obtain the free concentration $\hat{X}_b$ of ternary complex species b as an implicit function of all free concentrations $\hat{X}_a$ of ternary complex species a with $a = 1, \ldots, 43$, the concentration E^{fr} of free EF-Tu molecules as determined by Eq (S49) in the Supporting Information, the concentration $\mathcal{R}$ of ribosomes, the codon-dependent probabilities $\mathcal{P}_{c,co}$ and $\mathcal{P}_{c,nr}$ of cognate and near-cognate accommodation defined in Eqs (15) and (16), and the codon usages p_c

$$\hat{X}_b = X_b \left(1 + \frac{\omega^{dis}}{\kappa^{ass} \mathcal{E}^{fr}} + \mathcal{R} \left(\Phi_{co} \sum_{c \in \mathbf{C}_{co}(b)} \frac{\mathcal{P}_{c,co} p_c}{\sum_{a \in \mathbf{A}_{co}(c)} \hat{X}_a} + \Phi_{nr} \sum_{c \in \mathbf{C}_{nr}(b)} \frac{\mathcal{P}_{c,nr} p_c}{\sum_{a \in \mathbf{A}_{nr}(c)} \hat{X}_a} + \omega_{elo} \tau_{no} \sum_{c \in \mathbf{C}_{no}(b)} \frac{\mathcal{P}_{c,co} p_c}{\sum_{a \in \mathbf{A}_{co}(c)} \hat{X}_a} \right) \right)^{-1}, \tag{28}$$

where κ^{ass} and ω^{dis} are the binding rate constant and the dissociation rate, respectively, which govern ternary complex formation from free Ef-Tu molecules and aa-tRNAs. For the 2-3-2 pathway of tRNA release from the ribosomal E site, the dimensionless constants Φ_{co} and Φ_{nr} assume the values

$$\Phi_{co} = \Phi^{2-3-2}_{co} \equiv 2 + \omega_{elo} \left(\tau_{co} + \frac{1}{\omega^{re}} + \frac{1}{\kappa^{ass} \mathcal{E}^{fr}} \right), \tag{29}$$

$$\Phi_{nr} = \Phi^{2-3-2}_{nr} \equiv 2 + \omega_{elo} \left(\tau_{nr} + \frac{1}{\omega^{re}} + \frac{1}{\kappa^{ass} \mathcal{E}^{fr}} \right), \tag{30}$$

where ω^{re} is the rate governing the recharging of de-aminoacylated tRNAs by synthetases with

new amino acids, see Fig 2. For the 2-1-2 pathway we obtain

$$\Phi_{\mathrm{co}} = \Phi_{\mathrm{co}}^{2-1-2} \equiv 1 + \omega_{\mathrm{elo}} \left(\tau_{\mathrm{co}} + \frac{1}{\omega_{\mathrm{pro}}} + \frac{1}{\omega^{\mathrm{re}}} + \frac{1}{\kappa^{\mathrm{ass}} \mathcal{E}^{\mathrm{fr}}} \right), \tag{31}$$

$$\Phi_{\mathrm{nr}} = \Phi_{\mathrm{nr}}^{2-1-2} \equiv 1 + \omega_{\mathrm{elo}} \left(\tau_{\mathrm{nr}} + \frac{1}{\omega_{\mathrm{pro}}} + \frac{1}{\omega^{\mathrm{re}}} + \frac{1}{\kappa^{\mathrm{ass}} \mathcal{E}^{\mathrm{fr}}} \right), \tag{32}$$

with constant time scales

$$\tau_{\mathrm{co}} = \frac{1}{\omega_{\mathrm{rec}} \pi_{23} \pi_{45}} + \frac{1}{\omega_{23} \pi_{45}} + \frac{1}{\omega_{\mathrm{con}} \pi_{45}} + \frac{1}{\omega_{45}}, \tag{33}$$

$$\tau_{\mathrm{nr}} = \frac{1}{\omega_{\mathrm{rec}} \pi_{78} \pi_{9,10}} + \frac{1}{\omega_{78} \pi_{9,10}} + \frac{1}{\omega_{\mathrm{con}} \pi_{9,10}} + \frac{1}{\omega_{9,10}}, \tag{34}$$

$$\tau_{\mathrm{no}} = \frac{1}{\omega_{\mathrm{rec}} \pi_{23} \pi_{45}} + \frac{1}{\omega_{\mathrm{off}} \pi_{45}}, \tag{35}$$

where π_{ij} represents the probability of transition from state $(c|i)$ to state $(c|j)$, see definition (Eq 14). Since *E. coli* contains 43 different elongator tRNA species b, there are 43 different Eq (28).

For a specific growth rate of 2.5 h^{-1} and the 2-3-2 pathway of E-site tRNA release, Fig 6 shows the resulting concentrations $\hat{X}_b$ of free ternary complexes relative to the measured total concentrations X_b of the corresponding tRNA molecules given in S4 Table in the Supporting Information. We find that the relative concentrations of free ternary complexes strongly depend on the tRNA species: About 87% of all $tRNA^{Ile2}$ molecules are part of free ternary complexes. In contrast, the concentration of free ternary complexes containing $tRNA^{Lys}$ is only about 8% of the total concentration of $tRNA^{Lys}$ molecules. Therefore, the difference between tRNA concentration and free ternary complex concentration is highly significant and cannot be neglected. Concentrations of free ternary complexes for different specific growth rates and both E-site tRNA release mechanisms can be found in S5 and S6 Tables in the Supporting Information.

Influence of Codon Usage on Dynamics of Translation

In this section, we apply the theory developed in the previous sections to study the influence of codon usage on codon-specific elongation rates. The strong expression of artificially induced genes leads to a shift of the codon usages towards the frequencies that the codons have in the induced gene. Because the codon usages determine how many of the corresponding tRNAs and ternary complexes are bound to ribosomes, they also affect the concentrations of available ternary complexes and ultimately the codon-specific elongation rates as well as missense error frequencies. Therefore, overexpressing a gene can notably change the codon-specific elongation rates with potential consequences for protein folding or ribosome stalling.

To test this hypothesis, we calculated the codon-specific elongation rates using codon usages that are equal to the frequencies of the codons in the β-galactosidase gene *lacZ*. This corresponds for example to a cell in which the vast majority of ribosomes are translating *lacZ* mRNAs, i.e., a cell in the limit of very strong β-galactosidase overexpression. A direct comparison of the β-galactosidase codon frequencies with codon usages from wild type *E. coli* cells growing at a specific growth rate of 2.5 h^{-1} is given in Fig 7(A): the *lacZ* codon usages of codons AAA and UGG differ the most from their corresponding wild type values. The wild type *rare*

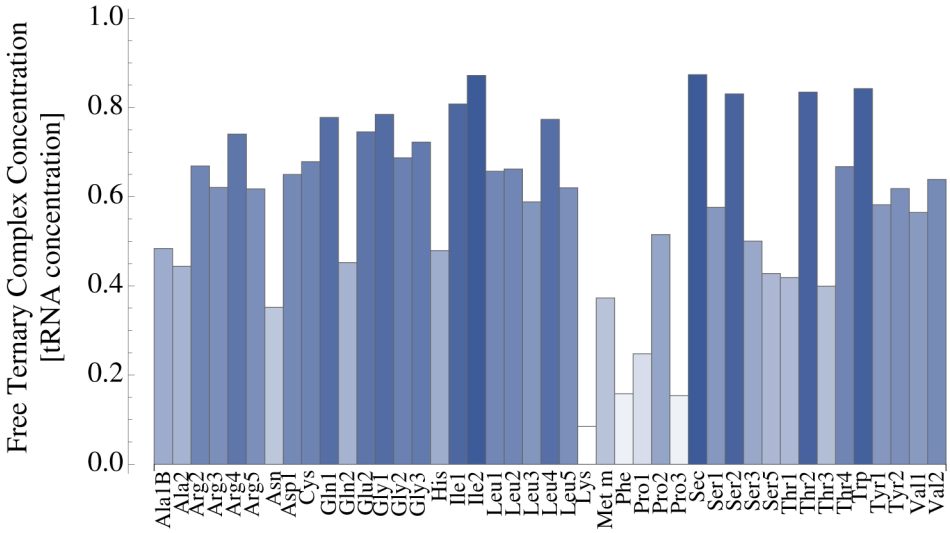

Fig 6. Concentrations of free ternary complexes relative to the total concentrations of the corresponding tRNAs. for all 43 tRNA species in *E. coli* cells growing at a specific rate of 2.5 h^{-1} under the assumption of the 2-3-2 pathway of E-site tRNA release. As an example, about 87% of all $tRNA^{Ile2}$ molecules are part of free ternary complexes. In contrast, only about 8% of all $tRNA^{Lys}$ molecules are contained within free ternary complexes.

doi:10.1371/journal.pone.0134994.g006

codon UGG is *abundant* in *lacZ*, which leads to a decrease of its elongation rate in cells overexpressing β-galactosidase by about a factor of two when E-site tRNA is released via the 2-1-2 pathway, see Fig 7(B). This codon usage related reduction of elongation rate is even more pronounced if one assumes a 2-3-2 pathway of E-site tRNA release: here, the elongation rate of UGG is decreased by about a factor of six when shifting from wild type to *lacZ* codon usage.

In contrast, the codon AAA is highly abundant in wild type cells but is not a frequent codon in lacZ. Therefore, the AAA elongation rate is increased in β-galactosidase overexpressing cells by factors of about two and six for the 2-1-2 and the 2-3-2 pathway of E-site tRNA release, respectively. In addition, the elongation rate of AAG is also increased in β-galactosidase overexpressing cells, although its codon usage is comparable to its wild type value. Both AAG and AAA are cognate to $tRNA^{Lys}$ and therefore an increase in free Lys-$tRNA^{Lys}$ ternary complex also increases the elongation rate of AAG.

To elucidate the interdependence of codon usage, ternary complex concentration, elongation rate and missense error frequency, we varied the codon usage of one particular codon while keeping the ratios of the codon usages of all other codons at their wild type values. All codon usages were rescaled to fulfill the normalization condition Eq (17). In particular, we have calculated the functional dependence of (i) the concentration of available Lys-$tRNA^{Lys}$ ternary complex, (ii) the codon-specific elongation rates of the corresponding cognate codons AAA and AAG, and (iii) the missense error frequencies of both codons as a function of AAA codon usage for the two alternative pathways of E-site tRNA release, see Fig 8. The concentration of free Lys-$tRNA^{Lys}$ ternary complexes decreases when the codon usage of its cognate codon AAA increases. This depletion of free Lys-$tRNA^{Lys}$ ternary complexes leads to decreasing elongation rates of the codons AAA and AAG. As an additional consequence of the decrease in free Lys-$tRNA^{Lys}$ ternary complex concentration, the near-cognate missense error frequencies of both codons AAA and AAG rise when the codon usage of AAA increases because of the increasing probability that a near-cognate ternary complex gets accommodated in the ribosomal A site.

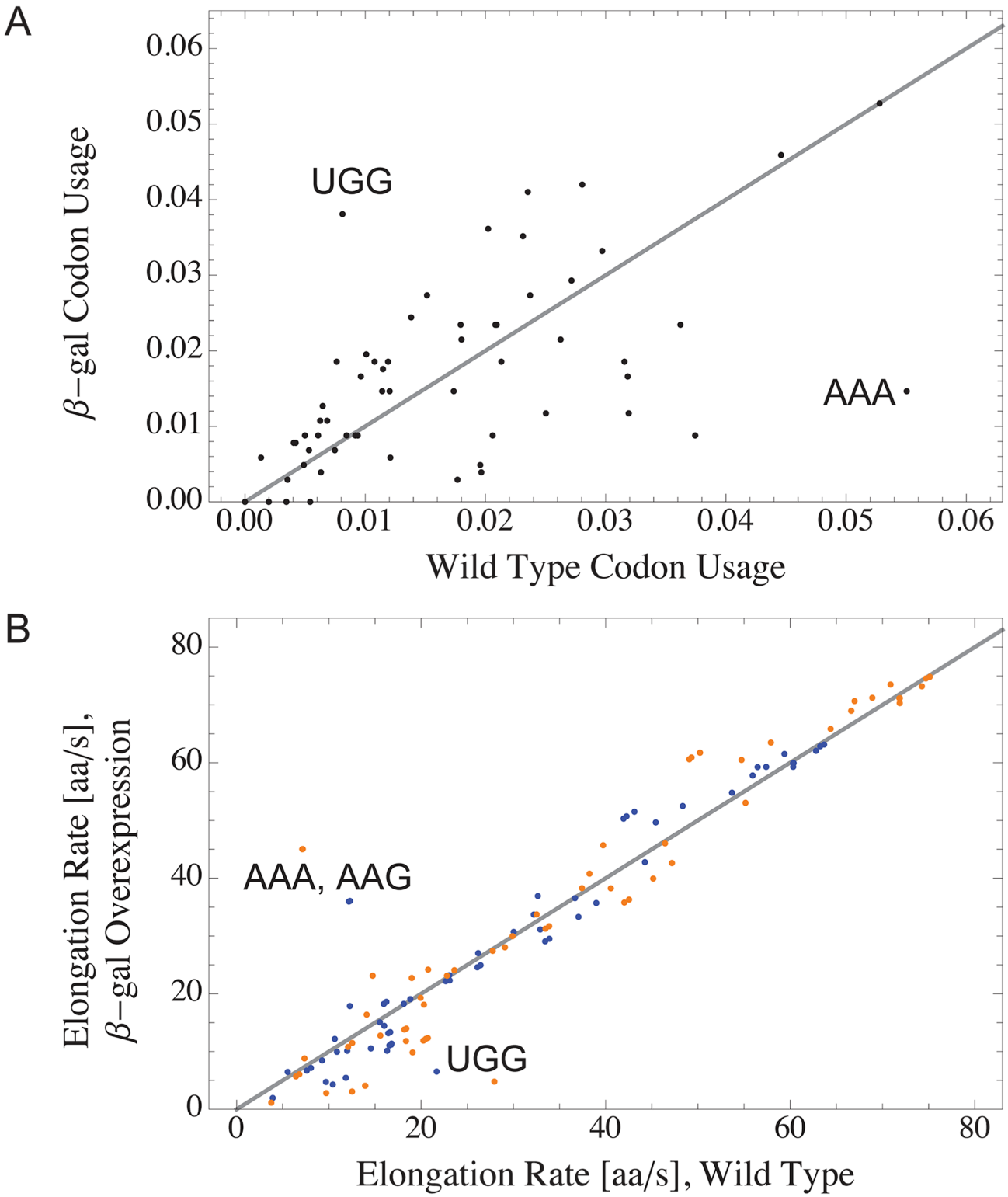

Fig 7. Influence of β-galactosidase overexpression on codon-specific elongation rates. (A) Codon usages of all sense codons in wild type *E. coli* cells growing at a specific rate of 2.5 h^{-1} compared to codon usages in *E. coli* cells overexpressing β-galactosidase. The largest changes are obtained for the codon usages of AAA and UGG. (B) Elongation rates of all 61 sense codons in wild type *E. coli* cells growing at a specific rate of 2.5 h^{-1} compared to codon-specific elongation rates in *E. coli* cells overexpressing β-galactosidase encoded by the lacZ gene, calculated for the 2-1-2 pathway (blue) and the 2-3-2 pathway (orange) of E-site tRNA release. Note that the elongation rates of the codons AAA and AAG are almost identical for both the 2-1-2 pathway (blue)

and the 2-3-2 pathway (orange). Except for the codon usages, all parameters used to calculate the elongation rates were identical in all four calculations; in particular, all calculations used the same set of total concentrations of all tRNA species.

doi:10.1371/journal.pone.0134994.g007

Our results show that the overexpression of a gene can strongly influence translation dynamics and fidelity due to changes in the codon usages. Altering the codon sequence of the desired gene by using synonymous codons that are apparently faster or less error-prone can in fact lead to slower or more erroneous translation if the resulting codon usages are much higher than in the wild type cell. Thus, instead of replacing all codons coding for a specific amino acid by the same apparently optimal synonymous codon, our results suggest that codon optimization may be improved by the possible substitution of a given codon by several synonymous codons that are cognate to different tRNAs.

Discussion

It is of great relevance for medical therapies and for the biotechnological production of protein-based substances to understand what determines the speed and accuracy of protein synthesis. To address this long-lasting puzzle, we developed a comprehensive theoretical framework on protein synthesis by ribosomes in bacteria based on current biochemical knowledge about ribosomal kinetics *in vitro* [7, 41–46]. We described translation elongation as a Markov process with 12 different ribosomal transition rates and considered two alternative pathways of tRNA release from the ribosomal E site. The ribosomal transition rates were determined by minimizing the kinetic distance to the measured *in-vitro* rates, a method that was recently introduced in [36]. A fundamental ingredient of our modeling is the dependence of codon-specific elongation rates and missense error frequencies on the concentrations of ternary complexes available for uptake by translating ribosomes. Thus, in contrast to earlier work, we distinguished the total concentrations of tRNA molecules from the concentrations of free ternary complexes and took the binding of aa-tRNAs and tRNAs to active ribosomes as well as the recharging of deacylated tRNAs with new amino acids into account. We analyzed the steady state of the translation process and found that for some tRNA species, e.g. $tRNA^{Lys}$, only a minor fraction of the tRNA molecules are actually incorporated into free ternary complexes and, thus, available for ribosomal uptake. Therefore, calculations of the speed of protein synthesis that are based on total tRNA concentrations instead of free ternary complex concentrations involve systematic errors.

Furthermore, we investigated how elongation rates and missense error frequencies are affected by codon usages. We showed that the overexpression of a single gene leads to altered codon-specific elongation rates such that codons that are slow in wild type cells can become fast in overexpressing cells and *vice versa*. This behavior is predicted to occur when the codon usages in the overexpressed gene are different from those in wild type cells.

The codon-specific Markov process introduced here can be used to study the dependence of the speed and accuracy of *in-vivo* and *in-vitro* protein synthesis on a variety of parameters which have not been considered in this work. For example, one may study how these quantities vary with changes in the overall ternary complex composition or how changes in internal transition rates, arising for example from ribosomal protein or rRNA mutagenesis, affect the synthesis rate.

In the study presented here, we chose to focus on translation in *E. coli* because of the extensive data base available for these cells. However, our theory is quite general and should in principle be applicable to all prokaryotic and eukaryotic cells.

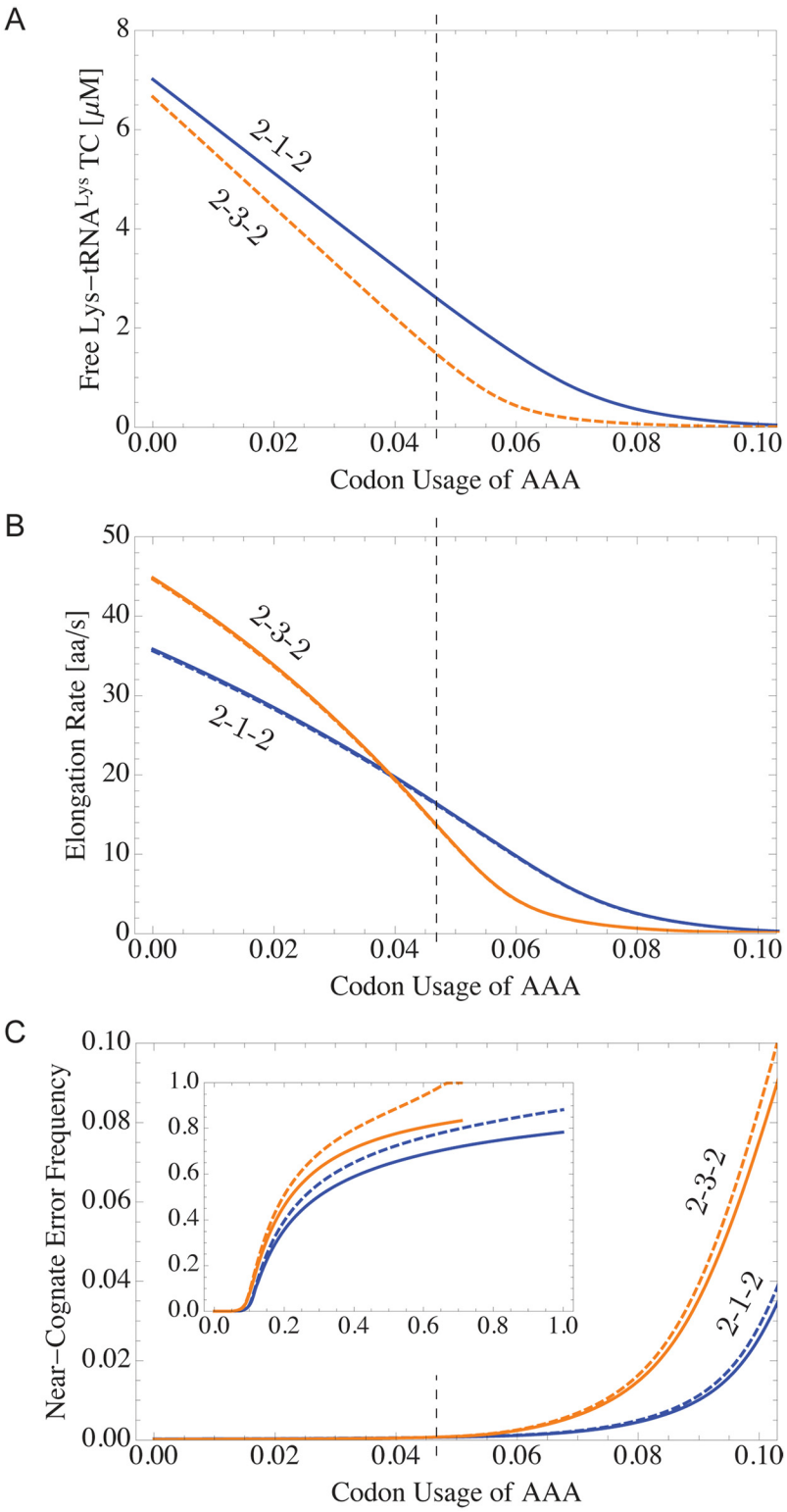

Fig 8. Changing the codon usage of AAA for fixed ratios of the codon usages for all other codons. (A) The concentration of free Lys-tRNA^{Lys} ternary complexes decreases when the codon usage of one of its cognate codons, AAA, increases. (B) Elongation rates of codons AAA (solid line) and AAG (dashed line), both of which are cognate to tRNA^{Lys}. The solid and dashed lines coincide almost perfectly. (C) Near-cognate missense error frequencies for both codons (AAA: solid; AAG: dashed). Results are shown for both the 2-1-2

pathway (blue) and the 2-3-2 pathway (orange) of E-site tRNA release. The vertical dashed lines (black) indicate the wild type value 0.0467 of AAA codon usage, see S7 Table in the Supporting Information.

doi:10.1371/journal.pone.0134994.g008

Supporting Information

S1 Text. Supporting Information Text.
(PDF)

S1 Table. *In-vivo* rates of ribosomal transitions. The values of the overall elongation rate ω_{elo} for the four specific growth rates 0.7, 1.07, 1.6, and 2.5 h^{-1} were obtained from the data in [53]. These growth rates have been chosen because the total tRNA concentrations have been measured for these conditions [37] as well. The *in-vivo* rates of ribosomal transitions (with relative standard deviations RSD) were obtained under the assumption of a 2-1-2 (top) or a 2-3-2 (bottom) pathway of tRNA release from the ribosomal E site by minimizing the kinetic distance of *in-vitro* and *in-vivo* rates as described in [36].
(PDF)

S2 Table. Codon-specific elongation rates $\omega_{c,\mathrm{elo}}$ for all sense codons c in *E. coli*, assuming a 2-1-2 pathway of tRNA release from the E site.
(PDF)

S3 Table. Codon-specific elongation rates $\omega_{c,\mathrm{elo}}$ for all sense codons c in *E. coli*, assuming a 2-3-2 pathway of tRNA release from the E site.
(PDF)

S4 Table. *In-vivo* concentrations of all tRNAs, actively translating ribosomes $\mathcal{R}$, and EF-Tu molecules $\mathcal{E}$ in *E. coli* for four different specific growth rates. Concentrations of tRNAs from Table 5 in [37], with $\mathrm{tRNA}^{\mathrm{Val2A}}$ and $\mathrm{tRNA}^{\mathrm{Val2B}}$ added (Val2), and individual concentrations of $\mathrm{tRNA}^{\mathrm{Gly1}}$ and $\mathrm{tRNA}^{\mathrm{Gly2}}$ as well as $\mathrm{tRNA}^{\mathrm{Ile1}}$ and $\mathrm{tRNA}^{\mathrm{Ile2}}$ obtained by using corresponding ratios given in [56]. The corresponding *in-vivo* concentrations $\mathcal{R}$ of active ribosomes are calculated from Table 3 in [37] by taking into account that only 85% of all ribosomes in the cell are active [53]. Furthermore, the total *in-vivo* concentration of EF-Tu can be estimated by interpolating the measured ratios of EF-Tu and ribosome concentrations for different specific growth rates, see [57], and multiplying these ratios by the ribosome concentration $\mathcal{R}$.
(PDF)

S5 Table. Concentrations of free ternary complexes in *E. coli* for four different specific growth rates, assuming a 2-1-2 pathway of tRNA release from the E site.
(PDF)

S6 Table. Concentrations of free ternary complexes in *E. coli* for four different specific growth rates, assuming a 2-3-2 pathway of tRNA release from the E site.
(PDF)

S7 Table. *In-vivo* codon usages p_c in percent for all sense codons c in *E. coli*. For the specific growth rate of 2.5 h^{-1}, the codon usages were determined from relative mRNA abundances of 4215 different genes [58]. In all other cases, data from [37] were used and rescaled to exclude stop codons.
(PDF)

Acknowledgments

We thank Angelo Valleriani for stimulating discussions during the initial phase of this study.

Author Contributions

Conceived and designed the experiments: SR RL. Performed the experiments: SR RL. Analyzed the data: SR RL. Wrote the paper: SR RL.

References

1. Ramakrishnan V. Ribosome structure and the mechanism of translation. Cell. 2002; 108(4):557–572. doi: 10.1016/S0092-8674(02)00619-0 PMID: 11909526
2. Schmeing TM, Ramakrishnan V. What recent ribosome structures have revealed about the mechanism of translation. Nature. 2009; 461(7268):1234–1242. doi: 10.1038/nature08403 PMID: 19838167
3. Marintchev A, Wagner G. Translation initiation: structures, mechanisms and evolution. Quarterly Reviews of Biophysics. 2004; 37(3–4):197–284. PMID: 16194295
4. Tomšic J, Vitali LA, Daviter T, Savelsbergh A, Spurio R, Striebeck P, et al. Late events of translation initiation in bacteria: a kinetic analysis. EMBO Journal. 2000; 19(9):2127–2136. doi: 10.1093/emboj/19.9.2127 PMID: 10790378
5. Rodnina MV, Gromadski KB, Kothe U, Wieden HJ. Recognition and selection of tRNA in translation. FEBS Letters. 2005; 579(4):938–942. doi: 10.1016/j.febslet.2004.11.048 PMID: 15680978
6. Savelsbergh A, Katunin VI, Mohr D, Peske F, Rodnina MV, Wintermeyer W. An elongation factor G-induced ribosome rearrangement precedes tRNA-mRNA translocation. Molecular Cell. 2003; 11(6):1517–1523. doi: 10.1016/S1097-2765(03)00230-2 PMID: 12820965
7. Fischer N, Konevega AL, Wintermeyer W, Rodnina MV, Stark H. Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy. Nature. 2010; 466(7304):329–333. doi: 10.1038/nature09206 PMID: 20631791
8. Semenkov YP, Rodnina MV, Wintermeyer W. The "allosteric three-site model" of elongation cannot be confirmed in a well-defined ribosome system from Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America. 1996; 93:12183–12188. doi: 10.1073/pnas.93.22.12183 PMID: 8901554
9. Uemura S, Aitken CE, Korlach J, Flusberg BA, Turner SW, Puglisi JD. Real-time tRNA transit on single translating ribosomes at codon resolution. Nature. 2010; 464:1012–1017. doi: 10.1038/nature08925 PMID: 20393556
10. Chen C, Stevens B, Kaur J, Smilansky Z, Cooperman BS, Goldman YE. Allosteric vs. spontaneous exit-site (E-site) tRNA dissociation early in protein synthesis. Proceedings of the National Academy of Sciences of the United States of America. 2011; 108(41):16980–16985. doi: 10.1073/pnas.1106999108 PMID: 21969541
11. Chen J, Petrov A, Tsai A, O'Leary SE, Puglisi JD. Coordinated conformational and compositional dynamics drive ribosome translocation. Nature Structural & Molecular Biology. 2013; 20(6):718–727. doi: 10.1038/nsmb.2567
12. Wilson DN, Nierhaus KH. The E-site story: the importance of maintaining two tRNAs on the ribosome during protein synthesis. Cellular and Molecular Life Sciences. 2006; 63:2725–2737. doi: 10.1007/s00018-006-6125-4 PMID: 17013564
13. Petry S, Weixlbaumer A, Ramakrishnan V. The termination of translation. Current Opinion in Structural Biology. 2008; 18(1):70–77. doi: 10.1016/j.sbi.2007.11.005 PMID: 18206363
14. Sørensen MA, Pedersen S. Absolute in vivo Translation Rates of Individual Codons in Escherichia coli—the 2 Glutamic-Acid Codons GAA and GAG Are Translated with a Threefold Difference in Rate. Journal of Molecular Biology. 1991; 222(2):265–280.
15. Curran JF, Yarus M. Rates of Aminoacyl-Trans-RNA Selection at 29 Sense Codons *in Vivo*. Journal of Molecular Biology. 1989; 209(1):65–77. doi: 10.1016/0022-2836(89)90170-8 PMID: 2478714
16. Spencer PS, Barral JM. Genetic code redundancy and its influence on the encoded polypeptides. Computational And Structural Biotechnology Journal. 2012; 1(1):e201204006. doi: 10.5936/csbj.201204006 PMID: 24688635
17. Tarrant D, von der Haar T. Synonymous codons, ribosome speed, and eukaryotic gene expression regulation. Cellular and Molecular Life Sciences. 2014 Nov; 71(21):4195–4206. doi: 10.1007/s00018-014-1684-2 PMID: 25038778

18. Wen JD, Lancaster L, Hodges C, Zeri AC, Yoshimura SH, Noller HF, et al. Following translation by single ribosomes one codon at a time. Nature. 2008; 452(7187):598–U2. doi: 10.1038/nature06716 PMID: 18327250
19. Chen C, Zhang H, Broitman SL, Reiche M, Farrell I, Cooperman BS, et al. Dynamics of translation by single ribosomes through mRNA secondary structures. Nature Structural & Molecular Biology. 2013; 20(5):582–588. Available from: doi: 10.1038/nsmb.2544
20. Li GW, Oh E, Weissman JS. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. Nature. 2012; 484(7395):538–U172. doi: 10.1038/nature10965 PMID: 22456704
21. Charneski CA, Hurst LD. Positively Charged Residues Are the Major Determinants of Ribosomal Velocity. PLoS Biology. 2013; 11(3). doi: 10.1371/journal.pbio.1001508 PMID: 23554576
22. Mitarai N, Sneppen K, Pedersen S. Ribosome collisions and translation efficiency: Optimization by codon usage and mRNA destabilization. Journal of Molecular Biology. 2008; 382(1):236–245. doi: 10.1016/j.jmb.2008.06.068 PMID: 18619977
23. Kramer G, Rauch T, Rist W, Vorderwülbecke S, Patzelt H, Schulze-Specking A, et al. L23 protein functions as a chaperone docking site on the ribosome. Nature. 2002; 419(6903):171–174. doi: 10.1038/nature01047 PMID: 12226666
24. Kramer G, Boehringer D, Ban N, Bukau B. The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. Nature Structural & Molecular Biology. 2009; 16 (6):589–597. doi: 10.1038/nsmb.1614
25. dos Reis M, Savva R, Wernisch L. Solving the riddle of codon usage preferences: a test for translational selection. Nucleic Acids Research. 2004; 32(17):5036–5044. doi: 10.1093/nar/gkh834 PMID: 15448185
26. Tuller T, Carmi A, Vestsigian K, Navon S, Dorfan Y, Zaborske J, et al. An Evolutionarily Conserved Mechanism for Controlling the Efficiency of Protein Translation. Cell. 2010; 141(2):344–354. doi: 10.1016/j.cell.2010.03.031 PMID: 20403328
27. Bilgin N, Ehrenberg M, Kurland C. Is Translation Inhibited by Noncognate Ternary Complexes. FEBS Letters. 1988; 233(1):95–99. doi: 10.1016/0014-5793(88)81362-0 PMID: 2454844
28. Wohlgemuth I, Pohl C, Rodnina MV. Optimization of speed and accuracy of decoding in translation. EMBO Journal. 2010; 29(21):3701–3709. doi: 10.1038/emboj.2010.229 PMID: 20842102
29. Johansson M, Bouakaz E, Lovmar M, Ehrenberg M. The kinetics of ribosomal peptidyl transfer revisited. Molecular Cell. 2008; 30(5):589–598. doi: 10.1016/j.molcel.2008.04.010 PMID: 18538657
30. Fluitt A, Pienaar E, Vijoen H. Ribosome kinetics and aa-tRNA competition determine rate and fidelity of peptide synthesis. Computational Biology and Chemistry. 2007; 31(5–6):335–346. doi: 10.1016/j.compbiolchem.2007.07.003 PMID: 17897886
31. Zouridis H, Hatzimanikatis V. Effects of codon distributions and tRNA competition on protein translation. Biophysical Journal. 2008; 95(3):1018–1033. doi: 10.1529/biophysj.107.126128 PMID: 18359800
32. Zouridis H, Hatzimanikatis V. A model for protein translation: Polysome self-organization leads to maximum protein synthesis rates. Biophysical Journal. 2007; 92(3):717–730. doi: 10.1529/biophysj.106.087825 PMID: 17098800
33. Zhang G, Ignatova Z. Generic Algorithm to Predict the Speed of Translational Elongation: Implications for Protein Biogenesis. PLoS ONE. 2009 Apr 3; 4(4):e5036. doi: 10.1371/journal.pone.0005036 PMID: 19343177
34. Zhang G, Fedyunin I, Miekley O, Valleriani A, Moura A, Ignatova Z. Global and local depletion of ternary complex limits translational elongation. Nucleic Acids Research. 2010; 38:4778–4787. doi: 10.1093/nar/gkq196 PMID: 20360046
35. Brackley CA, Romano MC, Thiel M. The Dynamics of Supply and Demand in mRNA Translation. PLoS Computational Biology. 2011; 7(10):e1002203. doi: 10.1371/journal.pcbi.1002203 PMID: 22022250
36. Rudorf S, Thommen M, Rodnina MV, Lipowsky R. Deducing the Kinetics of Protein Synthesis *in vivo* from the Transition Rates Measured *in vitro*. PLoS Computational Biology. 2014; 10(10):e1003909. Available from: doi: 10.1371/journal.pcbi.1003909 PMID: 25358034
37. Dong H, Nilsson L, Kurland CG. Co-variation of tRNA abundance and codon usage in Escherichia coli at different growth rates. Journal of Molecular Biology. 1996; 260(5):649–663. doi: 10.1006/jmbi.1996.0428 PMID: 8709146
38. Kramer EB, Farabaugh PJ. The frequency of translational misreading errors in E. coli is largely determined by tRNA competition. RNA-a Publication of the RNA Society. 2007; 13(1):87–96. doi: 10.1261/rna.294907
39. Garai A, Chowdhury D, Chowdhury D, Ramakrishnan TV. Stochastic kinetics of ribosomes: Single motor properties and collective behavior. Physical Review E. 2009 Jul; 80(1, 1).

40. Potapov I, Makela J, Yli-Harja O, Ribeiro AS. Effects of codon sequence on the dynamics of genetic networks. Journal of Theoretical Biology. 2012 Dec 21; 315:17–25. doi: 10.1016/j.jtbi.2012.08.029 PMID: 22960571

41. Rodnina MV, Pape T, Fricke R, Kuhn L, Wintermeyer W. Initial binding of the elongation factor Tu·GTP·aminoacyl-tRNA complex preceding codon recognition on the ribosome. Journal of Biological Chemistry. 1996; 271(2):646–652. doi: 10.1074/jbc.271.2.646 PMID: 8557669

42. Pape T, Wintermeyer W, Rodnina MV. Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the E-coli ribosome. EMBO Journal. 1998; 17(24):7490–7497. doi: 10.1093/emboj/17.24.7490 PMID: 9857203

43. Gromadski KB, Rodnina MV. Kinetic determinants of high-fidelity tRNA discrimination on the ribosome. Molecular Cell. 2004; 13(2):191–200. doi: 10.1016/S1097-2765(04)00005-X PMID: 14759365

44. Gromadski KB, Daviter T, Rodnina MV. A Uniform Response to Mismatches in Codon-Anticodon Complexes Ensures Ribosomal Fidelity. Molecular Cell. 2006; 21:369–377. doi: 10.1016/j.molcel.2005.12.018 PMID: 16455492

45. Kothe U, Rodnina MV. Delayed release of inorganic phosphate from elongation factor Tu following GTP hydrolysis on the ribosome. Biochemistry. 2006; 45(42):12767–12774. doi: 10.1021/bi061192z PMID: 17042495

46. Mittelstaet J, Konevega AL, Rodnina MV. A Kinetic Safety Gate Controlling the Delivery of Unnatural Amino Acids to the Ribosome. Journal of the American Chemical Society. 2013; 135:17031–17038. doi: 10.1021/ja407511q PMID: 24079513

47. Ingolia NT, Lareau LF, Weissman JS. Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. Cell. 2011; 147(4):789–802. doi: 10.1016/j.cell.2011.10.002 PMID: 22056041

48. Chu D, Thompson J, von der Haar T. Charting the dynamics of translation. Biosystems. 2014 May; 119:1–9. doi: 10.1016/j.biosystems.2014.02.005 PMID: 24631460

49. Wohlgemuth I, Pohl C, Mittelstaet J, Konevega AL, Rodnina MV. Evolutionary optimization of speed and accuracy of decoding on the ribosome. Philosophical Transactions of the Royal Society B: Biological Sciences. 2011; 366(1580):2979–2986. doi: 10.1098/rstb.2011.0138

50. Taylor HM, Karlin S. An Introduction to Stochastic Modeling. 3rd ed. San Diego: Academic Press; 1998.

51. Norris JR. Markov Chains. Cambridge University Press; 1997. Cambridge Books Online. Available from: http://dx.doi.org/10.1017/CBO9780511810633.

52. Hill TL. Interrelations between random walks on diagrams (graphs) with and without cycles. Proceedings of the National Academy of Sciences of the United States of America. 1988; 85(9):2879–2883. Available from: http://www.pnas.org/content/85/9/2879.abstract. doi: 10.1073/pnas.85.9.2879 PMID: 3362853

53. Liang ST, Xu YC, Dennis PP, Bremer H. mRNA composition and control of bacterial gene expression. Journal of Bacteriology. 2000; 182(11):3037–3044. doi: 10.1128/JB.182.11.3037-3044.2000 PMID: 10809680

54. Zhang J. Protein-length distributions for the three domains of life. Trends in Genetics. 2000; 16(3):107–109. Available from: http://www.sciencedirect.com/science/article/pii/S0168952599019228. doi: 10.1016/S0168-9525(99)01922-8 PMID: 10689349

55. Bremer H, Dennis PP. In: Neidhardt FC, editor. Modulation of Chemical Composition and Other Parameters of the Cell by Growth Rate. vol. 2. 2nd ed. Washington: ASM Press; 1996. p. 1553–1569.

56. Ikemura T. Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice that is optimal for the E. coli translational system. Journal of Molecular Biology. 1981; 151(3):389–409. Available from: http://www.sciencedirect.com/science/article/pii/0022283681900036. doi: 10.1016/0022-2836(81)90003-6 PMID: 6175758

57. Neidhardt FC, Bloch PL, Pedersen S, Reeh S. Chemical Measurement of Steady-state Levels of 10 Aminoacyl Transfer Ribonucleic-acid Synthetases In Escherichia coli. Journal of Bacteriology. 1977; 129(1):378–387. PMID: 318645

58. Zhang G, Fedyunin I, Kirchner S, Xiao C, Valleriani A, Ignatova Z. FANSe: an accurate algorithm for quantitative mapping of large scale sequencing reads. Nucleic Acids Research. 2012; 40(11). doi: 10.1093/nar/gks196