Deducing the Kinetics of Protein Synthesis In Vivo from the Transition Rates Measured In Vitro

Sophia Rudorf¹, Michael Thommen², Marina V. Rodnina², Reinhard Lipowsky¹*

¹ Theory and Bio-Systems, Max Planck Institute of Colloids and Interfaces, Potsdam, Germany, ² Physical Biochemistry, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

Abstract

The molecular machinery of life relies on complex multistep processes that involve numerous individual transitions, such as molecular association and dissociation steps, chemical reactions, and mechanical movements. The corresponding transition rates can be typically measured in vitro but not in vivo. Here, we develop a general method to deduce the in-vivo rates from their in-vitro values. The method has two basic components. First, we introduce the kinetic distance, a new concept by which we can quantitatively compare the kinetics of a multistep process in different environments. The kinetic distance depends logarithmically on the transition rates and can be interpreted in terms of the underlying free energy barriers. Second, we minimize the kinetic distance between the in-vitro and the in-vivo process, imposing the constraint that the deduced rates reproduce a known global property such as the overall in-vivo speed. In order to demonstrate the predictive power of our method, we apply it to protein synthesis by ribosomes, a key process of gene expression. We describe the latter process by a codon-specific Markov model with three reaction pathways, corresponding to the initial binding of cognate, near-cognate, and non-cognate tRNA, for which we determine all individual transition rates in vitro. We then predict the in-vivo rates by the constrained minimization procedure and validate these rates by three independent sets of in-vivo data, obtained for codon-dependent translation speeds, codon-specific translation dynamics, and missense error frequencies. In all cases, we find good agreement between theory and experiment without adjusting any fit parameter. The deduced in-vivo rates lead to smaller error frequencies than the known in-vitro rates, primarily by an improved initial selection of tRNA. The method introduced here is relatively simple from a computational point of view and can be applied to any biomolecular process, for which we have detailed information about the in-vitro kinetics.

Introduction

Life is based on the continuous synthesis, modification, and degradation of proteins and other macromolecules. These processes are performed by complex biomolecular machines that bind their ligands and transform them into product molecules. Examples are provided by the transcription of DNA by RNA polymerases, the translation of mRNA by ribosomes, or the degradation of proteins by proteasomes. Each of these processes involves several steps: the binding of the ligand molecules, chemical reactions catalyzed at the active sites, as well as specific conformational changes and directed mechanical movements of parts of the molecular machinery. In principle, the kinetics of such multistep processes can be understood in terms of the individual transitions and the associated transition rates, a well-established approach both for enzyme kinetics [1–3] and for free energy transduction by molecular motors [4,5]. In practice, the values of the individual transition rates can be typically measured in vitro but not in vivo, and the in-vitro rates depend on the composition of the buffer. Because the cytosol represents a rather complex buffer, it is difficult to assess whether a certain in-vitro assay provides a reliable description of the process in vivo. One important tool that is missing for such an assessment is a simple measure by which we can quantitatively compare the kinetics of a multistep process in different environments.

Here, we develop a general method that provides such a measure and allows the deduction of the in-vivo rates from their in-vitro values. Our method has two basic components. First, we introduce the ‘kinetic distance’, i.e., a distance metric for the kinetics, by which we can describe the similarity or dissimilarity of multistep processes in vitro and in vivo in a quantitative manner. The kinetic distance depends logarithmically on the rates and has an intuitive interpretation in terms of the associated free energy barriers. Second, we minimize the kinetic distance between the in-vitro and in-vivo processes, imposing the constraint that the deduced rates reproduce a known global property such as the overall in-vivo speed. Computationally, this constraint defines a hypersurface in the multi-dimensional space of transition rates. In order to demonstrate the predictive power of our method, we apply it to the elongation cycle of protein synthesis, a key process of gene expression.


Editor: Shi-Jie Chen, University of Missouri, United States of America

Received July 21, 2014; Accepted September 14, 2014; Published October 30, 2014

Copyright: © 2014 Rudorf et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. 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.

* Email: lipowsky@mpikg.mpg.de
Deducing the Kinetics In Vivo from the In-Vitro Rates

Author Summary

The proverb ‘life is motion’ also applies to the molecular scale. Indeed, if we looked into any living cell with molecular resolution, we would observe a large variety of highly dynamic processes. One particularly striking aspect of these dynamics is that all macromolecules within the cell are continuously synthesized, modified, and degraded by complex biomolecular machines. These ‘nanorobots’ follow intricate reaction pathways that form networks of molecular transitions or transformation steps. Each of these steps is stochastic and takes, on average, a certain amount of time. A fundamentally important question is how these individual step times or the corresponding transition rates determine the overall speed of the process in the cell. This question is difficult to answer, however, because the step times can only be measured in vitro but not in vivo. Here, we develop a general computational method by which one can deduce the individual step times in vivo from their in-vitro values. In order to demonstrate the predictive power of our method, we apply it to protein synthesis by ribosomes, a key process of gene expression, and validate the deduced step times by three independent sets of in-vivo data.

In all living cells, proteins are synthesized by ribosomes, which translate the codon sequences of mRNA into peptide chains of proteins. During the elongation cycle of this process, the ribosome translates one codon after another by binding a ternary complex consisting of aminoacyl-tRNA (aa-tRNA), elongation factor Tu (EF-Tu) and GTP. The amino acid is transferred from the tRNA to the nascent peptide chain, and the ribosome moves to the next codon with the help of elongation factor G (EF-G) allowing for the next elongation cycle. Translation elongation involves several individual states with rapid transitions between them [6,7]. The next elongation cycle. Translation elongation involves several individual states with rapid transitions between them [6,7]. The kinetic rates but, so far, it has not been possible to measure the rates of these pathways were measured in-vitro, both of which have been measured in the overall elongation rates and the tRNA concentrations taking into account two known properties of the phy and single molecule methods. The kinetic techniques: chemical probing methods, pre-steady state electron microscopy, X-ray crystallography, and single molecule methods. The kinetic measurements in vitro provided values for the individual transition rates but, so far, it has not been possible to measure the corresponding rates in the cell.

The different states and transitions of the elongation cycle are schematically shown in Fig. 1. When the ribosome dwells at a certain codon and binds a ternary complex, the tRNA within this complex can be cognate, near-cognate, or non-cognate to the codon, which implies that the elongation cycle contains three different reaction pathways corresponding to the three branches in Fig. 1. During each round of elongation, the ribosome typically explores all three pathways in order to select a cognate tRNA and to reject the near-cognate and non-cognate ones. The individual rates of these pathways were measured in vitro at 20°C and/or 37°C using the ribosomes and translation factors from Escherichia coli [6,13,25]. Here, we combine these results with new data on the overall elongation rates in vitro to first derive a complete set of individual in-vitro rates at both temperatures. We then minimize the kinetic distance between the in-vitro and in-vivo processes, taking into account two known properties of the in-vitro process: the overall elongation rates and the tRNA concentrations, both of which have been measured in E. coli for different growth conditions.

The in-vitro rates of the elongation cycle obtained in this way are then validated by three independent sets of in-vivo data [29–31]. First, we compute codon-specific elongation rates and show that these rates correlate well with relative translation rates as obtained experimentally by [29]. Second, we predict the time-dependent incorporation of radioactively labeled amino acids into proteins as studied in vivo by [30]. The time course of synthesis obtained theoretically is in excellent agreement with the experimental data. Third, using the same in-vivo rates, we also compute the missense error frequency and obtain good agreement with the experimental results of [31]. In all three cases, our computations do not involve any fit parameter and, thus, directly validate the derived set of in-vivo rates.

Results

Distance between in-vitro and in-vivo kinetics

In order to introduce a quantitative measure for the (dis)similarity of the in-vitro and in-vivo kinetics, we consider a generic multistep process within the cell and first focus on one of the individual transitions from state i to state j. The corresponding transition rates have the values \( \omega_{ij} \) and \( \omega_{ij}^* \) for a certain in-vitro assay and for specific in-vivo growth conditions, respectively. Instead of the rates, we can equally well consider the associated transition times \( \tau_{ij} = 1/\omega_{ij} \) and \( \tau_{ij}^* = 1/\omega_{ij}^* \). Thus, we require that the distance \( D_\ell(\omega_{ij},\omega_{ij}^*) \) between the rates \( \omega_{ij} \) and \( \omega_{ij}^* \) is equal to the distance \( D_\ell(\tau_{ij},\tau_{ij}^*) \) between the times \( \tau_{ij} \) and \( \tau_{ij}^* \), i.e., that

\[
D_\ell(\omega_{ij},\omega_{ij}^*) = D_\ell(\tau_{ij},\tau_{ij}^*) = D_\ell(1/\omega_{ij},1/\omega_{ij}^*). \tag{1}
\]

The simplest expression for \( D_\ell \) that fulfills this requirement is provided by

\[
D_\ell(\omega_{ij},\omega_{ij}^*) = |\ln(\omega_{ij}^*/\omega_{ij})| = |\Delta_\ell| \tag{2}
\]

with the logarithmic difference

\[
\Delta_\ell \equiv \ln(\omega_{ij}) - \ln(\omega_{ij}^*) = -\ln(\omega_{ij}^*/\omega_{ij}) = \ln(\tau_{ij}^*/\tau_{ij}). \tag{3}
\]

between the in-vitro and the in-vivo value of the individual transition rate.

The single transition distance \( D_\ell \) is dimensionless and does not involve any parameter apart from the two rates \( \omega_{ij} \) and \( \omega_{ij}^* \). In addition, this distance satisfies the two scaling relations \( D_\ell(\omega,b) = D_\ell(\omega_0,b) \) and \( D_\ell(\omega,b_0) = D_\ell(\omega_0,\omega_0) \) for any rescaling factor \( b > 0 \). The first scaling relation implies that \( b \omega \) and \( \omega/b \) have the same distance from \( \omega \), which agrees with our intuition. The second scaling relation implies that the distance \( D_\ell \) does not depend on the units used to measure the rates. For small deviations of \( \omega_{ij}^* \) from \( \omega_{ij} \), which are equivalent to small deviations of \( \tau_{ij}^* \) from \( \tau_{ij} \), the distance \( D_\ell = |\Delta_\ell| \) becomes asymptotically equal to both \( |\omega_{ij}^* - \omega_{ij} / \omega_{ij} | \) and \( |\tau_{ij}^* - \tau_{ij} / \tau_{ij} | \).

The in-vitro rate \( \omega_{ij} \) can be expressed in terms of the activation free energy or free energy barrier \( \Delta G_{ij} \) and the attempt frequency \( v_j \) which leads to

\[
\omega_{ij} = v_j \exp[-\Delta G_{ij}/k_BT], \tag{4}
\]

where the thermal energy \( k_BT \) provides the basic free energy scale. When we combine this expression with the analogous expression for the in-vitro rate, the logarithmic difference between the two rates becomes
\[ \Delta_{ij} = \frac{\Delta G_{ij} - \Delta G_{ji}}{k_B T} - \ln \left( \frac{v_{ij}^*}{v_{ij}} \right). \]  

Because the prefactors \( v_{ij}^* \) and \( v_{ij} \) are expected to have the same order of magnitude, the second term \( \ln \left( \frac{v_{ij}^*}{v_{ij}} \right) \) should usually be small compared to the first term which represents the shift of the free energy barrier between state \( i \) and state \( j \), see Fig. 2A. Therefore, for each individual transition along one of the reaction pathways, the logarithmic difference \( \Delta_{ij} \) can be interpreted as the shift of the free energy barrier that governs the transition from state \( i \) to state \( j \). In the following, we will use the intuitive terminology ‘single barrier shift’ for the quantity \( \Delta_{ij} \). It should be noted, however, that, in spite of this terminology, changes in the attempt frequency as described by the term \( \ln \left( \frac{v_{ij}^*}{v_{ij}} \right) \) in Eq. 5 are included in the logarithmic difference \( \Delta_{ij} \) and, thus, will be taken into account in all our calculations.

Next, we consider all individual transitions along the reaction pathways of the multistep process and regard the associated \( \textit{in-vitro} \) rates \( \omega_{ij}^* \) as unknown variables that can be visualized as the coordinates of a multi-dimensional space. These coordinates are somewhat impractical, however, because they are restricted to positive values. In order to eliminate this restriction, we perform a coordinate transformation from the \( \textit{in-vitro} \) rates \( \omega_{ij}^* \) to the single barrier shifts \( \Delta_{ij} \), which can attain both positive and negative values. This coordinate transformation is highly nonlinear but invertible with the inverse transformation given by \( \omega_{ij} = \omega_{ij}^* \exp \left( - \Delta_{ij} \right) \).

The overall distance \( D \) between the \( \textit{in-vitro} \) and \( \textit{in-vivo} \) kinetics is now defined by

\[ D \equiv \sqrt{\sum \Delta_{ij}^2} = \sqrt{\sum \left[ \ln \left( \frac{\omega_{ij}^*}{\omega_{ij}} \right) \right]^2}, \]  

where the summation under the square root runs over all individual transitions along the reaction pathways. As illustrated in Fig. 2B, the distance \( D \) represents the Euclidean distance within the multi-dimensional space defined by the single barrier shifts \( \Delta_{ij} \). Therefore, the distance \( D \) provides a genuine metric in the mathematical sense, which implies that it satisfies the triangle inequality if we compare three different \( \textit{in-vitro} \) and/or \( \textit{in-vivo} \) conditions.

If all \( \textit{in-vivo} \) rates are identical to their \( \textit{in-vitro} \) values, apart from a single one, \( \omega_{kl}^* \neq \omega_{kl} \), Eq. 6 for the kinetic distance \( D \)
reduces to Eq. 1 for the single transition distance $D_{ij}$. Because the choice of the two states $k$ and $l$ is arbitrary, this property of the kinetic distance applies to all individual transitions $v_{ij}$ that enter in Eq. 6. The latter property represents, in fact, a general requirement for any meaningful definition of the kinetic distance. Therefore, if we considered the more general expression $D_u := \sqrt{\sum u_{ij}^2 \Delta_{ij}^2}$ with dimensionless weight factors $u_{ij}$, this requirement would imply that all weight factors must assume the unique values $u_{ij} = 1$ and that $D_u$ must be equal to the kinetic distance $D$ as given by Eq. 6.

If we consider two different in-vitro assays, say $A$ and $A'$, the corresponding transition rates $\omega_{ij}$ and $\omega_{ij}'$ will, in general, be different and define two sets of single barrier shifts via $\Delta_{ij} = \ln (\omega_{ij}/\omega_{ij}')$ and

$$\Delta_{ij}' = \ln (\omega_{ij}'/\omega_{ij}') = \Delta_{ij} + d_{ij}$$

with the logarithmic differences $d_{ij} := \ln (\omega_{ij}'/\omega_{ij})$. The latter quantities determine the kinetic distance $D_{A,A'}$ between the two in-vitro assays, which is given by $D_{A,A'} = \sqrt{\sum d_{ij}^2}$. The two sets of barrier shifts, $\Delta_{ij}$ and $\Delta_{ij}'$, provide two different coordinates for the multi-dimensional barrier space. Because of the linear relations as given by Eq. 7, the primed coordinates are obtained from the unprimed ones by shifting the latter coordinates by the logarithmic differences $d_{ij}$. Therefore, the transformation from the unprimed

Figure 2. Kinetic distance based on the logarithmic differences $\Delta_{ij}$ between the in-vitro and in-vivo rates. (A) In-vitro free energy barrier $\Delta G_{ij}$ and in-vivo barrier $\Delta G_{ij}'$ for the transition from state $i$ to state $j$. When expressed in units of $k_B T$, the single barrier shift $\Delta G_{ij} = \Delta G_{ij}'$ determines the logarithmic difference $D_{ij}$, see Eq. 5; and (B) Three-dimensional section of the multi-dimensional barrier space with coordinates $\Delta_{12}, \Delta_{21}$, and $\Delta_{23}$. The origin of this space (light blue dot) corresponds to the in-vitro system. The surface (purple) represents a two-dimensional section of the hypersurface described by Eq. 8, corresponding to a fixed in-vivo value for the global kinetic quantity. Each point on this surface has a certain kinetic distance that is equal to the Euclidean distance of this point from the origin, as indicated by the three double arrows. The point with the shortest kinetic distance determines the predicted in-vivo rates $\omega_{ij}^{\text{min}}$. doi:10.1371/journal.pcbi.1003909.g002
to the primed coordinates, corresponding to a change from assay A to assay A', represents a Euclidean translation of the coordinate system, which preserves the shape of any geometric object within the multi-dimensional barrier space.

Constrained minimization of the kinetic distance

Next, we combine the kinetic distance as given by Eq. 6 with a minimization procedure to predict the unknown in-vivo rates from their known in-vitro values. Even though the rates of the individual transitions are difficult to study in the cell, one can usually measure some quantity that characterizes the overall kinetics of the intracellular process. One such quantity is provided by the average speed of the process. Any such global kinetic quantity, $Q$, depends on the individual transition rates, $Q = \{Q_{ij}\} = \{Q_{12}, Q_{21}, Q_{32}, \ldots\}$. The in-vitro values $Q_{ij}$ of the individual transition rates must reproduce the experimentally measured value $Q_{\text{exp}}$ of the global quantity. This requirement implies the equation

$$Q(\omega_{12}, \omega_{21}, \omega_{32}, \ldots) = Q_{\text{exp}},$$

which represents a constraint on the unknown in-vivo values $\omega_{ij}$. This constraint can be expressed in terms of the single barrier shifts $\Delta_\gamma$ using the inverse coordinate transformation $\omega_{ij} = \omega_{ij}\exp[-\Delta_\gamma]$ with the known in-vitro values $\omega_{ij}$. As a result, the constraint in Eq. 8 defines a hypersurface in the multi-dimensional barrier space as illustrated in Fig. 2B. Each point on this hypersurface is compatible with the measured value $Q_{\text{exp}}$ of the global kinetic quantity. In addition, the Euclidean distance of such a point from the origin is equal to the kinetic distance $\Delta$ between the (unknown) in-vivo and the (known) in-vitro values of the transition rates. Our prediction for the in-vitro values $\omega_{ij}$ is then obtained by minimizing this kinetic distance, i.e., by the point on the hypersurface that has the shortest distance from the origin. For clarity, the coordinate values of this point will be denoted by $\Delta_\gamma_{ij\text{, min}}$ in order to distinguish these values from the variable coordinates $\Delta_\gamma$.

Our approach involves the following assumptions. First, we make the usual assumption that the states of the biomolecular system that have been identified in vitro are also present in vivo. The molecular conformations of the corresponding in vitro and in-vivo states are expected to be somewhat different when viewed with atomic resolution, but the gross features of these conformations should be similar, in particular when the in-vitro assay is functional and has been optimized. It is then plausible to assume that the in-vitro and in-vivo values of the individual transition rates do not differ by many orders of magnitude, which implies that the point in the multi-dimensional barrier space that represents the true in-vivo rates is located 'in the neighborhood' of the origin of this space and, thus, characterized by a 'small' in-vivo environment on the individual transitions of the system. Such differences are indeed obtained when we apply our minimization approach to the kinetics of ribosomes as described in the next subsection. It is important to note that this approach leads to different scale factors even though the expression for the kinetic distance (Eq. 6) does not include any bias for one of the individual barrier shifts $\Delta_\gamma$. Therefore, the different scale factors $\omega_{ij, \text{min}}/\omega_{ij}$ follow from the imposed in-vivo constraint (Eq. 8) alone and do not involve any additional assumptions or expectations about the in-vivo conditions.

The minimization procedure described above represents an extremum principle with constraints. Such principles have been successfully applied in many areas of science, in particular in the context of optimization problems. One important and useful feature of extremum principles is that they provide global solutions for nonlinear systems. Thus, in the present context, we would obtain a prediction for the in-vivo rates even if the in-vitro assay were rather different from the in-vivo conditions. Another advantage of extremum principles is that they typically lead to a unique solution without any additional assumptions ('principle of least prejudice'). In some exceptional cases, one may find more than one solution, which then indicates that the system undergoes some kind of bifurcation. For the kinetics of ribosomes, see next subsection, we always found a unique solution and, thus, a unique set of predicted in-vivo rates.

The rates $\omega_{ij}$ of the in-vitro assay have only a known with a certain accuracy. As a consequence, the predicted in-vivo rates $\omega_{ij, \text{min}}$ have some uncertainty as well. As explained in the Methods section, this uncertainty reflects both the accuracy of the measured in-vitro rates and the associated changes in the location of the predicted in-vivo point. Furthermore, the latter location will also depend, in general, on the rates of the chosen in-vitro assay. Indeed, the change from assay A to assay A' corresponds to a Euclidean translation of the coordinate system (Eq. 7) while the shape of the hypersurface (Eq. 8 and Fig. 2B) remains unchanged. These two properties imply that the distance of the hypersurface from the origin of the $\Delta_\gamma$-coordinates may differ from the distance of this surface from the origin of the $\Delta_\gamma$-coordinates. Therefore, the validity of the predicted in-vivo rates $\omega_{ij, \text{min}}$ is difficult to assess a priori, but can be checked a posteriori in a self-consistent manner: we first deduce the unknown in-vivo rates from the known in-vitro rates via the minimization procedure and subsequently validate the deduced rates $\omega_{ij, \text{min}}$ by calculating some other quantities that have been experimentally studied in vivo. In the next two subsections, we will apply this two-step procedure to the kinetics of ribosome elongation based on the in-vitro assay developed in [6,25].

Our minimization procedure becomes computationally simpler if we have additional knowledge about some of the in-vitro values $\omega_{ij}$ of the individual transition rates. If we knew one of these rates, e.g., $\omega_{kl}$, we would restrict our minimization procedure to the subspace with constant $\Delta_{kl} = \ln(\omega_{kl}/\omega_{kl}')$. As a consequence, we would not vary the coordinate $\Delta_{kl}$ during the minimization and use the constant value of this coordinate in Eq. 6 for the kinetic distance $D$. On the other hand, if we knew only that the in-vitro rate $\omega_{ij}$ is located within the range $\Omega_i < \omega_{ij} < \Omega_j$, we would minimize the kinetic distance also with respect to $\Delta_{ij}$ but within the subspace defined by $\ln(\omega_{ij}/\Omega_i) < \Delta_{ij} < \ln(\omega_{ij}/\Omega_j)$. The latter procedure may lead to a boundary minimum, i.e., to a predicted in-vitro point that is located at the boundary of the...
considered subspace. Another simplification is obtained if the rates of two individual transitions, say from state \( k \) to state \( l \) and from state \( k' \) to state \( l' \), have the same values in vitro and in vivo, i.e., if \( \omega_{kl} = \omega_{k'l'} \). We will then reduce the multi-dimensional barrier space to the subspace with \( \Delta k = \Delta k' \), and the corresponding expression for the kinetic distance \( D \) in Eq. 6 will now contain the term \( \Delta k + \Delta k' = 2\Delta k \) under the square root. The latter reduction will be used in the next subsection on the kinetics of ribosomes for which different individual transition rates have the same in-vitro values.

**Kinetics of ribosomes during protein synthesis**

Our quantitative description of the translation elongation cycle is based on the codon-specific Markov process displayed in Fig. 3. This process can visit, for each sense codon \( c \), twelve ribosomal states, numbered from 0 to 11. After the ribosome has moved to the next sense codon, it dwells in state 0, until it binds a ternary complex with an elongator tRNA that may be cognate, near-cognate, or non-cognate to codon \( c \).

The genetic code involves 61 sense codons, which encode 20 proteinogenic amino acids and are decoded by a certain number of elongator tRNAs. The latter number depends on the organism but is always larger than 20 and smaller than 61 [32,33]. For *E. coli*, 43 distinct species of elongator tRNA have been identified [28]. The corresponding codon-tRNA relationships can be visualized by the large matrix in Fig. 4 with 61 rows and 43 columns. As shown by the color code in this figure, each sense codon defines a different decomposition of the total set of tRNA species into three subsets of cognates, near-cognates, and non-cognates. The corresponding molar concentrations \( X_{c,co} \), \( X_{c,nt} \), and \( X_{c,no} \) of cognate, near-cognate, and non-cognate ternary complexes determine the association rates

\[
\omega_{c,co} = k_{on} X_{c,co}, \quad \omega_{c,nt} = k_{on} X_{c,nt}, \quad \text{and} \quad \omega_{c,no} = k_{on} X_{c,no}
\]

for initial binding with the pseudo-first-order association rate constant \( k_{on} \). This constant is taken to be independent both of the codon and of the ternary complex as observed *in vitro* [10,25]. The latter experiments also imply that all ternary complexes dissociate with the same rate \( \omega_{off} \) from the initial binding site and that the cognate and near-cognate ternary complexes have the same recognition rate \( \omega_{rec} \).

After initial binding of a *non-cognate* ternary complex, this complex dissociates without visiting any other state, so that the ribosome returns back to state 0 with an empty initial binding site. Initial binding of a *cognate* ternary complex leads to state 1, from which the ternary complex can be released with rate \( \omega_{off} \) or can move into the A site to attain the cognon recognition state 2 with rate \( \omega_{rec} \). When the ternary complex is recognized as cognate, the ribosome undergoes a forward transition from state 2 to state 3, which corresponds to the combined process of GTPase activation of the cognate ternary complex and GTP hydrolysis, followed by the irreversible transition from state 3 to state 4, which describes phosphate release and conformational rearrangements of EF-Tu [6]. From state 4, the cognate ternary complex may either move to become fully accommodated into the A site via a transition from state 4 to state 5 or, with low probability, may be released from the A site via a transition from state 4 to state 0. After the cognate ternary complex has been fully accommodated, the ribosome/tRNA complex undergoes the final transition from state 5 into the empty state 0 at the next codon \( c' \). This transition describes the combined process of peptide bond formation and translocation, the corresponding processing rate is denoted by \( \omega_{pro} \).

Initial binding of a near-cognate ternary complex leads to state 6, from which the ternary complex can be released with rate \( \omega_{off} \) or move to the codon recognition state 7 with rate \( \omega_{rec} \). When the ternary complex is recognized as near-cognate, it is rejected and the ribosome undergoes a backward transition from state 7 to state 6, which provides the initial selection step during the decoding process. With low probability, the near-cognate ternary complex undergoes an irreversible transition from state 7 to state 8, corresponding to GTPase activation and GTP hydrolysis, as well as from state 8 to state 9, which describes phosphate release and conformational rearrangements of EF-Tu. From state 9, the near-

The elongation cycle starts in state 0 corresponding to a ribosome without any bound ternary complex. Initial binding of a cognate, near-cognate, or non-cognate ternary complex is indicated by the green, orange, and purple arrow, compare the color code in Fig. 4; the corresponding association rates are proportional to the association rate constant \( k_{on} \), as in Eq. 9. The black arrows represent the individual transitions along the reaction pathways. All ternary complexes dissociate initially with the same dissociation rate \( \omega_{off} \). Likewise, cognate and near-cognate ternary complexes are governed by the same recognition rate \( \omega_{rec} \), conformational rate \( \omega_{con} \), and processing rate \( \omega_{pro} \). The kinetic distinction between the cognate and near-cognate branches arises from initial selection at the states 2 and 7 as well as from proofreading at the states 4 and 9.

Figure 3. Codon-specific Markov process for translation elongation based on 12 ribosomal states for each codon \( c \). The elongation cycle starts in state 0 corresponding to a ribosome without any bound ternary complex. Initial binding of a cognate, near-cognate, or non-cognate ternary complex is indicated by the green, orange, and purple arrow, compare the color code in Fig. 4; the corresponding association rates are proportional to the association rate constant \( k_{on} \), as in Eq. 9. The black arrows represent the individual transitions along the reaction pathways. All ternary complexes dissociate initially with the same dissociation rate \( \omega_{off} \). Likewise, cognate and near-cognate ternary complexes are governed by the same recognition rate \( \omega_{rec} \), conformational rate \( \omega_{con} \), and processing rate \( \omega_{pro} \). The kinetic distinction between the cognate and near-cognate branches arises from initial selection at the states 2 and 7 as well as from proofreading at the states 4 and 9.

doi:10.1371/journal.pcbi.1003909.g003
cognate ternary complex is typically released again via a transition from state 9 to state 0, which provides the proofreading step during decoding. Very rarely, the near-cognate ternary complex is fully accommodated via a transition from state 9 to state 10. After a near-cognate tRNA has been fully accommodated, it is further processed via peptide bond formation and translocation and undergoes the transition from state 10 to state 0 with rate $v_{\text{pro}}$.

Apart from the association rate constant $k_{\text{on}}$, the kinetics of the elongation cycle then involves 12 different transition rates $v_{ij}$ for the 17 transitions along the cognate, near-cognate, and non-cognate branches of the Markov process. All of these transition rates have been determined in vitro for the high-fidelity buffer developed in [12,25,34]. The corresponding in-vitro values are reported in Table 1. A few individual rates were measured at both 20 and 37°C whereas most of these rates were obtained either at 20 or at 37°C. We used a variety of computational methods to obtain complete and consistent sets of individual rates at both temperatures as described in the Methods section. In addition, we measured the overall elongation rate $v_{\text{elo}}$ in vitro for a model protein, $v_{\text{elo}} \approx 0.8$ aa/s for 20°C and $v_{\text{elo}} \approx 6.9$ aa/s for 37°C.

Figure 4. Decoding pattern arising from the cognate (green), near-cognate (yellow), and non-cognate (purple) relationships between all 61 sense codons and the 43 elongator tRNA species of *E. coli* as identified in Ref. [28]. For each tRNA species, the near-cognate codons differ from the cognate ones by a mismatch in one position of the codon-anticodon complex.

doi:10.1371/journal.pcbi.1003909.g004
Because several transition rates of the Markov process considered has the explicit form distance to this subspace. After this reduction, the latter distance subspace and restrict the minimization procedure of the kinetic reduce the 17-dimensional barrier space to a 12-dimensional then used to compute, for both temperatures, the

Deducing the Kinetics In Vivo from the In-Vitro Rates

Table 1. In-vitro rates of ribosomal transitions.

<table>
<thead>
<tr>
<th>Rates</th>
<th>k-not.</th>
<th>20°C</th>
<th>37°C</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \kappa_{\text{con}} )</td>
<td>( k_1 )</td>
<td>( 140 \pm 20 )</td>
<td>( 175 \pm 25 )</td>
<td>( \frac{1}{\mu\text{M}s} )</td>
</tr>
<tr>
<td>( v_{\text{el}0} )</td>
<td>( k_{-1} )</td>
<td>( 85 \pm 25 )</td>
<td>( 700 \pm 270 )</td>
<td>( \text{/s} )</td>
</tr>
<tr>
<td>( v_{\text{rec}} )</td>
<td>( k_2 )</td>
<td>( 180 \pm 30 )</td>
<td>( 1500 \pm 450 )</td>
<td>( \text{/s} )</td>
</tr>
<tr>
<td>( v_{\text{off}} )</td>
<td>( k_{1,00} )</td>
<td>( 0.2 \pm 0.03 )</td>
<td>( 2 \pm 0.6 )</td>
<td>( \text{/s} )</td>
</tr>
<tr>
<td>( v_{\text{con}} )</td>
<td>( k_4 )</td>
<td>( 50 )</td>
<td>( 450 )</td>
<td>( \text{/s} )</td>
</tr>
<tr>
<td>( v_{\text{G45}} )</td>
<td>( k_{3,00} )</td>
<td>( 22 \pm 4 )</td>
<td>( 200 \pm 40 )</td>
<td>( \text{/s} )</td>
</tr>
<tr>
<td>( v_{\text{G40}} )</td>
<td>( k_{7,00} )</td>
<td>( 0.1 )</td>
<td>( 1 )</td>
<td>( \text{/s} )</td>
</tr>
<tr>
<td>( v_{\text{G16}} )</td>
<td>( k_{3,0R} )</td>
<td>( 140 \pm 20 )</td>
<td>( 1100 \pm 330 )</td>
<td>( \text{/s} )</td>
</tr>
<tr>
<td>( v_{\text{349}} )</td>
<td>( k_{3,AR} )</td>
<td>( 0.6 \pm 0.1 )</td>
<td>( 7 \pm 2 )</td>
<td>( \text{/s} )</td>
</tr>
<tr>
<td>( v_{\text{G10}} )</td>
<td>( k_{3,MR} )</td>
<td>( 0.06 \pm 0.006 )</td>
<td>( 0.26 \pm 0.04 )</td>
<td>( \text{/s} )</td>
</tr>
<tr>
<td>( v_{\text{G90}} )</td>
<td>( k_{7,AR} )</td>
<td>( 0.84 \pm 0.08 )</td>
<td>( 4 \pm 0.7 )</td>
<td>( \text{/s} )</td>
</tr>
<tr>
<td>( v_{\text{P26S}} )</td>
<td>( 3 \pm 1 )</td>
<td>( 150 \pm 50 )</td>
<td>( \text{/s} )</td>
<td></td>
</tr>
<tr>
<td>( v_{\text{Aho}} )</td>
<td>( 0.8 \pm 0.2 )</td>
<td>( 6.9 \pm 2.3 )</td>
<td>( \text{aa/s} )</td>
<td></td>
</tr>
</tbody>
</table>

Apart from the processing rate \( v_{\text{Aho}} \), all individual rates and the overall elongation rate \( v_{\text{el}0} \) have been measured in vitro at 20°C and/or 37°C. The processing rate \( v_{\text{P26S}} \) was calculated from the overall elongation rate \( v_{\text{el}0} \) via Eq. 22 in the Methods section. The column ‘k-not.’ provides the notation for the transition rates as used in Ref. [8].

doi:10.1371/journal.pcbi.1003909.t001

(Supporting Figure S1). As explained in the Methods section (Eq. 22), the measured value of the overall elongation rate \( v_{\text{el}0} \) was then used to compute, for both temperatures, the in-vitro value \( v_{\text{pro}} \) of the processing rate. The results of these computations are included in Table 1.

To predict the unknown in-vitro rates \( v_{\text{pro}} \) from the known in-vitro rates \( v_{\text{con}} \), we consider the multi-dimensional space of single barrier shifts as described by the coordinates \( \Delta_q = -\ln (v_{\text{pro}} / v_{\text{con}}) \). Because several transition rates of the Markov process considered here have the same values (Fig. 3), we use the resulting equalities for the associated coordinates as given by \( \Delta_{10} = 0 = \Delta_{11,0} = \Delta_{\text{off}}, \Delta_{12} = \Delta_{7} = \Delta_{\text{rec}}, \Delta_{13} = \Delta_{9} = \Delta_{\text{con}} \), and \( \Delta_{50} = \Delta_{\text{off}} = \Delta_{\text{pro}} \) to reduce the 17-dimensional barrier space to a 12-dimensional subspace and restrict the minimization procedure of the kinetic distance to this subspace. After this reduction, the latter distance has the explicit form

\[
D = \sqrt{3 \Delta_{\text{off}}^2 + 2 \Delta_{\text{rec}}^2 + 2 \Delta_{\text{con}}^2 + 2 \Delta_{\text{pro}}^2 + \sum \Delta_{ij}^2} \quad (10)
\]

where the sum \( \sum \Delta_{ij}^2 \) contains all the remaining transition rates of the Markov process in Fig. 3.

Because the in-vitro experiments are typically performed at 37°C, we use the in-vitro values \( v_{\text{con}} \) for the same temperature, see Table 1. Furthermore, we take into account the known in-vitro values of the overall elongation rate \( v_{\text{el}0} \) at different growth conditions [26,27]. For each growth condition, the constraint in Eq. 8 now has the explicit form as given by Eq. 23 in the Methods section. As a result of the constrained minimization procedure, we find the in-vitro rates \( v_{\text{pro}} \) as given in Table 2 and the single barrier shifts \( \Delta_q \) displayed in Fig. 5A, where we have again omitted the subscript ‘min’ for notational simplicity.

Validation of deduced in-vitro rates for translation elongation

Starting from the complete set of individual in-vitro rates (Table 2), we computed the codon-specific elongation rates \( v_{\text{el}0} \) as described in the Methods section (Eq. 21, Supporting Figure S3). We then compared the in-vitro rates \( v_{\text{el}0} \) calculated for a growth rate of 2.5 dlb/h to relative translation rates as estimated in Ref. [29] based on the frequencies of the measured +1 frameshifting vs. readthrough of different codons. As shown in Fig. 6A, we obtain reasonable overall agreement between both data sets with a Pearson correlation coefficient of 0.56. The deviations reflect both limitations of our model parametrization and uncertainties in the experimental method. First, the calculated elongation rates for CGA, CGC, and CGU appear to be overestimated. These codons are all read by tRNAAns2, which does not form a Watson-Crick base pair with any of its cognate codons because it carries inosine at the wobble position of its anticodon IGG. The corresponding reductions in the transition rates are not included in the parametrization of our model because we use only two different sets of values for these rates, corresponding to an average over all cognate and over all near-cognate ternary complexes, respectively. Second, for the experimental setup in [29], the UUU, UUC, UUG, UCC, and CCG codons, when located between a preceding CUU codon and a subsequent CXX codon, generate potential slippery sequences, which can lead to –1 frameshifting events. The latter events were not considered and, thus, not taken into account by [29], which implies that the frameshifting rates were underestimated and the translation rates were overestimated for the respective codons. When we exclude these two particular sets of codons, we obtain an increased correlation coefficient of 0.73 as shown in Fig 6A. Thus, the deduced values \( v_{\text{pro}} \) of the individual transition rates in vitro lead to a reliable description for the majority of codons.

To further validate these deduced values, we used the computed values \( v_{\text{el}0} \) of the codon-specific elongation rates (Supporting Figure S3), to model the time course of protein synthesis measured by [30]. In those experiments, the lacZ gene was expressed in E. coli at a growth rate of about 0.7 dlb/h, the cells were exposed to a 10-s pulse of radioactively labeled methionine, and the radioactivity of the synthesized proteins was measured over time. The calculated time course is in excellent agreement with the
experimental data (Fig. 6B). Furthermore, varying the values of the internal transition rates leads to significant deviations of the simulation curve from the data (Supporting Figure S4).

Figure 5. Quantitative comparison between the in-vitro kinetics of translation elongation at 37°C and the in-vivo kinetics as deduced for the growth condition of 2.5 dbl/h. (A) Single barrier shifts $\Delta_{ij}$ for the individual transition rates, see Eq. 2 and Fig. 2; and (B) Scale factors $\omega^*_{ij}/\omega_{ij} = \exp(-\Delta_{ij})$ for all individual transitions of the ribosomes. A barrier shift $\Delta_{ij} < 0$ implies that the in-vivo rate $\omega^*_{ij}$ is increased compared to the in-vitro rate $\omega_{ij}$. doi:10.1371/journal.pcbi.1003909.g005

Another quantity that can be used to validate the deduced in-vivo rates $\omega^*_{ij}$ is the missense error frequency arising from the accommodation of near-cognate ternary complexes with incorrect
amino acids. The calculated error frequency depends on all individual transitions for the accommodation of a cognate or near-cognate aa-tRNA and, in particular, on the concentrations of cognate and near-cognate ternary complexes whereas it is independent of the concentrations of non-cognates, see Eq. 36 in the Methods section. Using the deduced \textit{in-vivo} rates in Table 2 and the ternary complex concentrations as estimated from the measured tRNA concentrations for 0.7 dbl/h [28], we obtain an average missense error frequency of \(3 \times 10^{-4}\) for tRNA\textsuperscript{Lys} misreading codons, in good agreement with the measured value \(2 \times 10^{-4}\) [31].

**Discussion**

The theoretical approach described here involves two novel concepts. First, we introduced the kinetic distance to provide a quantitative measure for the similarity of the \textit{in-vitro} and \textit{in-vivo} kinetics. This distance has an intuitive interpretation in terms of the free energy barriers that govern the individual transition rates along the reaction pathways, and provides a genuine metric in the mathematical sense. Second, we constructed a constrained minimization procedure in order to deduce the unknown \textit{in-vivo} values of the individual transition rates from their known \textit{in-vitro} values.

It is instructive to compare our approach with flux control or sensitivity analysis, a widely used method for multistep reaction pathways [3,35–37], which has also been applied to protein synthesis [38]. The latter method explores the local vicinity of a given kinetics and describes the linear response of the overall flux to small changes in the individual transition rates in terms of flux control or sensitivity coefficients. In contrast, the theoretical approach introduced here is not restricted to the linear response regime but explores the space of transition rates in a global manner via an extremum principle (Fig. 2B). Furthermore, both the coordinate transformation from the individual transition rates \(k_i^v\) to the single barrier shifts \(\Delta_i\) and the constraint arising from the global \textit{in-vivo} property make our approach highly nonlinear.

When we applied our computational method to translation elongation by ribosomes, we obtained predictions for the individual \textit{in-vivo} rates \(k_i^v\) that could be validated by three independent sets of data for codon-dependent translation speeds, codon-specific translation dynamics and misense error frequencies of protein synthesis. In all cases, we found good agreement between theory and experiment without adjusting any fit parameter.

Even for the largest growth condition of 2.5 dbl/h, most of the deduced \textit{in-vivo} rates \(k_i^v\) are similar to the measured \textit{in-vitro} rates \(k_i^m\) (Fig. 5B) but three \textit{in-vivo} rates are significantly increased compared to their \textit{in-vitro} values: the rejection rate \(k_{76}\) for near-cognates, the dissociation rate \(k_{off}\) after initial binding, and the recognition rate \(k_{rec}\) for cognate and near-cognate ternary complexes. The largest difference is found for the rejection rate \(k_{76}\), which is increased \textit{in vivo} by a factor of 3.9, while the dissociation rate \(k_{off}\) and the recognition rate \(k_{rec}\) are increased by a factor 3.3 and 2.2, respectively.

For all transition rates of the elongation cycle, we find that the deviations between the \textit{in-vivo} and \textit{in-vitro} rates correspond to relatively small shifts of the corresponding free energy barriers (Fig. 5A). In fact, all single barrier shifts are predicted to be smaller than \(2k_BT\). Because the cytosol represents a rather complex buffer, such small changes in the free energy barriers can be easily envisaged, arising, e.g., from changes in the hydrogen bond networks around the ribosome or from changes in the flexibility of some parts of this complex. On the other hand, our results also show that the high-fidelity buffer at 37°C, used here and developed by [25] represents a good approximation to the cytosol as far as the ribosomal kinetics is concerned, in contrast to earlier estimates in Ref. [39].

**Table 2. \textit{In-vivo} rates of ribosomal transitions.**

<table>
<thead>
<tr>
<th>Rates</th>
<th>E. coli growth rates [dbl/h]</th>
<th>RSD</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.7</td>
<td>1.07</td>
<td>1.6</td>
</tr>
<tr>
<td>(k_{in}^v)</td>
<td>94</td>
<td>94</td>
<td>94</td>
</tr>
<tr>
<td>(k_{off}^v)</td>
<td>1400</td>
<td>1700</td>
<td>2100</td>
</tr>
<tr>
<td>(k_{on}^v)</td>
<td>2100</td>
<td>2500</td>
<td>3000</td>
</tr>
<tr>
<td>(k_{1/2}^v)</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>(k_{1/2}^v)</td>
<td>1600</td>
<td>1600</td>
<td>1700</td>
</tr>
<tr>
<td>(k_{on}^v)</td>
<td>490</td>
<td>500</td>
<td>530</td>
</tr>
<tr>
<td>(k_{4/4}^v)</td>
<td>270</td>
<td>300</td>
<td>340</td>
</tr>
<tr>
<td>(k_{8/0}^v)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(k_{9/8}^v)</td>
<td>2700</td>
<td>3100</td>
<td>3900</td>
</tr>
<tr>
<td>(k_{5/5}^v)</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>(k_{1/2}^v)</td>
<td>0.27</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>(k_{5/8}^v)</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>(k_{7/2}^v)</td>
<td>190</td>
<td>200</td>
<td>230</td>
</tr>
<tr>
<td>(k_{8/4}^v)</td>
<td>15</td>
<td>18</td>
<td>22</td>
</tr>
</tbody>
</table>

The values of the overall elongation rate \(k_{elo}^v\), for the four growth conditions 0.7, 1.07, 1.6, and 2.5 dbl/h were obtained from the data in Ref. [27]. These growth conditions have been chosen because, for these conditions, the total tRNA concentrations have been measured as well in Ref. [28]. The relative standard deviations (RSDs) in the sixth column were obtained from the errors of the \textit{in-vitro} rates in Table 1, and from the associated changes in the predicted \textit{in-vivo} point, see Methods section.

doi:10.1371/journal.pcbi.1003909.t002
Figure 6. Comparison with *in-vivo* experiments of translation elongation. (A) Codon-specific elongation rates $\omega^*_c,\text{elo}$ as determined here from the complete set of individual transition rates for *E. coli* at a growth rate of 2.5 dbl/h, see Eq. 21 and Supporting Figure S3, compared to relative translation rates as measured in Ref. [29] for 29 codons; highlighted symbols indicate the codons CGA, CGC, and CGU (orange) as well as the codons UUU, UUC, UUG, UCC, and CCC (cyan), see text. The Pearson correlation coefficient is 0.56 for all codons and 0.73 when the highlighted codons are
The free energy barriers considered here could be studied by Molecular Dynamics simulations. The latter method has been recently applied to explore the free energy landscape of tRNA translocation through the ribosome [40,41]. From such simulations, one can estimate the attempt frequencies for barrier crossing which are difficult to determine by other computational methods. In principle, these simulation techniques could also be used to investigate how the energy landscape changes as one varies the ambient buffer conditions in the simulations.

Even though the predicted shifts of the free energy barriers are relatively small, the associated changes of the transition rates have an interesting consequence for the relative importance of initial selection and proofreading for the error frequency of protein synthesis. For the codon-specific Markov process depicted in Fig. 3, the efficiency of initial selection and proofreading are described by the coefficients $\langle \omega_2 / \omega_1 \rangle$ and $\langle \omega_5 / \omega_4 \rangle$, respectively. The $\text{in-vitro}$ value of the initial selection coefficient is increased by a factor of 7.7 compared to the corresponding $\text{in-vivo}$ value whereas the proofreading coefficient is increased by a factor of 2.9. The combined improvement of initial selection and proofreading leads to a reduction of the $\text{in-vivo}$ error frequency by a factor of 6.7, a reduction that is primarily achieved by the improved initial selection of the bound ternary complexes.

In the present study, the codon-dependence of the elongation cycle arose from the initial binding rates that depend on the concentrations of cognate, near-cognate, and non-cognate tRNA, because we used the same transition rates along the reaction pathways for all cognate as well as for all near-cognate tRNAs. Thus, the values of the rates $\omega_{21}, \omega_{21}, \ldots$ of the cognate branch represent average values, obtained by averaging over all cognate tRNAs of all codons, and likewise for the internal rates $\omega_{67}, \omega_{76}, \ldots$ of the near-cognate branch. In $\text{in-vivo}$, the decoding rates of different cognate tRNAs were observed to be rather similar [14,25] whereas the GTPase activation rate $\omega_8$ was found to vary between 0.06/s and 1.3/s for different near-cognate codons of tRNA$^{\text{Phe}}$ [25]. Likewise, recent $\text{in-vitro}$ experiments provided evidence that the error frequency on 4 out of 14 near-cognate codons of tRNA$^{\text{Gly}}$ is much higher than the remaining 10 near-cognate codons [42]. Theoretically, it is straightforward to include codon-specific decoding and processing rates. Experimentally, it is, however, quite challenging to determine these rates $\text{in-vitro}$ for all codons and tRNA species.

Our theory for protein synthesis by ribosomes can be extended in a variety of ways. For example, one could study how the overall elongation rate or the missense error frequency vary with changes in the overall ternary complex composition or as a function of individual ternary complex concentrations. Likewise, one may investigate how changes in internal transition rates arising, e.g., from protein or tRNA mutagenesis, affect the speed and accuracy of translation elongation.

The computational method developed here to deduce the $\text{in-vivo}$ from the $\text{in-vitro}$ rates is relatively simple and can be applied, in general, to any multistep process or Markov model, for which one can estimate the $\text{in-vitro}$ rates. Simple examples are provided by the folding and unfolding of proteins, the catalytic activity of enzymes with one active site, or the motility of molecular motors. More complex examples are transcription by RNA polymerase, protein refolding by chaperones, or protein degradation by proteases. Our method can also be applied to the large number of biochemical processes that have been studied by flux control or sensitivity analysis. Furthermore, the similarity measure provided by the kinetic distance could be useful in the context of systems biology, where the importance of detailed kinetics has been recently emphasized [43]. One important target in systems biology is to standardize the experimental data for such networks. Using the kinetic distance introduced here, one could, in fact, compare the kinetic data obtained by different groups in a systematic and quantitative manner.

**Methods**

**Codon-specific accommodation times**

Using the general theory of stochastic processes [44,45], we derived explicit expressions for important dynamical quantities of the translation elongation cycle (Fig. 3) in terms of the individual transition rates. These quantities include the codon-specific accommodation times, i.e., the times that the ribosome needs to fully accommodate a cognate or near-cognate tRNA and, thus, to move from state 0 to state 5 or state 10 for the Markov process in Fig. 3. A straightforward but somewhat tedious computation leads to explicit, analytical expressions for these time scales in terms of the individual transition rates $\omega_{ij}$. These expressions can be decomposed into four different dwell times according to

$$t_{c,\text{acc}} = t_{c,\theta} + t_{c,\text{no}} + t_{c,\text{co}} + t_{c,\text{nr}} ,$$

a decomposition that directly reflects the state space of the Markov process in Fig. 3 and has the following intuitive interpretation.

The first dwell time $t_{c,\theta}$ represents the total time that the ribosome spends in state 0 during one complete elongation cycle at codon $c$. Because of the different dissociation and backward transitions, the ribosome typically visits the state 0 several times before it is fully accommodated in the states 5 or 10, see Fig. 3. The second dwell time $t_{c,\text{no}}$ in Eq. 11 corresponds to the total time that the ribosome binds a non-cognate ternary complex and, thus, dwells in state 11 during one complete elongation cycle at codon $c$. The third dwell time $t_{c,\text{co}}$ corresponds to the total time that the ribosome spends in the intermediate states 1, 2, 3, and 4 of the cognate branch during one complete elongation cycle at codon $c$. Finally, the fourth dwell time $t_{c,\text{nr}}$ in Eq. 11 represents the total time that the ribosome spends in the intermediate states 6, 7, 8, and 9 of the near-cognate branch.

These four dwell times can be expressed in a particularly compact and transparent manner if one uses the transition probabilities

$$\pi_i \equiv \frac{\omega_{ik}}{\sum_k \omega_{ik}} .$$

The dwell time $t_{c,\theta}$, which the ribosome spends in state 0 during one complete elongation cycle at codon $c$, then has the form

$$t_{c,\theta} = \frac{1}{k_{\text{on}} [X_{c,\text{co}} \rho_{c,\text{co}} + X_{c,\text{nr}} \rho_{c,\text{nr}}]} ,$$

and, thus, depends on the concentrations $X_{c,\text{co}}$ and $X_{c,\text{nr}}$ of free cognate and near-cognate ternary complexes as well as on the
dimensionless, concentration-independent ratios

\[ \rho_{co} = \frac{\pi_{12} \pi_{23} \pi_{45}}{\Gamma - \pi_{12} \pi_{21}} \]  

and

\[ \rho_{nt} = \frac{\pi_{57} \pi_{78} \pi_{9,10}}{\Gamma - \pi_{57} \pi_{76}}. \]  

The second dwell time \( t_{c,nt} \) for state 11 with a bound non-cognate ternary complex is given by the expression

\[ t_{c,nt} = t_{c,nt} k_{on} X_{c,nt} / \omega_{off}, \]  

and is, thus, proportional to the concentration \( X_{c,nt} \) of free non-cognate ternary complexes.

The third dwell time \( t_{c,co} \), which represents the sum of all dwell times for the intermediate states 1, 2, 3, and 4 of the cognate branch, can be written as

\[ t_{c,co} = \frac{X_{c,co} \rho_{co}}{X_{c,co} \rho_{co} + X_{c,nt} \rho_{nt}} \tau_{co} \]  

with the concentration-independent time scale

\[ \tau_{co} = \frac{1}{\omega_{12} \pi_{23} \pi_{45}} + \frac{1}{\omega_{23} \pi_{45}} + \frac{1}{\omega_{34} \pi_{45}} + \frac{1}{\omega_{45}} \]  

that depends only on transitions that emanate from the intermediate states of the cognate branch. Likewise, the fourth dwell time \( t_{c,nt} \) for the intermediate states 6, 7, 8, and 9 of the non-cognate branch has the form

\[ t_{c,nt} = \frac{X_{c,nt} \rho_{nt}}{X_{c,co} \rho_{co} + X_{c,nt} \rho_{nt}} \tau_{nt} \]  

with the concentration-independent time scale

\[ \tau_{nt} = \frac{1}{\omega_{57} \pi_{78} \pi_{9,10}} + \frac{1}{\omega_{78} \pi_{9,10}} + \frac{1}{\omega_{98} \pi_{9,10}} + \frac{1}{\omega_{99}} \]  

that depends only on transitions that emanate from the intermediate states of the near-cognate branch.

**Codon-specific and overall elongation rates**

The expression for the codon-specific accommodation time \( t_{c,acc} \) as given by Eq. 11 involves all individual rates \( \omega_{ij} \) apart from the processing rate \( \omega_{pro} \). When we add the processing time \( 1/\omega_{pro} \), we obtain the codon-specific elongation time \( t_{c,elo} = t_{c,acc} + 1/\omega_{pro} \) which the ribosome needs to complete a full elongation cycle at a certain codon \( c \). The codon-specific elongation rates are then given by

\[ \omega_{c,elo} = 1/t_{c,elo} = 1/[t_{c,acc} + 1/\omega_{pro}]. \]  

One important global property of protein synthesis is the average speed of the ribosomes, which defines the overall elongation rate \( \omega_{elo} \). The inverse of the overall elongation rate is equal to the average elongation time \( \langle t_{elo} \rangle \equiv \sum_c p_c t_c,elo = \sum_c p_c t_c,acc + 1/\omega_{pro} \), which is obtained by averaging the codon-specific elongation times \( t_{c,elo} \) over all codons \( c \) using the codon usage \( p_c \). For each codon \( c \), the quantity \( p_c \) represents the probability that the ribosome encounters this codon. These probabilities are normalized and satisfy \( \sum_c p_c = 1 \).

For the *in-vitro* assay, the relation between the overall elongation rate \( \omega_{elo} \) and the codon-specific accommodation times \( t_{c,acc} \) was rewritten in the form

\[ \omega_{pro} = \left[ \frac{1}{\omega_{elo}} - \sum_c p_c t_{c,acc} \right]^{-1} \]  

and then used to calculate the processing rate \( \omega_{pro} \) from the measured value of the overall elongation rate \( \omega_{elo} \) and the measured values of the individual rates \( \omega_{ij} \), which determine the codon-specific accommodation times \( t_{c,acc} \).

In *vivo*, the overall elongation rate \( \omega_{elo}^* \) is given by the analogous expression

\[ \omega_{elo}^* = \left[ \sum_c p_c t_{c,acc}^* + 1/\omega_{pro}^* \right]^{-1}, \]  

where the codon-specific accommodation times \( t_{c,acc}^* \) follow from the same expression as in Eq. 11 but with the in-*vivo* values \( \omega_{ij}^* \) replaced by the in-*vivo* values \( \omega_{ij}^* \). When we insert the known *in-vivo* value \( \omega_{pro}^* \) of the overall elongation rate into Eq. 23, we obtain a constraint on the (unknown) *in-vivo* values \( \omega_{ij}^* \) of the individual transition rates. This constraint can be expressed in terms of the single barrier shifts \( \Delta \gamma \) when we replace \( \omega_{ij}^* \) in Eq. 23 by \( \omega_{ij}^* \exp[-\Delta \gamma] \), see Eq. 2.

**In-vitro values of individual transition rates**

All *in-vitro* values of the individual transition rates as given in Table 1 have been obtained for the high-fidelity buffer as developed in [12], [25], and [34]. Most of these values are based on previous measurements as explained in the following paragraph. In addition, we also performed new experiments to measure the overall elongation rate \( \omega_{elo} \) at 20° C and at 37° C, see Supporting Figure S1, as well as the individual rates \( \omega_{90\rightarrow k_{5,nt}} \) and \( \omega_{90\rightarrow k_{7,nt}} \) at 20° C, see Supporting Figure S2, using the experimental protocols described previously [34,46,47].

The *in-vitro* values \( \omega_{con} \) of the association constant was previously measured at 20° C [12]. Its value at 37° C was obtained assuming an Arrhenius temperature dependence and using the previously determined activation energy of 2.4 kcal/mol for initial binding [10]. The dissociation rate \( \omega_{diss} \) at 20° C was taken from [12]. The decoding rates at 20° C were obtained by averaging over previously published values as measured for different codons of tRNA\(^{Phe}\). In particular, we averaged the rates as given in Table 1 of [25] for cognate as well as for near-cognate codons to obtain the rates \( \omega_{90\rightarrow k_{5,nt}} \) and \( \omega_{90\rightarrow k_{7,nt}} \) at 20° C, see Supporting Figure S2, using the experimental protocols described previously [34,46,47].

The *in-vitro* value \( \omega_{con} \) of the association constant was previously measured at 20° C [12]. Its value at 37° C was obtained assuming an Arrhenius temperature dependence and using the previously determined activation energy of 2.4 kcal/mol for initial binding [10]. The dissociation rate \( \omega_{diss} \) at 20° C was taken from [12]. The decoding rates at 20° C were obtained by averaging over previously published values as measured for different codons of tRNA\(^{Phe}\). In particular, we averaged the rates as given in Table 1 of [25] for cognate as well as for near-cognate codons to obtain the rates \( \omega_{90\rightarrow k_{5,nt}} \) and \( \omega_{90\rightarrow k_{7,nt}} \) at 20° C, see Supporting Figure S2, using the experimental protocols described previously [34,46,47].

Finally, we assumed an Arrhenius temperature dependence to estimate some of the *in-vitro* rates at 37° C from their values as measured at 20° C. These estimates are based on the following...
considerations. We start from Eq. 4 for the transition rates and use the decomposition \( \Delta G_{ij} = \Delta H_{ij} - T \Delta S_{ij} \) of the activation free energy \( \Delta G_{ij} \) into the activation enthalpy \( \Delta H_{ij} \) and the activation entropy \( \Delta S_{ij} \), which leads to

\[
\begin{align*}
\omega_{ij}(T) &= v_g \exp[\Delta S_{ij}/k_B] \exp[-\Delta H_{ij}/k_B T] \\
&= v_o \exp[\Delta S_{ij}/k_B] \exp[-\Delta H_{ij}/k_B T] \quad (24)
\end{align*}
\]

where the last expression involves the attempt frequency \( v_o = k_B T/h \) as obtained from transition-state theory [2]. In this way, any state-dependence of the attempt frequency \( v_g \) has been absorbed into the activation entropies \( \Delta S_{ij} = \Delta S_{ij} + k_B \ln(v_g/v_o) \). If one plots the logarithms \( \ln(\omega_{ij}/v_o) \) for small \( T \) vary significantly with the ribosomal states \( i \) and \( j \) [10,13,48,49]. Possible molecular mechanisms for this variation have been recently discussed based on atomistic molecular dynamics simulations [50].

Using the expression in Eq. 24 with \( T \)-independent enthalpies \( \Delta H_{ij} \) and entropies \( \Delta S_{ij} \), we now consider the ratios \( \omega_{ij}(T)/\omega_{45}(T) \) at the two temperatures of interest, \( T_1 = 20^\circ C = 293.15 K \) and \( T_2 = 37^\circ C = 310.15 K \). We take the accommodation rate \( \omega_{45} \) as a reference rate because the value of this rate has been measured at both temperatures. For each individual transition, we then obtain two equations, corresponding to the two temperatures \( T_1 \) and \( T_2 \), which can be combined to eliminate the enthalpy \( \Delta H_{ij} \). As a result, we obtain the relation

\[
\begin{align*}
\frac{\omega_{ij}(T_2)}{\omega_{45}(T_2)} &= \frac{\omega_{ij}(T_1)}{\omega_{45}(T_1)} \frac{T_1/T_2}{\exp \left[ \frac{\Delta S_{ij} - \Delta S_{45}}{k_B} \frac{T_2 - T_1}{T_2} \right]}. \quad (25)
\end{align*}
\]

At present, the entropy differences \( \Delta S_{ij} - \Delta S_{45} \) are difficult to estimate for all individual transitions from the available experimental information. However, these differences are multiplied by the relative temperature difference \( (T_2 - T_1)/T_2 \approx 0.055 \) which is rather small. Therefore, we used the approximate relation

\[
\begin{align*}
\frac{\omega_{ij}(37^\circ C)}{\omega_{45}(37^\circ C)} &\approx \frac{\omega_{ij}(20^\circ C)}{\omega_{45}(20^\circ C)} \approx 293.15/310.15 \\
&= 0.9418 \\
\end{align*}
\]

(26)

to estimate the values of the rates \( \omega_{45}, \omega_{421}, \omega_{21}, \omega_{223}, \omega_{76}, \omega_{78}, \omega_{83}, \) and \( \omega_{49} \) at \( 37^\circ C \) (Table 1) from the measured values of \( \omega_{ij}(20^\circ C), \omega_{45}(20^\circ C), \) and \( \omega_{45}(37^\circ C) \).

**In-vivo values of association rates**

The overall elongation rate \( \omega^*_c \) as given by Eq. 23 also depends on the association rates for initial binding, which are proportional to the pseudo-first-order rate constant \( \kappa_{on} \) and to the concentrations \( X^*_c \) of the ternary complexes as in Eq. 9. Therefore, in order to use Eq. 23 for the process \( m \) in vivo, we had to estimate the corresponding values \( \kappa_{on}^m \) and the ternary complex concentrations \( X^*_m \) in the cell.

The diffusion of ternary complexes and, thus, their binding to ribosomes is slowed down in vivo by molecular crowding. The time it takes a ternary complex to find a single ribosome depends on the cell volume, the diffusion constant of the ternary complex, and the ribosome size [51]. Using the diffusion constant of 2.57 \( \mu \)m\(^2\)/s [52,53] for a ternary complex in the cytosol, we found that the \( m \)-in vivo value \( \kappa_{on}^m \) of the bimolecular association rate constant is about 54% of the \( m \)-in vitro value \( \kappa_{on} \), compare Table 1 and Table 2.

For the \( m \)-in vivo concentrations \( X^*_m \) of the ternary complexes, we used the values of the tRNA concentrations as measured by [28] in \( E. coli \) for the growth conditions of 0.7, 1.07, 1.6, and 2.5 dBh/h. In the latter study, the authors determined the concentrations \( X^*_a \) of all 43 elongator tRNA species \( a \). These concentrations are then combined, for each codon \( c \), into the concentrations \( X^*_c, qos, \) \( X^*_c, enr, \) and \( X^*_c, no \), of cognate, near-cognate, and non-cognate ternary complexes within the cell. Thus, for each codon \( c \), we started from the corresponding row in Fig. 4, and added all concentrations \( X^*_a \) up that correspond to green (cognate), yellow (near-cognate), and purple (non-cognate) tRNA species, respectively.

**Uncertainty of predicted in-vivo rates**

To estimate the uncertainty of the predicted \( m \)-in vivo rates \( \omega^*_m, min \), we first simplify the notation. In this section, the internal transitions with distinct transition rates will be distinguished by the subscript \( m \) with \( m = 1,2,\ldots,M \). Thus, we now use the short-hand notation \( \omega_{m}, \omega^*_m, \) and \( \omega^*_m, min \) for the \( m \)-in vitro rates \( \omega^*_m \) of a certain assay, for the unknown \( m \)-in vivo rates \( \omega_m \), and for the predicted \( m \)-in vivo rates \( \omega^*_m, min \), respectively. For ribosome elongation as described by the Markov process in Fig 3, we distinguish \( M = 12 \) internal transitions.

The inaccuracy or error of the \( m \)-in vitro rates can be described by

\[
\omega_m = \omega_m \pm \delta_m \equiv \omega_m (1 \pm \epsilon_m) \quad (27)
\]

with the absolute error \( \delta_m \) and the relative error

\[
\epsilon_m \equiv \delta_m / \omega_m \quad (28)
\]

of the \( m \)-in vitro rate \( \omega_m \). Both the average values \( \omega_m \) and the absolute errors \( \delta_m \) are estimated from the experimental data for the \( m \)-in vitro assay under consideration.

When we apply the minimization procedure to the average values \( \omega_m \) of the \( m \)-in vitro rates, we use the coordinates

\[
\Delta_m \equiv \ln (\omega_m / \omega^*_m) \quad (29)
\]

for the multi-dimensional barrier space. We then determine the point \( \mathbf{H} \) that is located on the hypersurface defined by Eq. 8 and depicted in Fig. 2B and has the shortest distance from the origin of the \( \Delta_m \)-coordinates. The coordinate values \( \Delta_m, min \) of the point \( \mathbf{H} \) then lead to the predicted values \( \omega^*_m, min = \omega_m \exp[- \Delta_m, min] \) for the \( m \)-in vivo rates.

In order to estimate the uncertainty of these predictions, it is useful to consider an auxiliary ensemble of fictitious \( m \)-in vitro assays that is constructed ‘around’ the given assay as follows. For each transition \( \omega_m \), we introduce the binary variable \( s_m = \pm 1 \). The \( M \) binary variables \( s_m \) can assume \( 2^M \) different ‘configurations’ \( C \) as described by the different \( M \)-tuples

\[
\mathbf{s} = (s_1, s_2, \ldots, s_M). \quad (30)
\]

Each of these configurations defines a fictitious \( m \)-in vitro assay, again denoted by \( C \), with transition rates
\[ \omega_m^C = \omega_m + \frac{s_m^C}{\Delta_m} \delta_m = \omega_m(1 + s_m^C \epsilon_m) \]  
(31)

The rates of assay \( C \) define the coordinates

\[ \Delta_m^C = \ln \left( \frac{\omega_m^C}{\omega_m^A} \right) = \Delta_m + d_m^C \]  
(32)

for the multi-dimensional barrier space with

\[ d_m^C = \ln (1 + s_m^C \epsilon_m) \approx s_m^C \epsilon_m \]  
(33)

where the asymptotic equality applies to the limit of small relative errors \( \epsilon_m = \delta_m / \omega_m \) of the \( in\text{-}vitro \) rates. Therefore, if the origin of the multi-dimensional barrier space is defined by the coordinates \( \Delta_m \) in Eq. 29, corresponding to the average \( in\text{-}vitro \) rates \( \bar{\omega}_m \), the ensemble of the fictitious assays \( C \) forms the corners of a multi-dimensional ‘error polyhedron’ around this origin. For each corner, again labeled by \( C \), we can apply our minimization procedure and minimize the kinetic distance of the point \( C \) from the hypersurface as defined by Eq. 8 and depicted in Fig. 2B. The point \( H^C \) on the hypersurface with the shortest distance from the corner \( C \) has the coordinates \( \Delta_m^C, \min \).

The variations in the coordinate values \( \Delta_m^C, \min \) of the predicted \( in\text{-}vivo \) point as obtained for different corners \( C \) can be used to obtain an estimate for the absolute error \( \epsilon_m^C \) of these coordinate values. We then write the coordinate values of the predicted \( in\text{-}vivo \) point in the form

\[ \Delta_m^C = \bar{\Delta}_m^C \pm \epsilon_m^C \]  
(34)

where the values \( \bar{\Delta}_m^C \) correspond to the average \( in\text{-}vitro \) rates \( \bar{\omega}_m \). The predicted \( in\text{-}vitro \) rates are now given by

\[ \omega_m^C = \omega_m \exp \left[ - \Delta_m^C \right] = \omega_m(1 \pm \epsilon_m^C) \]  
(35)

Therefore the relative error of the predicted \( in\text{-}vitro \) rates \( \epsilon_m^C \) reflects both the relative error \( \epsilon_m^C \) of the \( in\text{-}vitro \) rates \( \omega_m \) (Eq. 27) and the absolute error \( \epsilon_m^C \) of the coordinate values \( \Delta_m, \min \) for the predicted \( in\text{-}vivo \) point (Eq. 34).

The Markov process for ribosome elongation considered here, see Fig. 3, involves \( M = 12 \) distinct transition rates, which implies that the corresponding barrier space has 12 dimensions. We first determined the coordinate values \( \bar{\Delta}_m^C \) of the \( in\text{-}vitro \) point as predicted from the average value \( \bar{\omega}_m \) of the \( in\text{-}vitro \) rates. The largest coordinate values \( \bar{\Delta}_m^C \) of the predicted \( in\text{-}vivo \) point were found for the three transition rates \( \bar{\omega}_{76}, \bar{\omega}_{87}, \) and \( \bar{\omega}_{97} \) (Fig. 5B). We then focused on the errors of these three \( in\text{-}vitro \) rates, which define 8 corners of the ‘error polyhedron’ around the origin of the \( \bar{\Delta}_m \)-coordinates. For each of these corners \( C \), we determined the closest point on the hypersurface and the coordinate values \( \Delta_m^C, \min \) of this point. We then estimated the absolute error \( \epsilon_m^C \) of the coordinate values \( \Delta_m^C, \min \) from the largest and smallest values of \( \Delta_m^C, \min \) as obtained for different corners \( C \). The errors \( \epsilon_m^C \) were finally used, together with the relative errors \( \epsilon_m \) of the measured \( in\text{-}vitro \) rates, to determine the relative standard deviations (RSDs) of the predicted \( in\text{-}vitro \) rates as displayed in Table 2.

**Missense error frequency**

Consider a certain tRNA species \( a \) and a codon \( c \) that is near-cognate to \( a \). The missense error frequency for misreading the codon \( c \) by the tRNA species \( a \) is equal to the probability \( P(a\rightarrow\text{c(nr)}) \) that \( a \) is fully accommodated at \( c \). For the multistep process considered here, this probability is given by

\[ P(a\rightarrow\text{c(nr)}) = \frac{X_a \rho_{\text{nr}}}{X_{\text{co}} \rho_{\text{co}} + X_{\text{nt}} \rho_{\text{nt}}}, \]  
(36)

which depends on the concentration \( X_a \) of the near-cognate ternary complex species \( a \), on the concentrations \( X_{\text{co}} \) and \( X_{\text{nt}} \) of all cognate and near-cognate ternary complexes as well as on the concentration-independent ratios \( \rho_{\text{co}} \) and \( \rho_{\text{nt}} \) as given by Eqs. 14 and 15.

The experimental study in [31] determined the error frequency for all codons that are near-cognate to \( a \equiv \text{tRNA}^{\text{Lys}} \). The average error frequency for misreading one of these codons is then obtained from

\[ \langle P(a\rightarrow\text{c(nr)}) \rangle = \sum_{c \in \text{tRNA}(a)} p_c P(a\rightarrow\text{c(nr)}) \left( \sum_{c \in \text{co}} p_c \right)^{-1} \]  
(37)

where the set \( \text{tRNA}(a) \) contains all codons \( c \) that are near-cognate to \( a \) and \( p_c \) denotes the codon usage as before.

**Supporting Information**

**Figure S1** Overall elongation rate as measured for a model protein \textit{in vitro}. Kinetics of CspA translation \textit{in vitro} at different temperatures. CspA mRNA, which codes for a 70 aa-long protein from \textit{E. coli}, was prepared by T7 RNA-polymerase transcription. Ribosomes were synchronized by forming an initiation complex consisting of 70S ribosomes, CspA mRNA and a fluorogenic derivative of initiator tRNA\textsuperscript{Met} carrying BodipyFL at the 2-aminooxy group of Met in the presence of initiation factors (IF1, IF2, and IF3) and GTP. Translation was carried out in a fully reconstituted translation system by adding initiation complexes (15 nM) to a mixture of EF-Tu-GTP-aminocyl-tRNA (40 µM aminocyl-tRNA, 100 µM EF-Tu in total), EF-G (3 µM), GTP (2 mM), phosphoenol pyruvate (6 mM), and pyruvate kinase (0.1 mg/ml) in HiFi buffer (50 mM Tris-HCl, pH 7.5, 30 mM KCl, 70 mM NH4Cl, 3.5 mM free MgCl\textsubscript{2}, 0.5 mM spermidine, and 8 mM putrescine) at the indicated temperatures [46]. In the absence of translation termination and ribosome recycling factors, translation was limited to a single round, i.e. at most one molecule of CspA was synthesized per ribosome. The reactions were stopped at the indicated time intervals and translation products separated on 16.5% Tris-Tricine-PAGE and visualized by the fluorescent reporter BODIPY-Fi at the N-terminus of the peptides [47] (left panels). The intensity of the full length product was quantified with ImageJ (right panel, circles). Average translation rates per codon, which depend on the elongation rates only, were determined by exponential fitting (fits in graphs of the right panel).

**Figure S2** \textit{In vitro} rates as measured for near-cognate accommodation and rejection after proofreading. \textit{In vitro} values of the rates \( \omega_{90,10} \equiv k_{\text{nr}} \) and \( \omega_{90} \equiv k_{\text{nt}} \) for near-cognate accommodation and rejection after proofreading at 20°C as determined by the experimental protocol described previously in Ref. [34]. The formation of [\textsuperscript{3}H]Met\textsuperscript{[14]C}]Phe was monitored.
under multiple-turnover conditions using initiation mixtures 70S–mRNA(AUG/GUC)-[^14]HMet-tRNA[^14]Met (0.14 µM) and varying concentrations of the ternary complex EF-Tu–GTP–[^14]CPhet-tRNA[^14]Glu, which is near-cognate to the CUC codon. For each concentration of the ternary complex, the rates were determined from the linear slopes of the time courses. From the hyperbolic dependence of the concentration dependence of $k_{\text{app}}$, we calculated $\omega_{0,10} = 0.060 \pm 0.006/s$ and $K_0 = 2.4 \mu M$. Using the previously measured efficiency $\omega_{0,10}/(\omega_{0,10} + \omega_{0,0}) = 1/15$ of the proofreading step [23], we then obtained the value $\omega_{0,0} = 0.84 \pm 0.08/s$ for near-cognate rejection after proofreading. (TIF)

**Figure S3** Codon-specific elongation rates in *in vitro* and *in vivo*. Codon-specific elongation rates $\omega_{\text{c},\text{elo}}$ in units of amino acids per second as calculated from Eq. 17, see Methods section in the main text, using the decomposition of the codon-specific elongation times in Eq. 7 and the complete sets of individual transition rates: (A) *in vitro* values $\omega_{\text{c},\text{elo}}$ for the high-fidelity buffer at 37°C, obtained from the individual rates in Table 1; (B, C) *in vivo* values $\omega_{\text{c},\text{elo}}$ for *E. coli* at growth conditions of (B) 0.7 d/b/h and (C) 2.5 d/b/h, calculated from the individual rates in Table 2. (TIF)

**References**


