OPEN ACCESS Freely available online

PLOS | COMPUTATIONAL BIOLOGY

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

Sophia Rudorf[1], Michael Thommen[2], Marina V. Rodnina[2], Reinhard Lipowsky[1]*

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.

Citation: Rudorf S, Thommen M, Rodnina MV, Lipowsky R (2014) Deducing the Kinetics of Protein Synthesis *In Vivo* from the Transition Rates Measured *In Vitro*. PLoS Comput Biol 10(10): e1003909. doi:10.1371/journal.pcbi.1003909

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

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

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

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.

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 different states have been studied by a variety of experimental techniques: chemical probing methods [8], pre-steady state kinetics [9–15], electron microscopy [16–19], X-ray crystallography [20–22], and single molecule methods [23,24]. 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-vivo* process: the overall elongation rates [26,27] and the tRNA concentrations [28], both of which have been measured in *E. coli* for different growth conditions.

The *in-vivo* 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-vivo* and *in-vitro* 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 ω_{ij} and $\omega_{ij}^{\star}$ 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 $\tau_{ij} \equiv 1/\omega_{ij}$ and $\tau_{ij}^{\star} \equiv 1/\omega_{ij}^{\star}$. Thus, we require that the distance $\mathcal{D}_{ij}(\omega_{ij},\omega_{ij}^{\star})$ between the rates ω_{ij} and $\omega_{ij}^{\star}$ is equal to the distance $\mathcal{D}_{ij}(\tau_{ij},\tau_{ij}^{\star})$ between the times τ_{ij} and τ_{ij}^{*}, i.e., that

$$\mathcal{D}_{ij}(\omega_{ij},\omega_{ij}^{\star}) = \mathcal{D}_{ij}(\tau_{ij},\tau_{ij}^{\star}) = \mathcal{D}_{ij}(1/\omega_{ij},1/\omega_{ij}^{\star}). \quad (1)$$

The *simplest* expression for $\mathcal{D}_{ij}$ that fulfills this requirement is provided by

$$\mathcal{D}_{ij}(\omega_{ij},\omega_{ij}^{\star}) \equiv |\ln(\omega_{ij}^{\star}/\omega_{ij})| = |\Delta_{ij}| \quad (2)$$

with the logarithmic difference

$$\Delta_{ij} \equiv \ln(\omega_{ij}) - \ln(\omega_{ij}^{\star}) = -\ln(\omega_{ij}^{\star}/\omega_{ij}) = \ln(\tau_{ij}^{\star}/\tau_{ij}), \quad (3)$$

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

The single transition distance $\mathcal{D}_{ij}$ is dimensionless and does not involve any parameter apart from the two rates ω_{ij} and $\omega_{ij}^{\star}$. In addition, this distance satisfies the two scaling relations $\mathcal{D}_{ij}(\omega,b\omega) = \mathcal{D}_{ij}(\omega,\omega/b)$ and $\mathcal{D}_{ij}(b\omega_1,b\omega_2) = \mathcal{D}_{ij}(\omega_1,\omega_2)$ for any rescaling factor $b>0$. The first scaling relation implies that $b\omega$ and ω/b have the same distance from ω, which agrees with our intuition. The second scaling relation implies that the distance $\mathcal{D}_{ij}$ does not depend on the units used to measure the rates. For small deviations of $\omega_{ij}^{\star}$ from ω_{ij}, which are equivalent to small deviations of $\tau_{ij}^{\star}$ from τ_{ij}, the distance $\mathcal{D}_{ij} = |\Delta_{ij}|$ becomes asymptotically equal to both $|\omega_{ij}^{\star} - \omega_{ij}|/\omega_{ij}$ and $|\tau_{ij}^{\star} - \tau_{ij}|/\tau_{ij}$.

The *in-vitro* rate ω_{ij} can be expressed in terms of the activation free energy or free energy barrier ΔG_{ij} and the attempt frequency ν_{ij} which leads to

$$\omega_{ij} = \nu_{ij} \exp[-\Delta G_{ij}/k_{\mathrm{B}}T], \quad (4)$$

where the thermal energy $k_{\mathrm{B}}T$ provides the basic free energy scale. When we combine this expression with the analogous expression for the *in-vivo* rate, the logarithmic difference between the two rates becomes

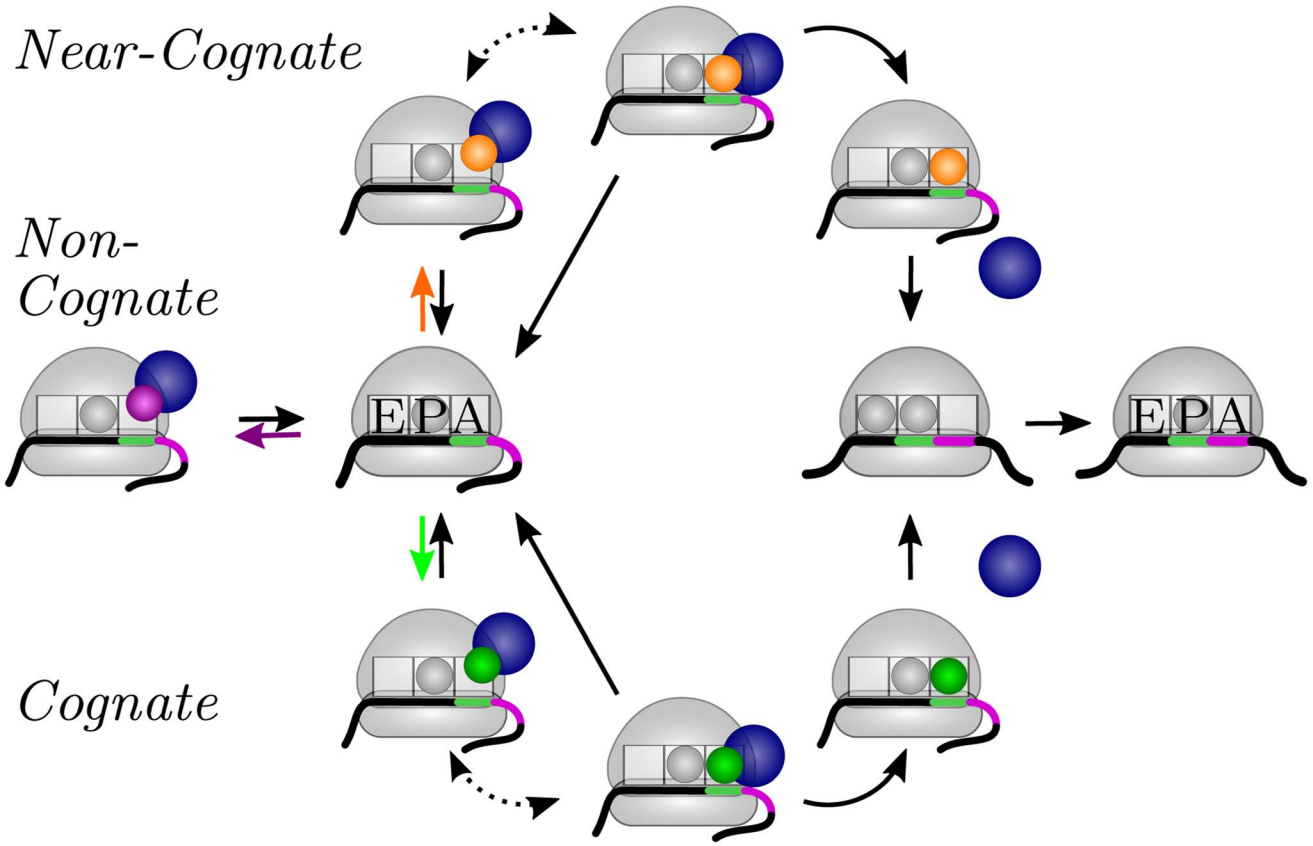

Figure 1. Elongation cycle of a ribosome (gray dome) translating an mRNA (black-green-purple line). Aminoacyl-tRNA (small gray, green, purple, or orange sphere) is delivered to the ribosome in a ternary complex with the elongation factor EF-Tu (larger blue sphere) and GTP (not shown). In addition to the initial binding site, the ribosome has three tRNA binding sites, the A, P, and E sites. The elongation cycle of translation starts when the A site of the ribosome has arrived at a new codon (green) of the mRNA. The ribosome then binds a ternary complex with a tRNA that may be cognate, near-cognate, or non-cognate to this codon. As a consequence, the elongation cycle exhibits three different branches corresponding to three different reaction pathways: (left) A non-cognate ternary complex is again released from the initial binding site of the ribosome; (top) A near-cognate ternary complex is usually rejected but is very rarely used to elongate the peptide chain; and (bottom) A cognate ternary complex may also be rejected but is typically used for elongation of the peptide chain. The two dotted arrows correspond to additional intermediate states and transitions as explained in more detail in Fig. 3 below.
doi:10.1371/journal.pcbi.1003909.g001

$$\Delta_{ij} = [\Delta G^{\star}_{ij} - \Delta G_{ij}]/k_{\mathrm{B}}T - \ln(\nu^{\star}_{ij}/\nu_{ij}). \qquad (5)$$

Because the prefactors $\nu^{\star}_{ij}$ and ν_{ij} are expected to have the same order of magnitude, the second term $\ln(\nu^{\star}_{ij}/\nu_{ij})$ 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 Δ_{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 Δ_{ij}. It should be noted, however, that, in spite of this terminology, changes in the attempt frequency as described by the term $\ln(\nu^{\star}_{ij}/\nu_{ij})$ in Eq. 5 are included in the logarithmic difference Δ_{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 *in-vivo* rates $\omega^{\star}_{ij}$ 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 *in-vivo* rates $\omega^{\star}_{ij}$ to the single barrier shifts Δ_{ij}, which can attain both positive and negative values. This coordinate transformation is highly nonlinear but invertible with the inverse transformation given by $\omega^{\star}_{ij} = \omega_{ij}\exp[-\Delta_{ij}]$.

The overall distance $\mathcal{D}$ between the *in-vitro* and the *in-vivo* kinetics is now defined by

$$\mathcal{D} \equiv \sqrt{\sum \Delta^2_{ij}} = \sqrt{\sum \left[\ln(\omega^{\star}_{ij}/\omega_{ij})\right]^2}, \qquad (6)$$

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

If all *in-vivo* rates are identical to their *in-vitro* values, apart from a single one, $\omega^{\star}_{kl} \neq \omega_{kl}$, Eq. 6 for the kinetic distance $\mathcal{D}$

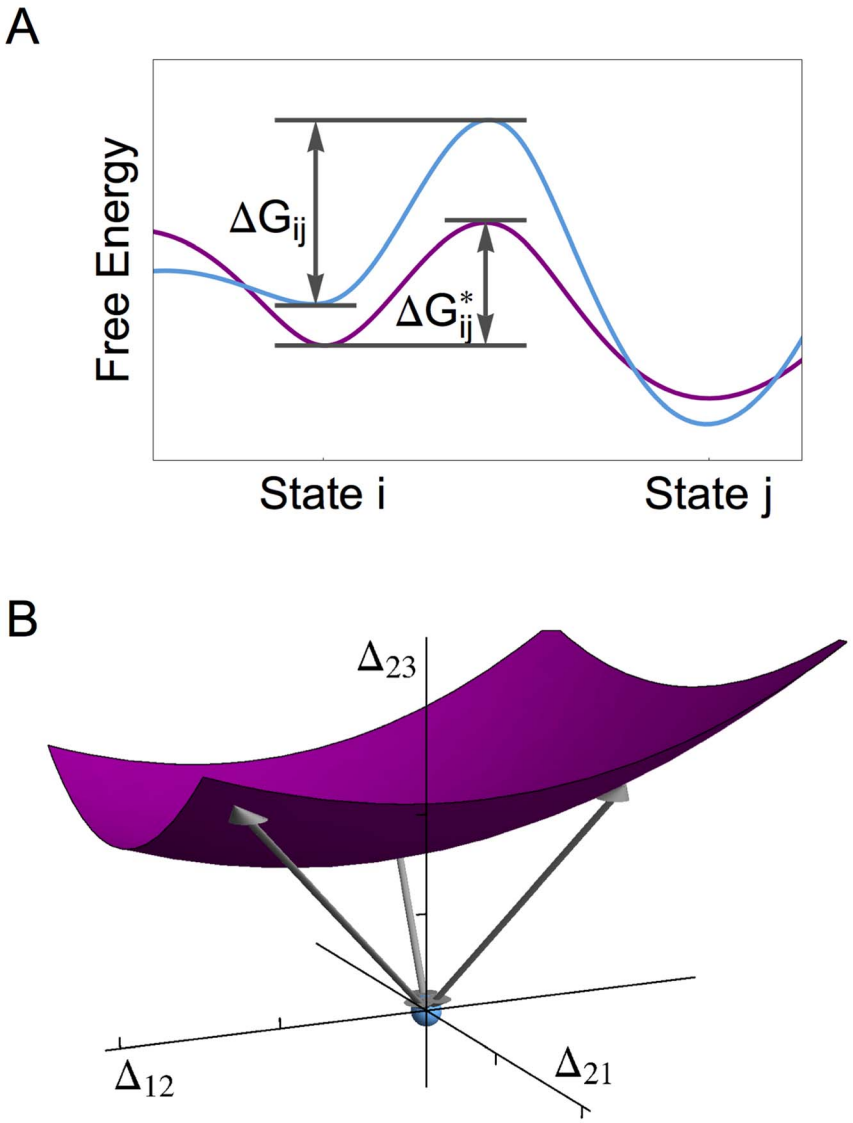

Figure 2. Kinetic distance based on the logarithmic differences Δ_{ij} between the *in-vitro* and *in-vivo* rates. (A) *In-vitro* free energy barrier ΔG_{ij} and *in-vivo* barrier ΔG^*_{ij} for the transition from state i to state j. When expressed in units of $k_B T$, the single barrier shift $\Delta G^*_{ij} - \Delta G_{ij}$ determines the logarithmic difference Δ_{ij}, see Eq. 5; and (B) Three-dimensional section of the multi-dimensional barrier space with coordinates Δ_{12},Δ_{21}, and Δ_{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 $\omega^*_{ij,\,\min}$.
doi:10.1371/journal.pcbi.1003909.g002

reduces to Eq. 1 for the single transition distance $\mathcal{D}_{kl}$. Because the choice of the two states k and l is arbitrary, this property of the kinetic distance applies to all individual transitions ω_{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 $\mathcal{D}_u \equiv \sqrt{\sum u_{ij}^2 \Delta_{ij}^2}$ with dimensionless weight factors u_{ij}, this requirement would imply that all weight factors must assume the unique values $u_{ij} = 1$ and that $\mathcal{D}_u$ must be equal to the kinetic distance $\mathcal{D}$ as given by Eq. 6.

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

$$\Delta'_{ij} \equiv \ln(\omega'_{ij}/\omega^\star_{ij}) = \Delta_{ij} + d_{ij} \tag{7}$$

with the logarithmic differences $d_{ij} \equiv \ln(\omega'_{ij}/\omega_{ij})$. The latter quantities determine the kinetic distance $\mathcal{D}_{A,A'}$ between the two *in-vitro* assays, which is given by $\mathcal{D}_{A,A'} = \sqrt{\sum d_{ij}^2}$. The two sets of barrier shifts, Δ_{ij} and Δ'_{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

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, $Q=Q(\omega_{12},\omega_{21},\omega_{23},\dots)$. The *in-vivo* values $\omega^{\star}_{ij}$ of the individual transition rates must reproduce the experimentally measured value $Q^{\star}_{exp}$ of the global quantity. This requirement implies the equation

$$Q(\omega^{\star}_{12},\omega^{\star}_{21},\omega^{\star}_{23},\dots)=Q^{\star}_{\rm exp}, \tag{8}$$

which represents a constraint on the unknown *in-vivo* values $\omega^{\star}_{ij}$. This constraint can be expressed in terms of the single barrier shifts Δ_{ij} using the inverse coordinate transformation $\omega^{\star}_{ij}=\omega_{ij}\exp[-\Delta_{ij}]$ with the known *in-vitro* values ω_{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 $Q^{\star}_{\rm exp}$ of the global kinetic quantity. In addition, the Euclidean distance of such a point from the origin is equal to the kinetic distance Δ between the (unknown) *in-vivo* and the (known) *in-vitro* values of the transition rates. Our prediction for the *in-vivo* values $\omega^{\star}_{ij}$ 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 $\Delta_{ij,\min}$ in order to distinguish these values from the variable coordinates Δ_{ij}.

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' kinetic distance $\mathcal{D}$. If the kinetic distance satisfied $\mathcal{D}<\mathcal{D}_o$, the true *in-vivo* point would be located within a sphere of radius $\mathcal{D}_o$ around the origin. The smallest sphere that is compatible with the *in-vivo* constraint as given by Eq. 8 is the one that touches the hypersurface depicted in Fig. 2B, and the radius $\mathcal{D}=\mathcal{D}_{\rm min}$ of this sphere is equal to the Euclidean distance of the hypersurface from the origin of the Δ_{ij}-coordinates. The associated contact point between $\mathcal{D}_{\rm min}$-sphere and hypersurface represents the predicted *in-vivo* point, and its coordinate values $\Delta_{ij,\min}$ lead to the predicted *in-vivo* rates $\omega^{\star}_{ij,\min}\equiv\Omega_{ij}\exp[-\Delta_{ij,\min}]$ based on the known *in-vitro* rates ω_{ij}.

For a general, nonlinear *in-vivo* constraint, the coordinate values $\Delta_{ij,\min}$ of the predicted *in-vivo* point will be different for different individual transitions. The minimization procedure then predicts different scale factors $\omega^{\star}_{ij,\min}/\omega_{ij}$ and, thus, different effects of the *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 Δ_{ij}. Therefore, the different scale factors $\omega^{\star}_{ij,\min}/\omega_{ij}$ 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 ω_{ij} of the *in-vitro* assay are only known with a certain accuracy. As a consequence, the predicted *in-vivo* rates $\omega^{\star}_{ij,\min}$ 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 Δ_{ij}-coordinates may differ from the distance of this surface from the origin of the Δ'_{ij}-coordinates. Therefore, the validity of the predicted *in-vivo* rates $\omega^{\star}_{ij,\min}$ 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 $\omega^{\star}_{ij,\min}$ 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-vivo* values $\omega^{\star}_{ij}$ of the individual transition rates. If we knew one of these rates, e.g., $\omega^{\star}_{kl}$, we would restrict our minimization procedure to the subspace with constant $\Delta_{kl}=\ln(\omega_{kl}/\omega^{\star}_{kl})$. As a consequence, we would not vary the coordinate Δ_{kl} during the minimization and use the constant value of this coordinate in Eq. 6 for the kinetic distance $\mathcal{D}$. On the other hand, if we knew only that the *in-vivo* rate $\omega^{\star}_{kl}$ is located within the range $\Omega_1<\omega^{\star}_{kl}<\Omega_2$, we would minimize the kinetic distance also with respect to Δ_{kl} but within the subspace defined by $\ln(\omega_{kl}/\Omega_2)<\Delta_{kl}<\ln(\omega_{kl}/\Omega_1)$. The latter procedure may lead to a boundary minimum, i.e., to a predicted *in-vivo* 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 $\omega_{kl}=\omega_{k'l'}$ and $\omega^{\star}_{kl}=\omega^{\star}_{k'l'}$. We will then reduce the multi-dimensional barrier space to the subspace with $\Delta_{kl}=\Delta_{k'l'}$, and the corresponding expression for the kinetic distance $\mathcal{D}$ in Eq. 6 will now contain the term $\Delta_{kl}+\Delta_{k'l'}=2\Delta_{kl}$ 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 $X_{c,\mathrm{co}}$, $X_{c,\mathrm{nr}}$, and $X_{c,\mathrm{no}}$ of cognate, near-cognate, and non-cognate ternary complexes determine the association rates

$$\omega_{c,\mathrm{co}}=\kappa_{\mathrm{on}}X_{c,\mathrm{co}},\omega_{c,\mathrm{nr}}=\kappa_{\mathrm{on}}X_{c,\mathrm{nr}},\textit{and }\omega_{c,\mathrm{no}}=\kappa_{\mathrm{on}}X_{c,\mathrm{no}} \quad (9)$$

for initial binding with the pseudo-first-order association rate constant κ_{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 ω_{off} from the initial binding site and that the cognate and near-cognate ternary complexes have the same recognition rate ω_{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 ω_{off} or can move into the A site to attain the codon recognition state 2 with rate ω_{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 ω_{pro}.

Initial binding of a near-cognate ternary complex leads to state 6, from which the ternary complex can be released with rate ω_{off} or move to the codon recognition state 7 with rate ω_{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-

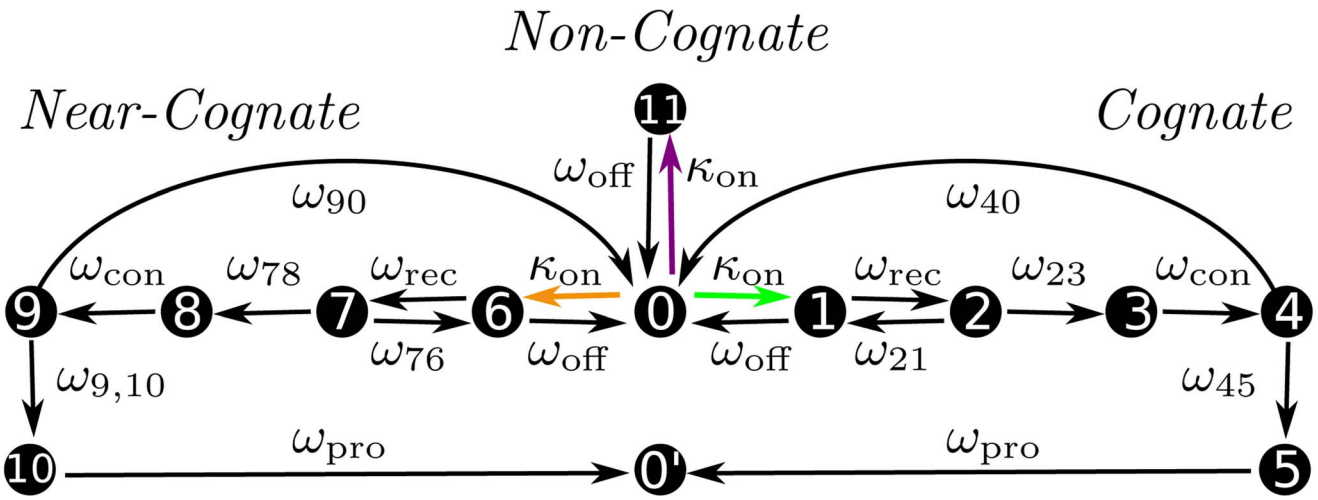

Figure 3. Codon-specific Markov process for translation elongation based on 12 ribosomal states for each codon c. The elongation cycle starts in state 0 corresponding to a ribosome without any bound ternary complex. Initial binding of a cognate, near-cognate, or non-cognate ternary complex is indicated 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 κ_{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 ω_{off}. Likewise, cognate and near-cognate ternary complexes are governed by the same recognition rate ω_{rec}, conformational rate ω_{con}, and processing rate ω_{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

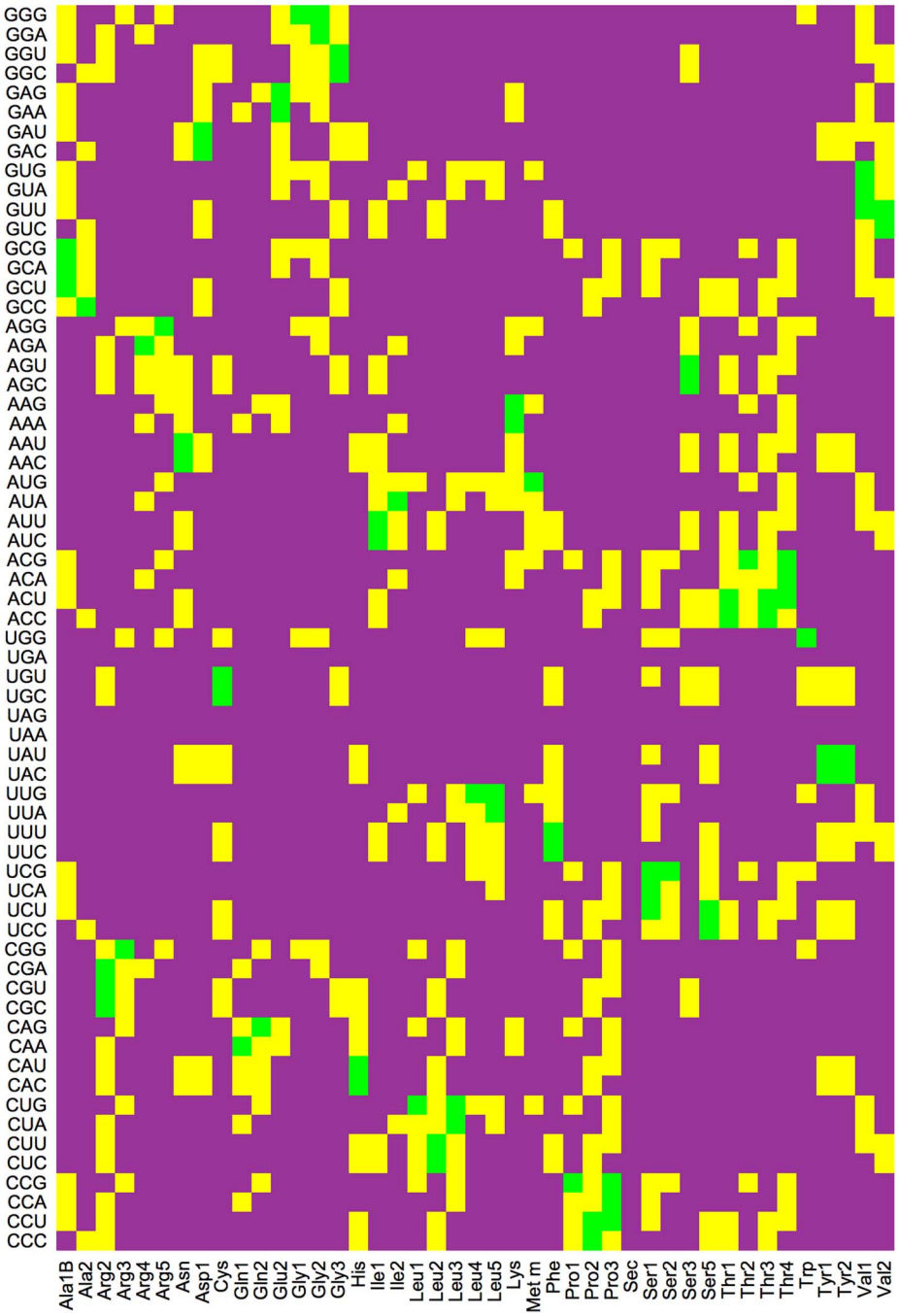

Figure 4. Decoding pattern arising from the cognate (green), near-cognate (yellow), and non-cognate (purple) relationships between all 61 sense codons and the 43 elongator tRNA species of *E. coli* as identified in Ref. [28]. For each tRNA species, the near-cognate codons differ from the cognate ones by a mismatch in one position of the codon-anticodon complex.
doi:10.1371/journal.pcbi.1003909.g004

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 $\omega_{\rm pro}$.

Apart from the association rate constant $\kappa_{\rm on}$, the kinetics of the elongation cycle then involves 12 different transition rates ω_{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 $\omega_{\rm elo}$ *in vitro* for a model protein, $\omega_{\rm elo} \simeq 0.8$ aa/s for 20°C and $\omega_{\rm elo} \simeq 6.9$ aa/s for 37°C

Table 1. *In-vitro* rates of ribosomal transitions.

Rates	*k*-not.	20°C	37°C	Units
κ_{on}	k_1	140 ± 20	175 ± 25	$\frac{1}{\mu M s}$
ω_{off}	k_{-1}	85 ± 25	700 ± 270	1/s
ω_{rec}	k_2	180 ± 30	1500 ± 450	1/s
ω_{21}	$k_{-2,co}$	0.2 ± 0.03	2 ± 0.6	1/s
ω_{23}	$k_{3,co}$	190 ± 30	1500 ± 450	1/s
ω_{con}	k_4	50	450	1/s
ω_{45}	$k_{5,co}$	22 ± 4	200 ± 40	1/s
ω_{40}	$k_{7,co}$	0.1	1	1/s
ω_{76}	$k_{-2,nr}$	140 ± 20	1100 ± 330	1/s
ω_{78}	$k_{3,nr}$	0.6 ± 0.1	7 ± 2	1/s
$\omega_{9,10}$	$k_{5,nr}$	0.06 ± 0.006	0.26 ± 0.04	1/s
ω_{90}	$k_{7,nr}$	0.84 ± 0.08	4 ± 0.7	1/s
ω_{pro}		3 ± 1	150 ± 50	1/s
ω_{elo}		0.8 ± 0.2	6.9 ± 2.3	aa/s

Apart from the processing rate ω_{pro}, all individual rates and the overall elongation rate ω_{elo} have been measured *in vitro* at 20°C and/or 37°C. The processing rate ω_{pro} was calculated from the overall elongation rate ω_{elo} via Eq. 22 in the *Methods* section. The column 'k-not.' provides the notation for the transition rates as used in Ref. [6].
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 ω_{elo} was then used to compute, for both temperatures, the *in-vitro* value ω_{pro} of the processing rate. The results of these computations are included in Table 1.

To predict the unknown *in-vivo* rates $\omega^\star_{ij}$ from the known *in-vitro* rates ω_{ij}, we consider the multi-dimensional space of single barrier shifts as described by the coordinates $\Delta_{ij} = -\ln(\omega^\star_{ij}/\omega_{ij})$. 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 $\Delta_{10} = \Delta_{60} = \Delta_{11,0} \equiv \Delta_{off}$, $\Delta_{12} = \Delta_{67} \equiv \Delta_{rec}$, $\Delta_{34} = \Delta_{89} \equiv \Delta_{con}$, and $\Delta_{50'} = \Delta_{10,0'} \equiv \Delta_{pro}$ 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

$$\mathcal{D} = \sqrt{3\Delta^2_{off} + 2\Delta^2_{rec} + 2\Delta^2_{con} + 2\Delta^2_{pro} + \sum \Delta^2_{kl}} \qquad (10)$$

where the sum $\sum \Delta^2_{kl}$ contains all the remaining transition rates of the Markov process in Fig. 3.

Because the *in-vivo* experiments are typically performed at 37°C, we use the *in-vitro* values ω_{ij} for the same temperature, see Table 1. Furthermore, we take into account the known *in-vivo* values of the overall elongation rate $\omega^\star_{elo}$ 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-vivo* rates $\omega^\star_{ij}$ as given in Table 2 and the single barrier shifts Δ_{ij} displayed in Fig. 5A, where we have again omitted the subscript 'min' for notational simplicity.

Validation of deduced *in-vivo* rates for translation elongation

Starting from the complete set of individual *in-vivo* rates (Table 2), we computed the codon-specific elongation rates $\omega^\star_{c,elo}$ as described in the *Methods* section (Eq. 21, Supporting Figure S3). We then compared the *in-vivo* rates $\omega^\star_{c,elo}$ calculated for a growth rate of 2.5 dbl/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 $tRNA^{Arg2}$, 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 ICG. 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 CCC 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 $\omega^\star_{ij}$ of the individual transition rates *in vivo* lead to a reliable description for the majority of codons.

To further validate these deduced values, we used the computed values $\omega^\star_{c,elo}$ 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 dbl/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

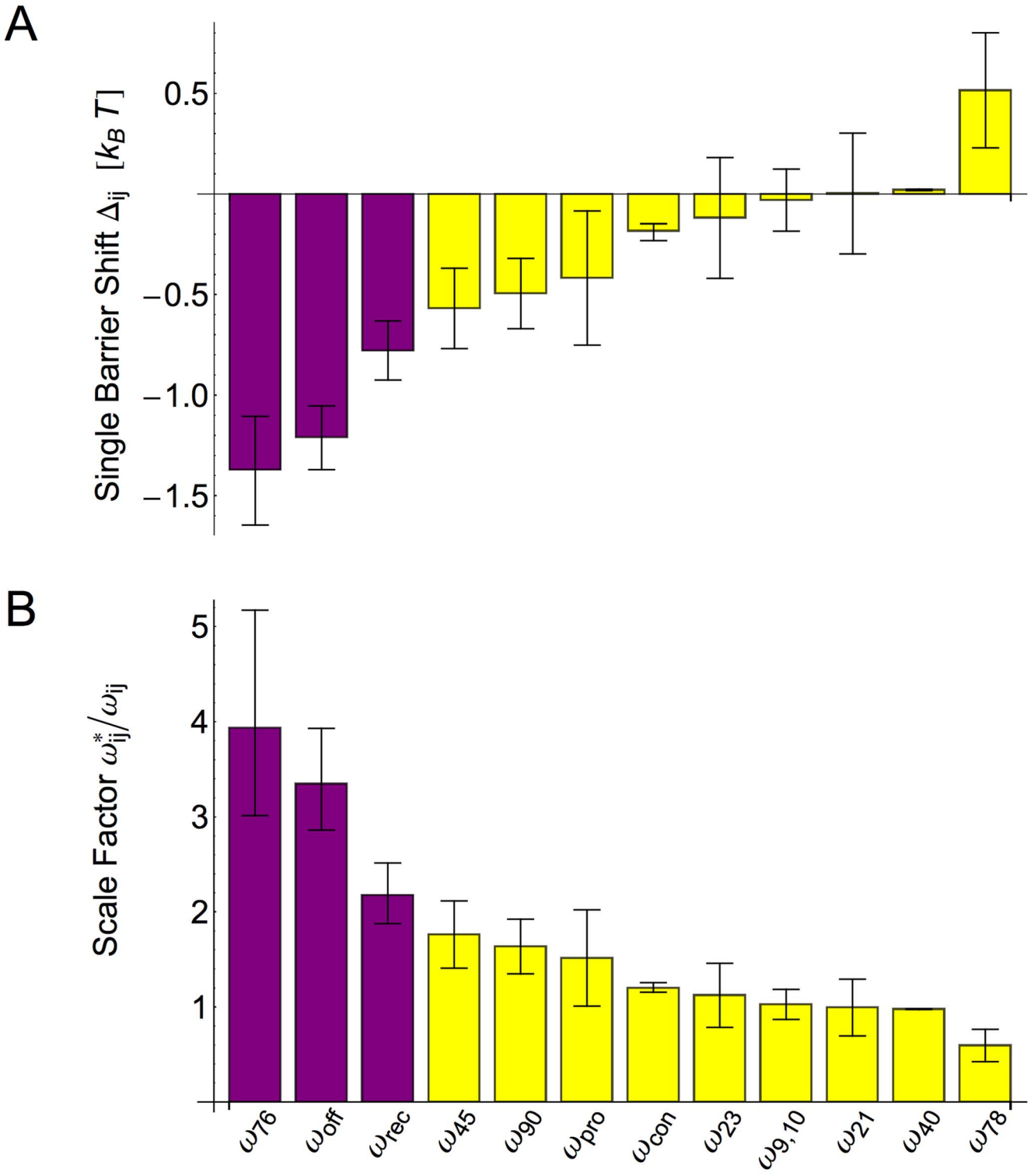

Figure 5. Quantitative comparison between the *in-vitro* kinetics of translation elongation at 37°C and the *in-vivo* kinetics as deduced for the growth condition of 2.5 dbl/h. (A) Single barrier shifts Δ_{ij} for the individual transition rates, see Eq. 2 and Fig. 2; and (B) Scale factors $\omega^*_{ij}/\omega_{ij}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

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).

Another quantity that can be used to validate the deduced *in-vivo* rates ω^*_{ij} is the missense error frequency arising from the accommodation of near-cognate ternary complexes with incorrect

Table 2. *In-vivo* rates of ribosomal transitions.

Rates	*E. coli* growth rates [dbl/h]				RSD	Units
	0.7	1.07	1.6	2.5		
$\kappa^{\star}_{\rm on}$	94	94	94	94	0.1	$\frac{1}{\mu {\rm M} s}$
$\omega^{\star}_{\rm off}$	1400	1700	2100	2300	0.4	1/s
$\omega^{\star}_{\rm rec}$	2100	2500	3000	3300	0.3	1/s
$\omega^{\star}_{21}$	2	2	2	2	0.3	1/s
$\omega^{\star}_{23}$	1600	1600	1700	1700	0.3	1/s
$\omega^{\star}_{\rm con}$	490	500	530	540		1/s
$\omega^{\star}_{45}$	270	300	340	350	0.2	1/s
$\omega^{\star}_{40}$	1	1	1	1		1/s
$\omega^{\star}_{76}$	2700	3100	3900	4300	0.3	1/s
$\omega^{\star}_{78}$	5	5	4	4	0.3	1/s
$\omega^{\star}_{9,10}$	0.27	0.27	0.27	0.27	0.2	1/s
$\omega^{\star}_{90}$	6	6	6	7	0.2	1/s
$\omega^{\star}_{\rm pro}$	190	200	230	230	0.5	1/s
$\omega^{\star}_{\rm elo}$	15	18	22	22		aa/s

The values of the overall elongation rate $\omega^{\star}_{\rm elo}$ 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 *in-vitro* rates in Table 1, and from the associated changes in the predicted in-vivo point, see *Methods* section.

doi:10.1371/journal.pcbi.1003909.t002

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 *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 3×10^{-4} for $\rm tRNA^{Lys}$ misreading codons, in good agreement with the measured value 2×10^{-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 *in-vitro* and *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 *in-vivo* values of the individual transition rates from their known *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 $\omega^{\star}_{ij}$ to the single barrier shifts Δ_{ij} and the constraint arising from the global *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 *in-vivo* rates $\omega^{\star}_{ij}$ that could be validated by three independent sets of data for codon-dependent translation speeds, codon-specific translation dynamics and missense 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 *in-vivo* rates $\omega^{\star}_{ij}$ are similar to the measured *in-vitro* rates ω_{ij} (Fig. 5B) but three *in-vivo* rates are significantly increased compared to their *in-vitro* values: the rejection rate ω_{76} for near-cognates, the dissociation rate $\omega_{\rm off}$ after initial binding, and the recognition rate $\omega_{\rm rec}$ for cognate and near-cognate ternary complexes. The largest difference is found for the rejection rate ω_{76}, which is increased *in vivo* by a factor of 3.9, while the dissociation rate $\omega_{\rm off}$ and the recognition rate $\omega_{\rm rec}$ 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 *in-vivo* and *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 $2k_{\rm B}T$. 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].

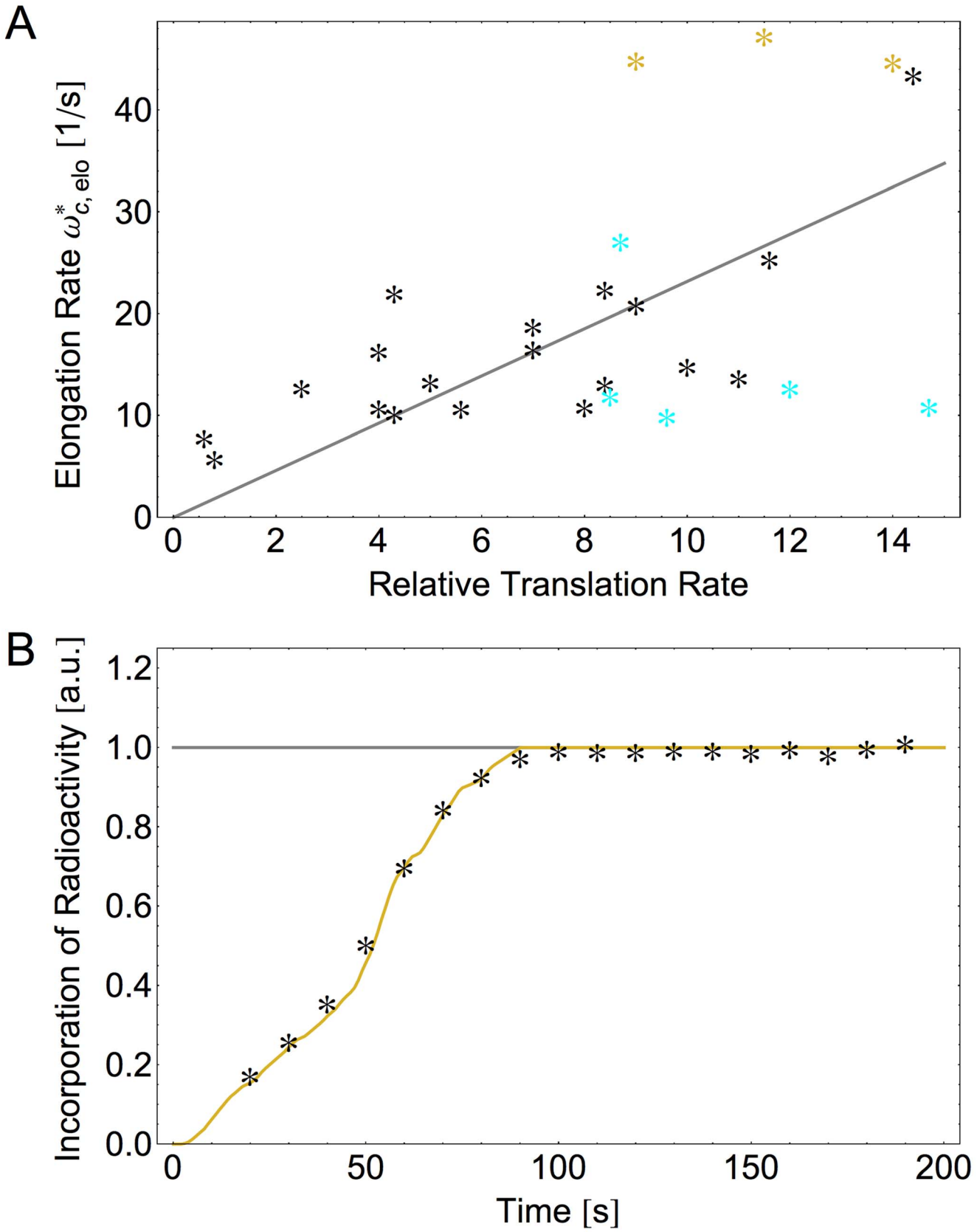

Figure 6. Comparison with *in-vivo* experiments of translation elongation. (A) Codon-specific elongation rates $\omega^{\star}_{c,\mathrm{elo}}$ as determined here from the complete set of individual transition rates for *E. coli* at a growth rate of 2.5 dbl/h, see Eq. 21 and Supporting Figure S3, compared 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

excluded (linear fit in gray). (B) For the incorporation of radioactively labeled amino acids as a function of time, we find very good agreement between the experimental data in Ref. [30] and the calculated curve (orange) based on the *in-vivo* rates ω_{ij}^{*} for 0.7 dbl/h in Table 2. For both (A) and (B), our computations do not involve any fit parameter.
doi:10.1371/journal.pcbi.1003909.g006

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 $(\omega_{23}/\omega_{21})(\omega_{76}/\omega_{78})$ and $(\omega_{45}/\omega_{40})(\omega_{90}/\omega_{9,10})$, respectively. The *in-vivo* value of the initial selection coefficient is increased by a factor of 7.7 compared to the corresponding *in-vitro* value whereas the proofreading coefficient is increased by a factor of 2.9. The combination of improved initial selection and proofreading leads to a reduction of the *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 ω_{12}, ω_{21}, …of the cognate branch represent average values, obtained by averaging over all cognate tRNAs of all codons, and likewise for the internal rates ω_{67}, ω_{76}, …of the near-cognate branch. *In vitro*, the decoding rates of different cognate codons were observed to be rather similar [14,25] whereas the GTPase activation rate ω_{78} was found to vary between 0.06/s and 1.3/s for different near-cognate codons of $tRNA^{Phe}$ [25]. Likewise, recent *in-vivo* experiments provided evidence that the error frequency on 4 out of 14 near-cognate codons of $tRNA^{Gly3}$ is much higher than on 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 *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 rRNA mutagenesis, affect the speed and accuracy of translation elongation.

The computational method developed here to deduce the *in-vivo* from the *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 *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 ω_{ij}. These expressions can be decomposed into four different dwell times according to

$$t_{c,\mathrm{acc}} = t_{c,0} + t_{c,\mathrm{no}} + t_{c,\mathrm{co}} + t_{c,\mathrm{nr}}\,, \tag{11}$$

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 $t_{c,0}$ 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 $t_{c,\mathrm{no}}$ 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 $t_{c,\mathrm{co}}$ 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 $t_{c,\mathrm{nr}}$ 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

$$\pi_{ij} \equiv \frac{\omega_{ij}}{\sum_k \omega_{ik}}\,. \tag{12}$$

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

$$t_{c,0} = \frac{1}{\kappa_{\mathrm{on}}\left[X_{c,\mathrm{co}}\,\rho_{\mathrm{co}} + X_{c,\mathrm{nr}}\,\rho_{\mathrm{nr}}\right]} \tag{13}$$

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

dimensionless, concentration-independent ratios

$$\rho_{\rm co} \equiv \frac{\pi_{12}\,\pi_{23}\,\pi_{45}}{1-\pi_{12}\,\pi_{21}} \tag{14}$$

and

$$\rho_{\rm nr} \equiv \frac{\pi_{67}\,\pi_{78}\,\pi_{9,10}}{1-\pi_{67}\,\pi_{76}}\,. \tag{15}$$

The second dwell time $t_{c,\rm no}$ for state 11 with a bound non-cognate ternary complex is given by the expression

$$t_{c,\rm no} = t_{c,0}\,\kappa_{\rm on}\,X_{c,\rm no}/\omega_{\rm off}\,, \tag{16}$$

and is, thus, proportional to the concentration $X_{c,\rm no}$ of free non-cognate ternary complexes.

The third dwell time $t_{c,\rm co}$, 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

$$t_{c,\rm co} = \frac{X_{c,\rm co}\,\rho_{\rm co}}{X_{c,\rm co}\,\rho_{\rm co}+X_{c,\rm nr}\,\rho_{\rm nr}}\,\tau_{\rm co} \tag{17}$$

with the concentration-independent time scale

$$\tau_{\rm co} \equiv \frac{1}{\omega_{12}\,\pi_{23}\,\pi_{45}}+\frac{1}{\omega_{23}\,\pi_{45}}+\frac{1}{\omega_{34}\,\pi_{45}}+\frac{1}{\omega_{45}} \tag{18}$$

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

$$t_{c,\rm nr} = \frac{X_{c,\rm nr}\,\rho_{\rm nr}}{X_{c,\rm co}\,\rho_{\rm co}+X_{c,\rm nr}\,\rho_{\rm nr}}\,\tau_{\rm nr} \tag{19}$$

with the concentration-independent time scale

$$\tau_{\rm nr} \equiv \frac{1}{\omega_{67}\,\pi_{78}\,\pi_{9,10}}+\frac{1}{\omega_{78}\,\pi_{9,10}}+\frac{1}{\omega_{89}\,\pi_{9,10}}+\frac{1}{\omega_{9,10}} \tag{20}$$

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

$$\omega_{c,\rm elo} \equiv 1/t_{c,\rm elo} = 1/[t_{c,\rm acc}+1/\omega_{\rm pro}]\,. \tag{21}$$

One important global property of protein synthesis is the average speed of the ribosomes, which defines the overall elongation rate $\omega_{\rm elo}$. The inverse of the overall elongation rate is equal to the average elongation time $\langle t_{\rm elo}\rangle \equiv \sum_c p_c t_{c,\rm elo} = \sum_c p_c t_{c,\rm acc} + 1/\omega_{\rm pro}$, which is obtained by averaging the codon-specific elongation times $t_{c,\rm elo}$ 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 $\omega_{c,\rm elo} = 1/\langle t_{\rm elo}\rangle$ and the codon-specific accommodation times $t_{c,\rm acc}$ was rewritten in the form

$$\omega_{\rm pro} = \left[\frac{1}{\omega_{\rm elo}} - \sum_c p_c t_{c,\rm acc}\right]^{-1} \tag{22}$$

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

In vivo, the overall elongation rate $\omega^\star_{\rm elo}$ is given by the analogous expression

$$\omega^\star_{\rm elo} = \left[\sum_c p_c t^\star_{c,\rm acc} + \frac{1}{\omega^\star_{\rm pro}}\right]^{-1}, \tag{23}$$

where the codon-specific accommodation times $t^\star_{c,\rm acc}$ follow from the same expression as in Eq. 11 but with the *in-vitro* values ω_{ij} replaced by the *in-vivo* values $\omega^\star_{ij}$. When we insert the known *in-vivo* value $\omega^\star_{\rm elo}$ of the overall elongation rate into Eq. 23, we obtain a constraint on the (unknown) *in-vivo* values $\omega^\star_{ij}$ of the individual transition rates. This constraint can be expressed in terms of the single barrier shifts Δ_{ij} when we replace $\omega^\star_{ij}$ in Eq. 23 by $\omega_{ij}\exp[-\Delta_{ij}]$, 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 $\omega_{\rm elo}$, both at 20°C and at 37°C, see Supporting Figure S1, as well as the individual rates $\omega_{90} \equiv k_{5,\rm nr}$ and $\omega_{9,10} \equiv k_{7,\rm nr}$ at 20°C, see Supporting Figure S2, using the experimental protocols described previously [34,46,47].

The *in-vitro* value $\kappa_{\rm on}$ of the association rate 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 $\omega_{\rm off}$ 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 $\rm tRNA^{Phe}$. 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 $\omega_{\rm rec}$, ω_{21}, ω_{23}, ω_{76}, and ω_{78}. The rate $\omega_{\rm con}$ has not been measured but estimated under the assumption that it is not rate-limiting. The rate $\omega_{9,10}$ at 37°C was reported previously and was used to determine the rate $\omega_{90} = \omega_{9,10}(1-0.06)/0.06$, i.e., using an error frequency of 0.06 for the proofreading step [34]. The rate ω_{45} has been measured both for 20°C and for 37°C [25,34]. The rate $\omega_{\rm pro}$ was calculated for both temperatures from the measured values of the overall elongation rate $\omega_{\rm elo}$ via Eq. 22.

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 ΔG_{ij} into the activation enthalpy ΔH_{ij} and the activation entropy ΔS_{ij}, which leads to

$$\begin{aligned}\omega_{ij}(T) &= \nu_{ij}\,\exp[\Delta S_{ij}/k_{\mathrm{B}}]\,\exp[-\Delta H_{ij}/k_{\mathrm{B}}T] \\ &\equiv \nu_o\,\exp[\Delta \tilde{S}_{ij}/k_{\mathrm{B}}]\,\exp[-\Delta H_{ij}/k_{\mathrm{B}}T]\end{aligned} \tag{24}$$

where the last expression involves the attempt frequency $\nu_o = k_{\mathrm{B}}T/h$ as obtained from transition-state theory [2]. In this way, any state-dependence of the attempt frequency ν_{ij} has been absorbed into the activation entropies $\Delta\tilde{S}_{ij} = \Delta S_{ij} + k_{\mathrm{B}}\ln(\nu_{ij}/\nu_o)$. If one plots the logarithms $\ln(\omega_{ij}/\nu_o)$ of the measured rates ω_{ij} as a function of the inverse temperature $1/T$ (conventional Arrhenius plots), one finds linear relationships [10,13,48,49], which imply that the two unknown parameters in Eq. 24, ΔH_{ij} and $\Delta\tilde{S}_{ij}$, do not depend on temperature over the experimentally studied temperature range. However, the activation entropies $\Delta\tilde{S}_{ij}$ as obtained from the behavior of $\ln(\omega_{ij}/\nu_o)$ for small $1/T$ 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 ΔH_{ij} and entropies $\Delta\tilde{S}_{ij}$, we now consider the ratios $\omega_{ij}(T)/\omega_{45}(T)$ at the two temperatures of interest, $T_1 = 20^{\circ}\mathrm{C} = 293.15\,\mathrm{K}$ and $T_2 = 37^{\circ}\mathrm{C} = 310.15\,\mathrm{K}$. We take the accommodation rate ω_{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 ΔH_{ij}. As a result, we obtain the relation

$$\frac{\omega_{ij}(T_2)}{\omega_{45}(T_2)} = \left(\frac{\omega_{ij}(T_1)}{\omega_{45}(T_1)}\right)^{T_1/T_2} \exp\left[\frac{\Delta\tilde{S}_{ij} - \Delta\tilde{S}_{45}}{k_{\mathrm{B}}}\,\frac{T_2 - T_1}{T_2}\right]. \tag{25}$$

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

$$\frac{\omega_{ij}(37^{\circ}\mathrm{C})}{\omega_{45}(37^{\circ}\mathrm{C})} \simeq \left(\frac{\omega_{ij}(20^{\circ}\mathrm{C})}{\omega_{45}(20^{\circ}\mathrm{C})}\right)^{293.15/310.15}. \tag{26}$$

to estimate the values of the rates ω_{off}, ω_{rec}, ω_{21}, ω_{23}, ω_{76}, ω_{78}, ω_{con}, and ω_{40} at 37°C (Table 1) from the measured values of $\omega_{ij}(20^{\circ}\mathrm{C})$, $\omega_{45}(20^{\circ}\mathrm{C})$, and $\omega_{45}(37^{\circ}\mathrm{C})$.

In-vivo values of association rates

The overall elongation rate $\omega^{\star}_{\mathrm{elo}}$ as given by Eq. 23 also depends on the association rates for initial binding, which are proportional to the pseudo-first-order rate constant $\kappa^{\star}_{\mathrm{on}}$ and to the concentrations $X^{\star}_a$ of the ternary complexes as in Eq. 9. Therefore, in order to use Eq. 23 for the process *in vivo*, we had to estimate the corresponding values $\kappa^{\star}_{\mathrm{on}}$ and the ternary complex concentrations $X^{\star}_a$ 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 $2.57\,\mu\mathrm{m}^2/\mathrm{s}$ [52,53] for a ternary complex in the cytosol, we found that the *in-vivo* value κ^{*}_{on} of the bimolecular association rate constant is about 54% of the *in-vitro* value κ_{on}, compare Table 1 and Table 2.

For the *in-vivo* concentrations $X^{\star}_a$ 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 dbl/h. In the latter study, the authors determined the concentrations $X^{\star}_a$ of all 43 elongator tRNA species a. These concentrations are then combined, for each codon c, into the concentrations $X^{\star}_{c,\mathrm{co}}$, $X^{\star}_{c,\mathrm{nr}}$, and $X^{\star}_{c,\mathrm{no}}$ 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 $X^{\star}_a$ 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 *in-vivo* rates $\omega^{\star}_{ij,\,\mathrm{min}}$, 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 ω_m, $\omega^{\star}_m$, and $\omega^{\star}_{m,\,\mathrm{min}}$ for the *in-vitro* rates ω_{ij} of a certain assay, for the unknown *in-vivo* rates $\omega^{\star}_{ij}$, and for the predicted *in-vivo* rates $\omega^{\star}_{ij,\,\mathrm{min}}$, 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 *in-vitro* rates can be described by

$$\omega_m = \bar{\omega}_m \pm \delta_m \equiv \bar{\omega}_m(1 \pm \epsilon_m) \tag{27}$$

with the absolute error δ_m and the relative error

$$\epsilon_m \equiv \delta_m/\bar{\omega}_m \tag{28}$$

of the *in-vitro* rate ω_m. Both the average values $\bar{\omega}_m$ and the absolute errors δ_m are estimated from the experimental data for the *in-vitro* assay under consideration.

When we apply the minimization procedure to the average values $\bar{\omega}_m$ of the *in-vitro* rates, we use the coordinates

$$\bar{\Delta}_m \equiv \ln(\bar{\omega}_m/\omega^{\star}_m) \tag{29}$$

for the multi-dimensional barrier space. We then determine the point $\overline{H}$ 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 $\bar{\Delta}_m$-coordinates. The coordinate values $\bar{\Delta}_{m,\,\mathrm{min}}$ of the point $\overline{H}$ then lead to the predicted values $\bar{\omega}^{\star}_{m,\,\mathrm{min}} = \bar{\omega}_m \exp[-\bar{\Delta}_{m,\,\mathrm{min}}]$ for the *in-vivo* rates.

In order to estimate the uncertainty of these predictions, it is useful to consider an auxiliary ensemble of fictitious *in-vitro* assays that is constructed 'around' the given assay as follows. For each transition ω_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

$$\underline{s}^C = (s^C_1, s^C_2, \ldots, s^C_M). \tag{30}$$

Each of these configurations defines a fictitious *in-vitro* assay, again denoted by C, with transition rates

$$\omega_m^C \equiv \bar{\omega}_m + s_m^C \delta_m = \bar{\omega}_m (1 + s_m^C \epsilon_m). \tag{31}$$

The rates of assay C define the coordinates

$$\Delta_m^C \equiv \ln(\omega_m^C/\omega_m^\star) = \bar{\Delta}_m + d_m^C \tag{32}$$

for the multi-dimensional barrier space with

$$d_m^C \equiv \ln(1 + s_m^C \epsilon_m) \approx s_m^C \epsilon_m \tag{33}$$

where the asymptotic equality applies to the limit of small relative errors $\epsilon_m = \delta_m/\bar{\omega}_m$ of the *in-vitro* rates. Therefore, if the origin of the multi-dimensional barrier space is defined by the coordinates $\bar{\Delta}_m$ in Eq. 29, corresponding to the average *in-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 $\Delta_{m,\min}^C$.

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

$$\Delta_{m,\min} = \bar{\Delta}_{m,\min} \pm \epsilon_m^\star \tag{34}$$

where the values $\bar{\Delta}_{m,\min}$ correspond to the average *in-vitro* rates $\bar{\omega}_m$. The predicted *in-vivo* rates are now given by

$$\begin{aligned} \omega_{m,\min}^\star &= \omega_m \exp[-\Delta_{m,\min}] = \bar{\omega}_{m,\min}^\star (1\epsilon_m) \\ \exp[\pm\epsilon_m^\star] &\approx \bar{\omega}_{m,\min}^\star (1 \pm \epsilon_m \pm \epsilon_m^\star). \end{aligned} \tag{35}$$

Therefore the relative error of the predicted *in-vivo* rates $\omega_{m,\min}^\star$ reflects both the relative error ϵ_m of the *in-vitro* rates ω_m (Eq. 27) and the absolute error $\epsilon_m^\star$ of the coordinate values $\Delta_{m,\min}$ for the predicted *in-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 $\bar{\Delta}_{ij,\min}$ of the *in-vivo* point as predicted from the average value $\bar{\omega}_{ij}$ of the *in-vitro* rates. The largest coordinate values $\bar{\Delta}_{ij,\min}$ of the predicted *in-vivo* point were found for the three transition rates $\bar{\omega}_{76}$, $\bar{\omega}_{\rm off}$, and $\bar{\omega}_{\rm rec}$ (Fig. 5B). We then focused on the errors of these three *in-vitro* rates, which define 8 corners of the 'error polyhedron' around the origin of the $\bar{\Delta}_{ij}$-coordinates. For each of these corners C, we determined the closest point on the hypersurface and the coordinate values $\Delta_{ij,min}^C$ of this point. We then estimated the absolute error $\epsilon_m^\star$ of the coordinate values $\Delta_{m,\min}$ from the largest and smallest values of $\Delta_{ij,min}^C$ as obtained for different corners C. The errors $\epsilon_m^\star$ were finally used, together with the relative errors ϵ_m of the measured *in-vitro* rates, to determine the relative standard deviations (RSDs) of the predicted *in-vivo* 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 $\mathcal{P}(a \to c|\mathrm{nr})$ that a is fully accommodated at c. For the multistep process considered here, this probability is given by

$$\mathcal{P}(a \to c|\mathrm{nr}) = \frac{X_a \rho_{\rm nr}}{X_{c,\rm co}\,\rho_{\rm co} + X_{c,\rm nr}\,\rho_{\rm nr}}, \tag{36}$$

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

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

$$\langle \mathcal{P}(a \to c|\mathrm{nr}) \rangle \equiv \sum_{c \in \mathbf{C}_{\rm nr}(a)} p_c\, \mathcal{P}(a \to c|\mathrm{nr}) \left[\sum_{c \in \mathbf{C}_{\rm nr}(a)} p_c \right]^{-1} \tag{37}$$

where the set $\mathbf{C}_{\rm nr}(a)$ 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 *in vitro*. Kinetics of CspA translation *in vitro* at different temperatures. CspA mRNA, which codes for a 70 aa-long protein from *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 fluorescence derivative of initiator $\mathrm{tRNA}^{\rm fMet}$ carrying BodipyFL at the α-amino 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–aminoacyl-tRNA (40 μM aminoacyl-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 NH_4Cl, 3.5 mM free $MgCl_2$, 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-Fl 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).
(TIF)

Figure S2 *In-vitro* rates as measured for near-cognate accommodation and rejection after proofreading. *In-vitro* values of the rates $\omega_{9,10} \equiv k_{5,\rm nr}$ and $\omega_{90} \equiv k_{7,\rm nr}$ 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 f[^{3}H]Met[^{14}C]Phe was monitored

under multiple-turnover conditions using initiation complexes 70S–mRNA(AUGCUC)–f[^{3}H]Met-tRNA^{fMet} (0.14 μM) and varying concentrations of the ternary complex EF-Tu–GTP–[^{14}C]Phe-tRNA^{Phe}, 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_{app}, we calculated $\omega_{9,10} = 0.060 \pm 0.006/s$ and $K_M = 2.4\,\mu M$. Using the previously measured efficiency $\omega_{9,10}/(\omega_{9,10}+\omega_{90}) = 1/15$ of the proofreading step [25], we then obtained the value $\omega_{90} = 0.84 \pm 0.08/s$ for near-cognate rejection after proofreading.
(TIF)

Figure S3 Codon-specific elongation rates *in vitro* and *in vivo*. Codon-specific elongation rates $\omega_{c,elo}$ 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 $\omega_{c,elo}$ for the high-fidelity buffer at 37°C, obtained from the individual rates in Table 1; (B, C) *In-vivo* values $\omega^{\star}_{c,elo}$ for *E. coli* at growth conditions of (B) 0.7 dbl/h and (C) 2.5 dbl/h, calculated from the individual rates in Table 2.
(TIF)

Figure S4 Incorporation of radioactively labeled amino acids for different dissociation rates. Experimental data (black stars) for the incorporation of radioactively labeled amino acids at a growth rate of 0.7 dbl/h [30] and simulation curves obtained for five different values of the initial dissociation rate ω_{off}. The orange simulation curve in the middle corresponds to $\omega_{off} = \omega^{\star}_{off} = 1400/s$, see Table 2. This value has been obtained from the minimization of the kinetic distance and provides an excellent fit to the data. The red, blue, green, and black curves have been obtained for simulations with $\omega_{off} = 2\,\omega^{\star}_{off}$, $1.2\,\omega^{\star}_{off}$, $0.8\,\omega^{\star}_{off}$, and $0.5\,\omega^{\star}_{off}$, respectively. Thus, changing the value of ω_{off} by 20% leads to a significant deviation of the simulation curve from the experimental data.
(TIF)

Acknowledgments

We thank Angelo Valleriani and Zoya Ignatova for stimulating discussions during the initial phase of this project.

Author Contributions

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

References

1. Michaelis L, Menten ML (1913) Die Kinetik der Invertinwirkung. Biochem Z 49: 333–369.
2. Cornish-Bowden A (1995) Fundamentals of Enzyme Kinetics. London: Portland Press.
3. Sauro HM (2012) Enzyme Kinetics for Systems Biology. Ambrosius Publishing and Future Skill Software.
4. Hill T (1977) Free Energy Transduction in Biology. New York: Academic Press, 1 edition.
5. Liepelt S, Lipowsky R (2007) Kinesin's network of chemomechanical motor cycles. Phys Rev Lett 98: 258102.
6. Wohlgemuth I, Pohl C, Mittelstaet J, Konevega AL, Rodnina MV (2011) Evolutionary optimization of speed and accuracy of decoding on the ribosome. Phil Trans R Soc B 366: 2979–2986.
7. Voorhees RM, Ramakrishnan V (2013) Structural basis of the translational elongation cycle. Annu Rev Biochem 82: 203–236.
8. Moazed D, Noller HF (1989) Intermediate states in the movement of transfer RNA in the ribosome. Nature 342: 142–148.
9. Dix D, Thompson RC (1989) Codon choice and gene expression - Synonymous codons differ in translational accuracy. Proc Natl Acad Sci 86: 6888–6892.
10. Rodnina MV, Pape T, Fricke R, Kuhn L, Wintermeyer W (1996) Initial Binding of the Elongation Factor Tu GTP Aminoacyl-tRNA Complex Preceding Codon Recognition on the Ribosome. J Biol Chem 271: 646–652.
11. Pape T, Wintermeyer W, Rodnina MV (1998) Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the E. coli ribosome. EMBO J 17: 7490–7497.
12. Gromadski KB, Rodnina MV (2004) Kinetic Determinants of High-Fidelity tRNA Discrimination on the Ribosome. Mol Cell 13: 191–200.
13. Johansson M, Bouakaz E, Lovmar M, Ehrenberg M (2008) The Kinetics of Ribosomal Peptidyl Transfer Revisited. Mol Cell 30: 589–598.
14. Ledoux S, Uhlenbeck OC (2008) Different aa-tRNAs Are Selected Uniformly on the Ribosome. Mol Cell 31: 114–123.
15. Zaher HS, Green R (2010) Hyperaccurate and Error-Prone Ribosomes Exploit Distinct Mechanisms during tRNA Selection. Mol Cell 39: 110–120.
16. Stark H, Rodnina MV, Wieden HJ, Zemlin F, Wintermeyer W, et al. (2002) Ribosome interactions of aminoacyl-tRNA and elongation factor Tu in the codon-recognition complex. Nat Struct Biol 9: 849–854.
17. Li W, Agirrezabala X, Lei J, Bouakaz L, Brunelle JL, et al. (2008) Recognition of aminoacyl-tRNA: a common molecular mechanism revealed by cryo-EM. EMBO J 27: 3322–3331.
18. Schuette JC, IV FVM, Kelley AC, Weir JR, Giesebrecht J, et al. (2009) GTPase activation of elongation factor EF-Tu by the ribosome during encoding. EMBO J 28: 755–765.
19. Agirrezabala X, Schreiner E, Trabuco LG, Lei J, Ortiz-Meoz RF, et al. (2011) Structural insights into cognate versus near-cognate discrimination during decoding. EMBO J 30: 1497–1507.
20. Ogle JM, Brodersen DE, Jr WMC, Tarry MJ, Carter AP, et al. (2001) Recognition of Cognate Transfer RNA by the 30S Ribosomal Subunit. Science 292: 897–902.
21. Schmeing TM, Voorhees RM, Kelley AC, Gao YG, IV FVM, et al. (2009) The Crystal Structure of the Ribosome Bound to EF-Tu and Aminoacyl-tRNA. Science 326: 688–694.
22. Voorhees RM, Schmeing TM, Kelley AC, Ramakrishnan V (2010) The Mechanism of Activation of GTP Hydrolysis on the Ribosome. Science 330: 835–838.
23. Rosenblum G, Chen C, Kaur J, Cui X, Zhang H, et al. (2013) Quantifying Elongation Rhythm during Full-Length Protein Synthesis. J Am Chem Soc 135: 11322–11329.
24. Johansson M, Chen J, Tsai A, Kornberg G, Puglisi JD (2014) Sequence-Dependent Elongation Dynamics on Macrolide-Bound Ribosomes. Cell Reports 7: 1534–1546.
25. Gromadski KB, Daviter T, Rodnina MV (2006) A Uniform Response to Mismatches in Codon-Anticodon Complexes Ensures Ribosomal Fidelity. Mol Cell 21: 369–377.
26. Bremer H, Dennis PP (1996) Modulation of Chemical Composition and Other Parameters of the Cell by Growth Rate, Washington: ASM Press, volume 2. 2nd edition, pp.1553–1569.
27. Liang S, Xu YC, Dennis PP, Bremer H (2000) mRNA composition and control of bacterial gene expression. J Bacteriol 182: 3037–3044.
28. Dong H, Nilsson L, Kurland CG (1996) Co-variation of tRNA Abundance and Codon Usage in Escherichia coli at Different Growth Rates. J Mol Biol 260: 649–663.
29. Curran JF, Yarus M (1989) Rates of Aminoacyl-Trans-RNA Selection at 29 Sense Codons In vivo. J Mol Biol 209: 65–77.
30. Sørensen MA, Pedersen S (1991) Absolute in Vivo Translation Rates of Individual Codons in Escherichia coli: The Two Glutamic Acid Codons GAA and GAG Are Translated with a Threefold Difference in Rate. J Mol Biol 222: 265–280.
31. Kramer EB, Farabaugh PJ (2007) The frequency of translational misreading errors in E. coli is largely determined by tRNA competition. RNA 13: 87–96.
32. Agris PF, Vendeix FAP, Graham WD (2007) tRNA's Wobble Decoding of the Genome: 40 Years of Modification. J Mol Biol 366: 1–13.
33. Grosjean H, de Crecy-Lagard V, Marck C (2010) Decipering synonymous codons in the three domains of life: Co-evolution with specific tRNA modification enzymes. FEBS Lett 584: 252–264.
34. Wohlgemuth I, Pohl C, Rodnina MV (2010) Optimization of speed and accuracy of decoding in translation. EMBO J 29: 3701–3709.
35. Kacser H, Burns J (1973) The Control of Flux. Symp Soc Exp Biol 32: 65–104.
36. Heinrich R, Rapoport T (1974) A linear steady-state treatment of enzymatic chains. Eur J Biochem 42: 89–95.
37. Fell DA (1992) Metabolic Control Analysis: a survey of its theoretical and experimental development. Biochem J 286: 313–330.
38. Zouridis H, Hatzimanikatis V (2007) A Model for Protein Translation: Polysome Self-Organization Leads to Maximum Protein Synthesis Rates. Biophys J 92: 717–730.
39. Johansson M, Lovmar M, Ehrenberg M (2008) Rate and accuracy of bacterial protein synthesis revisited. Curr Opin Microbiol 11: 141–147.

40. Whitford PC, Blanchard SC, Cate JHD, Sanbonmatsu KY (2013) Connecting the Kinetics and Energy Landscape of tRNA Translocation on the Ribosome. PLoS Comp Biol 9: e1003003.
41. Bock LV, Blau C, Schröder CF, Davydov II, Fischer N, et al. (2013) Energy barriers and driving forces in tRNA translocation through the ribosome. Nat Struct Mol Biol 20: 1390–1397.
42. Manickam N, Nabanita N, Patel AAK, Farabaugh PJ (2014) Studies of translational misreading in vivo show that the ribosome very efficiently discriminates against most potential errors. RNA 20: 9–15.
43. Tummler K, Lubitz T, Schelker M, Klipp E (2014) New types of experimental data shape the use of enzyme kinetics for dynamic network modeling. FEBS J 281: 549–571.
44. Norris JR (1997) Markov Chains. Cambrigde: Cambridge University Press.
45. Taylor HM, Karlin S (1998) An Introduction to Stochastic Modelling. San Diego: Academic Press, 3rd edition.
46. Mittelstaet J, Konevega AL, Rodnina MV (2013) A Kinetic Safety Gate Controlling the Delivery of Unnatural Amino Acids to the Ribosome. J Am Chem Soc 135: 17031–17038.
47. Doerfel LK, Wohlgemuth I, Kothe C, Peske F, Urlaub H, et al. (2013) EF-P Is Essential for Rapid Synthesis of Proteins Containing Consecutive Proline Residues. Science 339: 85–88.
48. Sievers A, Beringer M, Rodnina MV, Wolfenden R (2004) The ribosome as an entropy trap. PNAS 101: 7897–7901.
49. Kuhlenkoetter S, Wintermeyer W, Rodnina MV (2011) Different substrate-dependent transition states in the active site of the ribosome. Nature 476: 351–354.
50. Whitford PC, Geggier P, Altman RB, Blanchard SC, Onuchic JN, et al. (2010) Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways. RNA 16: 1196–1204.
51. Condamin S, Benichou O, Moreau M (2007) Random walks and brownian motion: A method of computation for first-passage times and related quantities in confined geometries. Phys Rev E 75: 021111.
52. Elowitz MB, Surette MG, Wolf PE, Stock JB, Leibler S (1999) Protein mobility in the cytoplasm of Escherichia coli. J Bacteriology 181: 197–203.
53. Fluitt A, Pienaar E, Viljoen H (2007) Ribosome kinetics and aa-tRNA competition determine rate and fidelity of peptide synthesis. Comput Biol Chem 31: 335–346.