Bacterial proteostasis balances energy and chaperone utilization efficiently

Mantu Santra[a], Daniel W. Farrell[a] and Ken A. Dill[a,b,1]

[a]Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, NY 11794, USA

[b]Department of Chemistry and Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA

[1]Correspondence: dill@laufercenter.org

Supporting Information

SI Text

The Model Details

The model described in this paper builds upon the FoldEco dynamical model of proteostasis (1). We modified FoldEco in several ways, as described below:

1. **Protein Synthesis, Folding/misfolding/aggregation, and growth:** The synthesis of an unfolded protein (U) in our model is represented as a single-step process (Fig. S1). This differs from the FoldEco in which the ribosome and corresponding protein translation are considered explicitly. The protein is synthesized with a rate σ. The newly synthesized protein either folds to its native state (N) or misfolds to an off pathway intermediate (M) which is the precursor of aggregates (A_2 and A). Two misfolded proteins associate to form an aggregate dimer (A_2) which is followed by the formation of larger aggregates (A) by addition of M and A_2. Explicit consideration of A_2 in the model is due to the fact that KJE binds to A_2 (Fig. S2). In FoldEco, aggregation is a series of ~50-200 elementary kinetic steps. This is computationally costly. Our simplification of aggregation increases the computational efficiency without much effect on the measurable physics, and it can successfully fit in vitro refolding data (Fig. S5). We go beyond FoldEco in treating the growth rate of the cell (λ). In steady state, the concentration of a protein is maintained because of a balance established by the synthesis of protein (σ) and its dilution by cell growth (λ). The relationship between growth and synthesis rates enables us to achieve steady-state concentrations of each protein (Eq. S9).

2. **Protein degradation by Lon:** There are multiple degradation systems in *E. coli*. Among them, ATP-dependent Lon protease is the most important. The activity of Lon starts with the binding of M and U forming protease-protein complex (Ln:M and Ln:U) (Fig. S1). The bound protein is then transferred to the proteolytic chamber via an ATP-driven process. Once it enters the proteolytic chamber, it undergoes degradation.

3. **Trigger Factor (TF):** After the synthesis of a client protein, TF binds to the client in either U or M states and prevents aggregation (Fig. S1). TF binding does not change the free energy landscape of folding/misfolding. Therefore, a TF-bound protein can fold and misfold according to its rates in the free state. This is different from the FoldEco in which TF is considered to bind to

both U and M but in which there is no folding/misfolding of a protein once it is bound to TF. Thus, FoldEco assumes a holdase activity of TF which not only prevents aggregation but also slows down the folding process. The folding/misfolding dynamics of a protein is assumed to be unperturbed by TF. So, we set the corresponding rate constants in the TF-bound state to be the same as in the free form. The current model does not consider the ribosome explicitly, or therefore translation, so it does not treat co-translational folding by TF.

4. **DnaK/DnaJ/GrpE (KJE) system:** The KJE system binds to non-native protein (U, M and A_2) and unfolds it, facilitating folding by giving the client protein a new chance to fold (Fig. S2). Both dimeric DnaJ (J_2) and ATP-bound DnaK (K_T) bind the client protein independently, followed by the binding of K_T and J_2, respectively. Thus, a ternary complex (J_2:K_T:U or J_2:K_T:M or J_2:K_T:A_2) is formed. Rapid ATP hydrolysis of the ternary complex causes a conformational change in the DnaK molecule, resulting in unfolding of the bound protein. The DnaK in the ADP-bound state exhibits high substrate affinity. After ATP hydrolysis, a GrpE dimer (E_2) binds to ADP bound DnaK-protein complex and accelerates nucleotide exchange process which reverts DnaK to its low affinity ATP bound state (K_T:U). Because of weak binding, the protein in the K_T:U complex is considered to fold and misfold according to its natural kinetics. The absence of holdase activity of ATP-bound DnaK (K_T) supports this contention (2). The model for KJE-mediated unfolding is based on that of Hu *et al.* and FoldEco (1, 3).

5. **ClpB/DnaK/DnaJ/GrpE (B+KJE)** is a disaggregation system. Small aggregates (A_2) can be disaggregated by the KJE system alone, as shown in Fig. S2, but disaggregation of big aggregates (A) necessarily needs ClpB protein in addition to KJE. There is a long-standing debate on different possible mechanisms of disaggregation. One of the proposed mechanism (mechanism 1) assumes that ClpB protein binds to the aggregate first. This step is followed by its binding to the KJE system, resulting in the extraction of unfolded protein from the aggregate (4). The second proposed mechanism (mechanism 2) considers the binding of KJE to the protein aggregate first followed by the transfer of the aggregate to ClpB which finally converts it to the unfolded state (5). These two mechanisms are mutually exclusive. FoldEco adopts the second mechanism. We explored the second mechanism (mechanism 2) here, but found that it results in a proteostasis crash in GroEL-depleted cell. We find that in the absence of GroEL, DnaK substrates aggregate; this is inconsistent with the observation that the solubility of DnaK substrates is not affected by the depletion of GroEL in *E. coli* (6). Our analysis reveals that the aggregation of the GroEL substrate due to GroEL depletion causes more KJE to be associated with their aggregates and therefore very few free KJE remain in solution to bind soluble non-native DnaK substrates (U, M and A_2). As a result, DnaK substrates also aggregate. On the other hand, mechanism 1 does not encounter such instability since the maximum amount of KJE that can be associated with aggregate (A) is the amount of ClpB protein present in the cell, and under normal conditions, the amount of ClpB is small compared to DnaK (Table S3). Therefore, even in the GroEL-depleted cell, most of KJE is available to bind DnaK substrates in their soluble non-native states (U, M and A_2) helping them to fold. We therefore choose to model the second mechanism of disaggregation by the B+KJE system (Fig. S4).

6. **GroEL/GroES (GroE) system:** GroEL encapsulates non-native proteins into its cavity and folds it in an isolated environment. Thus, it prevents proteins from aggregation by isolating them from solution, at the same time creating a folding environment inside the cavity after GroES binding. There are various models of GroE action in the literature. Here, we have adopted the "cage

model" in our protein folding framework as shown in Fig. S4 (7). In the present model, ATP bound GroEL molecule (GrL_T) first binds unfolded (U) or misfolded (M) proteins followed by closure of the GroEL cavity by rapid GroES (GrS) binding to GroEL. The misfolded protein inside the cavity undergoes forced unmisfolding and is converted to U. The unfolded protein inside GroEL is protected from further misfolding and it folds with the same rate constant as in the free state (k_f). The protein inside the cavity gets ~10 seconds to fold before ATP hydrolyzes and the protein is released from the cavity. Thus, GroE accelerates the folding process via forced unmisfolding and protection of unfolded state from further misfolding. In the absence of GroES, the protein inside the cavity is assumed to unmisfold with the rate same as that in the free state. Thus, GroEL alone only prevents aggregation without affecting its natural unmisfolding and folding kinetics. The current model can also capture the possibility of iterative annealing mechanism (IAM) (8, 9) for proteins with rate of folding (k_f) much slower than the rate of release from GroEL cavity.

The present model does not treat small heat-shock proteins IbpA and IbpB, ATP-dependent bacterial Hsp90 (HtpG) and degradation systems ClpXP, ClpAP and HsIUV. However, the effects of these systems on proteostasis are subtle. In healthy cells, these chaperones are expected to be less important (10). They seem to play significant roles mainly for cells that are stressed or old.

Parameterization and validation of the model against refolding kinetic data

Our model is highly constrained by extensive experimental data (6). Fig. S5 illustrates the constraining dynamical data that we used and it shows the degree to which the model fits the data. Fig. S5 is not a prediction of the model. It is just a validation that the rate coefficients given in Tables S1-S2 do adequately capture the observed dynamics.

Proteostasis is complex, so we require a big model to capture the diverse aspects of its health and working principles at various external conditions. A fair question is the common concern about simplifications, approximations and parameters that are invariably made in cell-scale modeling.

First, we note that parameter counts, per se, are not the main issue. Important models in biology often have more than 1000 of them — e.g. systems biology ODEs (where Mycoplasma genitalia modeling, for example, has required 1900+ parameters (11)), or even atomistic force fields. The bigger concern is the nature of how the parameters are obtained and the ratio of parameters to independent constraining data. We believe the model described above is well constrained. Because of the richness of experiments in the proteostasis field over the past 15 years, the experimental literature gives us 20 independent relaxation curves (such as those shown in Fig. S5), to learn the 1.05 parameters per full dynamical relaxation curve for this model.

Second, studies of the modeling process itself in systems biology show that, perhaps because biology itself is robust, models of complex biological circuits often achieve correct collective behavior irrespective of poor or sensitive parameters (12-16). Sethna states that "our complex world is understandable" because in physical or biological reality — and in models of that reality — only a few parameters are usually controlling (12-14). His study of 17 different models leads to the view that microscopic complexity often manifests as macroscopic simplicity. That the behaviors are more important than the parameters reflect Sethna's principle (17, 18). Of course, this is no guarantee for any one particular model, but we believe the present work follows the current best practices of bottom-up systems biology modeling (19).

Sensitivity analysis of the model parameters

We have estimated the effects of changes in the kinetic parameters around their best-fit values (Fig. S5 and Table S1) by performing a local sensitivity analysis. We computed the time-dependent change in the native yield, Δ(t), in response to the variations in each fitting parameter. The change, Δ(t), is computed by using:

$$\Delta(t)=\sum_i \frac{\delta N(t)}{\delta k_i}\Delta_{k_i} \tag{S1}$$

where Δ_{ki} is the variation in the kinetic parameter k_i and $(\delta N(t)/\delta k_i)$ is the derivative of the time-dependent native yield with respect to k_i. The magnitude of Δ(t) is an estimate of the variation in native yield in response to fluctuations in k_i by Δk_i. The magnitude of Δ(t) is an indication of the extent to which the parameter k_i affects N(t) and thus it measures the sensitivity of the model to its parameters. Δ(t) is computed by varying each rate constant by ±5% around the values given in Table S1. The derivatives are obtained using a simple finite difference scheme. Using all the values of $\delta N(t)/\delta k_i$, we obtain Δ(t) by summing them up [Eq. S1]. The time-dependent value of |Δ(t)|, which is the residual in the best fit of refolding data, is small (see shaded areas in Fig. S5). The maximum value, $|\Delta|_{max}$, is less than 5% of total protein (Table S4). The individual contribution of each parameter is estimated and they are also below 5%. Thus, overall the kinetic scheme is robust to variations in the kinetic parameters.

Relation between growth and synthesis rates in the steady state

Let us consider a situation where a cell is growing exponentially with a growth rate λ, and a protein is being synthesized with synthesis rate σ. Newly synthesized protein undergoes folding/misfolding/aggregation depending on its rate parameters (Fig. S1). Under this situation, the protein concentration will reach a steady state value at long times due to the opposing effects of synthesis (σ) and dilution by cell growth (λ). The time dependent concentrations of each state will be as follows:

$$\frac{d[U]}{dt}=\sigma-k_f[U]+k_u[N]-k_{um}[U]+k_{mu}[M]-\lambda[U] \tag{S2}$$

$$\frac{d[N]}{dt}=k_f[U]-k_u[N]-\lambda[N] \tag{S3}$$

$$\frac{d[M]}{dt}=k_{um}[U]-k_{mu}[M]-2k_a[M]^2-k_a[A_2][M]-\lambda[M] \tag{S4}$$

$$\frac{d[A_2]}{dt}=k_a[M]^2-k_a[A_2][M]-\lambda[A_2] \tag{S5}$$

$$\frac{d[A]}{dt}=3k_a[A_2][M]-\lambda[A] \tag{S6}$$

The total concentration of protein at a given time is a sum over each state's concentrations,

$$[tot]=[U]+[N]+[M]+2[A_2]+[A] \tag{S7}$$

Now, if we add Eqs. S2-S6 and insert Eq. S7 into the resulting equation, we get the time dependent total protein concentration as,

$$\frac{d[tot]}{dt} = \sigma - \lambda[tot] \tag{S8}$$

At steady state, total protein concentration will be constant which implies left hand side of Eq. S8 will be 0. Thus, at steady state we get a relationship between synthesis and growth rate,

$$\sigma = \lambda[tot] \tag{S9}$$

This is a general relation for any species that is simultaneously synthesized and diluted due to cell growth.

Derivation of a protein's dwell time in the misfolded state

Consider the folding/misfolding kinetics of a protein,

$$M \underset{k_{um}}{\overset{k_{mu}}{\rightleftarrows}} U \xrightarrow{k_f} N$$

The folding step ($U \to N$) is considered as unidirectional for the sake of computation ease. The small value of the unfolding rate constant compared to the folding rate for a standard protein under normal conditions makes this assumption valid. We seek the dwell time of a protein at M using the method of mean first-passage time, as described below (20, 21).

We assume that the total time for the protein to first get to the native state (N) starting from the misfolded state (M) is a random variable T_M that is governed by the first passage time distribution, $f_{T_M}(t)$. To obtain $f_{T_M}(t)$, we calculate the waiting time distribution in the unfolded state. The duration an unfolded protein needs to wait before reaching the native state (N) is also considered as a random variable denoted by $f_{T_U}(t)$.

The time of realizing each elementary step is assumed to be random and therefore we consider the waiting time distribution of each elementary step as a Poisson distribution:

$$f_{T_f}(t) = k_f\, e^{(-k_f t)}$$

$$f_{T_{um}}(t) = k_{um}\, e^{(-k_{um} t)}$$

$$f_{T_{mu}}(t) = k_{mu}\, e^{(-k_{mu} t)}$$

Suppose the protein is initially in the misfolded state. The total time needed to fold is the random variable T_M. The probability that T_M is realized within some time interval t, which we denote by $P(T_M < t)$, can be written as,

$$P(T_M < t) = P(T_U + T_{mu} < t) \tag{S10}$$

In general, for any random variable *X*, $f_X(x)=\frac{dP(X<x)}{dx}$. Eq. S10 is differentiated with respect to *t* to obtain the corresponding waiting time distribution

$$f_{T_M}(t)=f_{T_U+T_{mu}}(t)$$

Since the distribution of the sum of two random variables is the convolution of the distributions of two individual random variables

$$f_{T_M}(t)=\int_0^\infty dt_1\, f_{T_U}(t-t_1)\, f_{T_{mu}}(t_1)$$
$$=\int_0^\infty dt_1\, f_{T_U}(t-t_1)\, k_{mu}\exp(-k_{mu}\,t_1)$$

By taking the Laplace transform, we obtain

$$\hat{f}_{T_M}(s)=\frac{k_{mu}\hat{f}_{T_U}(s)}{s+k_{mu}} \qquad \text{(S11)}$$

where $\hat{f}_{T_M}(s)$ and $\hat{f}_{T_U}(s)$ are Laplace transforms of $f_{T_M}(t)$ and $f_{T_U}(t)$, respectively. In the case of U, there are two processes. The first is the transition back to the *M* state; and the other choice is the forward step to *N*. The probability that T_U is realized within a time *t* is

$$P(T_U<t)=P(T_{um}<T_f)\,P(T_M+T_{um}<t)+P(T_f<T_{um})\,P(T_f<t)$$

In this expression, the first term indicates that U first goes to M, then from there it goes to N. The second term corresponds to the forward path starting from U to N. The first passage time of U is

$$f_{T_U}(t)=P(T_{um}<T_f)\,f_{T_M+T_{um}}(t)+P(T_f<T_{um})\,f_{T_f}(t)$$

Expressing $f_{T_M+T_{um}}(t)$ as convolution of two terms

$$f_{T_U}(t)=\int_0^\infty dt_1\,P(T_f<t_1)\,f_{T_M}(t-t_1)\,f_{T_{um}}(t_1)+\int_0^\infty dt_1\,P(T_{um}<t_1)\,f_{T_f}(t_1)$$

We put the expressions of the probabilities and waiting time distribution of the elementary steps in right hand side of above equation and find the final expression

$$f_{T_U}(t)=\int_0^\infty dt_1\, f_{T_M}(t-t_1)\,k_{um}\exp\left[-(k_{um}+k_f)t_1\right]+\int_0^\infty dt_1\,k_f\exp\left[-(k_{um}+k_f)t_1\right]$$

Performing the Laplace transformation, we obtain

$$\hat{f}_{T_U}(s) = \frac{k_{um}\hat{f}_{T_M}(s)}{s + k_{um} + k_f} + \frac{k_f}{s + k_{um} + k_f} \tag{S12}$$

Eqs. S11 and S12 can be represented in matrix form as

$$s\hat{\mathbf{f}}(s) = \mathrm{Q}\hat{\mathbf{f}}(s) + \mathbf{r} \tag{S13}$$

where

$$\hat{\mathbf{f}}(s) = \begin{bmatrix} \hat{f}_{T_M}(s) \\ \hat{f}_{T_U}(s) \end{bmatrix}, \mathbf{Q} = \begin{bmatrix} k_{mu} & k_{mu} \\ -(k_{um} + k_f) & k_{um} \end{bmatrix} \text{ and } \mathbf{r} = \begin{bmatrix} 0 \\ k_f \end{bmatrix}.$$

Eq. S13 can be inverted to yield the following expression of $\hat{\mathrm{f}}(s)$:

$$\hat{\mathbf{f}}(s) = (s\mathbf{I} - \mathbf{Q})^{-1}\mathbf{r}. \tag{S14}$$

Once we obtain $\hat{\mathbf{f}}(s)$, the Laplace transform of the waiting-time distribution of *M*, $\hat{f}_{T_M}(s)$, can be obtained as

$$\hat{f}_{T_M}(s) = \begin{bmatrix} 1 & 0 \end{bmatrix}\hat{\mathbf{f}}(s). \tag{S15}$$

The mean first passage time (MFPT), $\tau_0 = \int_0^\infty dt\, t\, f_{T_M}(t)$ is given by $\tau_0 = -\frac{d\hat{f}_{T_M}(s)}{ds}\Big|_{s=0}$. Using Eqs. S14 and S15 we have

$$\tau_0 = \begin{bmatrix} 1 & 0 \end{bmatrix}\mathbf{Q}^{-2}\mathbf{r}\ .$$

Therefore, τ_0 (mean first passage time or the dwell time in non-native states) is

$$\tau_0 = \frac{k_f + k_{um} + k_{mu}}{k_f k_{mu}} \tag{S16}$$

The dwell time in the presence of KJE (τ_K) and GroE (τ_G), respectively, can be derived in similar way and the final expressions are

$$\tau_K = \frac{k_f + k_{um} + k_{mu} + k_K}{k_f(k_{mu} + k_K)} \tag{S17}$$

$$\tau_G = \frac{1}{k_f} \tag{S18}$$

where k_K is the rate of unmisfolding by KJE.

Additional information

GroE is essential for cell viability, whereas KJE and TF are not

How essential is each type of chaperone? We have computed the extent of folding of each class of protein under various conditions by deleting one or more chaperones (Fig. S7). As expected, class I proteins are not affected by the deletion of any chaperone (TF, KJE and GroE) (Fig. S7A). For class II proteins, deleting either TF or KJE results in ~50% loss in folding, but deleting them both together leads to a more significant reduction. Class III is not affected by deleting TF or KJE. However, the model shows that class III requires GroE. This is consistent with in vivo studies which find that TF or DnaK are not individually essential, but *E. coli* is not viable without GroE. Their simultaneous deletion causes synthetic lethality at normal temperatures (22). In our model, the effect of deletion of these chaperones (TF and KJE) is only on class II. Therefore, the cause of synthetic lethality is due to loss of activity of class II proteins. Since, class II proteins are frail and their misfolding rate is not extremely high, it is expected that at low temperature these proteins can fold without any chaperone. Experimental observation of retrieval of spontaneous folding of class II proteins supports this argument (6). Therefore, our study suggests that the cell should not suffer from synthetic lethality at low temperature due to simultaneous deletion of TF and DnaK. Indeed, cells at low temperatures without TF and DnaK chaperones do not suffer from synthetic lethality and can grow normally (22). Alternatively, over-expression of GroEL can prevent cell death. The over-expression of GroEL by two-fold increases native population of class II proteins in proteome without TF and/or DnaK (Fig. S7B). The effect is more in individual deletion of TF and DnaK. The GroEL dependent folding of class II proteins makes it possible for GroEL to help them in absence of other chaperones. Thus, we predict that GroEL alone can help the cell to survive in the absence of other chaperones provided GroEL level is sufficiently high.

Although the deletion of GroEL alone (Fig. S7A, panel ,TF+KJE) reduces the native population of class II proteins to some degree, the native population is still high enough for cell functionality. Downshifting GroE does not pose a threat to class II proteins because those proteins can use the KJE system. But, GroE is essential for cell viability because it is essential for folding class III proteins (Fig. S7).

SI Figures

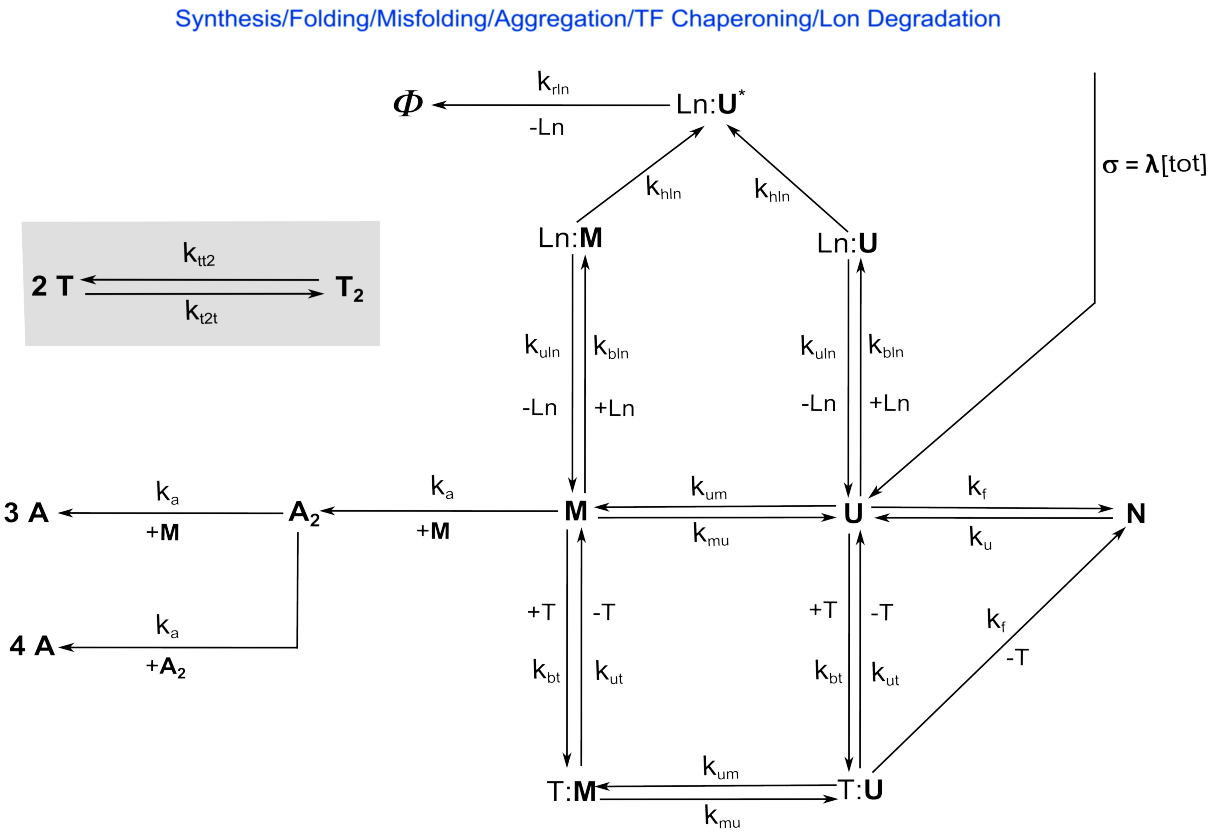

Fig. S1. Kinetic scheme representing synthesis of a new protein with flux σ, followed by folding, misfolding, aggregation, chaperoning by trigger factor and degradation by Lon. N, U, M, A_2, and A are native, unfolded, misfolded, aggregate dimer and post-critical big aggregate, respectively. 'T' and 'Ln' represent trigger factor and Lon proteins, respectively. 'U*' denotes protein committed to degradation and 'ϕ' represents degraded proteins. The values of the rate constants used in this study are given in Tables S1 and S2.

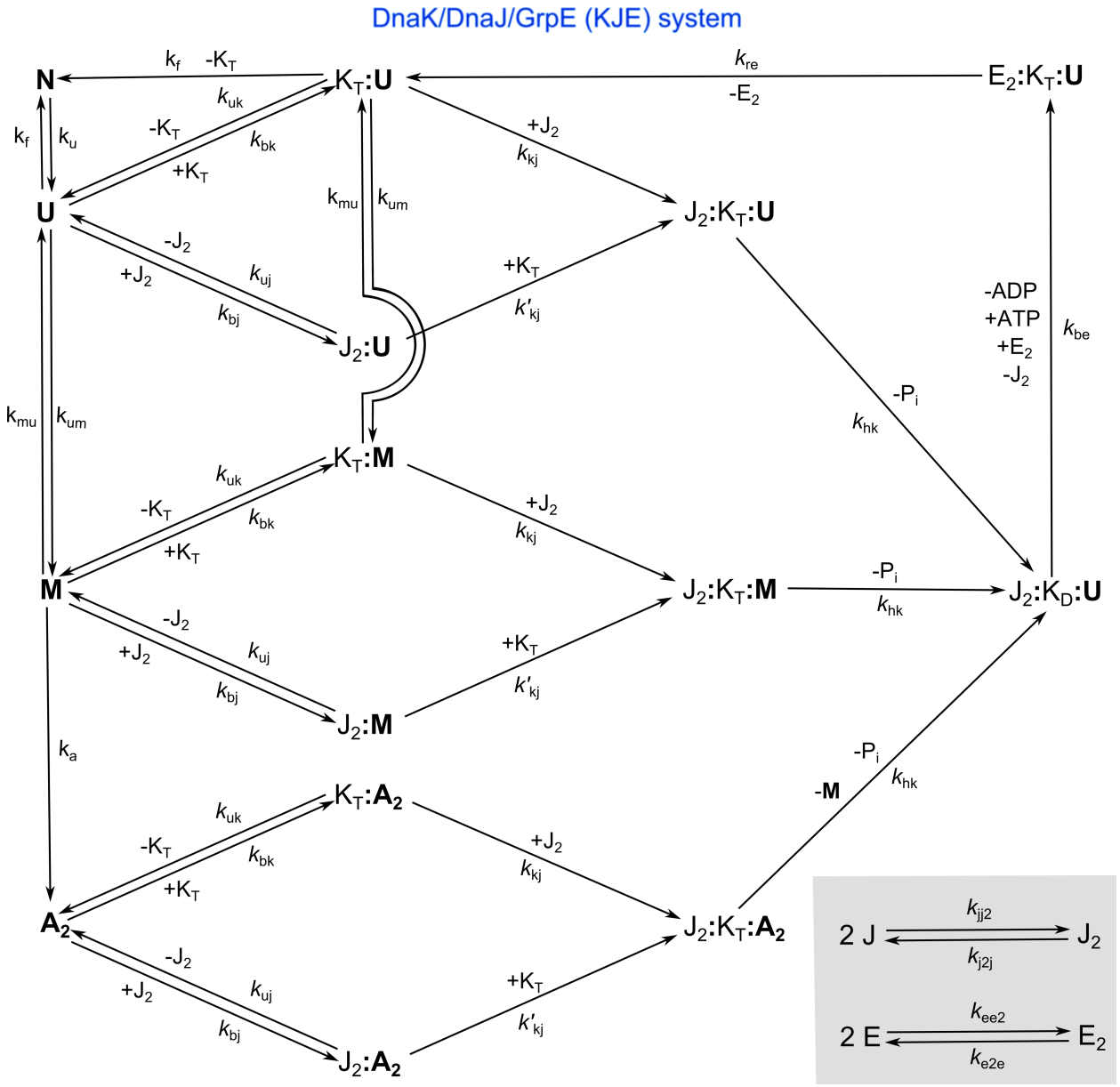

Fig. S2. Kinetic scheme of KJE (DnaK, DnaJ and GrpE) mediated unfolding of a client protein. K_T, J_2 and E_2 are ATP-bound states of DnaK, DnaJ-dimer and GrpE-dimer, respectively. ATP is implicit in the model. Subscripts 'T' and 'D' in 'K' indicate 'ATP' and 'ADP' bound states of DnaK, respectively. The values of the rate constants used in this study are given in Tables S1 and S2.

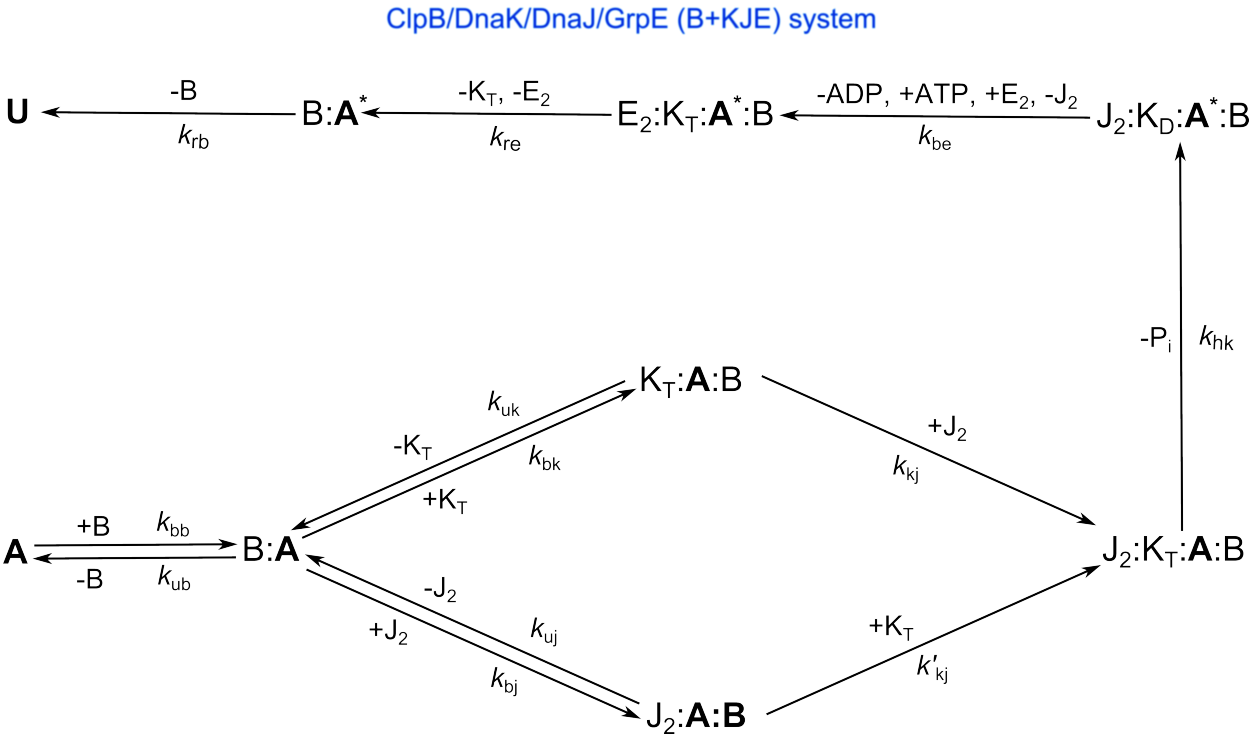

Fig. S3. Disaggregation kinetics mediated by ClpB protein with the help of KJE system. 'B' represents the hexameric active form of ClpB. 'A*' is a KJE modified state of a client protein that can undergo translocation through the ClpB pore and is released in the unfolded state. Other symbols are the same as in Fig. S2. The values of the rate constants used in this study are given in Tables S1 and S2.

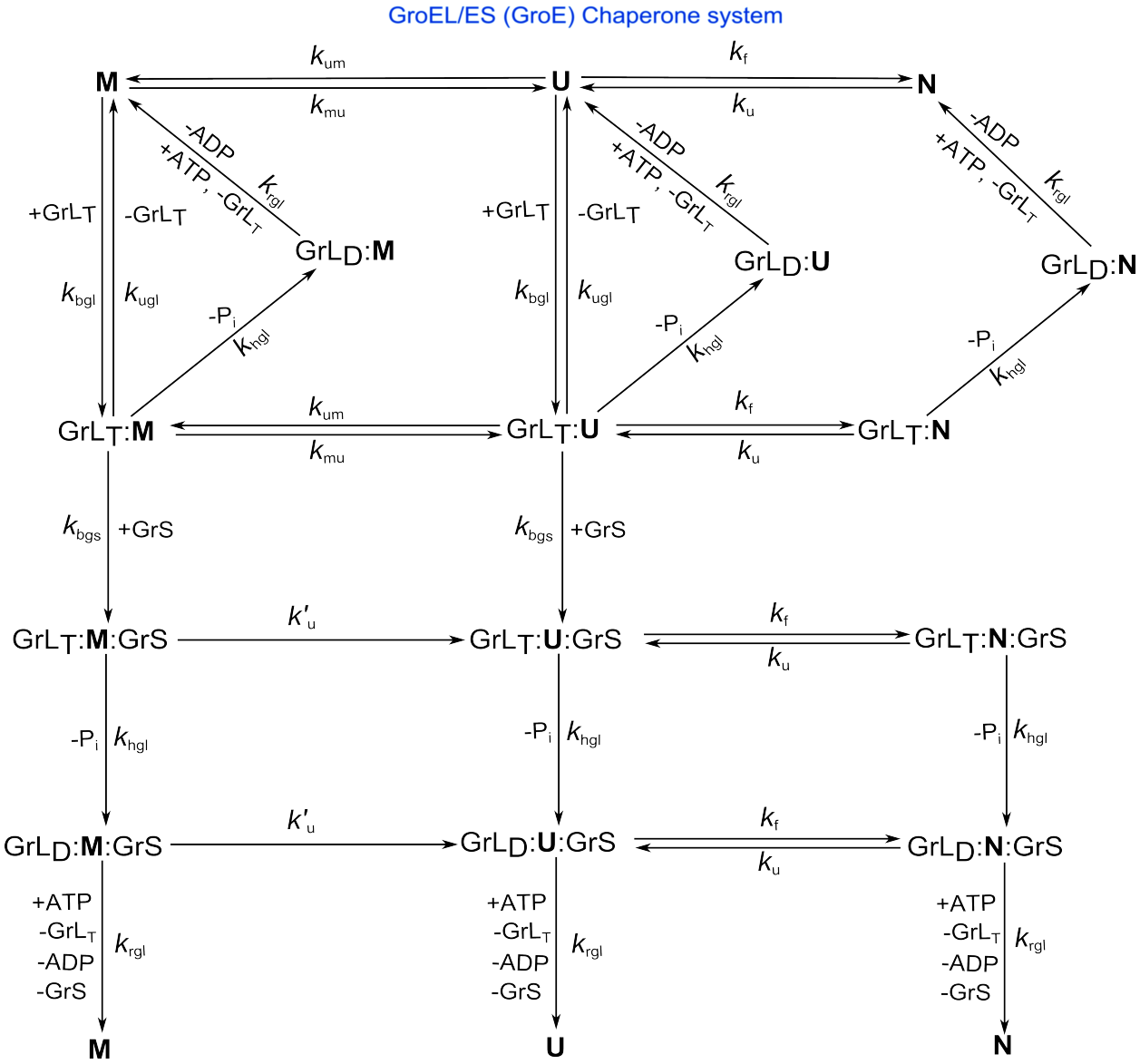

Fig. S4. Kinetic scheme of GroE (GroEL/GroES) chaperoning of a client protein. GrL and GrS indicate active 14-mer and 7-mer of GroEL and GroES molecules, respectively. Subscripts 'T' and 'D' indicate 'ATP' and 'ADP' bound states of GroEL, respectively. The values of the rate constants used in this study is given in Tables S1 and S2.

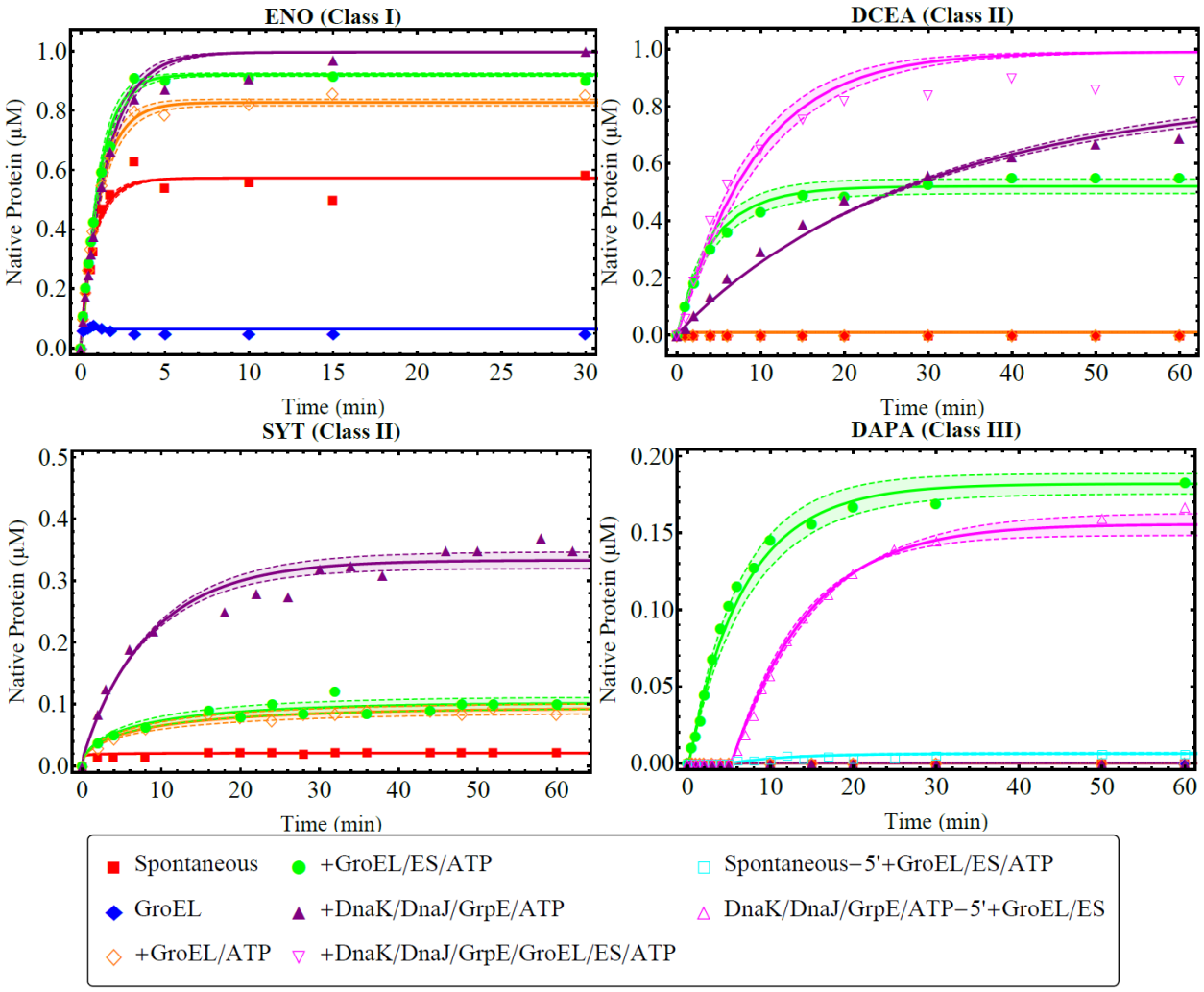

Fig. S5. Refolding kinetics of GroEL interacting proteins, ENO (class I), DCEA (class II), SYT (class II) and DAPA (class III), at 37°C under different conditions. Symbols represent experimental data (6) and the solid lines are corresponding theoretical fits. The stoichiometry of the components is 1 protein : 2 GroEL (14-mer) : 4 GroES (7-mer) : 5 DnaK : 2.5 DnaJ : 5 GrpE. Concentrations of ENO, DCEA, SYT and DAPA are 1.0 μM, 1.0 μM, 0.5 μM and 0.25 μM, respectively. "Spontaneous-5'+GroEL/ES/ATP": without any chaperone followed by addition of GroEL/ES+ATP at 5 minutes; "DnaK/DnaJ/GrpE/ATP-5'+GroEL/ES": in presence of DnaK/DnaJ/GrpE+ATP followed by addition of GroEL/ES at 5 minutes. Rate parameters of the model are given in Tables S1 and S2. The dashed lines with shaded area represent deviations from the best fits on changing each of the fitting parameters by ±5%. The maximum values of the deviations are listed in Table S4.

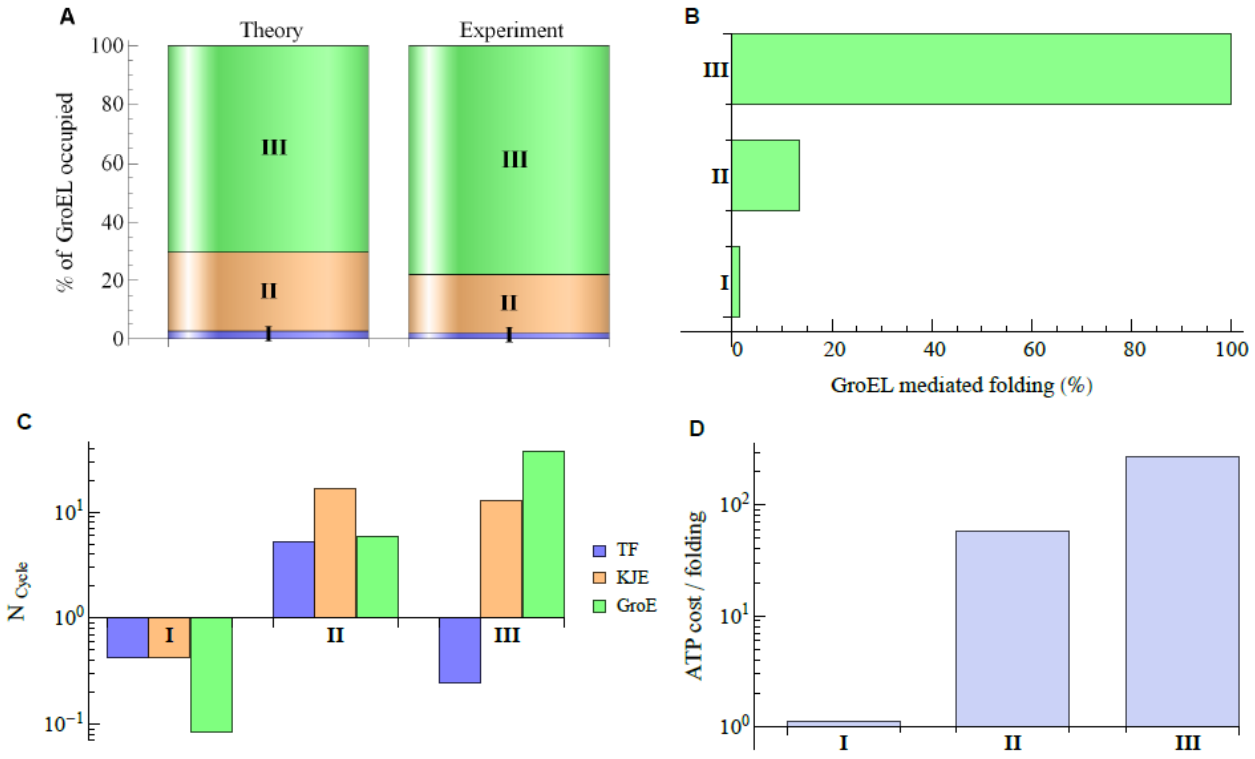

Fig. S6. GroEL dependence, number chaperone cycles and ATP cost of folding for three classes of GroE substrates. (A) Distribution of GroE occupancy by three classes of proteins and (B) extent of GroEL mediated folding. (C) Number of cycles through three chaperone systems (TF, KJE and GroE) before folding for each class of proteins. (D) Total ATP cost to fold each class of proteins.

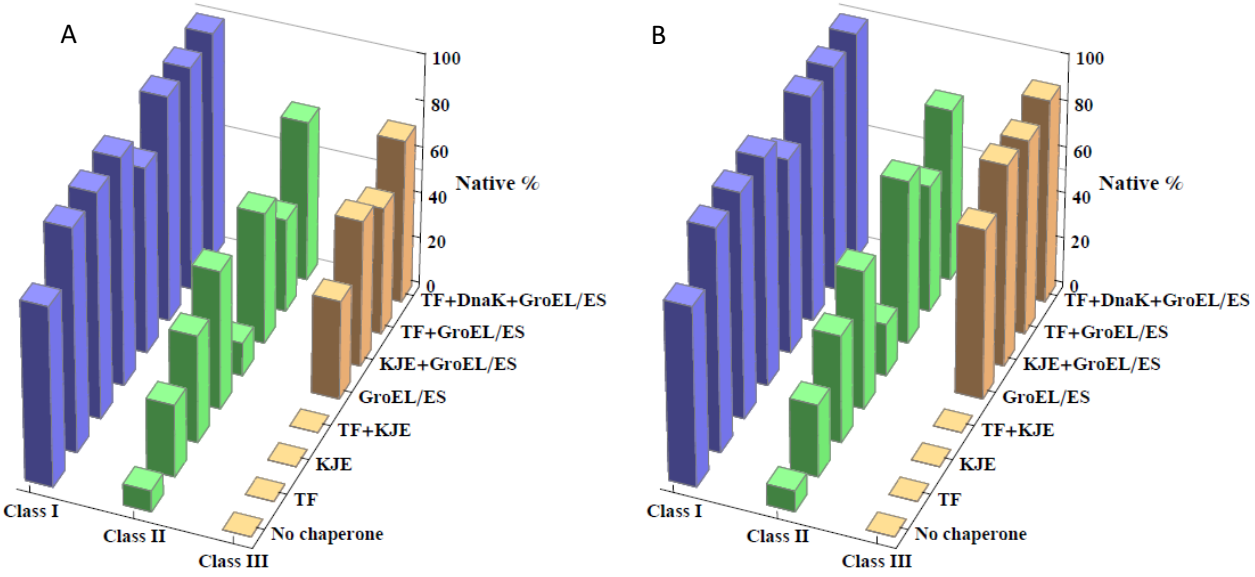

Fig. S7. Extent of folding in the absence or presence of different combinations of chaperones at 40-minute doubling time. (A) The concentration of chaperones when they are present are taken to be at cellular levels (Table S3). (B) Same as in (A) except the concentration of GroEL/ES is double compared to that in (A).

SI Tables

Table S1. Protein specific rate constants (extracted by fitting data in Fig. S5).

Parameter	Class I	Class II		Class III
	ENO	DCEA	SYT	DAPA
k_f	0.017 s^{-1}	0.004 s^{-1}	0.03 s^{-1}	0.003 s^{-1}
k_u	5.3×10^{-6} †	1.7×10^{-8} †	7.6×10^{-6} †	8.4×10^{-7} †
k_{um}	0.15 s^{-1}	0.42 s^{-1}	1 s^{-1}	3.1×10^{5} s^{-1}
k_{mu}	1.1 s^{-1}	7×10^{-8} s^{-1}	0.03 s^{-1}	0.02 s^{-1}
k_a	0.9 μM^{-1} s^{-1}	0.09 μM^{-1} s^{-1}	0.2 μM^{-1} s^{-1}	0.16 μM^{-1} s^{-1}
k_{bgl}	0.14 μM^{-1} s^{-1}	0.15 μM^{-1} s^{-1}	0.8 μM^{-1} s^{-1}	1.9 μM^{-1} s^{-1}
k_{bgs}	10 μM^{-1} s^{-1}*	10 μM^{-1} s^{-1}*	3×10^{-4} μM^{-1} s^{-1}	10 μM^{-1} s^{-1}*

† Computed from protein's stability formula reported in (23). *Obtained from (24).

Table S2. Protein non-specific rate constants.

TF		**KJE**	
Parameter	*Value*	*Parameter*	*Value*
k_{tt2}	3 μM^{-1} s^{-1} (1, 25)	k_{jj2}	0.1 μM^{-1} s^{-1} (3)
k_{t2t}	6 s^{-1} (1, 25)	k_{j2j}	10^{-4} s^{-1} (3)
k_{bt}	3 μM^{-1} s^{-1} (25)	k_{ee2}	0.1 μM^{-1} s^{-1} (1, 26)
k_{ut}	0.5 s^{-1} (25)	k_{e2e}	0.2 s^{-1} (1, 26)
Lon		k_{bk}	1 μM^{-1} s^{-1} (27, 28)
Parameter	*Value*	k_{uk}	2 s^{-1} (27, 28)
k_{bln}	0.1 μM^{-1} s^{-1} (29-31)	k_{bj}	0.3 μM^{-1} s^{-1} (32)
k_{uln}	1 s^{-1} (29-31)	k_{uj}	0.006 s^{-1} (32)
k_{hln}	5 s^{-1} (31)	k_{kj}	0.01 μM^{-1} s^{-1} (33-35)
k_{rln}	0.03 s^{-1} (36)	k'_{kj}	10 μM^{-1} s^{-1} (37)
GroELS		k_{hk}	2 s^{-1} (38)
Parameter	*Value*	k_{be}	0.3 μM^{-1} s^{-1} (39)
k_{ugl}	0.1 s^{-1} (40)	k_{re}	1 s^{-1} (27, 28)
k'_u	2.5 s^{-1} (41)	**ClpB**	
k_{hgl}	0.1 s^{-1} (42-44)	*Parameter*	*Value*
k_{rgl}	1 s^{-1} (42, 44)	k_{bb}	0.1 μM^{-1} s^{-1} (1)
		k_{ub}	0.002 s^{-1} (1)
		k_{rb}	0.03 s^{-1} (4)

Table S3. Concentrations of each component used in this study. The values represent concentrations of monomers.

Component	Concentration
TF	20 μM (1, 45-49)
DnaK	30 μM (1, 45-49)
DnaJ	1 μM (1, 46-48)
GrpE	15 μM (1, 45-47, 49)
ClpB	1.8 μM (1, 45-48)
GroEL	42 μM (1, 45-47, 49)
GroES	35 μM (1, 45-49)
Lon	1.8 μM (1, 45-49)
Class I	252 μM (50)
Class II	75.6 μM (50)
Class III	29.4 μM (50)

Table S4. The maximum values of deviations from best fits of refolding data represented in Fig. S5 with changes of best fitting parameter values (Table S1) by ±5%.

Deviation	Protein			
	ENO (I)	DCEA (II)	SYT (II)	DAPA (III)
$\lvert\Delta\rvert_{\max}$	3.5%	4.3%	2.7%	4.2%
$\lvert\Delta_{k_f}\rvert_{\max}$	3.6%	3.7%	3.2%	3.3%
$\lvert\Delta_{k_{um}}\rvert_{\max}$	3.8%	3.3%	3.2%	0%
$\lvert\Delta_{k_{mu}}\rvert_{\max}$	3.4%	0%	1.1%	0%
$\lvert\Delta_{k_a}\rvert_{\max}$	1.8%	1.9%	2.9%	3.0%
$\lvert\Delta_{k_{bgl}}\rvert_{\max}$	1.2%	4.3%	1.7%	2.9%
$\lvert\Delta_{k_{bgs}}\rvert_{\max}$	-	-	0.2%	-

1. Powers ET, Powers DL, & Gierasch LM (2012) FoldEco: a model for proteostasis in E. coli. *Cell Rep* 1(3):265-276.
2. Nunes JM, Mayer-Hartl M, Hartl FU, & Muller DJ (2015) Action of the Hsp70 chaperone system observed with single proteins. *Nat Commun* 6:6307.
3. Hu B, Mayer MP, & Tomita M (2006) Modeling Hsp70-mediated protein folding. *Biophys J* 91(2):496-507.
4. Goloubinoff P, Mogk A, Zvi AP, Tomoyasu T, & Bukau B (1999) Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. *Proc Natl Acad Sci U S A* 96(24):13732-13737.
5. Zietkiewicz S, Lewandowska A, Stocki P, & Liberek K (2006) Hsp70 chaperone machine remodels protein aggregates at the initial step of Hsp70-Hsp100-dependent disaggregation. *J Biol Chem* 281(11):7022-7029.
6. Kerner MJ*, et al.* (2005) Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. *Cell* 122(2):209-220.
7. Brinker A*, et al.* (2001) Dual function of protein confinement in chaperonin-assisted protein folding. *Cell* 107(2):223-233.
8. Todd MJ, Lorimer GH, & Thirumalai D (1996) Chaperonin-facilitated protein folding: optimization of rate and yield by an iterative annealing mechanism. *Proc Natl Acad Sci U S A* 93(9):4030-4035.
9. Todd MJ, Viitanen PV, & Lorimer GH (1994) Dynamics of the chaperonin ATPase cycle: implications for facilitated protein folding. *Science* 265(5172):659-666.
10. Cho Y*, et al.* (2015) Individual and collective contributions of chaperoning and degradation to protein homeostasis in E. coli. *Cell Rep* 11(2):321-333.
11. Karr JR*, et al.* (2012) A whole-cell computational model predicts phenotype from genotype. *Cell* 150(2):389-401.
12. Casey FP*, et al.* (2007) Optimal experimental design in an epidermal growth factor receptor signalling and down-regulation model. *IET Syst Biol* 1(3):190-202.
13. Gutenkunst RN*, et al.* (2007) Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput Biol* 3(10):1871-1878.
14. Transtrum MK*, et al.* (2015) Perspective: Sloppiness and emergent theories in physics, biology, and beyond. *J Chem Phys* 143(1):010901.
15. Li F, Long T, Lu Y, Ouyang Q, & Tang C (2004) The yeast cell-cycle network is robustly designed. *Proc Natl Acad Sci U S A* 101(14):4781-4786.
16. Ashyraliyev M, Fomekong-Nanfack Y, Kaandorp JA, & Blom JG (2009) Systems biology: parameter estimation for biochemical models. *FEBS J* 276(4):886-902.
17. Li D*, et al.* (2006) Protein interaction networks of Saccharomyces cerevisiae, Caenorhabditis elegans and Drosophila melanogaster: large-scale organization and robustness. *Proteomics* 6(2):456-461.
18. Wang J, Xu L, & Wang E (2008) Potential landscape and flux framework of nonequilibrium networks: robustness, dissipation, and coherence of biochemical oscillations. *Proc Natl Acad Sci U S A* 105(34):12271-12276.
19. Bruggeman FJ & Westerhoff HV (2007) The nature of systems biology. *Trends Microbiol* 15(1):45-50.
20. Santra M & Bagchi B (2012) Catalysis of tRNA aminoacylation: single turnover to steady-state kinetics of tRNA synthetases. *J Phys Chem B* 116(39):11809-11817.
21. Kou SC, Cherayil BJ, Min W, English BP, & Xie XS (2005) Single-molecule Michaelis-Menten equations. *J Phys Chem B* 109(41):19068-19081.
22. Vorderwulbecke S*, et al.* (2004) Low temperature or GroEL/ES overproduction permits growth of Escherichia coli cells lacking trigger factor and DnaK. *FEBS Lett* 559(1-3):181-187.

23. Ghosh K & Dill KA (2009) Computing protein stabilities from their chain lengths. *Proc Natl Acad Sci U S A* 106(26):10649-10654.
24. Motojima F, Chaudhry C, Fenton WA, Farr GW, & Horwich AL (2004) Substrate polypeptide presents a load on the apical domains of the chaperonin GroEL. *Proc Natl Acad Sci U S A* 101(42):15005-15012.
25. Maier R, Eckert B, Scholz C, Lilie H, & Schmid FX (2003) Interaction of trigger factor with the ribosome. *J Mol Biol* 326(2):585-592.
26. Mehl AF, Heskett LD, Jain SS, & Demeler B (2003) Insights into dimerization and four-helix bundle formation found by dissection of the dimer interface of the GrpE protein from Escherichia coli. *Protein Sci* 12(6):1205-1215.
27. Schmid D, Baici A, Gehring H, & Christen P (1994) Kinetics of molecular chaperone action. *Science* 263(5149):971-973.
28. Gisler SM, Pierpaoli EV, & Christen P (1998) Catapult mechanism renders the chaperone action of Hsp70 unidirectional. *J Mol Biol* 279(4):833-840.
29. Gur E & Sauer RT (2008) Recognition of misfolded proteins by Lon, a AAA(+) protease. *Genes Dev* 22(16):2267-2277.
30. Patterson-Ward J, Huang J, & Lee I (2007) Detection and characterization of two ATP-dependent conformational changes in proteolytically inactive Escherichia coli Lon mutants by stopped flow kinetic techniques. *Biochemistry* 46(47):13593-13605.
31. Vineyard D, Patterson-Ward J, Berdis AJ, & Lee I (2005) Monitoring the timing of ATP hydrolysis with activation of peptide cleavage in Escherichia coli Lon by transient kinetics. *Biochemistry* 44(5):1671-1682.
32. Gamer J*, et al.* (1996) A cycle of binding and release of the DnaK, DnaJ and GrpE chaperones regulates activity of the Escherichia coli heat shock transcription factor sigma32. *EMBO J* 15(3):607-617.
33. Suh WC*, et al.* (1998) Interaction of the Hsp70 molecular chaperone, DnaK, with its cochaperone DnaJ. *Proc Natl Acad Sci U S A* 95(26):15223-15228.
34. Mayer MP, Laufen T, Paal K, McCarty JS, & Bukau B (1999) Investigation of the interaction between DnaK and DnaJ by surface plasmon resonance spectroscopy. *J Mol Biol* 289(4):1131-1144.
35. Suh WC, Lu CZ, & Gross CA (1999) Structural features required for the interaction of the Hsp70 molecular chaperone DnaK with its cochaperone DnaJ. *J Biol Chem* 274(43):30534-30539.
36. Kenniston JA, Baker TA, Fernandez JM, & Sauer RT (2003) Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. *Cell* 114(4):511-520.
37. Wittung-Stafshede P, Guidry J, Horne BE, & Landry SJ (2003) The J-domain of Hsp40 couples ATP hydrolysis to substrate capture in Hsp70. *Biochemistry* 42(17):4937-4944.
38. Laufen T*, et al.* (1999) Mechanism of regulation of hsp70 chaperones by DnaJ cochaperones. *Proc Natl Acad Sci U S A* 96(10):5452-5457.
39. Chesnokova LS*, et al.* (2003) Deletion of DnaK's lid strengthens binding to the nucleotide exchange factor, GrpE: a kinetic and thermodynamic analysis. *Biochemistry* 42(30):9028-9040.
40. Fenton WA & Horwich AL (1997) GroEL-mediated protein folding. *Protein Sci* 6(4):743-760.
41. Lin Z, Madan D, & Rye HS (2008) GroEL stimulates protein folding through forced unfolding. *Nat Struct Mol Biol* 15(3):303-311.
42. Burston SG, Ranson NA, & Clarke AR (1995) The origins and consequences of asymmetry in the chaperonin reaction cycle. *J Mol Biol* 249(1):138-152.
43. Ranson NA, Burston SG, & Clarke AR (1997) Binding, encapsulation and ejection: substrate dynamics during a chaperonin-assisted folding reaction. *J Mol Biol* 266(4):656-664.

44. Rye HS*, et al.* (1999) GroEL-GroES cycling: ATP and nonnative polypeptide direct alternation of folding-active rings. *Cell* 97(3):325-338.
45. Lopez-Campistrous A*, et al.* (2005) Localization, annotation, and comparison of the Escherichia coli K-12 proteome under two states of growth. *Mol Cell Proteomics* 4(8):1205-1209.
46. Lu P, Vogel C, Wang R, Yao X, & Marcotte EM (2007) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat Biotechnol* 25(1):117-124.
47. Ishihama Y*, et al.* (2008) Protein abundance profiling of the Escherichia coli cytosol. *BMC Genomics* 9:102.
48. Taniguchi Y*, et al.* (2010) Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science* 329(5991):533-538.
49. Van Bogelen RA, Abshire, K.Z., Pertsemlidis, A., Clark, R.L., and Neidhardt, F.C. (1996) Gene-protein database of Escherichia coli K-12. *Escherichia coli and Salmonella : cellular and molecular biology*, ed Neidhardt FC, R. Curtiss III, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger (ASM Press, Washington, D.C.), 2nd Ed, pp 2067-2117.
50. Wang M*, et al.* (2012) PaxDb, a database of protein abundance averages across all three domains of life. *Mol Cell Proteomics* 11(8):492-500.