Title
Design principles for the glycoprotein quality control pathway

Permalink

Journal
PLOS Computational Biology, 17(2)

ISSN
1553-734X

Authors
Brown, Aidan I
Koslover, Elena F

Publication Date
2021

DOI
10.1371/journal.pcbi.1008654

Peer reviewed

RESEARCH ARTICLE

Design principles for the glycoprotein quality control pathway

Aidan I. Brown [ID][¤]**, Elena F. Koslover** [ID]*

Department of Physics, University of California, San Diego, San Diego, California, United States of America

¤ Current address: Department of Physics, Ryerson University, Toronto, Canada
* ekoslover@ucsd.edu

Abstract

Newly-translated glycoproteins in the endoplasmic reticulum (ER) often undergo cycles of chaperone binding and release in order to assist in folding. Quality control is required to distinguish between proteins that have completed native folding, those that have yet to fold, and those that have misfolded. Using quantitative modeling, we explore how the design of the quality-control pathway modulates its efficiency. Our results show that an energy-consuming cyclic quality-control process, similar to the observed physiological system, outperforms alternative designs. The kinetic parameters that optimize the performance of this system drastically change with protein production levels, while remaining relatively insensitive to the protein folding rate. Adjusting only the degradation rate, while fixing other parameters, allows the pathway to adapt across a range of protein production levels, aligning with *in vivo* measurements that implicate the release of degradation-associated enzymes as a rapid-response system for perturbations in protein homeostasis. The quantitative models developed here elucidate design principles for effective glycoprotein quality control in the ER, improving our mechanistic understanding of a system crucial to maintaining cellular health.

Author summary

We explore the architecture and limitations of the quality-control pathway responsible for efficient folding of secretory proteins. Newly-synthesized proteins are tagged by the attachment of a 'glycan' sugar chain which facilitates their binding to a chaperone that assists protein folding. Removal of a specific sugar group on the glycan ends the interaction with the chaperone, and not-yet-folded proteins can be re-tagged for another round of chaperone binding. A degradation pathway acts in parallel with the folding cycle, to remove those proteins that have remained unfolded for a sufficiently long time. We develop and solve a mathematical model of this quality-control system, showing that the cyclical design found in living cells is uniquely able to maximize folded protein throughput while avoiding accumulation of unfolded proteins. Although this physiological model provides the best performance, its parameters must be adjusted to perform optimally under different protein production loads, and any single fixed set of parameters leads to poor performance when production rate is altered. We find that a single adjustable

parameter, the protein degradation rate, is sufficient to allow optimal performance across a range of conditions. Interestingly, observations of living cells suggest that the degradation speed is indeed rapidly adjusted.

Introduction

The general principle of quality control is of critical importance to the maintenance, function, and growth of biological cells. Autophagy and the ubiquitin-proteasome system selectively remove damaged proteins and organelles to maintain the quality of cellular components [1, 2]. Fidelity is aided by proofreading processes during DNA copying [3], immune signaling [4], and external sensing [5]. Quality control is particularly important for proteins, with a high fraction of proteome mass across the kingdoms of life devoted to protein homeostasis and folding [6]. Unfolded and misfolded proteins often form aggregates, which can impede cellular processes and are associated with a variety of human diseases [7–10].

Protein quality control begins with transcriptional proofreading by RNA polymerase [11] and continues with proofreading of tRNA matching to mRNA codons during translation [12] to reduce errors in the polypeptide sequence. Quality control continues beyond production, throughout the lifetime of a protein [13–15]. We focus on post-translational quality control pathways that ensure nascent polypeptides fold into the correct or 'native' three-dimensional conformation, rather than roaming the cell in a misfolded state [13–15].

Nearly one-third of eukaryotic proteins, or ~ 8000 proteins in humans, are synthesized through the secretory pathway and begin as nascent polypeptides in the endoplasmic reticulum (ER) [16]. The majority of ER-manufactured proteins acquire branched carbohydrate chains, via N-linked glycosylation [17]. While these glycan chains can be important for protein function [18] and stabilization [19], the specific sugar residues in the glycan serve as a tunable barcode to direct the interactions that lead to further protein folding attempts or protein degradation [16]. Accordingly, glycans play a key role in the folding quality control of secretory proteins.

The quality control pathway, in deciding which proteins to degrade and which to continue folding, attempts to distinguish between three groups of proteins: natively folded, as yet unfolded, and terminally misfolded. Natively folded proteins can be distinguished by the lack of exposed hydrophobic residues and free thiols [20, 21], and are permitted to leave the ER to continue through the secretory pathway. It is less straightforward to distinguish between as yet unfolded proteins, which should be provided more time to fold; and terminally misfolded proteins, which should be targeted for degradation [22]. Newly-synthesized proteins and unfolded proteins are flagged by a monoglucosylated glycan chain, which facilitates chaperone binding to attempt folding. Proteins dissociate from the chaperone upon removal of this glucose moiety, which is not added back to proteins that have reached their native conformation. Proteins that fail to reach a native conformation will eventually experience trimming of other glycan moieties, leading to degradation via the ER-associated degradation (ERAD) pathway [16].

In this work we investigate how the design of the glycoprotein folding quality-control pathway facilitates decisions of whether nascent proteins may continue trying to fold, and how specific pathway features impact performance. Specifically, we seek to understand the advantages provided by the cyclic structure of the quality control pathway. We describe how the nonequilibrium driving of the chaperone binding cycle enables improved performance, reinforcing previous descriptions [15, 23, 24] of this process as kinetic proofreading [3, 25]. Overall, we

find that the consensus physiological model outperforms other designs, and describe how its kinetic parameters can be tuned to maintain performance across a broad range of conditions.

Model

Upon translation, glycoproteins enter the quality control pathway marked with a single glucose moiety [16, 26]. These monoglucosylated proteins can bind calnexin and calreticulin [27], chaperone proteins that assist protein folding. The interaction of a glycoprotein with a chaperone ends with trimming of the glucose by glucosidase II [22, 28–30]. A monoglucosylated protein dynamically binds to and unbinds from a chaperone, with glucosidase II likely trimming the glucose when the protein is unbound due to steric limitations [31–34]. We consider the monoglucosylated protein as remaining effectively in a state of association with the chaperone until it is deglucosylated [28, 29, 35]. This treatment of dynamic on/off binding as a single composite state follows previous modeling of proteins binding DNA [36] and molecular motors on their tracks [37].

Proteins that have reached a native conformation are eligible to be exported from the ER and to proceed down the secretory pathway [38]. However, not all proteins that are released from the chaperone are successfully folded. Uridine diphosphate-glucose:glycoprotein glucosyltransferase (UGGT) can reglucosylate incompletely folded glycoproteins to enable another round of chaperone binding that further facilitates folding [29, 39]. UGGT does not reglucosylate proteins that have reached a native conformation, and is thought to use indicators such as the availability of the entire glycan chain and hydrophobic patches to detect non-native conformations [16, 29, 39, 40]. There is some evidence that UGGT may prefer to reglucosylate unfolded glycoproteins rather than those that have misfolded into an incorrect conformation, but overall it is unclear if UGGT can distinguish between these two groups of non-natively folded proteins [22, 39, 40].

Glycoprotein interaction with chaperones, glucosidase II, and UGGT thus forms a cycle: a monoglucosylated protein binds a chaperone (calnexin or calreticulin) for folding assistance, the glucose is trimmed by glucosidase II to end the chaperone interaction, and UGGT restores the glucose to non-natively folded proteins to direct chaperone rebinding [39]. Folding time in the ER can vary from a few minutes to several hours [41], with some proteins natively folded after one round of chaperone binding, and others requiring multiple rounds of chaperone interaction [29].

In addition to departing the cycle by folding, proteins can be selected for ER-associated degradation (ERAD), a pathway involving removal from the ER followed by proteasomal degradation [29]. Commitment of a protein to the ERAD pathway for degradation can involve interaction with various enzymes, some of which irreversibly trim additional moieties off the glycan chains [16, 17, 22, 27, 42–50]. Unglucosylated glycans, which do not allow chaperone binding, are thought to be specifically vulnerable to the modifications that commit a protein to ERAD [49, 51, 52].

We represent the glycoprotein quality control cycle with three discrete states, along with an additional discrete state for chaperone-bound natively folded proteins (see Fig 1). Proteins enter the cycle in a monoglucosylated state (whose concentration is represented by P_g) with a production rate k_p. Monoglucosylated proteins bind to chaperones as a bimolecular reaction with rate constant k_c. The available chaperone concentration is represented by C_A and the concentration of chaperone-bound unfolded proteins by P_c. Proteins bound to the chaperone fold into their native conformation with rate constant k_f, and P_{cf} represents the concentration of folded proteins bound to the chaperone. Chaperone-bound proteins (both natively folded and not) end their chaperone interactions through glucose trimming with rate constant k_r, with
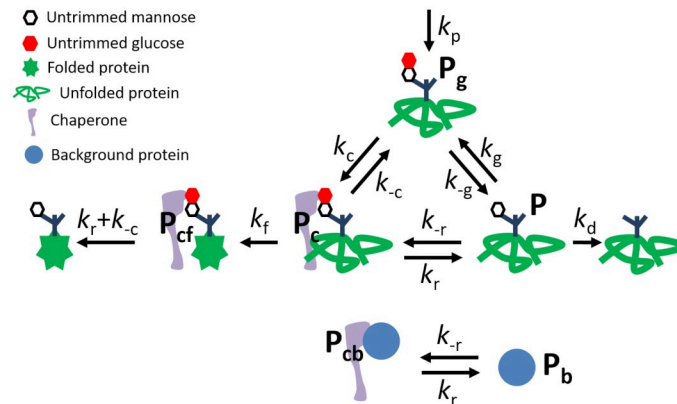
Fig 1. Model of glycoprotein quality control via the chaperone binding cycle. P_g represents the monoglucosylated proteins, P_c the unfolded chaperone-bound proteins, P_{cf} the folded chaperone-bound proteins, P the proteins lacking a glucose tag, P_b the background proteins, and P_{cb} the chaperone-bound background proteins.

https://doi.org/10.1371/journal.pcbi.1008654.g001

natively folded proteins then exiting the cycle. In a subsequent section, we address an alternate model where monoglucosylated protein unbinding would instead return the protein to state P_g, wherein it has no memory of prior chaperone interaction. We refer to this variant as the 'No tag memory' model. Glucose-trimmed proteins that are not natively folded (at concentration P) are lacking a glucose moiety, and can be reglucosylated with a rate constant k_g. Monoglucosylated proteins not bound to a chaperone can have their glucose removed with rate constant k_{-g}, serving as a "safety-valve" pathway when the concentration of proteins to be folded overwhelms the available chaperones. Deglucosylated proteins are vulnerable to degradation via ERAD [49, 51, 52]. Specifically, the sugar moiety to which the glucose attaches can be removed, irreversibly committing the protein to degradation via the ERAD pathway [16, 17, 41, 49, 50, 53–55]. We treat ERAD commitment and protein degradation as a single irreversible process with rate constant k_d.

For completeness we also consider termination of the chaperone interaction of monoglucosylated proteins without glucose trimming (rate constant k_{-c}) and rebinding of deglucosylated proteins back to the chaperone (rate constant k_{-r}). Because such a putative rebinding pathway does not rely on a glucosylation signal to recognize proteins in need of folding, it is assumed to be non-specific and to allow the general binding of some 'background' proteins onto the chaperones. Such background proteins could include ER-resident proteins, or folded proteins that have not yet been exported. The concentration of these additional background proteins is represented by P_b (for free background proteins) and P_{cb} for background proteins bound to the chaperone. We assume each chaperone can bind only one protein at a time.

Overall, the dynamics of the chaperone binding cycle are described by

$$\frac{dP_g}{dt} = k_g P + k_{-c} P_c - (k_c C_A + k_{-g}) P_g + k_p \ , \tag{1a}$$

$$\frac{dP_c}{dt} = (k_c P_g + k_{-r} P) C_A - (k_r + k_{-c} + k_f) P_c \ , \tag{1b}$$

$$\frac{dP}{dt} = k_r P_c + k_{-g} P_g - (k_g + k_{-r} C_A - k_d) P \ , \tag{1c}$$

$$\frac{\mathrm{d}P_{\mathrm{cf}}}{\mathrm{d}t} = k_{\mathrm{f}}P_{\mathrm{c}} - (k_{\mathrm{r}} + k_{-\mathrm{c}})P_{\mathrm{cf}} \;,$$ (1d)

$$\frac{\mathrm{d}P_{\mathrm{cb}}}{\mathrm{d}t} = k_{-\mathrm{r}}C_{\mathrm{A}}P_{\mathrm{b}} - k_{\mathrm{r}}P_{\mathrm{cb}} \;.$$ (1e)

Some proteins entering the chaperone binding cycle are unable to natively fold, as a result of translation errors or mutations [56]. Heat and oxidative stress can also cause proteins to enter states that cannot fold [56], and these stressors may have a differential impact on different proteins. We label these terminally misfolded, unfoldable proteins as simply 'misfolded'. Their dynamics are described by equations similar to Eqs 1a–1c, with analogous protein quantities P_{g}^{*}, P_{c}^{*}, and P^{*}. The misfolded protein production rate is defined as k_{p}^{*} and the folding rate is set to zero ($k_{\mathrm{f}}^{*} = 0$). All other rate constants are assumed to be identical for foldable and misfolded proteins.

Both the background proteins (P_{b}) and misfolded proteins (P_{i}^{*}) represent proteins capable of binding to and occupying the limited supply of total chaperone (C_{tot}) available in the cell. The concentration of available chaperones is then given by $C_{\mathrm{A}} = C_{\mathrm{tot}} - P_{\mathrm{c}} - P_{\mathrm{cf}} - P_{\mathrm{c}}^{*} - P_{\mathrm{cb}}$. In our model, background proteins represent those proteins that can bind weakly to the chaperone in the absence of a glucose moiety flagging them as newly-made proteins requiring folding. These can represent, for example, already folded proteins. They are not subject to the glucosylation and deglucosylation processes of the quality-control cycle. By contrast, 'misfolded' proteins represent those that move through the quality control cycle with the same rate constants as normal proteins but are ultimately incapable of folding. In other words, the enzymes of the quality control cycle cannot distinguish these unfoldable proteins from native proteins [22, 39, 40].

The total rate of proteins entering the cycle is defined as $k_{\mathrm{pt}} = k_{\mathrm{p}} + k_{\mathrm{p}}^{*}$ with a misfolded fraction $m_{\mathrm{f}} = k_{\mathrm{p}}^{*}/(k_{\mathrm{p}} + k_{\mathrm{p}}^{*})$ unable to fold. Eq 1 and the corresponding misfolded protein equations are non-dimensionalized by the timescale of glucose trimming for chaperone-bound proteins, setting $k_{\mathrm{r}} = 1$, and by total chaperone number, setting $C_{\mathrm{tot}} = 1$ (see Methods for details). Table 1 summarizes model concentrations, parameters, and performance metrics.

For a given set of rates k_i, the steady state protein concentrations P_i can be found as derived in the Methods. We will use this steady-state solution to evaluate performance, on the assumption that protein production and processing parameters remain constant over timescales much longer than the individual cycle time.

Results

Quality control efficiency and energy input

We begin by considering how the glycoprotein folding system illustrated in Fig 1 is governed by a trade-off between accuracy and speed. On the one hand, the system needs to achieve robust, error-free quality control. On the other hand, it needs to process incoming proteins sufficiently rapidly to keep up with production and avoid accumulation of unfolded proteins in the cell. We quantify system accuracy using the steady-state fraction of foldable proteins that successfully undergo folding rather than degradation,

$$f = \frac{k_{\mathrm{f}}P_{\mathrm{c}}}{k_{\mathrm{p}}} \;.$$ (2)

A higher folding fraction f indicates a more efficient folding process that produces more functional proteins per input of nascent unfolded proteins.

Table 1. Model protein concentrations, parameters, and performance metrics.

Quantity	Description
Steady-state protein concentrations	
P_g	Concentration of foldable proteins with glucose tag allowing chaperone binding. Proteins enter cycle in this state.
P_c	Concentration of foldable proteins bound to chaperone.
P_{cf}	Concentration of folded proteins bound to chaperone.
P	Concentration of foldable proteins lacking a glucose tag and vulnerable to degradation.
P_i^*	Concentration of unfoldable proteins (i.e. P_g^*, P_c^*, and P^*).
P_b	Concentration of 'background' proteins that are not part of the chaperone-binding cycle but can bind chaperone.
P_{cb}	Concentration of chaperone-bound background proteins.
C_{tot}	Total chaperone concentration.
C_A	Concentration of available chaperones.
Parameters	
k_p	Rate of foldable protein production.
k_p^*	Rate of unfoldable protein production.
k_{pt}	Total rate of protein production, $k_p + k_p^*$.
m_f	Fraction of produced proteins that cannot fold, $k_p^*/(k_p + k_p^*)$.
k_c	Rate constant of chaperone binding of glucose-tagged proteins.
k_{-c}	Rate constant of chaperone-bound proteins unbinding from chaperone and retaining glucose tag.
k_r	Rate constant of chaperone-bound proteins unbinding from chaperone and losing glucose tag.
k_{-r}	Rate constant of unfolded proteins without glucose tag binding to chaperone.
k_g	Rate constant of unfolded protein glucosylation.
k_{-g}	Rate constant of glucose removal from unfolded protein.
k_f	Rate constant of foldable protein folding.
k_d	Rate constant of non-glucosylated protein degradation.
Performance metrics	
f	Fraction of foldable proteins that fold.
f_{max}^*	Maximum fraction of foldable proteins that fold with $P_{unfolded} = 1$.
$P_{unfolded}$	Total concentration of unfolded proteins.

A second metric for processing efficiency is the total unfolded protein present in the cycle at steady-state: $P_{unfolded} = P_g + P_g^* + P_c + P_c^* + P + P^*$. Low values of $P_{unfolded}$ correspond to rapid processing of individual nascent proteins that prevents their accumulation in the system. High concentrations of unfolded proteins can lead to protein aggregation, which impede cellular function and health [7]. With a typical influx to the ER of 0.1–1 million proteins per minute in each cell [57], proteins accumulate rapidly if the folding system cannot keep up with production. High protein concentrations also induce ERAD and the unfolded protein response to limit the accumulation of protein aggregates, curtailing the throughput of functional proteins [58].

Overall, we aim to understand how glycoprotein quality control can achieve both efficient shunting of foldable proteins towards folding rather than degradation, and rapid processing that limits the accumulation of unfolded proteins. To assess this interplay, we determine the maximum folding fraction for each fixed value of total unfolded proteins, generating a phase-diagram of achievable values for these two metrics (Fig 2A). For fixed values of the production rate k_{pt}, misfolded fraction m_f folding rate k_f and background protein level P_b, the cycle rate constants k_c, k_{-c}, k_r, k_{-r}, k_g, and k_{-g} are allowed to vary (details in Methods) to map out the
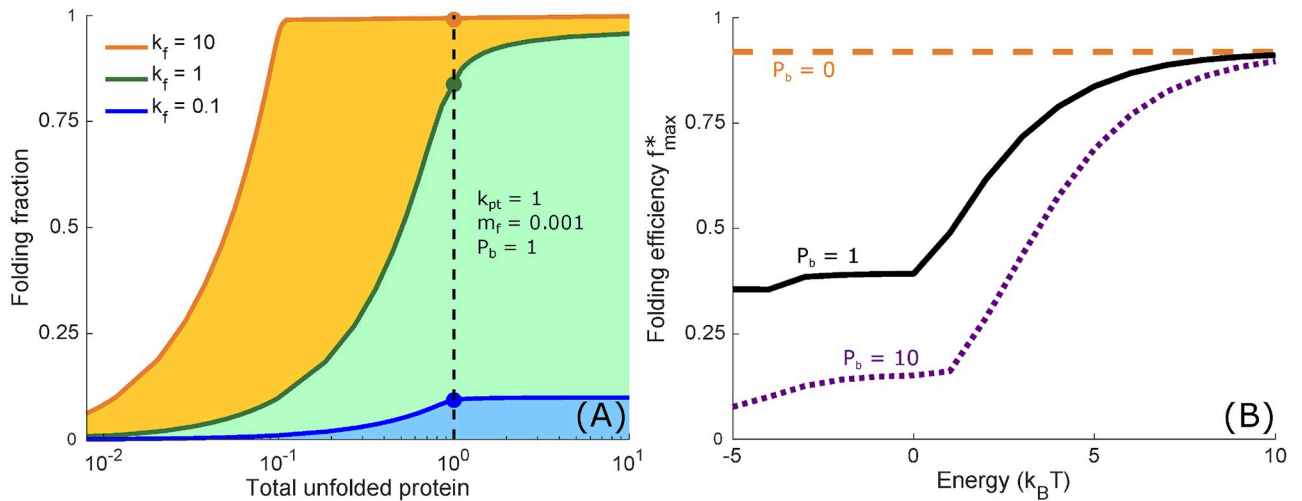
Fig 2. Quality control engenders a trade-off between folding accuracy, speed, and energy. (A) Shaded regions in the phase diagram represent all combinations of folding fraction f and total steady-state unfolded protein $P_{unfolded}$ that can be achieved by varying cycle parameters k_c, k_{-c}, k_r, k_{-r}, k_g, k_{-g}, and k_d while keeping a fixed folding rate k_f, production rate k_{pt}, misfolded fraction m_f, and background protein concentration P_b. Solid lines represent the maximal achievable folding fraction f_{max}. Dots represent the efficiency metric f^*_{max}. (B) Each curve adjusts k_c, k_{-c}, k_r, k_{-r}, k_g, k_{-g}, and k_d to maximize the folding fraction (Eq 2) while the total cycle energy (Eq 3) is varied and the total unfolded protein $P_{unfolded} = P_g + P^*_g + P_c + P^*_c + P + P^*$ is constrained to equal one. Each curve shows a distinct level of background proteins P_b, with fixed $k_{pt} = 0.1$, $k_f = 0.1$, and $m_f = 0.001$ for all curves.

space of accessible efficiency metrics. The curves of maximum folding fraction vs. total unfolded protein represent a Pareto frontier [59] of folding cycle performance, where performance above or to the left of the curves in Fig 2A is not achievable. In Fig 2A, protein production (k_{pt}), misfolded fraction (m_f), and background protein concentration (P_b) are fixed for all curves, and each curve has a different protein folding speed (k_f). Faster folding speeds allow for more efficient folding at each given value for the total unfolded protein. The Pareto frontier has a characteristic shape of an increasing f_{max} at low $P_{unfolded}$, followed by a plateau in f_{max} at high $P_{unfolded}$. These curves demonstrate the trade-off between the two measures for efficient quality control, showing that maximization of folding fraction and minimization of steady-state total unfolded protein cannot be simultaneously achieved.

The characteristic curve shape in Fig 2A for f_{max} vs. $P_{unfolded}$ suggests it is not always feasible to operate the glycoprotein quality control pathway at or near the maximum folding fraction as these high folding fractions can require a very high concentration of unfolded proteins. To assess pathway performance, we choose to limit the total unfolded protein quantity to $P_{unfolded} = 1$, corresponding to a total unfolded protein concentration equal to the concentration of chaperones. We then define the folding efficiency (f^*_{max}) as the maximum folding fraction at $P_{unfolded} = 1$, serving as an overall utility function to evaluate the performance of the glycoprotein quality control pathway. This metric represents the best efficiency that can be achieved by the pathway without accumulating so many unfolded proteins as to overwhelm the binding capacity of the chaperones.

The consensus physiological model of the glycoprotein quality control pathway forms a cycle (Fig 1), with proteins proceeding through the various states in a directed fashion. This directed protein flux requires free-energy dissipation [60], representing a cost to cellular resources. To evaluate the impact of this free-energy dissipation on pathway performance, we consider how the folding efficiency depends on the free energy input, for fixed values of protein production rate k_{pt}, misfolded fraction m_f and protein folding speed k_f. The free energy

driving the quality control cycle is given by [60]

$$E = k_{\text{B}} T \log \frac{k_{\text{c}} k_{\text{r}} k_{\text{g}}}{k_{-\text{c}} k_{-\text{r}} k_{-\text{g}}} \quad . \tag{3}$$

For each value of this driving energy, the cycle rate constants are allowed to vary so as to maximize the folding efficiency f_{max}^*.

Fig 2B shows that the folding efficiency can increase with the cycle driving energy. In the absence of chaperone-binding background proteins ($P_{\text{b}} = 0$), the optimal folding fraction is independent of the energy input into the system, because this energy allows discrimination between proteins that participate in the full cycle (non-background proteins) and proteins that do not participate in the full cycle (background proteins). When $P_{\text{b}} = 0$, there is no need for such discrimination between background and non-background proteins, and the same optimal folding fraction can be reached independent of cycle energy. However, when there are background proteins present ($P_{\text{b}} > 0$), increasing the energy driving the quality control cycle enables more efficient allocation of chaperone resources specifically to foldable rather than background proteins. For example, reducing the rebinding rate of deglucosylated proteins ($k_{-\text{r}}$) would decrease the fraction of chaperones occupied by background proteins. In the extreme limit $k_{-\text{r}} \to 0$, background proteins no longer contribute to the system, and the maximal folding efficiency is achieved. However, fully eliminating binding of unglucosylated proteins would require an infinite energy input to provide a fully irreversible process. S1 Fig explores the extent to which energetic driving enhances folding efficiency under different input conditions.

Comparison of performance between models

A finite driving energy for the quality control cycle implies the presence of reverse processes for all the cycle transitions. We proceed to consider how the presence of the non-physiological reverse transitions for chaperone rebinding $k_{-\text{r}}$ and unbinding $k_{-\text{c}}$ modulate the pathway efficiency.

Fig 3A shows that f_{max}^* monotonically decreases as $k_{-\text{r}}$ increases, for all cases where background proteins are present ($P_{\text{b}} > 0$). This result suggests that removing untagged chaperone binding (i.e. setting $k_{-\text{r}} = 0$) improves the performance of the chaperone cycle, allowing higher folded protein throughput. Removing untagged binding allows only those proteins recognized as foldable to occupy the chaperone. For the moderate level of background proteins assumed here ($P_{\text{b}} = 1$), this effect becomes small when $k_{-\text{r}} < 1$ (corresponding to a rebinding rate smaller than the rate of deglucosylation and chaperone unbinding). However, its importance increases for higher values of P_{b} (see $P_{\text{b}} = 10$ curve in Fig 3A). Removing the untagged rebinding process entirely can protect the quality control system from potential fluctuations in the total levels of untagged background protein that can result in unproductive chaperone occupation. Having demonstrated the detrimental effects of untagged rebinding, we hereafter set $k_{-\text{r}} = 0$, removing this process from the cycle.

We next turn our attention to how quality control efficiency varies with $k_{-\text{c}}$, the rate of protein detachment from the chaperone without removal of the glucose tag. For low production and slow folding rates, the folding fraction is maximized or nearly maximized when $k_{-\text{c}}$ is kept low (Fig 3B). In this regime, it is advantageous for the quality control cycle to operate slowly, and high values of $k_{-\text{c}} \gtrsim 5$ lead to a reduction in the folding fraction by allowing proteins to escape the chaperones before they have a chance to fold. By contrast, at high production and fast folding rates, the folding fraction peaks at an intermediate $k_{-\text{c}}$ value as rapid turnover through the quality control cycle is advantageous. In this regime, low values of $k_{-\text{c}}$ would result
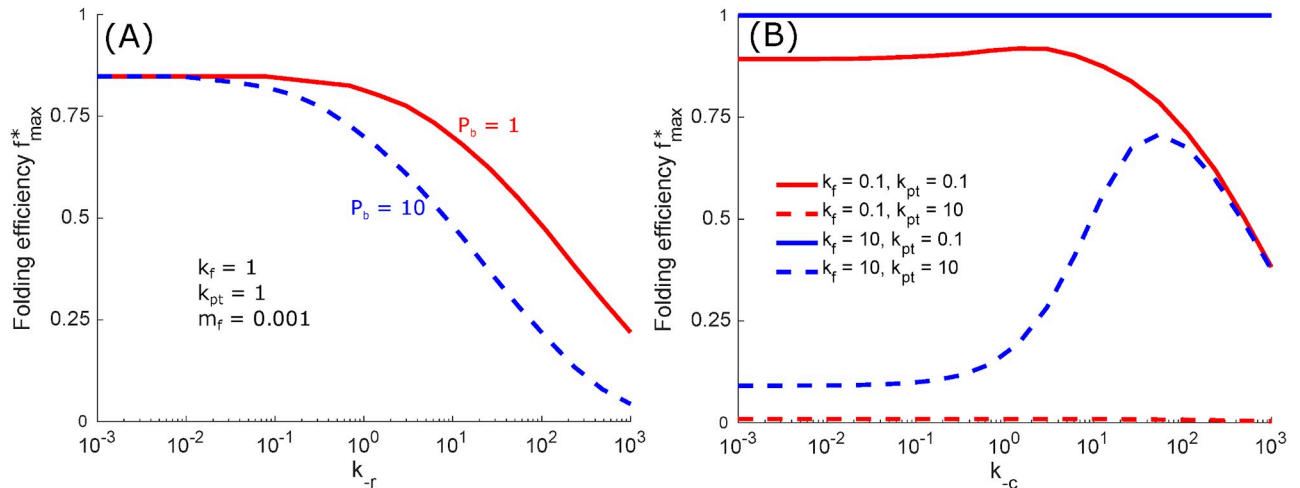
Fig 3. Untagged protein binding and reversible chaperone binding can be disadvantageous. (A) Maximal achievable folding fraction at fixed unfolded protein, $P_{\text{unfolded}} = 1$, plotted versus the untagged rebinding rate k_{-r}, as cycle parameters k_c, k_{-c}, k_r, k_g, k_{-g}, and k_d are free to vary. Other curves show similar behavior when folding and production rates are altered. (B) Folding efficiency is plotted as a function of unbinding rate k_{-c}. Cycle parameters k_c, k_r, k_g, k_{-g}, and k_d are free to vary. Folding rate k_f and production rate k_{pt} are held constant as indicated. The misfolding fraction is set to $m_f = 0.001$.

in the chaperones becoming overloaded by already folded proteins (see S2 Fig), decreasing folding efficiency.

Fig 3B suggests that different rates of chaperone unbinding (k_{-c}) become optimal in different regimes, depending on whether protein production is sufficiently high and folding is sufficiently slow to overwhelm the available quantity of chaperones. Chaperones in the ER, such as BiP, are thought to be present in excess quantities [61–63], to facilitate rapid chaperone binding of nascent proteins. This suggests that the glycoprotein quality control pathway typically operates in the regime of relatively low production $k_{pt} \lesssim 1$, so that protein release from chaperones can keep up with the incoming proteins and the chaperones do not become overwhelmed. At low protein production, Fig 3B shows that raising the chaperone release rate k_{-c} primarily decreases the maximum folding fraction. Removing the ability of a protein to completely detach from a chaperone without glucose trimming should thus improve the performance of the chaperone binding cycle, and folding proteins should remain closely associated with the chaperone until glucose removal. This tight association may have additional functional importance, such as facilitating recruitment of other enzymes important for folding [40, 64] or as a by-product of the high specificity of chaperone-glucose interaction [65].

Fig 3 demonstrates that removing non-specific chaperone binding (k_{-r}) and detachment of proteins from the chaperone without glucose trimming (k_{-c}) improves the performance of the chaperone binding cycle by increasing the maximum folding fraction with a limited accumulation of unfolded protein. The consensus physiological model, with these two processes absent, is thus shown to be more efficient (in the low-production regime) than the full model illustrated in Fig 1.

We now explore further glycoprotein quality control pathway model variations, including those that are not cyclic (Fig 4A). The non-cyclic models include all possible variations of a three-state model that lack untagged binding (no k_{-r}) and are capable of producing a finite steady-state solution. We compare the performance of these models to the consensus physiological model in terms of the efficiency metric f^*_{max}, at varying levels of protein production (Fig 4B).
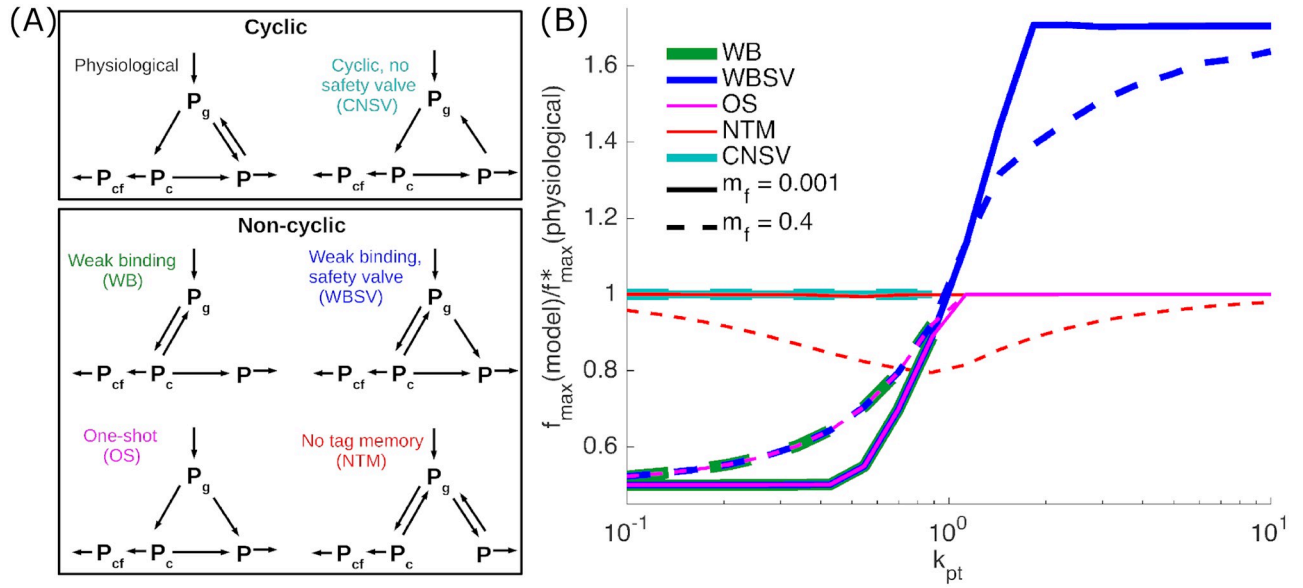
Fig 4. Physiological model for glycoprotein quality control outperforms other models. (A) Schematic of cyclic and non-cyclic models. The physiological model corresponds to the consensus description of the glycoprotein quality control pathway. (B) Ratios of folding efficiency f_{max}^* comparing performance of all models to the physiological model. Curve color indicates the model being compared, with solid lines for $m_f = 0.001$ and dashed lines for $m_f = 0.4$. For all curves, $k_f = 1$.

The WB (weak binding) model allows proteins to bind and unbind from the chaperone, until the glucose tag is removed. The WBSV (weak binding, safety valve) model introduces an additional "safety-valve" pathway where the glucose tag can be removed without chaperone binding. The OS (one shot) model treats chaperone binding and glucose trimming as irreversible, so that each protein only has one chance to attempt folding. These three models share a common feature—they lack the ability to restore a glucose tag once it is removed, irreversibly committing deglucosylated proteins (P) to degradation. Each of these models performs worse than the physiological system in the regime of low protein production ($k_{pt} < 1$), with the folding efficiency f_{max}^* dropping by approximately a factor of 2 (Fig 4B). In the regime of high production, the WBSV model is capable of more effectively funneling proteins into a degradation-committed state, allowing it to significantly outperform the physiological model (Fig 4B). However, as discussed previously, cells are believed to typically operate in a regime of limited protein production levels and excess chaperone capacity, so that we focus largely on model performance at low k_{pt}.

In the physiological model, a deglucosylated protein (P) is more likely to have first passed through chaperone binding than a monoglucosylated protein (P_g). This feature allows glucose moieties to serve as a form of molecular memory—the presence of a glucose tag means the protein is more likely to be newly made; the absence of the tag means the protein is more likely to have already attempted folding. A contrasting non-cyclic model is the NTM (no tag memory) model, which allows chaperone binding and glucose removal to function as independent processes (Fig 4A). When the fraction of misfolded proteins (m_f) is low, the NTM model performs equivalently to the physiological model. However, when a substantial number of proteins entering the quality control cycle are incapable of being folded (high m_f), the NTM model is at a disadvantage to the physiological system (Fig 4B). In the presence of such defective unfoldable proteins, the cyclic addition and removal of glucose tags allows the physiological model to have a memory of which proteins already attempted (and failed) folding and thus should be

made vulnerable to degradation. Overall, the physiological model outperforms all non-cyclic models in the low-production regime.

The cyclic model with no safety valve (CNSV) exhibits the same cycle as the physiological model: of chaperone binding, deglucosylation upon release, and subsequent reglucosylation (Fig 4A). However, it lacks the direct transition from the tagged state P_g to the vulnerable state P. In the absence of this safety valve pathway, the CNSV model matches the performance of the physiological model at low production rates (Fig 4B). However, for $k_{pt} > 1$, proteins cannot be released from the chaperones fast enough to keep up with new protein production, and the CNSV model cannot reach a steady state. In this regime, all chaperones would become clogged with protein and the protein would accumulate indefinitely. A similar behavior is observed for the non-cyclic WB model, which also lacks the safety-valve (Fig 4B).

Performance and robustness of the physiological model

We now explore the performance of the physiological model, as well as the optimal kinetic parameter values under different conditions. Performance is quantified in terms of the folding efficiency f^*_{max} (the maximum folding fraction at a total protein content $P_{unfolded} = 1$). We treat the total production rate k_{pt}, protein folding rate k_f, and misfolding fraction m_f as external input conditions for the system. As always, these rates are expressed relative to the rate of chaperone removal ($k_r = 1$ for non-dimensionalization), which is also treated as fixed. The quality control pathway is then allowed to adjust all other kinetic rate constants to optimize the folding efficiency—the resulting optimal folding fraction and the optimized parameters are plotted in Fig 5.

When the overall production rate is low, the optimal folding fraction approaches one (blue curve in Fig 5A), indicating that nearly all the foldable proteins that enter the quality control cycle are successfully folded. At higher production ($k_p \gtrsim 1$), the removal of proteins from chaperones cannot keep up with the flux of incoming proteins. In this regime, the available chaperones in the system are overwhelmed and the folding efficiency drops.

The optimal parameters (red curves in Fig 5A) describe how the optimized quality control system adjusts to changing production rates. For all conditions explored, binding rate constant (k_c) is always maximized, allowing nascent or reglucosylated proteins to bind to chaperones as quickly as possible. For low k_{pt}, the reglucosylation rate constant k_g is high and the rate constant k_{-g} for glucose removal from free (not chaperone-associated) proteins is low. High k_g and low k_{-g} indicate that the cycle is quickly removing proteins from the vulnerable state P to prevent degradation. This is expected for low protein production (k_{pt}) and low misfolded fraction
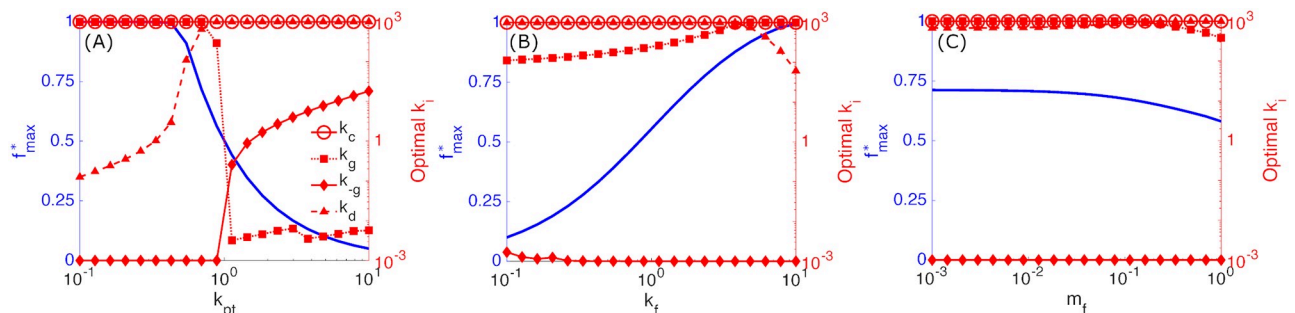
Fig 5. Optimal performance and corresponding parameters. Maximum folding fraction f_{max} at $P_{unfolded} = 1$, (blue curves, left blue vertical axis), and corresponding optimal rate constants, k_i (red curves with markers, right red vertical axis) as cycle conditions are varied for the physiological model. (A) varies protein production rate k_{pt} for fixed $m_f = 10^{-3}$ and $k_f = 1$, (B) varies protein folding rate constant k_f for fixed $m_f = 10^{-3}$ and $k_{pt} = 0.9$, and (C) varies misfolded fraction m_f for fixed $k_{pt} = 0.7$ and $k_f = 1$.

(m_f) as proteins will then usually be provided multiple rounds of chaperone binding. In this regime, the optimal degradation rate (k_d) rises gradually with increasing production in order to maintain a constant amount of unfolded protein $P_{unfolded} = 1$. Eventually (when $k_{pt} \rightarrow 1$) there will not be sufficient chaperones to fold all proteins, and protein degradation must increase sharply to maintain a fixed level of total unfolded protein.

As the production rate passes $k_{pt} \approx 1$, protein reglucusylation (k_g) steeply decreases and glucose removal (k_{-g}) increases. This switch indicates the activation of the 'safety valve' pathway which moves excess proteins directly into the degradation-vulnerable state P to avoid accumulation of unfolded proteins. As protein production continues to increase, glucose removal via k_{-g} further increases to enhance this safety valve. Overall, there are two regimes: low protein production, where chaperones are available and proteins are quickly tagged for chaperone rebinding to prioritize folding; and high protein production, where chaperones are overwhelmed and rapid deglucosylation and degradation is prioritized.

Fig 5B shows how performance and optimal parameters change as protein folding speed k_f is varied. As expected, the folding fraction increases with folding speed. The increased folding speed does not cause significant changes in the optimal parameters, with a modest increase in reglucosylation (k_g) and decreases in glucose removal (k_{-g}) and degradation (k_d) as faster folding frees up chaperones. Fig 5C shows that increasing the misfolded fraction m_f modestly decreases the folding efficiency while leaving optimal parameters largely unchanged.

Overall, Fig 5 demonstrates that maintaining maximum folding efficiency requires large variation in parameters if the protein production level changes, but limited variation in parameters is needed to respond to changes in protein folding speed and misfolded protein fraction. We next proceed to explore how well the cycle can perform under changing production levels if a single fixed parameter set is used across all values of k_{pt}. The goal is to assess the robustness of this quality control system to fluctuation in protein production rates, for the case where other parameters cannot be adjusted sufficiently rapidly to keep up with such changes.

We consider the robustness of a fixed quality control system as follows. The rate constants are optimized to give maximal folding efficiency for a given value of input conditions k_{pt}, k_f, m_f. For those parameters and input conditions, the system gives the highest folding fraction f^*_{max} that maintains a fixed protein content $P_{unfolded} = 1$. When the input production rate k_{pt} is varied and all remaining parameters are held fixed, the folding fraction will decrease below this optimum value (Fig 6A) and the total protein content $P_{unfolded}$ will also change (Fig 6B). The values plotted in Fig 6A and 6B are given relative to the folding fraction and protein content at the point where the system was optimized.

If the parameters are optimized at low protein production ($k_{pt} = 0.1$), the system continues to achieve close to the optimal folding fraction when the protein production is increased (Fig 6A, red curves). However, the total accumulated protein increases by orders of magnitude even for a modest rise in the production rate (Fig 6B, red curves). If the parameters are optimized at high protein production ($k_{pt} = 10$), and the production rate is lowered significantly, then the folding efficiency is reduced to roughly half of the optimal amount and the accumulated protein is also decreased (Fig 6A and 6B, blue curves). A system optimized at intermediate production ($k_{pt} = 1$) exhibits analogous behaviors (Fig 6A and 6B, green curves). If the production rate is lowered, the folded fraction drops below optimal values. If raised, then a massive increase in accumulated protein is observed.

These results highlight a general principle: the quality control cycle can be optimized to operate in one of two regimes: a regime with excess chaperone capacity, and one where the chaperones are overwhelmed. Optimizing for the former requires shutting off the safety-valve, and prioritizing reglucosylation over degradation. Optimizing for the latter requires enhancing degradation and deglucosylation. The transition between the two regimes occurs when the
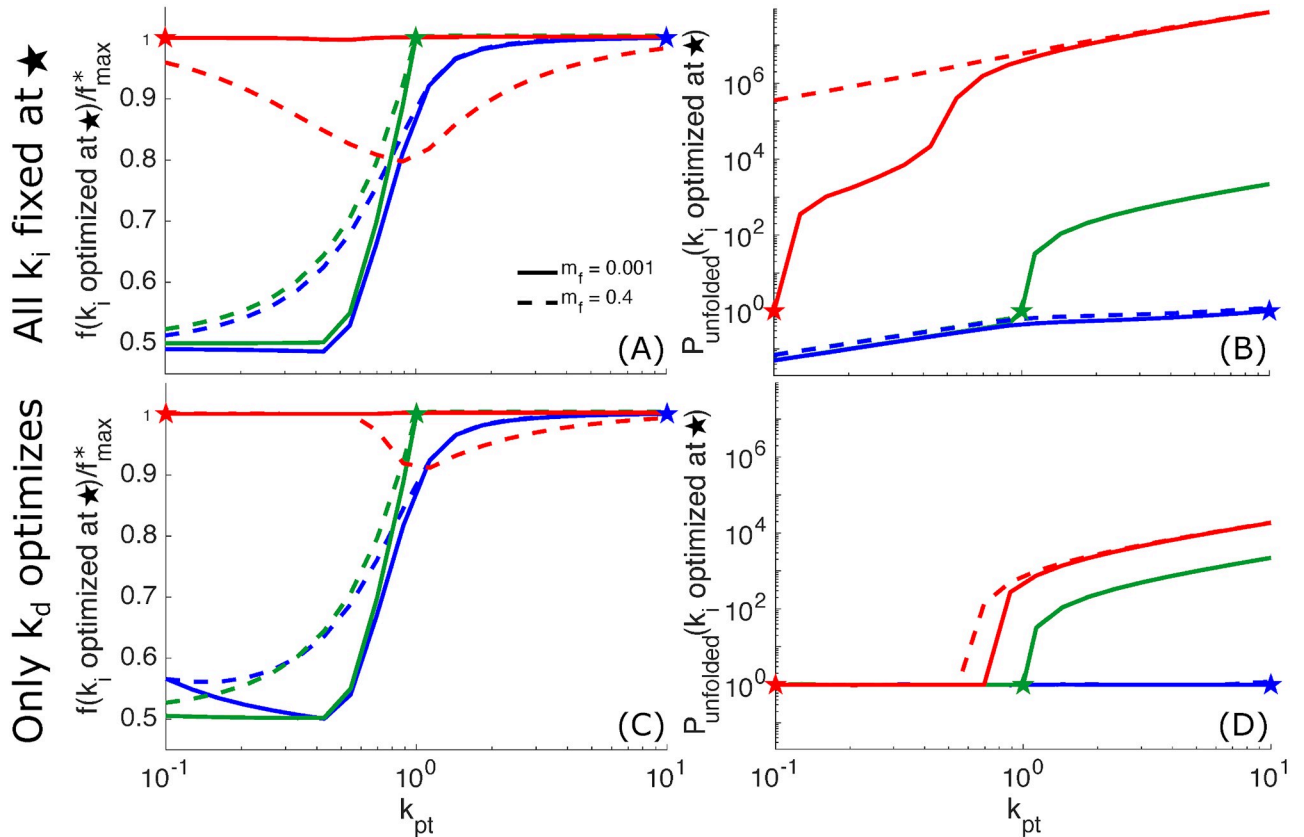
Fig 6. Robustness of quality control cycle to changing production rates. (A) The folding fraction achieved with fixed rate constants is plotted as a fraction of the maximum achievable folding fraction f_{max} for $P_{unfolded} = 1$ as the protein production rate k_{pt} is varied. Fixed rate constants k_c, k_g, k_{-g}, and k_d are those that achieve f_{max} with $m_f = 0.001$ at various protein production levels: $k_{pt} = 0.1$ (red curves and star), $k_{pt} = 1$ (green), and $k_{pt} = 10$ (blue). (B) The $P_{unfolded}$ corresponding to the folding fractions in (A). (C) Analogous plot to (A), with rate constants fixed at the optimal values for specific k_{pt} values, except that the degradation rate constant k_d adjusts to maintain $P_{unfolded} = 1$. If k_d adjustment cannot achieve $P_{unfolded} = 1$, then k_d adjustment minimizes the difference from $P_{unfolded} = 1$. (D) The $P_{unfolded}$ corresponding to the folding fractions in (C).

rate of protein production becomes comparable to the rate at which chaperone-bound proteins detach from chaperones (i.e.: at $k_{pt} = 1$). A system optimized for low production will result in large-scale protein accumulation if the production rate is increased by even a modest amount. A system optimized for high production yields suboptimal folding throughput if shifted to the low-production regime.

Without any flexibility to adjust cycle parameters, the glycoprotein quality control system will perform poorly in one of the two regimes. A natural question is to what extent adjusting a single kinetic parameter will allow the system to compensate for changing production rates and to perform well across a broad range of conditions. Fig 5A shows that the optimal degradation rate (k_d) continuously changes across a range of low protein production levels (k_{pt}), suggesting k_d as a good candidate for an adjustable parameter. Thus we choose to treat the degradation rate k_d as capable of adapting to changing production levels, while all other rate constants in the cycle are held fixed. At each value of the production rate, k_d is adjusted to maintain a total protein content $P_{unfolded} = 1$ whenever possible, with the resulting folding fraction shown in Fig 6C. For a system optimized at low protein production, an adjustable degradation rate allows the optimum folding fraction to be maintained across all production rates.

Even when the fraction of misfolded proteins is increased (dashed curves in Fig 6C), the optimum folding fraction can be maintained up to intermediate production levels.

Strikingly, a system optimized at low protein production can also maintain a fixed total protein $P_{\text{unfolded}} = 1$ up to intermediate production levels ($k_{\text{p}} \lesssim 0.7$) by adjusting the degradation rate k_{d} (Fig 6D). The ability of this system to maintain fixed total protein content over a broad range of low to intermediate production values is in sharp contrast to the rapidly increasing protein levels that arise when all parameters are held fixed (Fig 6B). At higher production rates, there is no value of the degradation rate that can maintain the fixed total protein content and we adjust k_{d} as needed to minimize P_{unfolded}. Allowing k_{d} to adjust in a system optimized for intermediate or high protein production has little impact on both folding fraction and protein accumulation compared to the fully fixed system (Fig 6C and 6D).

This analysis establishes that the quality control pathway can perform well at typical low production rates, yet be capable of adapting to moderate surges in protein production. Such robust behavior requires only for the protein degradation rate to be rapidly adjustable to changing conditions. Other parameters in the quality control cycle can be held constant while allowing near-optimal system performance over an order of magnitude range in protein production. Interestingly, there is evidence that cellular quality control systems do in fact control protein degradation throughput in response to perturbations in protein homeostasis. Specifically, cells maintain a reservoir of ERAD enzymes in ER-associated vesicles that can fuse with the ER lumen in response to an accumulation of unfolded proteins, rapidly upregulating protein degradation [66–68].

Discussion

We have investigated the impact of pathway architecture and kinetic parameters on the performance of the glycoprotein quality control cycle in the endoplasmic reticulum. Two metrics are used to evaluate steady-state performance. The fraction of foldable proteins that are successfully folded measures the accuracy of the system. The total quantity of unfolded proteins measures processing speed, with lower steady-state protein levels corresponding to more rapid processing.

Broadly, we find that a cyclic quality control process, with protein substrates driven in a preferred direction through three quality control states, leads to improved performance. Energy is required for cyclic driving, and increased driving energy per cycle allows higher protein folding fractions (Fig 2B). A higher folding fraction is achieved by eliminating reverse transitions that are absent from the consensus physiological model of the glycoprotein quality control pathway (Fig 3). This matches the directed, cyclic behavior commonly described as occurring for physiological glycoprotein quality control.

'Kinetic proofreading' processes increase biochemical specificity by harnessing cyclic free energy consumption coupled with distinct rates for different substrate types [15, 23, 24]. DNA replication is famously driven out of equilibrium to perform kinetic proofreading that increases replication accuracy [3, 69]. Similar cyclic nonequilibrium processes increase the accuracy of T-cell signaling [4] and sensing of external concentrations [70]. Our work highlights analogous behavior of the glycoprotein quality control cycle as a kinetic proofreading process, consistent with previous descriptions [15, 23, 24].

By exhaustively considering all remaining cyclic and non-cyclic variations of the glycoprotein quality control pathway, we show that the consensus physiological model outperforms all other viable models (Fig 4). Models lacking a 'safety valve', or a path for protein degradation without chaperone binding, will dangerously accumulate unfolded proteins at high protein production levels. This safety valve requirement aligns with the only two-way transition in the

consensus physiological model, which allows glucose tags to be removed from proteins that are not bound the chaperone, facilitating their degradation.

We find that the optimal tuning of the consensus physiological model varies substantially with protein production level (Fig 5A). If the cell must choose a particular set of rate constants, it will either sacrifice folded protein throughput at low protein production levels, or induce massive unfolded protein accumulation at higher production levels (Fig 6A and 6B). A particularly robust system design requires optimizing parameters for low protein production and allowing a single rate constant (the degradation rate) to adapt to changing production levels. Such a system can successfully maintain both maximum folding efficiency and low unfolded protein accumulation across a range of low-to-intermediate production rates (Fig 6C and 6D).

The adjustable degradation rate, which alone can maintain both high folding throughput and low unfolded protein accumulation, corresponds to the dynamic behavior observed for some ERAD enzymes that remove proteins from the ER for degradation. Certain mannosidases, important for ERAD targeting, are largely sequestered to quality control vesicles in the absence of ER stress [66–68]. When the ER becomes stressed (i.e. unfolded proteins accumulate), these mannosidases converge on the ER, rapidly increasing degradation targeting [66–68].

Our results suggest that an adjustable degradation rate, implemented by mannosidase convergence to the ER, can maintain optimal protein folding conditions across a range of relatively low protein folding loads. *In vivo* glycoprotein folding in the ER is thought to largely operate in a low protein production regime, matching the robust system design. Namely, there is excess protein folding capacity in the ER under basal conditions [61], with abundant chaperones that exceed the requirements of the protein folding load [62, 63]. Under conditions of high protein folding load, chaperones are overwhelmed [71, 72]. Gene expression changes are then triggered through a spectrum of pathways collectively termed the unfolded protein response (UPR) [73], to return the system to an effectively low protein folding load.

The timescale of mannosidase convergence to the ER following proteasome inhibition is approximately a couple hours [67], allowing ERAD-mediated degradation adjustments to curtail the accumulation of unfolded proteins relatively quickly. Gene expression stimulated by the UPR is comparatively slow, and starts to take effect in the ER after approximately 5 to 10 hours [61]. The difference in these timescales suggests that for mild unfolded protein stress the earlier ERAD-mediated adjustments may be sufficient without triggering the slower UPR. UPR stimulation, by contrast, may require substantial accumulation of unfolded proteins to overwhelm the basal protein folding capacity [61].

For more severe unfolded protein stress, our results show that simply adjusting degradation is not sufficient. This could occur for a large increase in protein folding load (such as occurs in insulin-producing beta cells [74]), cell differentiation that results in a permanent increase in protein folding load [75], or for cellular stress that inhibits protein folding [56]. The UPR then triggers a broad range of perturbations to restore homeostasis. In particular, it reduces the nascent protein influx into the ER, increases ER chaperone concentrations, upregulates ERAD, modifies glycosylation enzyme levels, and increases the ER size [73, 76, 77]. These actions would affect much of the ER quality control system. Quantitatively addressing how the many facets of the UPR return glycoprotein quality control to homeostasis, both for short-term non-steady-state and long-term adaptive steady-state effects (such as those for differentiating cells [75]), is a promising future endeavor.

In contrast to the large variations in optimal rate constants with protein production level, changes in protein folding speed require relatively little variation to the optimal rate constants (Fig 5B). The glycoprotein quality control pathway must simultaneously process a variety of proteins, which can have folding times ranging from a few minutes to several hours [41]. The

ability of a single pathway to near-optimally process this variety of folding speeds appears to be a strength of its design. The efficiency of protein throughput can approach 100% and range down to 25% or lower for slow-folding proteins or proteins with mutations [41]. Fig 5B suggests that these low efficiencies (ranging down to 25% or lower) are not the result of a poorly-tuned quality control process, but instead that the low efficiencies are an unavoidable consequence of slow folding.

The optimal rate constants also change little with the fraction of produced proteins which are inherently misfolded or unfoldable (Fig 5C). This suggests the design of the quality control pathway is robust to the onset of systematic misfolding, which may arise from translation errors, environmental stress, or mutations [56], so long as the total protein production levels remain relatively unchanged.

Effective quality control of glycoprotein folding in the endoplasmic reticulum ensures an adequate supply of functional natively-folded proteins and limits the accumulation of misfolded proteins. The failure to provide sufficient natively-folded proteins [41] and the formation of misfolded protein aggregates [9] can both contribute to the onset of disease. Our modeling quantitatively demonstrates how the performance of this pathway under a broad range of conditions is modulated by key kinetic parameters that serve as potential targets for pharmacological or genetic perturbations. This quantitative framework serves as a basic foundation for understanding the glycoprotein quality control pathway, which can be further expanded in future work to account for more complex aspects, such as sequential glycan sugar moiety removal [50, 67] and the spatial organization of quality control activities [66].

Methods

Non-dimensionalization. We non-dimensionalize all times by the timescale of protein removal from the chaperone via glucose trimming, k_r^{-1}, and all concentrations by total chaperone concentration, C_{tot}. For conciseness of notation, all kinetic parameters in the text refer to non-dimensionalized values. The dimensionless dynamic equations for foldable proteins are then:

$$\frac{dP_g}{dt} = k_g P + k_{-c} P_c - (k_c C_A + k_{-g}) P_g + k_p \tag{4a}$$

$$\frac{dP_c}{dt} = (k_c P_g + k_{-r} P) C_A - (1 + k_{-c} + k_f) P_c \tag{4b}$$

$$\frac{dP}{dt} = P_c + k_{-g} P_g - (k_g + k_{-r} C_A - k_d) P \tag{4c}$$

$$\frac{dP_{cf}}{dt} = k_f P_c - (1 + k_{-c}) P_{cf} \ , \tag{4d}$$

$$\frac{dP_{cb}}{dt} = k_{-r} C_A P_b - k_r P_{cb} \ . \tag{4e}$$

The dynamics of misfolded proteins are described by

$$\frac{dP_g^*}{dt} = k_g P^* + k_{-c} P_c^* - (k_c C_A + k_{-g}) P_g^* + k_p^* \tag{5a}$$

$$\frac{dP_c^*}{dt} = (k_c P_g^* + k_{-r} P^*) C_A - (1 + k_{-c}) P_c^* \tag{5b}$$

$$\frac{dP^*}{dt} = P_c^* + k_{-g} P_g^* - (k_g + k_{-r} C_A - k_d) P^* . \tag{5c}$$

Note that most rate constants are the same for both foldable and misfolded proteins, except k_p changes to k_p^* to allow different production rates of foldable and misfolded proteins, $k_f^* = 0$ (misfolded proteins cannot fold), and $P_b^* = 0$ (as only a single comprehensive population of background proteins is considered). The available amount of chaperone is $C_A = 1 - P_c - P_{cf} - P_c^* - P_{cb}$, where $C_{tot} = 1$ is the dimensionless total chaperone concentration.

Steady-state solution of chaperone cycle dynamics. Eqs 4 and 5 describe the dynamics of the chaperone folding cycle. In steady state each of the time derivatives must equal zero. By summing together Eqs 4a–4d, we get the steady-state condition for the total flux of foldable proteins through the cycle:

$$k_p = k_f P_c + k_d P . \tag{6}$$

Rearranged, this gives P in terms of parameters and P_c,

$$P = \frac{k_p}{k_d} - \frac{k_f}{k_d} P_c . \tag{7}$$

Applying $dP_{cf}/dt = 0$ gives

$$P_{cf} = \frac{k_f}{k_r + k_{-c}} P_c , \tag{8}$$

and $dP_{cb}/dt = 0$ gives

$$P_{cb} = k_{-r} P_b C_A , \tag{9}$$

where the available chaperone is

$$C_A \equiv 1 - P_c - P_{cf} - P_c^* - P_{cb} \tag{10a}$$

$$= 1 - \left(1 + \frac{k_f}{k_r + k_{-c}}\right) P_c - P_c^* - k_{-r} P_b C_A . \tag{10b}$$

Substituting Eqs 7, 8, and 9 into Eqs 4a and 4b gives

$$\frac{dP_c}{dt} = C_A \left(k_c P_g + \frac{k_{-r} k_p}{k_d} - \frac{k_{-r} k_f}{k_d} P_c\right) - (k_f + 1 + k_{-c}) P_c \tag{11a}$$

$$\frac{dP_g}{dt} = k_p + k_{-c} P_c + \frac{k_g k_p}{k_d} - \frac{k_g k_f}{k_d} P_c - k_{-g} P_g - C_A k_c P_g . \tag{11b}$$

Similarly for misfolded proteins, which have $k_f^* = 0$, the steady state condition for protein fluxes entering and exiting the cycle is

$$P^* = \frac{k_p^*}{k_d} \quad . \tag{12}$$

Substituting Eq 12 into Eqs 5a and 5b gives

$$\frac{dP_c^*}{dt} = C_A \left(k_c P_g^* + \frac{k_{-r} k_p^*}{k_d} \right) - (1 + k_{-c}) P_c^* \tag{13a}$$

$$\frac{dP_g^*}{dt} = k_p^* + k_{-c} P_c^* + \frac{k_g k_p^*}{k_d} - k_{-g} P_g^* - C_A k_c P_g^* \quad . \tag{13b}$$

Eq 11 can be rewritten as $\mathbf{M}\vec{P} = \vec{b}$:

$$\mathbf{M}\vec{P} = \begin{bmatrix} C_A k_c & C_A m_1 + n_1 \\ -C_A k_c - k_{-g} & n_2 \end{bmatrix} \begin{bmatrix} P_g \\ P_c \end{bmatrix} = \begin{bmatrix} C_A b_1 \\ b_2 \end{bmatrix} = \vec{b} \quad , \tag{14}$$

with $m_1 = -k_{-r} k_f / k_d$, $n_1 = -(k_f + 1 + k_{-c})$, $n_2 = k_{-c} - k_g k_f / k_d$, $b_1 = -k_{-r} k_p / k_d$, and $b_2 = -(k_p + k_g k_p / k_d)$. The determinant $|\mathbf{M}| = k_c m_1 C_A^2 + (k_{-g} m_1 + k_c n_1 + k_c n_2) C_A + k_{-g} n_1 = r_2 C_A^2 + r_1 C_A + r_0$. Rearranging gives

$$\begin{bmatrix} P_g \\ P_c \end{bmatrix} = \frac{1}{r_2 C_A^2 + r_1 C_A + r_0} \begin{bmatrix} p_1 C_A + p_0 \\ q_2 C_A^2 + q_1 C_A \end{bmatrix} \quad , \tag{15}$$

with $p_1 = b_1 n_2 - m_1 b_2$, $p_0 = -b_2 n_1$, $q_2 = k_c b_1$, $q_1 = k_{-g} b_1 + k_c b_2$.

Similarly, Eq 13 can be rewritten as $\mathbf{M}^* \vec{P}^* = \vec{b}^*$:

$$\mathbf{M}^* \vec{P}^* = \begin{bmatrix} C_A k_c & n_1^* \\ -C_A k_c - k_{-g} & k_{-c} \end{bmatrix} \begin{bmatrix} P_g^* \\ P_c^* \end{bmatrix} = \begin{bmatrix} C_A b_1^* \\ b_2^* \end{bmatrix} = \vec{b}^* \quad , \tag{16}$$

with $n_1^* = -(1 + k_{-c})$, $b_1^* = -k_{-r} k_p^* / k_d$, and $b_2^* = -k_p^* - k_g k_p^* / k_d$. The determinant $|\mathbf{M}^*| = C_A(k_c k_{-c} + n_1^* k_c) + n_1^* k_{-g} = C_A r_1^* + r_0^*$. Rearranging gives

$$\begin{bmatrix} P_g^* \\ P_c^* \end{bmatrix} = \frac{1}{C_A r_1^* + r_0^*} \begin{bmatrix} p_1^* C_A + p_0^* \\ q_2^* C_A^2 + q_1^* C_A \end{bmatrix} \quad , \tag{17}$$

with $p_1^* = k_{-c} b_1^*$, $p_0^* = -b_2^* n_1^*$, $q_2^* = k_c b_1^*$, and $q_1^* = k_{-g} b_1^* + k_c b_2^*$.

We now insert P_c from Eq 15 and P_c^* from Eq 17 into Eq 10b,

$$
\begin{aligned}
C_A &= 1 - \left(1 + \frac{k_f}{1 + k_{-c}} \right) \frac{q_2 C_A^2 + q_1 C_A}{r_2 C_A^2 + r_1 C_A + r_0} \\
&\quad - \frac{q_2^* C_A^2 + q_1^* C_A}{r_1^* C_A + r_0^*} - \frac{k_{-r}}{k_r} P_b C_A \quad .
\end{aligned}
\tag{18}
$$

Rearranging,

$$
\begin{aligned}
& C_{\mathrm{A}}(1 + k_{-\mathrm{r}}P_{\mathrm{b}})[r_2 r_1^* C_{\mathrm{A}}^3 + (r_2 r_0^* + r_1 r_1^*)C_{\mathrm{A}}^2 \\
+ \quad & (r_1 r_0^* + r_0 r_1^*)C_{\mathrm{A}} + r_0 r_0^*] \\
- \quad & [r_2 r_1^* C_{\mathrm{A}}^3 + (r_2 r_0^* + r_1 r_1^*)C_{\mathrm{A}}^2 + (r_1 r_0^* + r_0 r_1^*)C_{\mathrm{A}} + r_0 r_0^*] \\
+ \quad & [1 + k_{\mathrm{f}}/(k_{-\mathrm{c}} + 1)](q_2 C_{\mathrm{A}}^2 + q_1 C_{\mathrm{A}})(C_{\mathrm{A}} r_1^* + r_0^*) \\
+ \quad & (q_2^* C_{\mathrm{A}}^2 + q_1^* C_{\mathrm{A}})(r_2 C_{\mathrm{A}}^2 + r_1 C_{\mathrm{A}} + r_0) = 0 \ .
\end{aligned}
\tag{19}
$$

This forms a quartic equation for C_{A}, which can be solved with standard root-finding algorithms. Once C_{A} is obtained, Eqs 15 and 17 give steady state P_{c}, P_{g}, P_{c}^*, and P_{g}^*, from which Eq 7 gives steady state P. Eq 12 gives steady state P^* once k_{p}^* and k_{d} are selected, without needing other information.

Optimization of cycle efficiency. For the results in Fig 2A, the maximum folding fraction independent of total unfolded protein was first found by allowing $k_i = k_{\mathrm{c}}$, $k_{-\mathrm{c}}$, k_{g}, $k_{-\mathrm{g}}$, $k_{-\mathrm{r}}$, and k_{d} to vary to maximize the folding fraction using the Matlab routine fmincon, with $k_i \in [10^{-3}, 10^3]$. The minimum total unfolded protein is found using the Matlab routine fmincon for each fixed folding fraction (at a value less than or equal to the maximum folding fraction), constrained with the nonlinear constraints option, and with $k_i \in [10^{-3}, 10^3]$.

For the results in Fig 2B and S1 Fig, $k_i = k_{\mathrm{c}}$, $k_{-\mathrm{c}}$, $k_{-\mathrm{r}}$, k_{g}, $k_{-\mathrm{g}}$, and k_{d} are allowed to vary to maximize the folding fraction (Eq 2), while fixing the energy (Eq 3) at a specific value, and fixing the total unfolded protein $P_{\mathrm{unfolded}} = 1$. The folding fraction maximization was performed using the Matlab routine fmincon with energy and total unfolded protein fixed using the nonlinear constraints option. The k_i were free within the range $k_i \in [10^{-3}, 10^3]$.

The results in Fig 3A are found similarly to those of Fig 2B, with the fixed folding fraction varied using the bisection method until a $P_{\mathrm{unfolded}} \in (0.99, 1.01)$ is found. Results in Figs 3B and 5 and S2 Fig are found with the same method as Fig 3A with the appropriate k_i set to zero and the appropriate k_i allowed to vary within $k_i \in [10^{-3}, 10^3]$. Almost all results in Fig 4 are also found with the method of Figs 3B and 5. The exception in Fig 4 is the no tag memory (NTM) model, which lacks the transition represented with rate constant k_{r}, and instead sets $k_{-\mathrm{c}} = 1$.

The f_{max}^* and optimizing k_i^* at particular k_{pt} in Fig 6 are found with the same method as Figs 3, 4 and 5. The optimal parameters that achieve f_{max}^* are then used as the fixed parameters in Eq 19 to determine the folding fraction and total unfolded protein in Fig 6A and 6B as the protein production is varied. The folding fraction and total unfolded protein in Fig 6C and 6D with only k_{d} free is found by using the bisection method to vary k_{d} to attempt to find a k_{d} value with $P_{\mathrm{unfolded}} = 1$. If $P_{\mathrm{unfolded}} = 1$ cannot be achieved with $k_{\mathrm{d}} \in [10^{-3}, 10^3]$ then $k_{\mathrm{d}} = 10^3$ is chosen to minimize P_{unfolded}.

Supporting information

S1 Fig. Nonequilibrium driving improves performance. Each curve adjusts k_{c}, $k_{-\mathrm{c}}$, k_{r}, $k_{-\mathrm{r}}$, k_{g}, $k_{-\mathrm{g}}$, and k_{d} to maximize the folding fraction (Eq 2) while the total cycle energy (Eq 3) is varied and the total unfolded protein is constrained to $P_{\mathrm{unfolded}} = P_{\mathrm{g}} + P_{\mathrm{g}}^* + P_{\mathrm{c}} + P_{\mathrm{c}}^* + P + P^* = 1$. Other parameters are fixed for each curve as indicated, and background protein levels are set to $P_{\mathrm{b}} = 1$. Colored curves show effects of increasing protein folding speed (red dashed), increasing protein production (blue dotted), and increasing misfolded fraction (green dashed-dotted) relative to the black curve, which is identical to the corresponding curve in Fig 2B. Increased k_{f} leads to a higher folding efficiency, because faster folding can better compete with

degradation, allowing folded proteins to rapidly leave the cycle and free up chaperones. Both high protein production (k_{pt}) and high misfolded fraction (m_f) lead to decreased folding efficiency because fewer chaperones are unoccupied and available for foldable protein binding. (TIF)

S2 Fig. Protein population variation with reversible chaperone binding. Steady-state concentrations of chaperone-bound foldable proteins (P_c, orange), and already-folded proteins (P_{cf}, green) vs. unbinding rate k_{-c}. Two regimes are shown, corresponding to two curves shown in Fig 3B. Solid lines correspond to low production, slow folding. Dashed lines show high production, rapid folding. $m_f = 0.001$ for all curves. Increased k_{-c} allows already-folded proteins to be rapidly removed from chaperones, freeing chaperones to bind other nascent proteins. For fast folding, already-folded proteins (P_{cf}, green dashed line) can occupy a significant fraction of the available chaperones, leading to a decrease in efficiency for low unbinding rates k_{-c}—this occupation of significant chaperone by already unfolded proteins at low k_{-c} is not seen for slow-folding proteins (green solid line). If the unbinding rate becomes much higher than the folding rate, proteins will frequently detach from the chaperone before they can fold, manifesting as low values of P_c (orange lines), with rapid unbinding reducing efficiency at all folding rates. (TIF)

Author Contributions

Conceptualization: Aidan I. Brown, Elena F. Koslover.

Investigation: Aidan I. Brown, Elena F. Koslover.

Methodology: Aidan I. Brown, Elena F. Koslover.

Software: Aidan I. Brown.

Writing – original draft: Aidan I. Brown.

Writing – review & editing: Aidan I. Brown, Elena F. Koslover.

References

1. Murrow L, Debnath J. Autophagy as a stress-response and quality-control mechanism: implications for cell injury and human disease. Annu Rev Pathol. 2013; 8:105–137. https://doi.org/10.1146/annurev-pathol-020712-163918 PMID: 23072311

2. Pohl C, Dikic I. Cellular quality control by the ubiquitin-proteasome system and autophagy. Science. 2019; 366(6467):818–822. https://doi.org/10.1126/science.aax3769 PMID: 31727826

3. Hopfield JJ. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. Proc Natl Acad Sci USA. 1974; 71(10):4135–4139. https://doi.org/10.1073/pnas.71.10.4135 PMID: 4530290

4. McKeithan TW. Kinetic proofreading in T-cell receptor signal transduction. Proc Natl Acad Sci USA. 1995; 92(11):5042–5046. https://doi.org/10.1073/pnas.92.11.5042 PMID: 7761445

5. Hartich D, Barato AC, Seifert U. Nonequilibrium sensing and its analogy to kinetic proofreading. New J Phys. 2015; 17(5):055026. https://doi.org/10.1088/1367-2630/17/5/055026

6. Müller JB, Geyer PE, Colaço AR, Treit PV, Strauss MT, Oroshi M, et al. The proteome landscape of the kingdoms of life. Nature. 2020; 582(7813):592–596. https://doi.org/10.1038/s41586-020-2402-x PMID: 32555458

7. Mogk A, Bukau B, Kampinga HH. Cellular handling of protein aggregates by disaggregation machines. Mol Cell. 2018; 69(2):214–226. https://doi.org/10.1016/j.molcel.2018.01.004 PMID: 29351843

8. Dobson CM. Protein misfolding, evolution and disease. Trends Biochem Sci. 1999; 24(9):329–332. https://doi.org/10.1016/S0968-0004(99)01445-0 PMID: 10470028

9. Soto C. Unfolding the role of protein misfolding in neurodegenerative diseases. Nat Rev Neurosci. 2003; 4(1):49–60. https://doi.org/10.1038/nrn1007 PMID: 12511861

10. Gregersen N, Bross P, Vang S, Christensen JH. Protein misfolding and human disease. Annu Rev Genomics Hum Genet. 2006; 7:103–124. https://doi.org/10.1146/annurev.genom.7.080505.115737 PMID: 16722804

11. Mellenius H, Ehrenberg M. Transcriptional accuracy modeling suggests two-step proofreading by RNA polymerase. Nucleic Acids Res. 2017; 45(20):11582–11593. https://doi.org/10.1093/nar/gkx849 PMID: 29036494

12. Sharma AK, Chowdhury D. Distribution of dwell times of a ribosome: effects of infidelity, kinetic proofreading and ribosome crowding. Phys Biol. 2011; 8(2):026005. https://doi.org/10.1088/1478-3975/8/2/026005 PMID: 21263169

13. Rodrigo-Brenni MC, Hegde RS. Design principles of protein biosynthesis-coupled quality control. Dev Cell. 2012; 23(5):896–907. https://doi.org/10.1016/j.devcel.2012.10.012 PMID: 23153486

14. Wolff S, Weissman JS, Dillin A. Differential scales of protein quality control. Cell. 2014; 157(1):52–64. https://doi.org/10.1016/j.cell.2014.03.007 PMID: 24679526

15. Shao S, Hegde RS. Target selection during protein quality control. Trends Biochem Sci. 2016; 41 (2):124–137. https://doi.org/10.1016/j.tibs.2015.10.007 PMID: 26628391

16. Caramelo JJ, Parodi AJ. A sweet code for glycoprotein folding. FEBS Lett. 2015; 589(22):3379–3387. https://doi.org/10.1016/j.febslet.2015.07.021 PMID: 26226420

17. Adams BM, Oster ME, Hebert DN. Protein quality control in the endoplasmic reticulum. Protein J. 2019; 38(3):317–329. https://doi.org/10.1007/s10930-019-09831-w PMID: 31004255

18. Dwek RA. Glycobiology: toward understanding the function of sugars. Chem Rev. 1996; 96(2):683–720. https://doi.org/10.1021/cr940283b PMID: 11848770

19. Shental-Bechor D, Levy Y. Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. Proc Natl Acad Sci USA. 2008; 105(24):8256–8261. https://doi.org/10.1073/pnas.0801340105 PMID: 18550810

20. Ellgaard L, Molinari M, Helenius A. Setting the standards: quality control in the secretory pathway. Science. 1999; 286(5446):1882–1888. https://doi.org/10.1126/science.286.5446.1882 PMID: 10583943

21. Anelli T, Sitia R. Protein quality control in the early secretory pathway. EMBO J. 2008; 27(2):315–327. https://doi.org/10.1038/sj.emboj.7601974 PMID: 18216874

22. Tannous A, Patel N, Tamura T, Hebert DN. Reglucosylation by UDP-glucose: glycoprotein glucosyltransferase 1 delays glycoprotein secretion but not degradation. Mol Biol Cell. 2015; 26(3):390–405. https://doi.org/10.1091/mbc.E14-08-1254 PMID: 25428988

23. Määttänen P, Gehring K, Bergeron JJ, Thomas DY; Elsevier. Protein quality control in the ER: the recognition of misfolded proteins. Semin Cell Dev Biol. 2010; 21:500–511. https://doi.org/10.1016/j.semcdb.2010.03.006 PMID: 20347046

24. Bieberich E. Synthesis, processing, and function of N-glycans in N-glycoproteins. Adv Neurobiol. 2014; 9:47–70. https://doi.org/10.1007/978-1-4939-1154-7_3 PMID: 25151374

25. Murugan A, Huse DA, Leibler S. Speed, dissipation, and error in kinetic proofreading. Proc Natl Acad Sci USA. 2012; 109(30):12034–12039. https://doi.org/10.1073/pnas.1119911109 PMID: 22786930

26. Aebi M, Bernasconi R, Clerc S, Molinari M. N-glycan structures: recognition and processing in the ER. Trends Biochem Sci. 2010; 35(2):74–82. https://doi.org/10.1016/j.tibs.2009.10.001 PMID: 19853458

27. Molinari M. N-glycan structure dictates extension of protein folding or onset of disposal. Nat Chem Biol. 2007; 3(6):313–320. https://doi.org/10.1038/nchembio880 PMID: 17510649

28. Hebert DN, Foellmer B, Helenius A. Glucose trimming and reglucosylation determine glycoprotein association with calnexin in the endoplasmic reticulum. Cell. 1995; 81(3):425–433. https://doi.org/10.1016/0092-8674(95)90395-X PMID: 7736594

29. Lamriben L, Graham JB, Adams BM, Hebert DN. N-Glycan-based ER Molecular Chaperone and Protein Quality Control System: The Calnexin Binding Cycle. Traffic. 2016; 17(4):308–326. https://doi.org/10.1111/tra.12358 PMID: 26676362

30. Shenkman M, Lederkremer GZ. Compartmentalization and selective tagging for disposal of misfolded glycoproteins. Trends Biochem Sci. 2019. https://doi.org/10.1016/j.tibs.2019.04.012 PMID: 31133362

31. Bosis E, Nachliel E, Cohen T, Takeda Y, Ito Y, Bar-Nun S, et al. Endoplasmic reticulum glucosidase II is inhibited by its end products. Biochemistry. 2008; 47(41):10970–10980. https://doi.org/10.1021/bi801545d PMID: 18803404

32. Zapun A, Petrescu SM, Rudd PM, Dwek RA, Thomas DY, Bergeron JJ. Conformation-independent binding of monoglucosylated ribonuclease B to calnexin. Cell. 1997; 88(1):29–38. https://doi.org/10.1016/S0092-8674(00)81855-3 PMID: 9019402

33. Schrag JD, Bergeron JJ, Li Y, Borisova S, Hahn M, Thomas DY, et al. The structure of calnexin, an ER chaperone involved in quality control of protein folding. Mol Cell. 2001; 8(3):633–644. https://doi.org/10. 1016/S1097-2765(01)00318-5 PMID: 11583625

34. Totani K, Ihara Y, Matsuo I, Ito Y. Substrate specificity analysis of endoplasmic reticulum glucosidase II using synthetic high mannose-type glycans. J Biol Chem. 2006; 281(42):31502–31508. https://doi.org/ 10.1074/jbc.M605457200 PMID: 16940048

35. Vanoni O, Paganetti P, Molinari M. Consequences of individual N-glycan deletions and of proteasomal inhibition on secretion of active BACE. Mol Biol Cell. 2008; 19(10):4086–4098. https://doi.org/10.1091/ mbc.E08-05-0459 PMID: 18632981

36. Koslover EF, de la Rosa MAD, Spakowitz AJ. Theoretical and computational modeling of target-site search kinetics in vitro and in vivo. Biophys J. 2011; 101(4):856–865. https://doi.org/10.1016/j.bpj.2011. 06.066 PMID: 21843476

37. Schnitzer MJ, Visscher K, Block SM. Force production by single kinesin motors. Nat Cell Biol. 2000; 2(10):718–723. https://doi.org/10.1038/35036345 PMID: 11025662

38. Budnik A, Stephens DJ. ER exit sites–localization and control of COPII vesicle formation. FEBS Lett. 2009; 583(23):3796–3803. https://doi.org/10.1016/j.febslet.2009.10.038 PMID: 19850039

39. Kiuchi T, Izumi M, Mukogawa Y, Shimada A, Okamoto R, Seko A, et al. Monitoring of Glycoprotein Quality Control System with a Series of Chemically Synthesized Homogeneous Native and Misfolded Glycoproteins. J Am Chem Soc. 2018; 140(50):17499–17507. PMID: 30475607

40. Tannous A, Pisoni GB, Hebert DN, Molinari M; Elsevier. N-linked sugar-regulated protein folding and quality control in the ER. Semin Cell Dev Biol. 2015; 41:79–89. https://doi.org/10.1016/j.semcdb.2014. 12.001 PMID: 25534658

41. Hebert DN, Molinari M. In and out of the ER: protein folding, quality control, degradation, and related human diseases. Physiol Rev. 2007; 87(4):1377–1408. https://doi.org/10.1152/physrev.00050.2006 PMID: 17928587

42. Brambilla Pisoni G, Molinari M. Five questions (with their Answers) on ER-associated degradation. Traffic. 2016; 17(4):341–350. https://doi.org/10.1111/tra.12373

43. Spiro RG, Zhu Q, Bhoyroo V, Söling HD. Definition of the lectin-like properties of the molecular chaperone, calreticulin, and demonstration of its copurification with endomannosidase from rat liver Golgi. J Biol Chem. 1996; 271(19):11588–11594. https://doi.org/10.1074/jbc.271.19.11588 PMID: 8626722

44. Sousa MC, Ferrero-Garcia MA, Parodi AJ. Recognition of the oligosaccharide and protein moieties of glycoproteins by the UDP-Glc: glycoprotein glucosyltransferase. Biochemistry. 1992; 31(1):97–105. https://doi.org/10.1021/bi00116a015 PMID: 1531024

45. Quan EM, Kamiya Y, Kamiya D, Denic V, Weibezahn J, Kato K, et al. Defining the glycan destruction signal for endoplasmic reticulum-associated degradation. Mol Cell. 2008; 32(6):870–877. https://doi. org/10.1016/j.molcel.2008.11.017 PMID: 19111666

46. Christianson JC, Shaler TA, Tyler RE, Kopito RR. OS-9 and GRP94 deliver mutant α1-antitrypsin to the Hrd1–SEL1L ubiquitin ligase complex for ERAD. Nat Cell Biol. 2008; 10(3):272–282. https://doi.org/10. 1038/ncb1689 PMID: 18264092

47. Hosokawa N, Kamiya Y, Kamiya D, Kato K, Nagata K. Human OS-9, a lectin required for glycoprotein endoplasmic reticulum-associated degradation, recognizes mannose-trimmed N-glycans. J Biol Chem. 2009; 284(25):17061–17068. https://doi.org/10.1074/jbc.M809725200 PMID: 19346256

48. Clerc S, Hirsch C, Oggier DM, Deprez P, Jakob C, Sommer T, et al. Htm1 protein generates the N-glycan signal for glycoprotein degradation in the endoplasmic reticulum. J Cell Biol. 2009; 184(1):159–172. https://doi.org/10.1083/jcb.200809198 PMID: 19124653

49. Avezov E, Frenkel Z, Ehrlich M, Herscovics A, Lederkremer GZ. Endoplasmic reticulum (ER) mannosidase I is compartmentalized and required for N-glycan trimming to Man5–6GlcNAc2 in glycoprotein ER-associated degradation. Mol Biol Cell. 2008; 19(1):216–225. https://doi.org/10.1091/mbc.E07-05-0505 PMID: 18003979

50. Shenkman M, Ron E, Yehuda R, Benyair R, Khalaila I, Lederkremer GZ. Mannosidase activity of EDEM1 and EDEM2 depends on an unfolded state of their glycoprotein substrates. Commun Biol. 2018; 1(1):1–11. https://doi.org/10.1038/s42003-018-0174-8

51. Frenkel Z, Gregory W, Kornfeld S, Lederkremer GZ. Endoplasmic reticulum-associated degradation of mammalian glycoproteins involves sugar chain trimming to Man6–5GlcNAc2. J Biol Chem. 2003; 278(36):34119–34124. https://doi.org/10.1074/jbc.M305929200 PMID: 12829701

52. Ellgaard L, Helenius A. Quality control in the endoplasmic reticulum. Nat Rev Mol Cell Bio. 2003; 4(3):181–191. https://doi.org/10.1038/nrm1052 PMID: 12612637

53. Ermonval M, Kitzmüller C, Mir AM, Cacan R, Ivessa NE. N-glycan structure of a short-lived variant of ribophorin I expressed in the MadIA214 glycosylation-defective cell line reveals the role of a

mannosidase that is not ER mannosidase I in the process of glycoprotein degradation. Method Enzymol. 2001; 11(7):565–576.

54. Olivari S, Cali T, Salo KE, Paganetti P, Ruddock LW, Molinari M. EDEM1 regulates ER-associated degradation by accelerating de-mannosylation of folding-defective polypeptides and by inhibiting their covalent aggregation. Biochem Bioph Res Co. 2006; 349(4):1278–1284. https://doi.org/10.1016/j.bbrc.2006.08.186 PMID: 16987498

55. Caramelo JJ, Parodi AJ. Getting in and out from calnexin/calreticulin cycles. J Biol Chem. 2008; 283 (16):10221–10225. https://doi.org/10.1074/jbc.R700048200 PMID: 18303019

56. Tyedmers J, Mogk A, Bukau B. Cellular strategies for controlling protein aggregation. Nat Rev Mol Cell Bio. 2010; 11(11):777–788. https://doi.org/10.1038/nrm2993 PMID: 20944667

57. Karagöz GE, Acosta-Alvear D, Walter P. The unfolded protein response: detecting and responding to fluctuations in the protein-folding capacity of the endoplasmic reticulum. Cold Spring Harb Perspect Biol. 2019; 11(9):a033886. https://doi.org/10.1101/cshperspect.a033886 PMID: 30670466

58. Hwang J, Qi L. Quality Control in the Endoplasmic Reticulum: Crosstalk between ERAD and UPR pathways. Trends Biochem Sci. 2018; 43(8):593–605. https://doi.org/10.1016/j.tibs.2018.06.005 PMID: 30056836

59. Shoval O, Sheftel H, Shinar G, Hart Y, Ramote O, Mayo A, et al. Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. Science. 2012; 336(6085):1157–1160. https://doi.org/10.1126/science.1217405 PMID: 22539553

60. Brown AI, Sivak DA. Theory of nonequilibrium free energy transduction by molecular machines. Chem Rev. 2019. PMID: 31411455

61. Bakunts A, Orsi A, Vitale M, Cattaneo A, Lari F, Tade L, et al. Ratiometric sensing of BiP-client versus BiP levels by the unfolded protein response determines its signaling amplitude. eLife. 2017; 6:e27518. https://doi.org/10.7554/eLife.27518 PMID: 29251598

62. Crofts AJ, Leborgne-Castel N, Pesca M, Vitale A, Denecke J. BiP and calreticulin form an abundant complex that is independent of endoplasmic reticulum stress. Plant Cell. 1998; 10(5):813–823. https://doi.org/10.1105/tpc.10.5.813 PMID: 9596639

63. Kopp MC, Larburu N, Durairaj V, Adams CJ, Ali MM. UPR proteins IRE1 and PERK switch BiP from chaperone to ER stress sensor. Nat Struct Mol Biol. 2019; 26(11):1053–1062. https://doi.org/10.1038/s41594-019-0324-9 PMID: 31695187

64. Frickel EM, Riek R, Jelesarov I, Helenius A, Wüthrich K, Ellgaard L. TROSY-NMR reveals interaction between ERp57 and the tip of the calreticulin P-domain. Proc Natl Acad Sci USA. 2002; 99(4):1954–1959. https://doi.org/10.1073/pnas.042699099 PMID: 11842220

65. Eaton BE, Gold L, Zichi DA. Let's get specific: the relationship between specificity and affinity. Chem Biol. 1995; 2(10):633–638. https://doi.org/10.1016/1074-5521(95)90023-3 PMID: 9383468

66. Benyair R, Ogen-Shtern N, Lederkremer GZ; Elsevier. Glycan regulation of ER-associated degradation through compartmentalization. Semin Cell Dev Biol. 2015; 41:99–109. https://doi.org/10.1016/j.semcdb.2014.11.006 PMID: 25460542

67. Benyair R, Ogen-Shtern N, Mazkereth N, Shai B, Ehrlich M, Lederkremer GZ. Mammalian ER mannosidase I resides in quality control vesicles, where it encounters its glycoprotein substrates. Mol Biol Cell. 2015; 26(2):172–184. https://doi.org/10.1091/mbc.E14-06-1152 PMID: 25411339

68. Ogen-Shtern N, Avezov E, Shenkman M, Benyair R, Lederkremer GZ. Mannosidase IA is in quality control vesicles and participates in glycoprotein targeting to ERAD. J Mol Biol. 2016; 428(16):3194–3205. https://doi.org/10.1016/j.jmb.2016.04.020 PMID: 27108681

69. Ninio J. Kinetic amplification of enzyme discrimination. Biochimie. 1975; 57(5):587–595. https://doi.org/10.1016/S0300-9084(75)80139-8 PMID: 1182215

70. Mehta P, Schwab DJ. Energetic costs of cellular computation. Proc Natl Acad Sci USA. 2012; 109 (44):17978–17982. https://doi.org/10.1073/pnas.1207814109 PMID: 23045633

71. Bergmann TJ, Molinari M. Three branches to rule them all? UPR signalling in response to chemically versus misfolded proteins-induced ER stress. Biol Cell. 2018; 110(9):197–204. https://doi.org/10.1111/boc.201800029 PMID: 29979817

72. Oikonomou C, Hendershot LM. Disposing of misfolded ER proteins: A troubled substrate's way out of the ER. Mol Cell Endocrinol. 2020; 500:110630. https://doi.org/10.1016/j.mce.2019.110630 PMID: 31669350

73. Hetz C. The unfolded protein response: controlling cell fate decisions under ER stress and beyond. Nat Rev Mol Cell Bio. 2012; 13(2):89–102. https://doi.org/10.1038/nrm3270 PMID: 22251901

74. Wang S, Kaufman RJ. The impact of the unfolded protein response on human disease. J Cell Biol. 2012; 197(7):857–867. https://doi.org/10.1083/jcb.201110131 PMID: 22733998

75. Federovitch CM, Ron D, Hampton RY. The dynamic ER: experimental approaches and current questions. Curr Opin Cell Biol. 2005; 17(4):409–414. https://doi.org/10.1016/j.ceb.2005.06.010 PMID: 15975777

76. Schröder M, Kaufman RJ. The mammalian unfolded protein response. Annu Rev Biochem. 2005; 74:739–789. https://doi.org/10.1146/annurev.biochem.73.011303.074134 PMID: 15952902

77. Prados MB, Caramelo JJ, Miranda SE. Progesterone regulates the expression and activity of two mouse isoforms of the glycoprotein folding sensor UDP-Glc: Glycoprotein glucosyltransferase (UGGT). Biochim Biophys Acta. 2013; 1833(12):3368–3374. https://doi.org/10.1016/j.bbamcr.2013.09.022 PMID: 24140206