



Optimal control strategies for inhibition of protein aggregation

Thomas C. T. Michaels^{a,1}, Christoph A. Weber^{a,1}, and L. Mahadevan^{a,b,c,2}

^aSchool of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138; ^bDepartment of Physics, Harvard University, Cambridge, MA 02138; and ^cDepartment of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138

Edited by William A. Eaton, National Institutes of Health, Bethesda, MD, and approved June 3, 2019 (received for review March 08, 2019)

Protein aggregation has been implicated in many medical disorders, including Alzheimer's and Parkinson's diseases. Potential therapeutic strategies for these diseases propose the use of drugs to inhibit specific molecular events during the aggregation process. However, viable treatment protocols require balancing the efficacy of the drug with its toxicity, while accounting for the underlying events of aggregation and inhibition at the molecular level. To address this key problem, we combine here protein aggregation kinetics and control theory to determine optimal protocols that prevent protein aggregation via specific reaction pathways. We find that the optimal inhibition of primary and fibril-dependent secondary nucleation require fundamentally different drug administration protocols. We test the efficacy of our approach on experimental data for the aggregation of the amyloid- β (1-42) peptide of Alzheimer's disease in the model organism *Caenorhabditis elegans*. Our results pose and answer the question of the link between the molecular basis of protein aggregation and optimal strategies for inhibiting it, opening up avenues for the design of rational therapies to control pathological protein aggregation.

amyloids | optimal control | protein aggregation

Over 50 current human diseases, including Alzheimer's disease, Parkinson's disease, and type II diabetes, are intimately connected with the aggregation of precursor peptides and proteins into pathological fibrillar structures known as amyloids (1–5). However, the development of effective therapeutics to prevent protein aggregation-related diseases has been very challenging, in part, due to the complex nature of the aggregation process itself, which involves several microscopic events operating at multiple timescales (6–8).^{*} A promising and recent approach is the use of molecular inhibitors designed to target selectively different types of aggregate species, including the mature amyloid fibrils, or the intermediate oligomeric species, and, in this manner, interfere directly with specific microscopic steps of aggregation (9–12). Examples of such compounds include small chemical molecules, such as the anticancer drug Bexarotene (10), molecular chaperones (13, 14), antibodies, or other organic or inorganic nanoparticles (15). Just as large quantities of the aggregates are toxic, in large doses the inhibitors themselves are also toxic, suggesting the following questions: what is the optimal control strategy (dose of inhibitor and timing of its administration) for the inhibition of aggregation that arises from a balance between the degree of inhibition and the toxicity of the inhibitor? Furthermore, most importantly, how does this optimal control strategy depend on the detailed molecular pathways involved in aggregation and its inhibition?

To address these questions, we combine kinetic theory of protein aggregation (16) with control theory (17) to devise optimal treatment protocols that emerge directly from an understanding of the molecular basis of aggregation and its inhibition. To test our theory, we consider the example of the inhibition of amyloid- β (1-42) ($A\beta_{42}$) aggregation by 2 compounds, Bexarotene (10) and DesAb_{29–35} (15), that selectively target different microscopic events of aggregation and qualitatively

confirm the theoretically predicted efficacy of the drug protocol in a model organism, *Caenorhabditis elegans*.

Results

Kinetic Theory of Protein Aggregation Inhibition. The microscopic mechanisms of irreversible protein aggregation involve a number of steps (Fig. 1A), including primary nucleation, followed by fibril elongation (18). Once a critical quantity of fibrils is formed, however, aggregation is accelerated by secondary nucleation pathways, where the rate of formation of new aggregates depends on the existing aggregate population, leading to exponential growth (19–25); examples of such secondary nucleation pathways include fibril fragmentation (19) and surface-catalyzed secondary nucleation (20–25), which is active in $A\beta_{42}$ aggregation (22). The combined action of these diverse microscopic aggregation mechanisms on the concentration $f(t, j)$ of aggregates of size j at time t can be quantified via a master equation (SI Appendix, Eq. S4 and subsequent discussion) (14, 16):

$$\begin{aligned} \frac{df(t, j)}{dt} = & 2k_+ M_m(t)f(t, j-1) - 2k_+ M_m(t)f(t, j) \\ & + 2k_- \sum_{i=j+1}^{\infty} f(t, i) - k_-(j-1)f(t, j) \\ & + k_1 M_m(t)^{n_1} \delta_{j, n_1} + k_2 M_m(t)^{n_2} \delta_{j, n_2} \sum_{i=n_2}^{\infty} if(t, i), \end{aligned} \quad [1]$$

Significance

A range of medical conditions, such as Alzheimer's disease, Parkinson's disease, and type II diabetes, are linked to protein aggregation. Thus, it is imperative to develop effective therapeutic strategies to combat protein aggregation. Here, we lay out a general approach for optimizing inhibition strategies based on small molecules that suppress nucleation or growth of aggregates. Our model reveals that the optimal timing of drug administration crucially depends on whether the compound inhibits primary nucleation, secondary nucleation, or the growth of aggregates. This approach could guide the rational design of therapeutic strategies to target protein aggregation diseases.

Author contributions: L.M. conceived the research and approach; T.C.T.M. and C.A.W. performed research; T.C.T.M., C.A.W., and L.M. contributed new reagents/analytic tools; T.C.T.M., C.A.W., and L.M. analyzed data; and T.C.T.M., C.A.W., and L.M. wrote the paper.

This article is a PNAS Direct Submission.

Published under the PNAS license.

¹T.C.T.M. and C.A.W. contributed equally to this work.

²To whom correspondence may be addressed. Email: lmahadev@g.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1904090116/-DCSupplemental.

Published online June 28, 2019.

^{*}We note that another key step in the development of Alzheimer's disease in humans, which precedes the aggregation process, is cleavage of amyloid precursor protein (APP) (8).

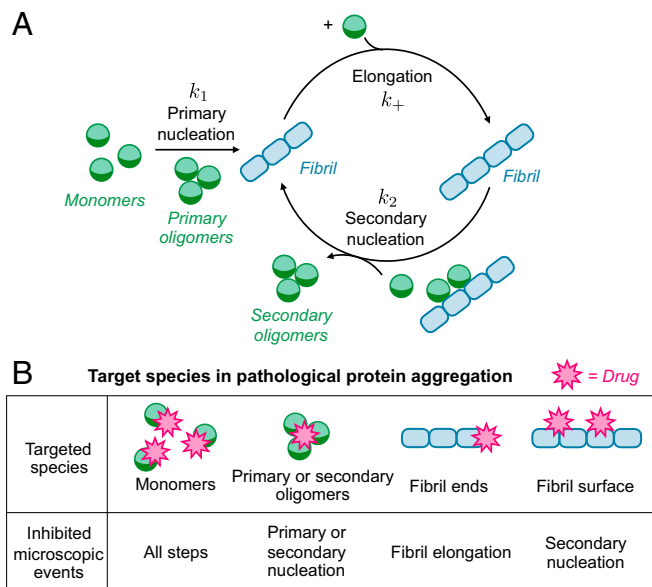


Fig. 1. Elementary molecular events of pathological protein aggregation and the diversity of mechanisms by which a drug can inhibit protein aggregation. (A) Fibrillar aggregates are formed through an initial primary nucleation step followed by elongation. Once a critical concentration of aggregates is reached, secondary nucleation (in the form of fragmentation or, as illustrated in the figure here, surface-catalyzed secondary nucleation) introduces a positive feedback cycle leading to exponential growth of aggregate concentration. (B) A drug can bind monomers; in addition, it can bind primary or secondary oligomers to inhibit primary or surface-catalyzed secondary nucleation. Alternatively, the drug can bind to the fibril ends or the fibril surface to suppress elongation, fragmentation or surface-catalyzed secondary nucleation.

where $M_m(t)$ is the monomer concentration; k_1, k_+, k_-, k_2 are, respectively, the rate constants for primary nucleation, elongation, fragmentation, and surface-catalyzed secondary nucleation; and n_1 and n_2 are the reaction orders of the primary and secondary nucleation steps. Summation of Eq. 1 over aggregate size j leads to a set of moment equations (SI Appendix, Eq. S7) for key experimental observables, including the total number concentration of aggregates $c_a(t) = \sum_j f(t, j)$. Solutions to such moment equations lead to a characteristic sigmoidal profile for the aggregate number concentration, with an initial lag phase followed by a saturation phase due to monomer depletion (SI Appendix, Fig. S2). During the initial lag phase, the monomer concentration is approximately constant, $M_m(t) \approx M_m^{\text{tot}}$, where M_m^{tot} is the total concentration of monomers. It can be shown (SI Appendix, Eq. S12) that, in this limit, the number concentration of aggregates increases exponentially with time (positive feedback), $c_a(t) \simeq (\alpha_0/\kappa_0) e^{\kappa_0 t}$ (25), where $\alpha_0 = k_1 (M_m^{\text{tot}})^{n_1}$ is the rate of generation of new aggregates through primary nucleation and $\kappa_0 = \sqrt{2k_+ M_m^{\text{tot}} [k_2 (M_m^{\text{tot}})^{n_2} + k_-]}$ is an effective aggregate proliferation rate arising from the combined effect of aggregate growth and multiplication through the secondary nucleation pathways. In the context of inhibiting protein aggregation, a key interest is to block this positive feedback mechanism observed during the early-time exponential growth of aggregates; we will thus focus on the early stages of aggregation (rather than on the saturation phase) and assume a constant concentration for the available soluble monomers throughout. This constant-monomer concentration scenario may also be relevant in vivo, where the monomeric protein concentration is likely to be maintained at constant levels by the action of external mechanisms such as protein synthesis (26).

Protein aggregation kinetics can be inhibited in its onset or progression by the presence of a drug through 5 pathways (Fig. 1B) (14): 1) binding to free monomers, 2) binding to oligomers produced by primary nucleation (primary oligomers), 3) binding to oligomers generated by secondary nucleation (secondary oligomers), 4) binding to aggregate ends to block elongation, and 5) binding to the fibril surface to suppress fragmentation or block the production of toxic species through surface-catalyzed secondary nucleation. Since the progression of aggregation is relatively slow compared with the binding rate of drugs, an explicit treatment of the full nonlinear master equation in the presence of a drug shows that, in the limit of constant monomer concentration, the aggregate number concentration $c_a(t)$ satisfies (see SI Appendix, section S2 for a derivation)

$$\frac{dc_a(t)}{dt} = \alpha(c_d) + \kappa(c_d) c_a(t), \quad [2a]$$

where the drug concentration c_d affects the rate parameters according to

$$\alpha(c_d) = \alpha_0 \left(\frac{1}{1 + K_m^{\text{eq}} c_d} \right)^{n_1} \left(\frac{1}{1 + K_{\text{olig},1}^{\text{eq}} c_d} \right), \quad [2b]$$

$$\kappa(c_d) = \kappa_0 \left(\frac{1}{1 + K_m^{\text{eq}} c_d} \right)^{\frac{n_2}{2}} \left(\frac{1}{1 + K_{\text{ends}}^{\text{eq}} c_d} \right)^{\frac{1}{2}} \left(\frac{1}{1 + K_{\text{surf}}^{\text{eq}} c_d} \right)^{\frac{1}{2}} \left(\frac{1}{1 + K_{\text{olig},2}^{\text{eq}} c_d} \right). \quad [2c]$$

Note that the kinetic equation for aggregate concentration and the drug-dependent rate parameters (Eq. 2) can be explicitly derived from a microscopic description of aggregation inhibition through a nonlinear master equation describing the time evolution of the entire aggregate size distribution. They provide a link between microscopic mechanisms of aggregation and inhibition to macroscopic aggregation measurements. The complex interplay between the multiple aggregation pathways and the drug is captured explicitly by renormalized kinetic parameters $\alpha(c_d)$ and $\kappa(c_d)$, which depend on the drug concentration c_d and are specific functions of the kinetic parameters of aggregation as well as the equilibrium binding constant of the drug to the targeted species, K_x^{eq} . Here, x is a placeholder for the target species and the respective pathway, i.e., monomers (m), primary or secondary oligomers (olig,1 and olig,2), fibril ends (ends), and fibril surface sites (surf). In Eq. 2, we have focused on the total aggregate particle concentration; it has, however, been shown that low molecular weight oligomers are key cytotoxic species linked to protein aggregation (27–29). To account for this situation, in SI Appendix, Eq. S15, we show that, in the constant-monomer concentration limit, a linear proportionality relationship links $c_a(t)$ to the concentration of oligomers. Thus, after appropriate rescaling of concentration, the same Eq. 2 can be used to describe oligomeric populations as well. Throughout this paper, we thus use the generic term “aggregate” to refer to the relevant population of toxic aggregate species.

Optimal Control of Protein Aggregation. To find the optimal therapeutic treatment that inhibits the formation of toxic aggregate species requires a cost functional that balances aggregate toxicity against drug toxicity:

$$C = \text{Cost} [c_a(t), c_d(t)] = \int_0^T dt \mathcal{L} (c_a(t), c_d(t)), \quad [3]$$

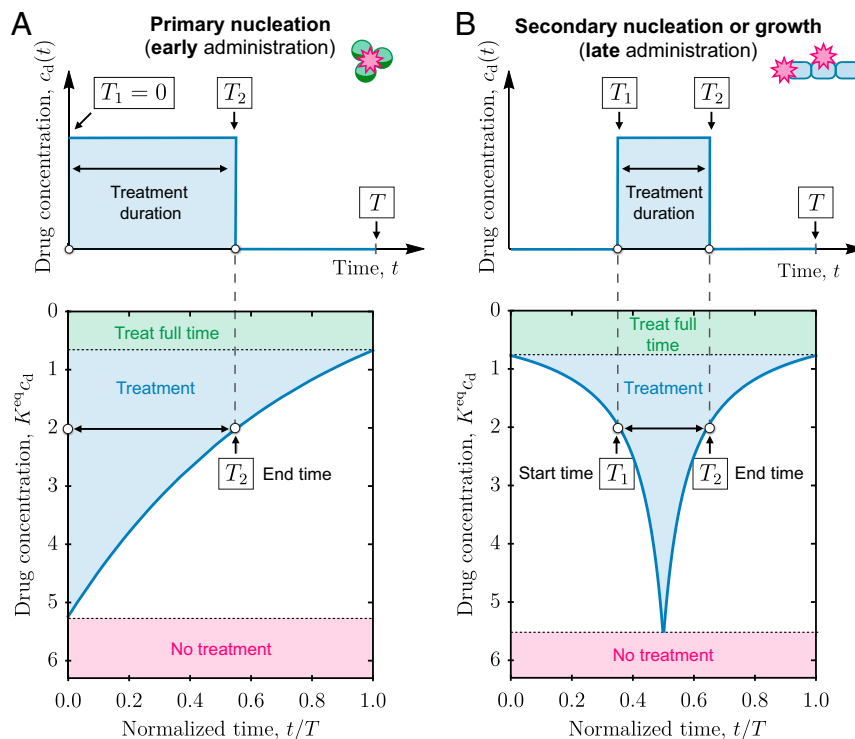


Fig. 2. Distinct optimal treatment protocols characterize the timing of drug administration for compounds that inhibit primary or secondary nucleation processes. (A) Optimal treatment protocol for the administration of a drug that inhibits primary nucleation (Top). In this case, the drug must be administered as early as possible ($T_1 = 0$) and for a duration T_2 . Increasing drug concentration decreases the overall duration T_2 of the optimal treatment (Bottom) but without affecting the need for an early administration. When the drug concentration is large, no treatment is favorable (pink), while at low drug concentrations, the optimal treatment can take the full available time T (green). (B) For a drug that inhibits either fibril elongation or secondary nucleation, a late, rather than early, administration of the drug is required (Top). The optimal treatment protocol is thus characterized by 2 switching times, T_1 and T_2 , that define the start and the end of drug administration, respectively (Bottom). The duration of the treatment, $T_2 - T_1$, decreases with increasing concentration of the drug. The parameters used in the plots are: $\zeta\kappa_0/(\alpha_0 K_{\text{oligo},1}^{\text{eq}}) = 0.6$, $\kappa_0 T = 1.3$ (A); and $\zeta\kappa_0/(\alpha_0 K_{\text{surf}}^{\text{eq}}) = 10$, $\kappa_0 T = 4.5$ (B).

where T is the total available time for treatment, and \mathcal{L} is a function that characterizes the cost rate that increases for larger aggregate and drug concentrations. \mathcal{L} is expected to be a non-linear and monotonically increasing function of drug and aggregate concentrations. In the absence of detailed experimental insights into the form of \mathcal{L} , we linearize and use the following monotonous form of the cost function $\mathcal{L} = c_a(t) + \zeta c_d(t)$, where $\zeta > 0$ quantifies the relative toxicity of aggregate and drug molecules. In *SI Appendix, section S3.7*, we show that the predictions from the linearized cost function remain qualitatively valid also in the case of a nonlinear cost function of the form $\mathcal{L} = c_a(t)^n + \zeta c_d(t)^n$. Future experiments may provide detailed insights into the specific form of the cost function allowing to scrutinize our prediction quantitatively. The optimal drug administration protocol $c_d(t)$ minimizes the cost functional Eq. 3 given the aggregation dynamics governed by Eq. 2, thus enabling us to couch our problem within the realm of classical optimal control theory (17) that allows for bang-bang control solutions, given the linear nature of the cost function.

Indeed, the optimal treatment protocol consists of using piecewise constant concentration levels of the drug over varying time spans of the treatment (Fig. 2A and B) determined by the drug toxicity, the aggregation kinetic parameters, and the mechanism of inhibition (*SI Appendix, section S3*). In this protocol, T_1 is the waiting time for drug administration, $T_2 - T_1$, denotes the time period during which the drug is applied, and $T - T_2$, is a drug-free period after treatment. We find that, depending on whether the drug suppresses selectively primary nucleation or secondary nucleation and growth at the ends of the aggregates, the optimal

protocol for drug administration is fundamentally distinct (30).[†] When the drug inhibits primary nucleation ($\alpha = \alpha(c_d)$, $\kappa = \kappa_0$; Fig. 2A), there is no waiting period for drug administration ($T_1 = 0$, “early administration”), and the optimal treatment duration reads

$$T_2 = T - \frac{1}{\kappa_0} \ln \left(\frac{\zeta c_d \kappa_0}{\alpha_0 - \alpha} \right). \quad [4]$$

When the drug affects secondary nucleation or elongation ($\kappa = \kappa(c_d)$, $\alpha = \alpha_0$; Fig. 2B), the optimal protocol is qualitatively different: the drug must be administered after a waiting period T_1 (“late administration”) and the optimal treatment duration is

$$T_2 - T_1 = \frac{\kappa_0}{\kappa_0 - \kappa} \left[T - \frac{1}{\kappa_0} \ln \left(\frac{\zeta c_d \kappa_0^2}{\alpha_0 (\kappa_0 - \kappa)} \right) \right]. \quad [5]$$

In either case, the optimal treatment time decreases with increasing drug concentration or toxicity. Moreover, at low drug concentrations, there is a regime where the drug must be administered for the full time period T , while if the drug concentration exceeds a critical threshold, $c_d > (\alpha_0 / \zeta \kappa_0) e^{\kappa_0 T}$, the preferable

[†]Note that the distinction between “early” and “late” administration is relative to the overall, macroscopic timescale of aggregation, κ_0^{-1} , and available time, T ; it is thus not related to the time required for secondary nucleation to dominate over primary nucleation the production of new aggregates, which occurs very early in the lag phase (30). In fact, secondary nucleation dominates the production of new aggregates both during an early and a late administration of the drug.

choice is no treatment. The optimal treatment duration corresponds to a minimum in cost and reflects the competition between drug-induced suppression of aggregates and drug toxicity (Fig. 3A). The achievability of optimal treatment conditions is determined by the curvature of the cost function at the optimal treatment, which approximately reads $(\kappa_0 - \kappa)\zeta c_d$ (SI Appendix, section S3.4.4). Lower curvatures around the optimal treatment parameters facilitates a robust possibility to find mostly optimal treatment conditions.

Our optimization approach allows to use the cost function to compare quantitatively different inhibition strategies and to identify the regions in the parameter space where a certain strategy is to be preferred over another; we illustrate this idea by comparing the costs for inhibition of primary or secondary nucleation (Fig. 3B and SI Appendix, section S3.4.6). We find that at large drug concentrations, and short available times $\kappa_0 T$, the inhibition of primary nucleation represents the optimal treatment strategy compared with the inhibition of secondary nucleation or elongation, as the former strategy exhibits lower cost. Indeed, a drug that inhibits primary nucleation must be administered from the beginning. Hence, preventing aggregation over a longer time $\kappa_0 T$ necessarily requires longer periods of drug administration, eventually making the inhibition of primary nucleation costlier than blocking secondary nucleation at later stages. A boundary line, corresponding to equal costs for both strategies, separates the regimes of optimal treatment. The position of the boundary line depends on the relative affinity of the drug to the primary oligomers compared with secondary

oligomers, fibril ends, or fibril surfaces. Increasing drug toxicity shifts the boundary line to the right, hence favoring inhibition of primary nucleation. Another interesting parameter to consider is the initial level of aggregates, which provides a measure of preaging of the system; we find (SI Appendix, section S3.6) that increasing the initial level of aggregates shifts the boundary line to the left, hence disfavoring the inhibition of primary nucleation. Overall, for known values of the relative toxicity, our approach suggests how to select specific drugs corresponding to different mechanisms of action either in an early or late stage of the detection of protein aggregation disorders and depending on experimentally accessible parameters, such as drug affinity.

We next use the cost function to characterize longevity gain as a function of the parameters of drug-induced inhibition of aggregation (Fig. 3C and SI Appendix, section S3.4.5). We define the life time as the time at which the cost reaches a critical value corresponding to the cost that a cell or an organism can tolerate before it dies. In the absence of any drug treatment, the cost function grows exponentially with available time T , i.e., $\text{Cost}(c_d = 0) \simeq (\alpha_0/\kappa_0^2) e^{\kappa_0 T}$. Crucially, the addition of a drug following the optimized treatment protocol lowers the cost down to a linear increase in time, $\text{Cost}_{\text{opt}} \simeq \zeta c_d T$. Hence, the difference in life times between an optimized treatment and the situation when no treatment is applied can be significant. The expected life time as a function of treatment duration displays a distinct maximum where the gain in longevity is maximal in correspondence of the optimal treatment protocol (Fig. 3D). The maximal life expectancy decreases with increasing drug concentration.

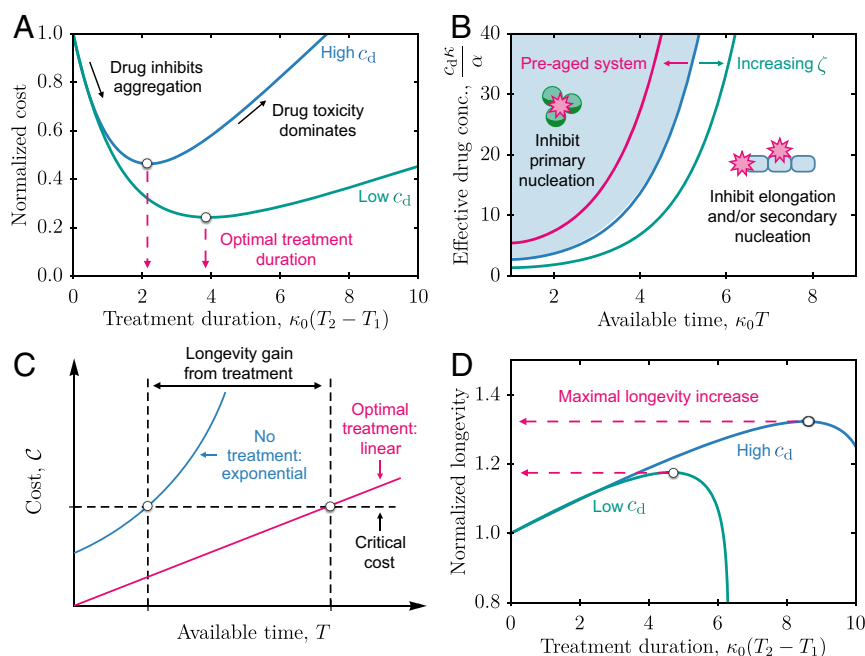


Fig. 3. Comparison between different inhibition strategies and predictions of lifetime gain due to optimal treatment. (A) The normalized cost, $\text{Cost}/\text{Cost}(c_d = 0)$, for the inhibition of secondary nucleation, has a minimum (Eq. 5) as function of the dimensionless treatment duration $\kappa_0(T_2 - T_1)$. At lower drug concentration (green line), the minimum of the cost becomes broader, indicating an easier access to the optimal protocol in the presence of fluctuations or limited knowledge of cellular kinetic parameters or concentrations. (B) Phase diagram indicating the region of parameter space where inhibition of primary nucleation has a lower cost than inhibition of secondary nucleation or growth. The green line indicates how the boundary line shifts when drug toxicity is increased by a factor of 2. Similarly, the pink line indicates how the boundary line shifts when the system is preaged, i.e., an increased concentration of aggregates initially. Note that $c_d \kappa / \alpha \simeq (\kappa_0 / \alpha_0) c_d^3 / 2 \sqrt{K_{10}^{\text{eq}} / K_{2nd}^{\text{eq}}}$, where K_{10}^{eq} and K_{2nd}^{eq} are the binding constants (affinities) for the inhibition, respectively, of primary and secondary nucleation. Thus, decreasing K_{10}^{eq} or increasing K_{2nd}^{eq} favors the inhibition of secondary nucleation over primary nucleation. (C) Cost without drug (blue) and optimal cost (pink) as a function of available time $\kappa_0 T$. Note the dramatic difference in the time dependence of the cost for the optimal treatment (linear in T) and without treatment (exponential in T). (D) Expected life expectancy as a function of treatment duration when the gain in life time is maximal in correspondence of the optimal treatment protocol. The parameters used in the plots are $\alpha_0/\kappa_0 = 2 \times 10^{-8}$, $\zeta = 200$, $K_{\text{surf}}^{\text{eq}} = 5 \mu\text{M}^{-1}$, $\kappa_0 T = 13$, $c_d = 2 \mu\text{M}$ (green), $c_d = 6 \mu\text{M}$ (blue) (A); and $\kappa_0 \text{Cost}_c = 10^{-3.5}$ M, $\alpha_0/\kappa_0 = 10^{-7}$, $\zeta = 10$, $K_{\text{surf}}^{\text{eq}} = 1 \mu\text{M}^{-1}$, $c_d = 3 \mu\text{M}$ (green), $c_d = 5 \mu\text{M}$ (blue) (D).

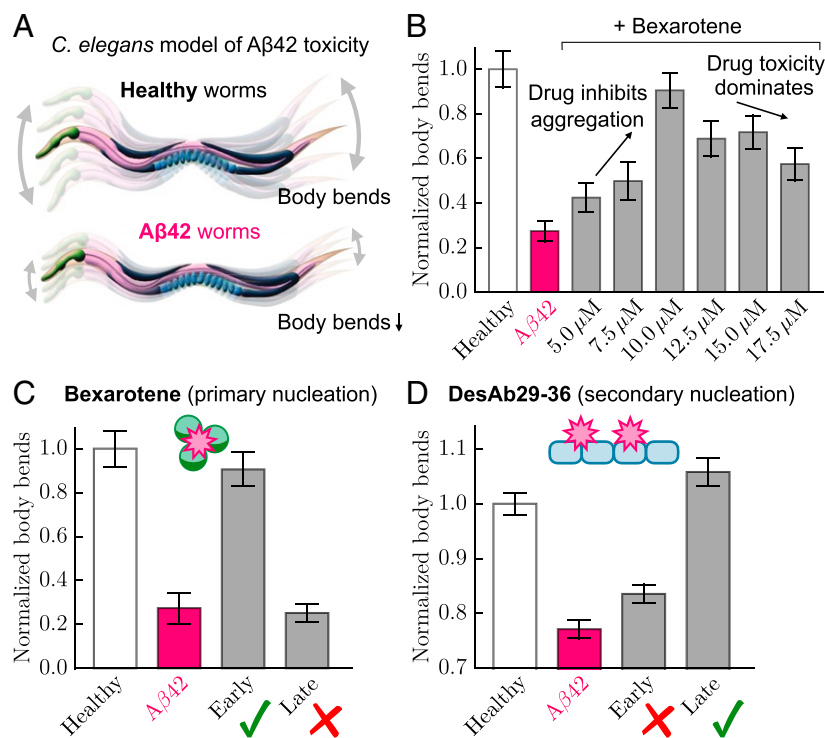


Fig. 4. Application to the inhibition of Alzheimer's A β ₄₂ aggregation in *C. elegans* model of A β ₄₂-mediated toxicity. (A) Expression of A β ₄₂ in the worm's muscle cells leads to age-progressive paralysis detected through the reduction in the frequency of body bends relative to healthy worms, which do not express A β ₄₂. (B) Low drug (Bexarotene) concentration, which selectively inhibits primary nucleation, improves worm fitness due to the inhibition of protein aggregation; however, too large drug concentrations decrease worm fitness due to toxicity of the drug (data from ref. 10). (C) Effect of early (72 h before day 0 of adulthood) and late (day 2 of adulthood) administration of Bexarotene (10 μ M) show that early administration is significantly more effective in alleviating the symptoms of worm paralysis compared with the late administration of the same drug. In the latter case, there was no observable improvement of worm fitness compared with untreated A β ₄₂ worms (data from ref. 10). (D) Effect of early (day 1 of adulthood) and late (day 6 of adulthood) administration of a selective inhibitor of secondary nucleation (DesAb_{29–36}) (data from ref. 15) show that a late administration of DesAb_{29–36} is more effective than an early administration in causing worm recovery. In B and C, the effect on fitness (body bends per second) was measured at day 6 of adulthood and compared with healthy worms and untreated A β ₄₂ worms, while in D, the effect on fitness was measured at day 7 of adulthood. Error bars indicate the SEM; the sample size was $n = 200$ worms for Bexarotene experiments (10) and $n = 500$ worms for DesAb_{29–36} experiments (15).

Comparison with Experiments. We finally tested qualitatively the efficacy of the optimal protocol in practice by considering previous data (10, 15) on the inhibition of A β ₄₂ amyloid fibril formation of Alzheimer's disease using the drug Bexarotene in a *C. elegans* model of A β ₄₂-induced dysfunction (Fig. 4A) (10, 31). Fig. 4B shows the effect of administering increasing concentrations of Bexarotene to A β ₄₂ worms in their larval stages on the frequency of body bends, a key parameter that indicates the viability of worms. At low drug concentrations, increasing Bexarotene concentration has beneficial effects on worm fitness, but too large drug concentrations decrease worm fitness. Thus, there is an optimal dose of Bexarotene (10 μ M) that leads to maximal the recovery of the worms. This optimal dose emerges from the competition between the inhibition of protein aggregation by Bexarotene (SI Appendix, Fig. S6A and B) and its toxicity (SI Appendix, Fig. S6C), as anticipated by our cost function (SI Appendix, section S3.5). At a mechanistic level, Bexarotene has been shown to affect protein aggregation by inhibiting selectively primary oligomers and hence reduce primary nucleation both in vitro (10) and in the *C. elegans* model of A β ₄₂-induced toxicity (10) (SI Appendix, Fig. S6A). Thus, the key prediction from our model is that Bexarotene would be most effective with an early administration protocol. This prediction is in line with the experimental observations (Fig. 4C) (10) that show that the administration of Bexarotene following a late administration protocol at day 2 of worm adulthood does not induce any observable improvement in fitness relative

to untreated worms. In contrast, administering Bexarotene at the onset of the disease in the larval stages (early administration), leads to a significant recovery of worm mobility. To further support our predictions, we consider in Fig. 4D the inhibition of A β ₄₂ aggregation by another compound, DesAb_{29–36}, which has previously been shown to inhibit selectively secondary nucleation (15). The data in this case show that DesAb_{29–36} is more efficacious when administered at late times than during the early stages of aggregation, an observation that is in line with the theoretical predictions of our model.

Conclusions

We have introduced a framework for estimating optimal control protocols for inhibition of irreversible protein aggregation. Overall, our results highlight and rationalize the fundamental importance of understanding the relationship between the mechanistic action, at the molecular level, of an inhibitor and the optimal timing of its administration during macroscopic profiles of protein aggregation. This understanding could have important implications in drug design against pathological protein aggregation. For example, using the cost function could provide a new platform for systematically ranking drugs in terms of their efficiency to inhibit protein aggregation measured under optimal conditions.

Our optimal protocols do not account for spatial heterogeneities, crowding, and fluctuations. Spatially dependent optimal protocols could be determined for instance by extending

our theory to reaction-diffusion systems or by including spatial organization effects, e.g., from liquid compartments (32). More generally, optimal protocols depend on the measurement accuracy of aggregate and drug concentrations and the nature and type of the cost functional. Stochasticity effects could be accounted using Kalman filter-based approaches (33). More generally, accounting explicitly in the cost function for additional factors such as organismal absorption, distribution, and clearance of the drug or its degradation over time in our theory could allow extrapolating most effective protocols from a model system, such as *C. elegans*, to clinically relevant conditions. This procedure may help to efficiently design future medical trials and would also suggest moving toward optimal drug cocktails or oscillatory protocols.

Methods

Determination of Optimal Protocol for Inhibition of Protein Aggregation. To obtain the optimal inhibition protocol, we use the Pontryagin minimum principle of optimal control theory (17). In particular, the cost functional $\text{Cost}[c_a(t), c_d(t)]$ (Eq. 3) must be minimized subject to a dynamic constraint of the form $dc_d(t)/dt = f(c_a(t), c_d(t))$ (Eq. 2). This variational problem can be solved most conveniently by introducing a time-dependent Lagrange multiplier $\lambda(t)$ (also known as costate variable in the context of optimal control theory) and considering the extended functional

$$\mathcal{F}[c_a(t), c_d(t)] = \text{Cost}[c_a(t), c_d(t)] + \int_0^T dt \lambda(t) \left[\frac{dc_d(t)}{dt} - f(c_a(t), c_d(t)) \right], \quad [6]$$

where the second term ensures that the kinetic equation $dc_d(t)/dt = f(c_a(t), c_d(t))$ is satisfied for all times t . The optimal inhibition protocol is

then determined by solving the dynamic equation $dc_a(t)/dt = f(c_a(t), c_d(t))$ together with the Euler-Lagrange equations for \mathcal{F}

$$\frac{\delta \mathcal{F}}{\delta c_a} = \frac{\partial \mathcal{L}}{\partial c_a} - \lambda(t) \frac{\partial f}{\partial c_a} - \frac{d\lambda(t)}{dt} = 0 \quad [7a]$$

$$\frac{\delta \mathcal{F}}{\delta c_d} = \frac{\partial \mathcal{L}}{\partial c_d} - \lambda(t) \frac{\partial f}{\partial c_d} = 0, \quad [7b]$$

subject to the initial condition $c_a(0) = 0$ and the constraint $\lambda(T) = 0$ (transversality condition). Eq. 7a describes the dynamics of the Lagrange multiplier $\lambda(t)$; once $\lambda(t)$ is known, Eq. 7b yields the optimal protocol.

Since the drug concentration is constant in the case of fast drug binding (SI Appendix), the optimal control consists of discrete jumps, yielding a bang-bang control of the form $c_d = c_d^{\max} [\theta(t - T_1) - \theta(t - T_2)]$, where $\theta(x)$ is the Heaviside function and T_1 and T_2 are the switching times (Eq. 4). For the choices $f(c_a(t), c_d(t)) = \alpha(c_d(t)) + \kappa(c_d(t))c_a(t)$ and $\mathcal{L}(c_a(t), c_d(t)) = c_a(t) + \zeta c_d(t)$ discussed in the main text, the evolution equation for the Lagrange multiplier, Eq. 7a, reads $d\lambda(t)/dt = -1 - \kappa(c_d(t))\lambda(t)$, while the optimal control can be calculated from

$$\lambda(T_i) [\alpha' + \kappa' c_a(T_i)] = \zeta, \quad i = 1, 2, \quad [8]$$

where continuous derivatives with respect to c_d in Eq. 7b have been replaced by discrete jumps $\kappa' = (\kappa_0 - \kappa(c_d^{\max})) / c_d^{\max}$ and $\alpha' = (\alpha_0 - \alpha(c_d^{\max})) / c_d^{\max}$. Eq. 8 determines the optimal values for the times to begin, T_1 , and to end the drug treatment, T_2 . Finally, considering the cases $\alpha' = 0$ and $\kappa' = 0$ separately, and, in the latter case, exploiting the fact that $T_i \gg \kappa^{-1}$, we arrive at the analytical results presented in Eqs. 4 and 5.

ACKNOWLEDGMENTS. We acknowledge support from the Swiss National Science Foundation (T.C.T.M.), the German Research Foundation (C.A.W.). We thank Michele Parni and Christopher M. Dobson (Center for Misfolding Diseases, University of Cambridge, Cambridge, UK) for useful discussions and for providing the extended experimental data on *C. elegans* from ref. 10.

1. T. P. J. Knowles, M. Vendruscolo, C. M. Dobson, The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15**, 384–396 (2014).
2. F. Chiti, C. M. Dobson, Protein misfolding, functional amyloid, and human disease: A summary of progress over the last decade. *Annu. Rev. Biochem.* **86**, 27–68 (2017).
3. C. M. Dobson, The amyloid phenomenon and its links with human disease. *Cold. Spring. Harb. Perspect. Biol.* **9**, a023648 (2017).
4. C. M. Dobson, Protein folding and misfolding. *Nature* **426**, 884–890 (2003).
5. D. J. Selkoe, J. Hardy, The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol. Med.* **8**, 595–608 (2016).
6. J. Hardy, D. J. Selkoe, The amyloid hypothesis of Alzheimer's disease: Progress and problems on the road to therapeutics. *Science* **297**, 353–356 (2002).
7. E. Karran, J. Hardy, A critique of the drug discovery and phase 3 clinical programs targeting the amyloid hypothesis for Alzheimer disease. *Ann. Neurol.* **76**, 185–205 (2014).
8. R. J. O'Brien, P. C. Wong, Amyloid precursor protein processing and Alzheimer's disease. *Annu. Rev. Neurosci.* **34**, 185–204 (2011).
9. P. Arosio, M. Vendruscolo, C. M. Dobson, T. P. J. Knowles, Chemical kinetics for drug discovery to combat protein aggregation diseases. *Trends Pharmacol. Sci.* **35**, 127–135 (2014).
10. J. Habchi et al., An anticancer drug suppresses the primary nucleation reaction that initiates the production of the toxic A β 42 aggregates linked with Alzheimer's disease. *Sci. Adv.* **2**, e1501244 (2016).
11. J. Habchi et al., Systematic development of small molecules to inhibit specific microscopic steps of A β 42 aggregation in Alzheimer's disease. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E200–E208 (2017).
12. S. Chia et al., SAR by kinetics for drug discovery for protein misfolding diseases. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 10245–10250 (2018).
13. S. I. A. Cohen et al., A molecular chaperone breaks the catalytic cycle that generates toxic A β oligomers. *Nat. Struct. Mol. Biol.* **22**, 207–213 (2015).
14. P. Arosio et al., Kinetic analysis reveals the diversity of microscopic mechanisms through which molecular chaperones suppress amyloid formation. *Nat. Commun.* **7**, 10948 (2016).
15. F. A. Aprile et al., Selective targeting of primary and secondary nucleation pathways in A β 42 aggregation using a rational antibody scanning method. *Sci. Adv.* **3**, e1700488 (2017).
16. T. C. T. Michaels et al., Chemical kinetics for bridging molecular mechanisms and macroscopic measurements of amyloid fibril formation. *Annu. Rev. Phys. Chem.* **69**, 273–298 (2018).
17. L. M. Hocking, "The Pontryagin maximum principle" in *Optimal Control: An Introduction to the Theory with Applications* (Oxford University Press, Oxford, UK, 1991), pp. 85–88.
18. F. Oosawa, S. Asakura, "Kinetics of polymerization" in *Thermodynamics of the Polymerization of Protein* (Academic, London, UK, 1975), pp. 41–55.
19. T. P. J. Knowles et al., An analytical solution to the kinetics of breakable filament assembly. *Science* **326**, 1533–1537 (2009).
20. G. Ramachandran, J. B. Udgaonkar, Evidence for the existence of a secondary pathway for fibril growth during the aggregation of tau. *J. Mol. Biol.* **421**, 296–314 (2012).
21. A. M. Ruschak, A. D. Miranker, Fiber-dependent amyloid formation as catalysis of an existing reaction pathway. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 12341–12346 (2007).
22. S. I. A. Cohen et al., Proliferation of amyloid- β 42 aggregates occurs through a secondary nucleation mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 9758–9763 (2013).
23. G. Meisl et al., Differences in nucleation behavior underlie the contrasting aggregation kinetics of the A β 40 and A β 42 peptides. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 9384–9389 (2014).
24. A. Sarić et al., Physical determinants of the self-replication of protein fibrils. *Nat. Phys.* **12**, 874–880 (2016).
25. F. A. Ferrone, J. Hofrichter, W. A. Eaton, Kinetics of sickle hemoglobin polymerization. II. A double nucleation mechanism. *J. Mol. Biol.* **183**, 611–631 (1985).
26. S. I. A. Cohen et al., Nucleated polymerization with secondary pathways. I. Time evolution of the principal moments. *J. Chem. Phys.* **135**, 065105 (2011).
27. I. Benilova, E. Karran, B. De Strooper, The toxic A β oligomer and Alzheimer's disease: An emperor in need of clothes. *Nat. Neurosci.* **15**, 349–357 (2012).
28. S. Campioni et al., A causative link between the structure of aberrant protein oligomers and their toxicity. *Nat. Chem. Biol.* **6**, 140–147 (2010).
29. G. Fusco et al., Structural basis of membrane disruption and cellular toxicity by α -synuclein oligomers. *Science* **358**, 1440–1443 (2017).
30. P. Arosio, T. P. J. Knowles, S. Linse, On the lag phase in amyloid fibril formation. *Phys. Chem. Chem. Phys.* **17**, 7606–7618 (2015).
31. G. McColl et al., Utility of an improved model of amyloid-beta (A β 1–42) toxicity in *Caenorhabditis elegans* for drug screening for Alzheimer's disease. *Mol. Neurodegener.* **7**, 57 (2012).
32. C. A. Weber, T. C. T. Michaels, L. Mahadevan, Spatial control of irreversible protein aggregation. *eLife* **8**, e4231 (2019).
33. R. F. Stengel, *Optimal Control and Estimation* (Dover Publications, 1994).