The AAPS Journal, Vol. 15, No. 3, July 2013 (© 2013)
DOI: 10.1208/s12248-013-9485-3

Review Article

Theme: Aggregation and Interactions of Therapeutic Proteins
Guest Editor: Craig Svensson

Non-Arrhenius Protein Aggregation

Wei Wang[1,3] and Christopher J. Roberts[2,3]

Received 1 December 2012; accepted 2 April 2013; published online 25 April 2013

Abstract. Protein aggregation presents one of the key challenges in the development of protein biotherapeutics. It affects not only product quality but also potentially impacts safety, as protein aggregates have been shown to be linked with cytotoxicity and patient immunogenicity. Therefore, investigations of protein aggregation remain a major focus in pharmaceutical companies and academic institutions. Due to the complexity of the aggregation process and temperature-dependent conformational stability, temperature-induced protein aggregation is often non-Arrhenius over even relatively small temperature windows relevant for product development, and this makes low-temperature extrapolation difficult based simply on accelerated stability studies at high temperatures. This review discusses the non-Arrhenius nature of the temperature dependence of protein aggregation, explores possible causes, and considers inherent hurdles for accurately extrapolating aggregation rates from conventional industrial approaches for selecting accelerated conditions and from conventional or more advanced methods of analyzing the resulting rate data.

KEY WORDS: Arrhenius; non-linear; prediction techniques; protein aggregation; shelf life.

INTRODUCTION

In the late nineteenth century, Svante Arrhenius proposed the following equation based on the collision theory in gaseous state to describe the relationship of a fundamental rate coefficient with temperature for a chemical reaction:

$$k = A\exp(-E_a/\mathrm{RT}) \tag{1}$$

Where k is the rate coefficient, T is the absolute temperature, E_a is the energy of activation for the reaction, R is the universal gas constant, and A is a pre-exponential factor related to steric effect and collision frequency. In this equation, the rate coefficient increases exponentially with the absolute temperature T. If the natural logarithm is taken on both sides,

$$\ln k = \ln A - E_a/\mathrm{RT} \tag{2}$$

Equation 2 shows a linear relationship between ln$k$ and the inverse of absolute temperature, assuming E_a is independent of temperature. If rate coefficients are experimentally determined at different temperatures, a linear plot of ln$k$ *vs.* $1/T$ can be obtained, and E_a and A can be regressed as the slope and intercept. The rate coefficient at any particular temperature can be then be calculated, provided the linear behavior is preserved when one considers extrapolation to temperatures outside the temperature range of the original experiment.

With development of the transition state theory, the Eyring equation was proposed to describe the relationship for the temperature dependence of a rate coefficient in a form that is mathematically similar to the Arrhenius equation:

$$k = \frac{k_B T}{h}\exp\left(-\Delta G^*/\mathrm{RT}\right) \tag{3}$$

Where k is the reaction rate coefficient, k_B is Boltzmann's constant ($R=N_A k_B$, with N_A denoting Avogadro's number), h is Planck's constant, and ΔG^* is the Gibbs free energy of activation. If one takes the natural log of Eq. 3, and considers the derivative of ln$k$ with respect to $1/T$, a form that is effectively the same as Eq. 2 is recovered, as one can typically neglect the temperature dependence due the factor of ln$T$ when dealing with the relatively small range of temperatures (on an absolute scale) that are relevant for real-time and accelerated pharmaceutical product stability. In this case, E_a from Eq. 2 is replaced with $\Delta H^{\ddagger}$, the activation enthalpy for the reaction, which follows from the Gibbs-Helmholtz relationship for $\Delta G/T$ with respect to $1/T$ (1).

Notably, the above relations were developed for relatively simple reactions, and the transition state is that for a single rate-limiting, elementary molecular step within a given mechanism (1). In either case, E_a or $\Delta H^{\ddagger}$ are treated as being independent of temperature over the temperature window of interest, and thus a plot of ln$k$ *vs.* $1/T$ (i.e., an Arrhenius plot)

[1] Pfizer Inc., BioTherapeutics Pharmaceutical Sciences, 700 Chesterfield Parkway West, Chesterfield, Missouri 63017, USA.
[2] Department of Chemical and Biomolecular Engineering, University of Delaware, 150 Academy St., Newark, Delaware 19716, USA.
[3] To whom correspondence should be addressed. (e-mail: wei.2.wang@pfizer.com; cjr@udel.edu)

results in a straight line. In the remainder of this review, this will be termed Arrhenius behavior while non-Arrhenius will denote systems in which significantly nonlinear behavior is observed in an Arrhenius plot.

The significance of the Arrhenius equation for pharmaceutical companies is its ability to predict shelf lives of drug products based on short-term, accelerated storage stability studies at elevated temperatures. Such predictions can significantly shorten the development time, and products can be put on the market faster. Therefore, the Arrhenius equation has been (implicitly or explicitly) used widely for rapid and accurate assessment of stability of certain types of pharmaceutical dosage forms through accelerated aging studies (2).

The Arrhenius equation, or its mathematical form, has been able to describe the temperature-dependent behavior of many chemical reactions in both liquid and solid states. Chemical reactions (i.e., those involving making/breaking covalent bonds) for even large and complex protein molecules have been shown to follow this relationship reasonably well. Examples include oxidation of several Met residues in recombinant human interleukin-1 receptor antagonist between 5 and 45 C (3), and several Met residues in recombinant human granulocyte colony-stimulating factor between 4 and 45 C (4), Asn deamidation in recombinant human interleukin-15 (IL-15) between 6 and 40 C (5), hydrolysis of a Fc fusion protein at Ser-Met and Asp-Glu between 15 and 40 C (6), and even multiple degradations in oxytocin from pH 2–9 between 40 and 80 C (7).

Some physical reactions of protein molecules (i.e., without changes in covalent bonds) were also shown to follow the Arrhenius or Eyring relationships, perhaps most notably the process of protein unfolding when it occurs as a cooperative step with a net rate coefficient k_u (8). Examples include unfolding of lysozyme between 5 and 45 C (9), RNase A between 45 and 63 C (10), α-chymotrypsinogen (aCgn) between 26 and 38 C (11), chitinase F1 between 42 and 55 C (12), several porcine PYY molecules (helical hairpin; 36 aa) between 25 and 65 C (13), β-galactosidase between 45 and 55 C (14,15), and ovalbumin between 68.5 and 80 C (16). The values of E_a or $\Delta H^{\ddagger}$ one obtains for the unfolding rate coefficient are large—typically of the order of 10^2 kcal/mol—and are of similar magnitude to the enthalpy of unfolding. To place this in context, the largest activation energy or enthalpy values one often observes for chemical reactions (e.g., deamidation) are of the order of only a few tens of kilocalories per mole (2).

Interestingly, different types of non-Arrhenius behaviors have been observed in a number of cases. A relatively simple case is one in which ln$k$ *vs.* $1/T$ is composed of two straight-line segments that cross each other at a specific overlap temperature. Examples include temperature-dependent concanavalin A-dextran precipitin reaction at pH 7.4 between 8.6 and 37 C (overlap point at 18 C) (17) and loss of enzymatic activity of adenylosuccinate synthetase between 20 and 85 C (overlap point at 40 C) (18). The causes for these behaviors are often either a change in reaction mechanisms (e.g., the first example above) or a change in the rate determining step (second example above).

In other cases, non-Arrhenius ln$k$ *vs.* $1/T$ clearly follow a continuous, nonlinear functional form. These types of curves can be simply divided into two major categories—concave up and concave down, as shown in Fig. 1. If high-temperature data are used to predict the rate of degradation at a lower temperature assuming a linear relationship, the rate can be under- or overestimated, respectively, for concave-up or concave-down types of behavior. Examples of each type have been reported in the literature. Concave-down behavior was demonstrated for: refolding of hen egg lysozyme between 2 and 50 C (19), refolding of trypsin inhibitor 2 between 10 and 80 C (20), and unfolding of several mutants of λ-repressor fragment (λ_{6-85}) in different temperature ranges (21). Concave-up behavior was demonstrated for de-excitation of Trp (triplet state) in LADH and α-crystallin between 4 and 60 C (22) (two activation energies), and unfolding of mutant T4 lysozyme in 3 M GdnHCl at pH 5 between −10 and 18 C (23). Protein aggregation often displays concave-up behavior (see the following section).

An extreme form of the non-Arrhenius behavior is an increase in the observed rate coefficient with an increase in inverse temperature. This type of curve would have a positive slope, leading to an apparent negative activation energy. This can be an extension of the concave-down type at high temperature end or the concave-up type at the low temperature end (dotted lines in Fig. 1). It was found that the observed folding rate coefficient for chymotrypsin inhibitor 2 shows a concave-down increase with increasing temperature from 25 to 50 C but decreased with increasing temperature above 50 C (24). Similar results were also observed for thermal unfolding of hen egg white lysozyme, where the observed folding rate increased from 5 to 40 C but decreased above 40 C (19), and for a RNase A mutant, where the observed unfolding rate coefficient decreased with increasing temperature in the low temperature range between 45 and 63 C (10). A rare case of nonlinearity is the zig–zag relationship for the inactivation of α-chymotrypsin between 45 and 85 C (25). This behavior was due to the presence of two stable native forms—one at a low temperature and one at a high temperature.

There are many possible causes for nonlinear Arrhenius behavior. Some include temperature-dependent changes in reaction conditions (such as phase, pH, and relative humidity or water content), reaction mechanisms, or Arrhenius equation parameters (2). Changing slopes could mean a gradual change in one or more of these aspects. The concave behavior for the refolding of trypsin inhibitor 2 is due to the activation

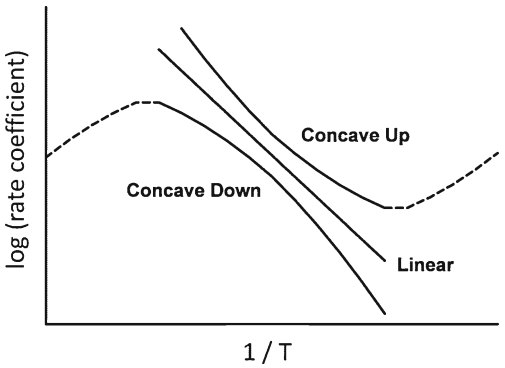

Fig. 1. Linear and nonlinear Arrhenius behaviors. The *dashed lines* represent extreme nonlinear behavior, where a maximal or minimal rate exists

enthalpy being temperature dependent—that is, the heat capacity difference between the "ground" state and transition state is significantly non-zero, at least putatively due to their differences in structure and solvent exposure (20).

The generalization of this observation is that an observed or effective rate coefficient k may account for a single elementary (chemical) reaction or may be a convolution of multiple steps. If any of the steps that contribute significantly to k involve a large difference in heat capacity (Δc_p), then the effective E_a or $\Delta H^{\ddagger}$ for k will change with temperature and non-Arrhenius behavior will be observed over a sufficiently large temperature range. Of course, if Δc_p or the temperature range of interest is not large then for practical purposes k may be approximately Arrhenius (8,26). As discussed in more detail below, one way of phrasing the problem of extrapolating nonlinear behavior is that one seeks a way to "linearize" k in some simple, yet reliable, manner that would not require one to have more than the high-temperature k values—i.e., this would allow a true prediction of low-T values of k, rather than needing to have the low-T data in order correct for the non-Arrhenius behavior.

PROTEIN AGGREGATION AND ITS TEMPERATURE DEPENDENCE

Proteins have a high tendency to aggregate via a variety of mechanisms/pathways. The major pathways are shown in Fig. 2 and can be roughly divided into: (1) aggregation through unfolding intermediates and/or unfolded states; (2a) aggregation through protein self-association (colloidal interactions) in native states; (2b) aggregation upon chemical degradation directly linking protein monomers; and (3) aggregation upon chemical degradation indirectly changing the aggregation tendency of the protein. The arrows in Fig. 2 often constitute multiple steps along a given pathway and are shown as single arrows only to give simple graphical representation. Also, path 2a (native aggregation) is typically reversible, and as such requires additional steps such as structural rearrangement and/or misfolding to provide the net irreversible non-native protein–protein contacts that are more typical of the aggregates relevant to the remainder of the discussion and analysis here.

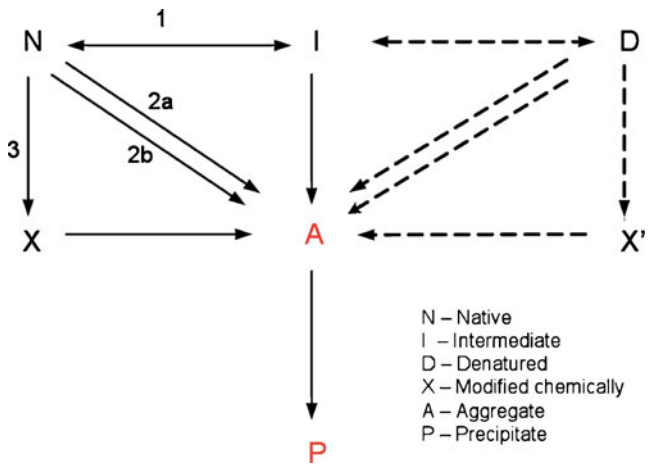

Fig. 2. Common pathways of protein aggregation

It appears that a majority of (non-native) protein aggregation pathways involve formation of an aggregation nucleus and hence, is nucleation dependent (27). However, for therapeutic proteins, the nucleus may be as small as a dimer, and the kinetics of nucleation do not need to include a "lag" phase such as observed in phase transitions (28). The process from N (or I/D) to A in Fig. 2 can be considered as the nucleation step, where A is the smallest stable (net irreversible) aggregate species. All the aggregation pathways noted earlier may lead to formation of a nucleus. The steps leading up to the nucleus are often reversible, as nucleation is a rate-limiting step in such cases. That is, nucleation is by definition the creation of the smallest stable (net irreversible) aggregate. Self-association and/or unfolding steps prior to nucleation are therefore reversible, although they are not necessarily in equilibrium (29). The next stage after nucleation is growth in aggregate size. Several terms have often been used interchangeably in the literature to describe the growth stage(s), including aggregation, elongation, fibrillation, polymerization, and condensation. More precise definitions to delineate these mechanisms have been suggested (8,30) but are beyond the scope of this report.

Nucleation-dependent protein aggregation originating from unfolded or partially unfolded intermediates has been incorporated mathematically into a number of models (for a more extensive review, see e.g., (29,31)). Except in cases where an empirical or heuristic approach is taken, most if not all of these models highlight that there are multiple steps that contribute to a measured rate or rate coefficient for aggregation. One example studied in detail by one of us is the Lumry–Eyring nucleated polymerization (LENP) model (32–34). The final working equation for the observed rate coefficient of monomer loss (k_{obs}) incorporates multiple steps in the aggregation pathway—including protein conformational change, reversible self-association prior to nucleation, nucleation, and growth via different pathways. Aggregation of several proteins appear to follow the LENP model reasonably well. These include the non-native aggregation of α-chymotrypsinogen (aCgn) and anti-streptavidin IgG1 over a range of pH and (NaCl) conditions (11,30) (Kim *et al.*, under review), a model IgG1 at pH 4.5 and 5.5 (35,36), and gamma-D crystallin (37). This type of aggregation can be unfolding limited, association limited, or nucleation limited (8). Depending on the rate-limiting step in this process, aggregation (monomer loss) can display first-, second-, or higher-order kinetics when monitoring monomer concentration over multiple half lives. In addition, the apparent reaction order of aggregation may vary with protein concentrations (36), and this may occur because of change in the identity of key intermediates or pre-nuclei (11).

Arguably, the most widely studied and reported example of nucleation-dependent protein/peptide aggregation is the fibrillation of β-amyloid forming peptides. The behavior of amyloid-type aggregation covers many peptides and small proteins. Their aggregation kinetics are often reported in terms of a sigmoidal curve, characterized by a so-called "lag phase," a growth phase and a plateau phase (38,39). This type of observation is typical when one employs assays such as turbidity (optical density) or dyes that are specific to only certain aggregate states. In this case, the "lag phase" is a

misnomer, as it can include a number of steps where aggregates form and rearrange, but are undetected simply due to limitations of the experimental assay in question. When one employs more sensitive techniques such as laser and neutron and/or X-ray scattering (40,41), it appears that the peptides/proteins form intermediate oligomeric aggregates before they reorganize and grow into much larger or structurally ordered species such as fibrils or macroscopically detectable aggregates (39,42).

Depending on the solution conditions, lag phases have been reported to range from a few minutes to several days. However, the relevance of true lag phases for aggregation in the context of pharmaceutical protein shelf life may be questionable once one accounts for limitations of a given assay, and because the time scales of observation in this context are extremely long ($\sim 10^5$ to 10^8 s) compared with time scales for molecular events such as (un)folding or misfolding, and molecular diffusion, unless one considers extremely small concentrations or sample volumes. In the latter scenario, one must then also observe stochastic kinetics, meaning that a set of identical samples will display aggregation rates that vary widely (factors of ten or more), and that such variations cannot be attributed to differences in the sample preparation, container defects, or other inherent sources of variability within product manufacture (29).

Elevated temperature is commonly used to probe protein conformational stability as well as to help accelerate aggregation. While increasing temperature does increase diffusion rates, this may be a relatively weak effect compared with the observed temperature dependence of k_{obs}. With the possible exception of very high concentration protein solutions (~100–300 g/L), the diffusion rate scales as $RT/\eta R_h$, where R_h is the hydrodynamic radius of the protein, and η is the solvent viscosity (other symbols defined previously). R_h depends weakly on T, while (on an absolute scale) T changes by less than 30% for the most extreme range one might practically select for testing with accelerated and real-time storage conditions. Therefore the primary source of changes in diffusion rates with temperature is the exponential dependence of η on temperature ($\eta \sim exp(-E_{a,\eta}/\mathrm{RT})$). Using the viscosity of water as an example, the effective activation energy is $E_{a,\eta} \approx 4$ kcal/mol. Recall from above that "large" activation energy values for covalent reactions in water are of the order of 10–20 kcal/mol, and that protein aggregation E_a values are often ~100 kcal/mol or higher. As such, it is difficult to attribute the temperature dependence of protein aggregation to that for bulk diffusion in solution.

Of course, for solid-state protein stability, or possibly for diffusion of proteins in near-glassy films at solid–liquid or air-liquid interfaces, diffusion may be sufficiently slow to be rate-limiting and more temperature sensitive. In addition, at high protein concentrations the interactions between proteins become important, and viscosity rises (albeit still remaining many orders of magnitude below that for solid or near-glassy systems). It remains an open question as to the magnitude of the temperature dependence of diffusion, and its relevance to aggregation rates, when high-concentration protein solutions are considered.

More generally, changing temperature can have an effect on each of the multiple aggregation steps involved, and thus the effective or observed activation energy (or enthalpy) is a combination of the temperature dependence of each of those steps. If the step is fast compared with k_{obs} then it equilibrates quickly (i.e., it "pre-equilibrates") relative to the progression of aggregation. In this case, its temperature dependence is described by the equilibrium enthalpy change (ΔH) of that process, rather than its activation energy. Because we are considering $\ln k_{obs}$ when referring to activation energies, the overall or observed E_a is therefore a sum of the activation energy value for the rate-limiting step, and the ΔH values of the pre-equilibrated step(s) that contribute to k_{obs}.

If one only considers a relatively small range of temperature, or if none of the steps that contribute to k_{obs} have an E_a or ΔH that changes significantly with T, then one will observe Arrhenius behavior (to within statistical/experimental uncertainty in $\ln k_{obs}$) values. There are a number of published cases where protein aggregation rates fit reasonably well to the Arrhenius equation. Examples include: nucleation rates for fibrillogenesis of β-Amyloid peptide (A_β) between 29 and 45 C (38), elongation rates for fibrillogenesis of A_β between 4 and 40 C (43), and between 28 and 68 C (42), fibrillation (covering both elongation and thickening, and association of fibrils into bundles or clusters) of human insulin at low pH and between 50 and 70 C, and between 60 and 80 C (44,45) both nucleation and elongation rates for the (agitation-induced) fibrillation of HET-s(218–289) between 20 and 50 C (46), elongation rates for fibrillation of α-synuclein, a natively unfolded protein, between 27 and 57 C (47), and fibril nucleation rates of N47A α-spectrin SH3 domain between 47 and 55 C (48). Notably, many of these examples are proteins or polypeptides that either are natively unfolded or are unable to adopt folded monomer structures under the solution conditions of interest. Based on the discussion below regarding the effects of conformational changes on the Arrhenius or non-Arrhenius behavior of k_{obs}, it is perhaps not surprising that the above examples show Arrhenius behavior.

For large and folded proteins, a small number of Arrhenius-like examples exist regarding solution-phase aggregation. For example, the aggregation growth rate coefficient of a human IgG follows a linear Arrhenius relationship in a narrow temperature range between 47 and 62 C at several concentrations from 1.9 to 15.4 mg/mL (49), and Yoshioka *et al.* (15) observed reasonably Arrhenius behavior for beta-galactosidase aggregation. However, in the former case the rate coefficient was established based on the growth of hydrodynamic radius of the aggregates. Care must be taken when using changes in aggregate size as a relevant measure of k_{obs} in the present context, as aggregates may have higher fractal dimension, and may grow via mechanisms that do not involve monomers. In the latter case, the rates were considered over a rather limited temperature range, and that can obscure non-Arrhenius behavior (see below).

In the solid state, protein aggregation appears more often to follow the Arrhenius relationship (50,51). It was found that the rate of aggregation for a mAb in lyophilized formulations at several different moisture levels followed Arrhenius kinetics between 5 and 50 C in the glassy state (well below the effective glass transition temperature (T_g)) (52). Also, it was found that the initial rate for tetramer formation for β-lactoglobulin showed an Arrhenius temperature dependence between 60 and 100 C in pure lyophililized protein or protein-trehalose mixture formulations (53).

In the liquid state examples above, the estimated activation energies for protein aggregation seem to vary significantly. The activation energy for insulin fibrillation between 50 and 70 C was estimated to be 25 kcal/mol (45), similar to the E_a values for nucleation (17.9 kcal/mol) and elongation (20.1 kcal/mol) during fibrillation of α-synuclein (47), as well as the fibrillogenesis of A_β (23 kcal/mol) between 4 and 40 C (43), and the nucleation (E_a=14.4–17.0 kcal/mol) of the HET-s(218–289) between 20 and 50 C (46). Under different agitation rates, E_a for elongation of fibrils of the HET-s(218–289) at physiological pH between 20 and 50 C was only 3.3–4.3 kcal/mol (46), which is likely a diffusion-controlled process (<<10 kcal/mol). In contrast, the nucleation activation energy for fibrillogenesis of A_β is 74.4 kcal/mol between 29 and 45 C (38). In the solid state, the E_a values for aggregation of a lyophilized mAb were 9.9 and 10.3 kcal/mol for samples containing 2% and 5% moisture, respectively, between 5 and 50 C (52), but 23 kcal/mol for the initial rate for tetramer formation for lyophilized β-lactoglobulin between 60 and 100 C (53).

These examples show a large variation in quantitative temperature dependence for protein aggregation both in liquid and solid states. In comparison, the activation energy for the oxidation of several Met residues was 9.2–13.2 kcal/mol in recombinant human interleukin-1 receptor antagonist (3), E_a ~9.1 to 18.8 kcal/mol for Met oxidation in recombinant human granulocyte colony-stimulating factor (4). While E_a for Asn deamidation was 13.3 kcal/mol in a model peptide (54) and 22.9 kcal/mol in IL-15 (5). The much larger variation in activation energy for aggregation when compared with chemical degradation indicates the complexity of the aggregation process and potential source(s) of its temperature dependence.

The large E_a found for aggregation in the liquid state suggests a large temperature dependence and possible contributions from multiple aggregations steps—including unfolding, growth and/or nucleation. For example, most the examples above with low E_a values (~30 kcal/mol or lower) are for natively unfolded proteins or polypeptides. In contrast, the effective E_a value for aggregation of the four-helix bundle protein bG-CSF is well over 100 kcal/mol over a range of temperatures (55), E_a for the alpha-beta protein aCgn is over 200 kcal/mol under accelerated (high T) conditions below its midpoint unfolding temperature (11), and E_a for an IgG1 is between 100 and 200 kcal/mol under typical formulation pH and salt conditions (56).

As noted above, it is more often than not that aggregation of folded proteins does not follow the Arrhenius relationship. The degree of non-linearity may vary, depending on the solution conditions (pH, salts, etc.) (26). In most cases, upward curvature is observed for temperature-dependent protein aggregation, as seen for different types of monoclonal antibodies (26,57,58). Upward curvature for $\ln k_{obs}$ *vs.* $1/T$ leads to an overestimation of E_a values when extrapolated to low temperatures, and thus to an overestimation of the shelf life that will be gained by cooling to low-temperature conditions (26). In practical terms, this type of nonlinearity makes shelf life prediction much more difficult and less reliable based on high-temperature stability data.

POSSIBLE REASONS FOR NONLINEAR PROTEIN AGGREGATION

The temperature-dependent non-linearity in protein aggregation can be attributed to several possible factors. These factors primarily include: (1) an intrinsically temperature-dependent contribution to E_a (e.g., a ΔH that depends on temperature) or (2) a competition between different pathways such that there is a change in pathway or rate-limiting step as temperature changes. The latter was already discussed above, and some specific examples of (1) are discussed next.

Temperature-Dependent Change in Protein Stability

For aggregation in which unfolding is not rate limiting, but some degree of unfolding is required within the pathway, $\ln k_{obs}$ will be a sum of at least two factors. One is that for the rate-limiting step (e.g., that for diffusion or for nucleating an irreversible species from an otherwise reversible oligomer). If this is diffusion then this will typically be a pseudo-Arrhenius process unless one is approaching a glass transition for the system. If this is nucleation, then it is unclear whether one should expect an intrinsically non-Arrhenius contribution to k_{obs}, as theories regarding the nucleation event itself, at a molecular level, are not well developed in general for non-native protein aggregation.

The other factor is a free energy for reversible conformational change or (partial) unfolding, ΔG_{un}. This is a pre-equilibrated step and so should be thought of as creating a reactive intermediate in the process, rather than creating a transition state. This ultimately gives a mathematical form that is analogous to that for the Eyring equation (Eq. 3) but now with an extra term for $\ln k_{obs}$, proportional to $\Delta G_{un}/RT$. As with the Eyring equation, when one takes the derivative of this term with respect to $1/T$ it gives a factor of ΔH_{un} (the equilibrium unfolding enthalpy) in the sum of terms that make up E_a for aggregation. ΔH_{un} is temperature dependent because the heat capacity change upon unfolding ($\Delta c_{p,\,un}$) is large and positive for protein folding. Therefore the overall E_a for aggregation will have a term that is proportional to the product of $\Delta c_{p,\,un}$ and T. If we denote the range of temperatures over which aggregation is tested as ΔT, then one should see non-Arrhenius behavior for k_{obs} if the product of $\Delta c_{p,\,un}$ and ΔT is of similar magnitude to the average value of ΔH_{un} over this range ΔT (8). This same effect can be illustrated by considering the temperature-dependent protein conformational stability, expressed by integrating the Gibbs–Helmholtz equation with a constant $\Delta c_{p,\,un}$ (59,60),

$$\Delta G_{un}(T) = \Delta H_{un,\,0}(1-T/T_m) - \Delta c_{p,\,un}(T_m - T + T\ln(T/T_m)), \quad (4)$$

where $\Delta H_{un,\,0}$ is the unfolding enthalpy evaluated at the midpoint unfolding temperature (T_m or T_0). The graphical relationship between ΔG_{un} and T (protein conformational stability curve) is shown in Fig. 3 (61–63). The stability curve shows three characteristic temperatures—that for maximal stability (T_{ms}), cold denaturation (T_{cd}) and heat denaturation (T_{hd} or T_m). For mesophilic globular proteins, the predicted or measured T_{ms} values typically range from −25 to +35 C

(64). Other analyses show that the majority of reversible two-state proteins are maximally stable around room temperature with an average of 20±8 C at or near neutral pH, irrespective of their structural properties, the melting temperature, or the living temperatures of their source organisms (63). This is in agreement with a recent suggestion that the maximal stability of a typical protein is around 17 C with ΔG_{un} about 7.7 kcal/mol (65). Note that the maximum K_{un} is not where $\Delta G_{un}=0$, but rather where $\Delta H_{un}=0$. In practice, this temperature differs from T_{ms} by a few degrees (61), but does not affect the discussion below.

The particular shape of the stability curve in Fig. 3 is the result of the effect of temperature on multiple aspects of the protein structure, manifest as a large $\Delta c_{p,\ un}$. Physically, a large difference in heat capacity for unfolding is attributed to the exposure of primarily hydrophobic side chains from the interior of folded domain(s) of the protein (66,67). Therefore it appears reasonable to argue that the hydrophobic effect, manifesting via the temperature dependence of protein unfolding, is a strong contribution to the non-Arrhenius behavior of aggregation rates. Note that simple models of protein aggregation often pose that aggregation is due primarily to hydrophobic attractions, however additional factors are clearly involved when one considers the importance of complementary steric packing and van der Waals forces, as well backbone–backbone hydrogen bonding (27,68,69). Here, the analysis simply is focused on why a large non-Arrhenius effect is seen, rather than assuming whether hydrophobic interactions are the dominant factor in folding or aggregation.

The above analysis is also consistent with the earlier examples based on natively unfolded proteins and other polypeptides that do not have a well-defined folded monomer state, and thus are either relatively hydrophilic or too short to fold effectively. In either case, they do not appear to require an unfolding step to enable aggregation, and therefore are lacking both the large contribution from ΔH_{un} ($\sim 10^2$ kcal/mol near T_m) as well as the large value of the product $\Delta c_{p,\ un}\Delta T$.

Temperature-Dependent Change in Aggregation Mechanisms or Solution Conditions

Temperature can enhance the rate of chemical degradation. Degraded proteins may have altered protein structure and stability for protein aggregation. It was shown that

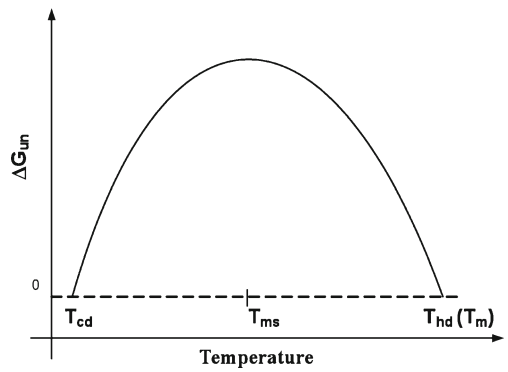

Fig. 3. Protein stability curve shows three characteristic temperatures—maximal stability (T_{ms}), cold denaturation (T_{cd}), and heat denaturation (T_{hd} or T_m)

methionine oxidation in a human IgG1 Fc altered secondary and tertiary structure, reduced melting temperature (T_m) of the C(H)2 domain, and led to enhanced aggregation at 45 C (70). The free energy change between the folded and unfolded states of lysozyme decreased by the accumulation of chemical reactions (isomerization of Asp-Gly, deamidation of Asn, racemization of Asp and Asn, and cleavage of the Asp-X-peptide bond) induced at high temperature (71). Fragmentation can generate smaller proteins/peptides of different sizes. It has been shown that the length of simple polypeptides can change the aggregation rate—e.g., as occurs for polyalanine (72). Indirectly, temperature-induced hydrolysis of sucrose, a commonly used protein excipient, led to formation of protein glycation products, increasing the rate of protein aggregation (73).

It is well known that protein aggregation tendency or mechanism in solution is strongly influenced by pH. Changing pH can significantly alter the mechanism of aggregation and size/solubility of protein aggregates such as IgG1 (36). In a relatively narrow pH range, IgG1 aggregates were shown to grow primarily by monomer addition and remained small and soluble at pH 3.5 but grew by both chain polymerization and condensation and become large and insoluble at pH 4.5 and higher pH (36). Changing temperature also changes the pH of certain formulation buffers due to temperature-dependent changes in the degree of ionization (74). Chemical degradation could lead to formation of basic or acidic species, altering the solution pH, such as deamidation of asparagines (75).

POSSIBLE SOLUTIONS FOR NON-ARRHENIUS AGGREGATION KINETICS

In the development of protein biotherapeutics, accelerated conditions are often used to facilitate the assessment of protein aggregation tendencies and to attempt to predict the low-temperature aggregation behavior based on an Arrhenius-like relationship. If the actual relationship between ln$k$ and $1/T$ is not linear, one may elect to consider alternatives for more accurate predictions.

As mentioned above, one type of non-linear Arrhenius curve is composed of two linear segments. The entire curve can be fit with two linear equations with two activation energies. It is noted that the overlap or transition temperature may not be easily discernable by visual inspection, especially when the number of data points is limited. For this purpose, a program has been developed for use to identify the ideal breaking point (76). Nevertheless, in accelerated stability studies, only a limited number of temperatures are tested, due to limited materials and/or resources. Therefore, it is not practical to use this method to predict low-temperature aggregation kinetics.

If the rate–temperature curves are slightly curved upward, the Erying equation or the modified Arrhenius equation could be used to describe the non-linear relationship: (77)

$$k = AT^n e^{-E_a/RT} \tag{5}$$

In this case, one more temperature term T^n with a floating n is included to the equation in fitting the nonlinear data and for extrapolating the rate coefficients (78). Addition of this extra term could improve the linearity of the curve but significant non-linear data would not fit well. The additional

parameter n is purely empirical, and as such is just a an attempt to "linearize" the data, assuming that one has sufficient data that show such nonlinear behavior to be able to obtain a reliable estimate of the value of n (see also, discussion below).

For significantly nonlinear ln$k$ *vs.* $1/T$ behavior, one option is to use a mechanism-based model for the change in E_a as a function of T that explicitly includes non-Arrhenius effects such as the effects of $\Delta c_{p,\,un}$. Such a nonlinear function has been included in the analysis of the LENP model, and gives a predicted concave-up ln$k$ *vs.* $1/T$ curve based on separately measured values of $\Delta c_{p,\,un}$ and the values of ln$k$ at high temperature (but well below T_m). In this scenario, one measures ln$k$ under accelerated conditions where E_a does not have a discernable dependence on T, and then ln$k$ is extrapolated to lower T where it is non-Arrhenius. All of the non-Arrhenius behavior is therefore presumed to result from the shifts in unfolding equilibrium with temperature (26,34). In prior work, this approach was successful to predict low-temperature rates from just accelerated data (i.e., without knowledge of the low-temperature rates) (34), but more recent work highlights that it can be difficult to directly determine $\Delta c_{p,\,un}$ because aggregation occurs during the measurement of unfolding (e.g., in calorimetry or other thermal scanning measurements) (36).

Alternatively, the Vogel–Fulcher–Tammann (VFT) equation could also be used to "straighten" the non-linear curve: (57)

$$k = A_{VFT} e^{-E_{VFT}/(T-T_0)} \tag{6}$$

where T_0 is an adjustable reference temperature. Addition of this parameter can significantly linearize the data (see below) but for this approach to work a sufficiently large temperature range must be used, such that the non-Arrhenius behavior is already apparent.

Figure 4 reproduces two literature examples where ln$k$ (from monomer loss data) *vs.* $1/T$ show significant non-Arrhenius behavior. One is for a single-domain four-helix bundle protein, bovine granulocyte-colony stimulating factor (bG-CSF) (55). The other is for a monoclonal antibody (MAb) (57). The main panel shows the data over the full set of temperatures reported, which include both accelerated (high-T) and real-time conditions (room T and below). The inset shows the data over just the high-T conditions that correspond to temperature ranges that are typical of industrial practice for choosing accelerated conditions—i.e., 40 C plus and additional few temperature values: for bG-CSF this is 30 to 40 C, while for the MAb it is 40 to 60 C. Notably, the set of four accelerated temperature conditions in the inset for either data set illustrate that non-Arrhenius behavior (i.e., significant curvature *vs.* $1/T$) is not apparent. That is, one must consider a wider range of temperature and/or have additional data points in order to observe statistically significant deviations from linearity (see also discussion below).

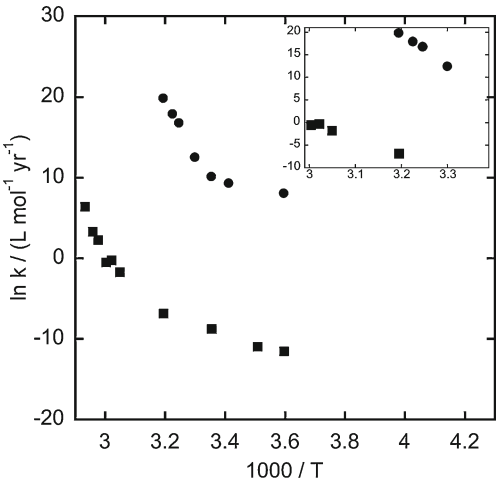

Fig. 4. Arrhenius plot of observed rate coefficient values for aggregation of bG-CSF (*circles*) and a MAb (*squares*). See main text for additional details

Rather than conventionally fitting high-T data to obtain a nonlinear extrapolation, Kayser *et al.* (57) identified an alternative approach guided by the mathematical form of the VFT equation. They observed that by choosing a T_0 value that lies at high temperature ($T>T_m$) their otherwise non-Arrhenius rate data for MAb aggregation would display a linear dependence if plotted on VFT temperature axis—i.e., when ln$k$ is plotted *vs.* $1/(T-T_0)$, rather than plotted *vs.* $1/T$. Figure 5a illustrates this behavior using the same MAb data as in Fig. 4. In each panel, the data are not regressed, but rather are simply replotted using $1{,}000/(T-T_0)$ on the X-axis, with a different value of T_0 for each panel. The case of T_0= 0 is of course equivalent to an Arrhenius plot of the data such as shown in Fig. 4.

Interestingly, if one selects T_0 lower than the experimental range of the k_{obs} data (i.e., 5 C in the present case) then the data show upward curvature. If one instead selects T_0 above the experimental range of k_{obs} then the data show downward curvature. As T_0 approaches the range of T values for the k_{obs} data, the curvature becomes increasingly pronounced. If one chooses a T_0 value within the range of the data, then there is an unphysical "break" in the data trend due to the mathematical artifact of the switch in curvature that this transform of the temperature axis causes. As such, one must select T_0 such that is either far above or far below the experimental data range if one is to employ the VFT-inspired equation to linearize the data. Physically, this is not surprising since mechanistic models that display VFT behavior identify T_0 as the point at which the configurational entropy of the system vanishes (e.g., as a liquid vitrifies as it is cooled extremely rapidly). In those systems, T_0 is always below the experimental range, as it is physically unrealistic for the system to exist at $T<T_0$.

Figure 5b (inset) quantifies the deviation from linearity as a function of the choice of T_0 for this data set. It shows the value of the R^2 goodness-of-fit parameter for regressing a straight line to lnk_{obs} plotted *vs.* $1/(T-T_0)$. The values that are close to 1 are reasonable choices for T_0 if one seeks to linearize non-Arrhenius data in this manner. The inset shows the results for fits where one uses the full set of temperatures from Fig. 4. Clearly, one must avoid T_0 values that lie close to the experimental range of temperatures, for the reasons noted above. Choosing high values of T_0 gives the most linear results, as noted by Kayser *et al.* (57), and as illustrated in Fig. 5a by the T_0=368 *K* example. However, choosing T_0 too high induces artificial downward curvature (e.g., T_0=400 *K* or 450 *K* in Fig. 5a). Figure 5b shows that there is a broad maximum in R^2 *vs.* T_0 if one considers high T_0 values, and so

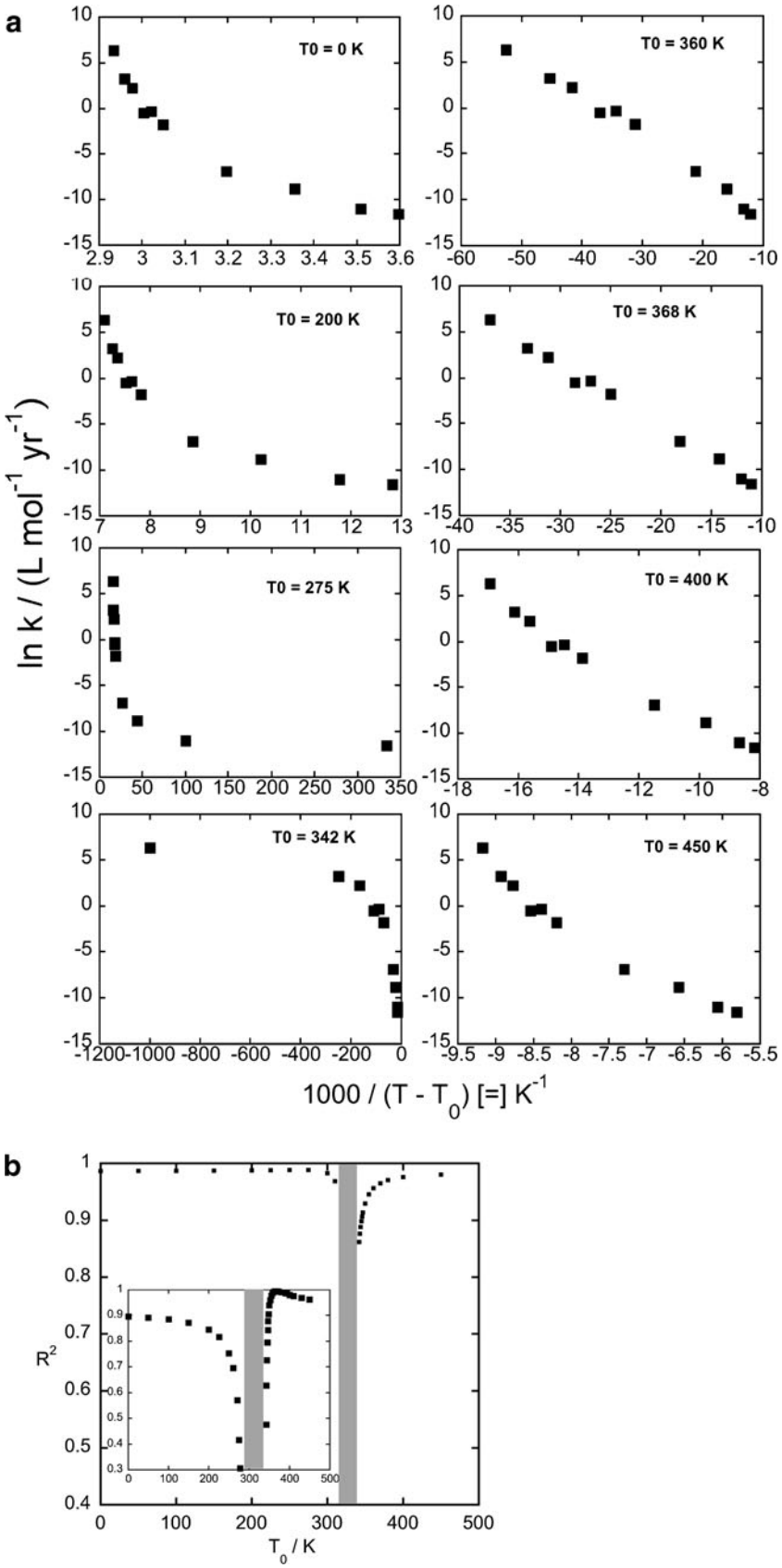

Fig. 5. **a** Illustration of the effect of choosing T_0 to attempt to linearize the non-Arrhenius MAb data from Fig. 4. **b** R^2 *vs.* T_0 for a linear fit of the results illustrated in **a**; *inset* corresponds to fitting with the full set of experimental temperature values in Fig. 4, while main panel is for fitting using only the four accelerated conditions from the inset to Fig. 4

it could be that one could choose a T_0 within this region somewhat arbitrarily and still obtain useful results.

However, one must also consider that this example employed data over the full set of temperatures, including many at much lower and higher temperatures than conventionally tested in formal accelerated stability programs in industrial laboratories. The main panel of Fig. 5b shows the analogous results if one uses only the four high-T values of k_{obs} from the inset to Fig. 4. Again, one must avoid choosing T_0 within the experimental data range (gray-shaded regions in Fig. 5b), but otherwise essentially any T_0 value gives equivalent results, and thus one could easily select a T_0 value that is inappropriate for linearizing the data for lower temperatures. This follows naturally because with only a small number of data points, over too small of a temperature range, the data will not show sufficiently non-Arrhenius behavior to allow any non-Arrhenius model to regress its additional parameter(s). Any approach based on regression or adjustment of all key model parameters using just the values of k *vs.* temperature will encounter this issue. As noted above in the context of the LENP approach, this is not an issue if there is an alternative way to estimate the parameter(s) that describe the non-Arrhenius behavior, but making such estimates may involve different difficulties (56).

Can a simpler linearization method be created for improved rate predictions, and one that can be done without have to introduce a third adjustable parameter (as in Eq. 6), or a third parameter that may be difficult to obtain from independent experiments (as in the LENP model)? A simple approach that at first glance does not involve more than two adjustable or fitted parameters is to try to empirically "linearize" the ln$k$ data by a mathematical transform such as taking an additional logarithm, i.e., consider ln(lnk_{obs}) *vs.* 1/T based on the following proposed relationship:

$$\ln\ k = Ae^{-E_a/\mathrm{RT}} \tag{7}$$

Unless one has a fundamental (e.g., mechanistic) reason for choosing a particular transform, this is a purely empirical approach to help eliminate curvature in lnk_{obs} without requiring a third parameter that introduces non-linearity in the model. However, choosing a particular transform could introduce bias or hidden additional arbitrary parameters, as will be illustrated below.

To illustrate this point, consider the simple example of attempting to "linearize" the data from Fig. 4 by taking the logarithm of ln$k$. Figure 6 (insert) shows the k data for the same monoclonal antibody (57) as in Fig. 4, but now plotted as ln(ln$k$) *vs.* 1/T. Inspection of the data shown as squares in Fig. 6 gives the impression that the transformed rate data is indeed more linear. However, in order to be able to take this second logarithm, the k data had to be scaled to assure none of the values of ln$k$ were below zero—i.e., mathematically, k must be above 1 in order to be able to take its logarithm more than once. As such, what is plotted in Fig. 6 is ln[ln(kY)], with Y a scaling constant. By judicious choice of units for k, one may make Y=1, but ultimately Y is a new, semi-arbitrary constant or fitting parameter. In order for this transform of k to be reliable in practice, the choice for Y (or for the units of k) must not affect the resulting prediction when extrapolating from high T to lower T.

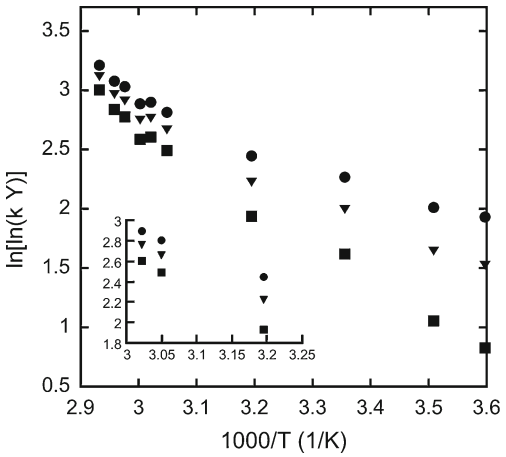

Fig. 6. Logarithm of lnk_{obs} *vs.* 1/T, with k_{obs} scaled by Y to assure the argument of ln() remains non-negative. The data are from Kayser *et al.* (57), with k_{obs} given in units of liters per mole per year. The symbols correspond to Y values of 10^6 (*squares*), 10^7 (*triangles*), and 10^8 (*circles*). The inset uses zoom-in of the same axes as in the main panel, to show just three high temperature conditions (40, 55, and 58 C) to illustrate uncertainties

The other data sets (different symbols) in Fig. 6 show the effect of changing Y on the resulting behavior of ln[ln(kY)] *vs.* 1/T. Clearly, the choice of Y has a pronounced effect on the degree of linearity of the transformed k data at low T. But if one inspects only the high-T regime, any choice of Y seems to produce reasonably linear behavior. Therefore the choice of Y is effectively arbitrary when one has only the high-T data available (as would occur during product development). This illustrates that even simple transformations of k data suffer from significant uncertainties when one seeks to extrapolate inherently non-linear data if one does not know the inherent form of the non-linearity. Thus, transforming either the k or the T values involves similar issues when one seeks to extrapolate rather than interpolate.

Returning briefly to the question of the number of accelerated (high-T) temperature points that are used in common practice in industrial laboratories, one must realize that for linear behavior one needs two parameters (e.g., intercept and slope) to mathematically describe the behavior—and these are the fitted E_a (slope) and A (intercept) values. If one has nonlinear Arrhenius behavior, then one needs at least three (and possibly more) parameters to have any rational chance of extrapolating accurately to lower temperatures. In the above examples, the "third parameter" was, respectively, $\Delta c_{p,\,un}$, n, T_0, or Y. If one is to fit any three-parameter model to experimental data, then there will be statistical uncertainty or variability in that data.

Therefore, in practice it is unrealistic to use only three temperature values if one is to expect reliable predictions from such a fit. Rather, one will necessarily need significantly more conditions if one is to reliably predict low-temperature behavior from just the high-temperature kinetics, if there is significantly non-Arrhenius behavior. Purely mathematically, if one is fit three parameters then a bare minimum of four temperatures must be used if there is to be any statistical relevance to the fitted parameters, and even then the uncertainties in the statistical parameters will be so large as to make them of little use. Given that the data themselves (e.g.,

the k_{obs} values) will also have statistical uncertainty, even four temperatures are not realistically enough to obtain reliable results. Secondly, one must consider that it is not just the number of temperatures that is essential; it is also the range of temperatures. As shown above, if one is to fit k_{obs} *vs.* T or $1/T$, then the data must extend over a large enough range of temperature that non-Arrhenius behavior is apparent within the uncertainty of the rate data. It is not possible to give a general "rule" as to how many temperatures this will require, as that will depend on the protein and underlying mechanism of the non-Arrhenius behavior. In the examples above, on the order of five to ten different temperatures, spanning a range of at least ~20°C, were needed to achieve this.

Pragmatically, the two simplest solutions to this problem are: (1) to obtain sufficient data points and/or supplemental data from complementary experiments to capture the non-linear behavior under accelerated conditions to make a reliable low-T prediction more practical; (2) develop improved means to accurately and quickly obtain aggregation rates at conditions that do not require one to extrapolate so far in terms of the extrapolating variable (in this case, temperature). The former is illustrated by each of the approaches presented earlier in this section, while the latter has been discussed in some detail elsewhere by one of us (8,56). In either case, our argument is that current practice regarding experimental design in the industry may mimic too closely what is done for small-molecule shelf life prediction. Doing so will inherently limit, or even preclude one from accurately accounting for, and accurately predicting, non-Arrhenius behavior for protein aggregation rates.

SUMMARY

Protein aggregation is a complex, multi-step process. Nonlinear or non-Arrhenius aggregation rates are most typically observed for proteins that exist primarily in folded states in their monomeric form, although it also occurs when there is a change in rate-limiting step (e.g., via chemical modifications) with a change in temperature. In the former case, temperature influences protein conformational stability effectively "parabolically," and this is a natural consequence of the large heat capacity changes upon unfolding and exposure of (primarily) hydrophobic residues. By analogy, the conformational changes within dimers or larger oligomers that are needed to stabilize aggregates may play a role in non-Arrhenius aggregation rates if they are involved in the rate-limiting step(s) for aggregation. Similar arguments apply for aggregation occurring in bulk solution and at interfaces. Different approaches have been offered in the literature and in this report to enable more accurate extrapolation of accelerated (high temperature) protein aggregation rates to lower temperature, but each has its limitations. Truly quantitative prediction of low-temperature aggregation rates, based on only accelerated data, remains an outstanding challenge. However, the factors and approaches reviewed here may prove useful in at least semi-quantitative extrapolation to low temperature based on accelerated stability studies. However, doing so may require researchers to incorporate significantly more temperatures and/or a wider range of temperature for accelerated stability programs than is the current the norm in their organization.

ACKNOWLEDGMENTS

The authors thank B. Trout, V. Kayser, V. Voynov, and N. Chennamsetty for their suggestions and careful reading of this review in manuscript form. It is a pleasure to submit this review as part of the issue honoring Professor Garnet Peck.

REFERENCES

1. Laidler KJ. Chemical Kinetics. 2nd ed. New York: McGraw-Hill; 1965.
2. Waterman KC, Adami RC. Accelerated aging: prediction of chemical stability of pharmaceuticals. Int J Pharm. 2005;293(1–2):101–25.
3. Thirumangalathu R, Krishnan S, Bondarenko P, Speed-Ricci M, Randolph TW, Carpenter JF, *et al.* Oxidation of methionine residues in recombinant human interleukin-1 receptor antagonist: implications of conformational stability on protein oxidation kinetics. Biochemistry. 2007;46(21):6213–24.
4. Pan B, Abel J, Ricci MS, Brems DN, Wang DI, Trout BL. Comparative oxidation studies of methionine residues reflect a structural effect on chemical kinetics in rhG-CSF. Biochemistry. 2006;45(51):15430–43.
5. Nellis DF, Michiel DF, Jiang MS, Esposito D, Davis R, Jiang H, *et al.* Characterization of recombinant human IL-15 deamidation and its practical elimination through substitution of asparagine 77. Pharm Res. 2011;29(3):722–38.
6. Crowley T. Case studies: utilizing arrhenius kinetics to predict stability of recombinant proteins. Informa Life Science's Non-Antibody Protein Production, June 8–9, ed. Berlin, Germany; 2011.
7. Hawe A, Poole R, Romeijn S, Kasper P, van der Heijden R, Jiskoot W. Towards heat-stable oxytocin formulations: analysis of degradation kinetics and identification of degradation products. Pharm Res. 2009;26(7):1679–88.
8. Weiss WF, Young TM, Roberts CJ. Principles, approaches, and challenges for predicting protein aggregation rates and shelf life. J Pharm Sci. 2009;98(4):1246–77.
9. Plaza del Pino IM, Ibarra-Molero B, Sanchez-Ruiz JM. Lower kinetic limit to protein thermal stability: a proposal regarding protein stability *in vivo* and its relation with misfolding diseases. Proteins Struct, Funct, Bioinform. 2000;40(1):58–70.
10. Torrent J, Marchal S, Ribo M, Vilanova M, Georges C, Dupont Y, *et al.* Distinct unfolding and refolding pathways of ribonuclease a revealed by heating and cooling temperature jumps. Biophys J. 2008;94(10):4056–65.
11. Andrews JM, Roberts CJ. Non-native aggregation of α-chymotrypsinogen occurs through nucleation and growth with competing nucleus sizes and negative activation energies. Biochemistry. 2007;46(25):7558–71.
12. Falconer RJ, Marangon M, Van Sluyter SC, Neilson KA, Chan C, Waters EJ. Thermal stability of thaumatin-like protein, chitinase, and invertase isolated from Sauvignon blanc and Semillon juice and their role in haze formation in wine. J Agric Food Chem. 2010;58(2):975–80.
13. Waegele MM, Gai F. Infrared study of the folding mechanism of a helical hairpin: porcine PYY. Biochemistry. 2010;49(35):7659–64.
14. Yoshioka S, Tajima S, Aso Y, Kojima S. Inactivation and aggregation of beta-galactosidase in lyophilized formulation described by Kohlrausch–Williams–Watts stretched exponential function. Pharm Res. 2003;20(10):1655–60.
15. Yoshioka S, Aso Y, Izutsu K, Kojima S. Is stability prediction possible for protein drugs? Denaturation kinetics of beta-galactosidase in solution. Pharm Res. 1994;11(12):1721–5.
16. Weijers M, Barneveld PA, Cohen SMA, Visschers RW. Heat-induced denaturation and aggregation of ovalbumin at neutral pH described by irreversible first-order kinetics. Protein Sci. 2003;12(12):2693–703.
17. Waseem A, Salahuddin A. Anomalous temperature-dependence of the specific interaction of concanavalin A with a multivalent ligand-dextran. Biochim Biophys Acta. 1983;746(1–2):65–71.

18. Mehrotra S, Balaram H. Methanocaldococcus jannaschii adenylosuccinate synthetase: studies on temperature dependence of catalytic activity and structural stability. Biochim Biophys Acta. 2008;1784(12):2019–28.
19. Matagne A, Jamin M, Chung EW, Robinson CV, Radford SE, Dobson CM. Thermal unfolding of an intermediate is associated with non-Arrhenius kinetics in the folding of hen lysozyme. J Mol Biol. 2000;297(1):193–210.
20. Oliveberg M, Tan Y-J, Silow M, Fersht AR. The changing nature of the protein folding transition state: implications for the shape of the free-energy profile for folding. J Mol Biol. 1998;277(4):933–43.
21. Yang WY, Gruebele M. Rate–temperature relationships in lambda-repressor fragment lambda 6–85 folding. Biochemistry (Mosc). 2004;43(41):13018–25.
22. Tölgyesi F, Ullrich B, Fidy J. Tryptophan phosphorescence signals characteristic changes in protein dynamics at physiological temperatures. Biochim Biophys Acta Protein Struct Mol Enzymol. 1999;1435(1–2):1–6.
23. Chen BL, Baase WA, Schellman JA. Low-temperature unfolding of a mutant of phage T4 lysozyme. 2. Kinetic investigations. Biochemistry (Mosc). 1989;28(2):691–9.
24. Oliveberg M, Tan Y-J, Fersht AR. Negative activation enthalpies in the kinetics of protein folding. Proc Natl Acad Sci U S A. 1995;92(19):8926–9.
25. Levitsky V, Melik-Nubarov NS, Siksnis VA, Grinberg V, Burova TV, Levashov AV, *et al*. Reversible conformational transition gives rise to 'zig–zag' temperature dependence of the rate constant of irreversible thermoinactivation of enzymes. Eur J Biochem. 1994;219(1–2):219–30.
26. Brummitt RK, Nesta DP, Roberts CJ. Predicting accelerated aggregation rates for monoclonal antibody formulations, and challenges for low-temperature predictions. J Pharm Sci. 2013;(in press)
27. Stefani M. Structural polymorphism of amyloid oligomers and fibrils underlies different fibrillization pathways: immunogenicity and cytotoxicity. Curr Protein Pept Sci. 2010;11(5):343–54.
28. Ferrone F. Analysis of protein aggregation kinetics. Methods Enzymol. 1999;309:256–74.
29. Roberts CJ. Non-native protein aggregation kinetics. Biotechnol Bioeng. 2007;98(5):927–38.
30. Li Y, Ogunnaike BA, Roberts CJ. Multi-variate approach to global protein aggregation behavior and kinetics: effects of pH, NaCl, and temperature for α-chymotrypsinogen A. J Pharm Sci. 2010;99(2):645–62.
31. Morris AM, Watzky MA, Finke RG. Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature. Biochim Biophys Acta Proteins Proteomics. 2009;1794(3):375–97.
32. Andrews JM, Roberts CJ. A Lumry–Eyring nucleated polymerization model of protein aggregation kinetics: 1. Aggregation with pre-equilibrated unfolding. J Phys Chem B. 2007;111(27):7897–913.
33. Li Y, Roberts CJ. Lumry-eyring nucleated-polymerization model of protein aggregation kinetics. 2. Competing growth via condensation and chain polymerization. J Phys Chem B. 2009;113(19):7020–32.
34. Roberts CJ. Kinetics of irreversible protein aggregation: analysis of extended Lumry–Eyring models and implications for predicting protein shelf life. J Phys Chem B. 2003;107(5):1194–207.
35. Brummitt RK, Nesta DP, Chang L, Chase SF, Laue TM, Roberts CJ. Nonnative aggregation of an IgG1 antibody in acidic conditions: part 1. Unfolding, colloidal interactions, and formation of high-molecular-weight aggregates. J Pharm Sci. 2011;100(6):2087–103.
36. Brummitt RK, Nesta DP, Chang L, Kroetsch AM, Roberts CJ. Nonnative aggregation of an IgG1 antibody in acidic conditions, part 2: nucleation and growth kinetics with competing growth mechanisms. J Pharm Sci. 2011;100(6):2104–19.
37. Sahin E, Jordan JL, Spatara ML, Naranjo A, Costanzo JA, Weiss WF, *et al*. Computational design and biophysical characterization of aggregation-resistant point mutations for γD crystallin illustrate a balance of conformational stability and intrinsic aggregation propensity. Biochemistry (Mosc). 2011;50(5):628–39.
38. Sabaté R, Gallardo M, Estelrich J. Temperature dependence of the nucleation constant rate in β amyloid fibrillogenesis. Int J Biol Macromol. 2005;35(1–2):9–13.
39. Auer S, Dobson CM, Vendruscolo M. Characterization of the nucleation barriers for protein aggregation and amyloid formation. HFSP J. 2007;1(2):137–46.
40. Lee CC, Walters RH, Murphy RM. Reconsidering the mechanism of polyglutamine peptide aggregation. Biochemistry (Mosc). 2007;46(44):12810–20.
41. Nayak A, Sorci M, Krueger S, Belfort G. A universal pathway for amyloid nucleus and precursor formation for insulin. Proteins Struct Funct Bioinforma. 2009;74(3):556–65.
42. Carrotta R, Manno M, Bulone D, Martorana V, San Biagio PL. Protofibril formation of amyloid beta-protein at low pH via a non-cooperative elongation mechanism. J Biol Chem. 2005;280(34):30001–8.
43. Kusumoto Y, Lomakin A, Teplow DB, Benedek GB. Temperature dependence of amyloid beta-protein fibrillization. Proc Natl Acad Sci U S A. 1998;95(21):12277–82.
44. Smith MI, Sharp JS, Roberts CJ. Nucleation and growth of insulin fibrils in bulk solution and at hydrophobic polystyrene surfaces. Biophys J. 2007;93(6):2143–51.
45. Mauro M, Craparo EF, Podesta A, Bulone D, Carrotta R, Martorana V, *et al*. Kinetics of different processes in human insulin amyloid formation. J Mol Biol. 2007;366(1):258–74.
46. Sabate R, Castillo V, Espargaro A, Saupe SJ, Ventura S. Energy barriers for HET-s prion forming domain amyloid formation. FEBS J. 2009;276(18):5053–64.
47. Uversky VN, Li J, Fink AL. Evidence for a partially folded intermediate in alpha-synuclein fibril formation. J Biol Chem. 2001;276(14):10737–44.
48. Morel B, Varela L, Azuaga AI, Conejero-Lara F. Environmental conditions affect the kinetics of nucleation of amyloid fibrils and determine their morphology. Biophys J. 2010;99(11):3801–10.
49. Rosenqvist E, Jossang T, Feder J. Thermal properties of human IgG. Mol Immunol. 1987;24(5):495–501.
50. Wang B, Tchessalov S, Cicerone MT, Warne NW, Pikal MJ. Impact of sucrose level on storage stability of proteins in freeze-dried solids: II. Correlation of aggregation rate with protein structure and molecular mobility. J Pharm Sci. 2008. doi:10.1002/jps.21622.
51. Duddu SP, Dal Monte PR. Effect of glass transition temperature on the stability of lyophilized formulations containing a chimeric therapeutic monoclonal antibody. Pharm Res. 1997;14(5):591–5.
52. Breen ED, Curley JG, Overcashier DE, Hsu CC, Shire SJ. Effect of moisture on the stability of a lyophilized humanized monoclonal antibody formulation. Pharm Res. 2001;18(9):1345–53.
53. Pérez-Moral N, Adnet C, Noel TR, Parker R. Characterization of the rate of thermally-induced aggregation of β-lactoglobulin and its trehalose mixtures in the glass state. Biomacromolecules. 2010;11(11):2985–92.
54. Stratton LP, Kelly RM, Rowe J, Shively JE, Smith DD, Carpenter JF, *et al*. Controlling deamidation rates in a model peptide: effects of temperature, peptide concentration, and additives. J Pharm Sci. 2001;90(12):2141–8.
55. Roberts CJ, Darrington RT, Whitley MB. Irreversible aggregation of recombinant bovine granulocyte-colony stimulating factor (bG-CSF) and implications for predicting protein shelf life. J Pharm Sci. 2003;92(5):1095–111.
56. Brummitt RK, Nesta DP, Roberts CJ. Predicting accelerated aggregation rates for monoclonal antibody formulations, and challenges for low-temperature predictions. J Pharm Sci. 2011;100:4234–43.
57. Kayser V, Chennamsetty N, Voynov V, Helk B, Forrer K, Trout BL. Evaluation of a non-arrhenius model for therapeutic monoclonal antibody aggregation. J Pharm Sci. 2011;100(7):2526–42.
58. Perico N, Purtell J, Dillon TM, Ricci MS. Conformational implications of an inversed pH-dependent antibody aggregation. J Pharm Sci. 2009;98(9):3031–42.
59. Talla-Singh D, Stites WE. Refinement of noncalorimetric determination of the change in heat capacity, DeltaC(p), of protein unfolding and validation across a wide temperature range. Proteins. 2008;71(4):1607–16.
60. Rees DC, Robertson AD. Some thermodynamic implications for the thermostability of proteins. Protein Sci. 2001;10(6):1187–94.

61. Becktel WJ, Schellman JA. Protein stability curves. Biopolymers. 1987;26(11):1859–77.
62. Wong HJ, Stathopulos PB, Bonner JM, Sawyer M, Meiering EM. Non-linear effects of temperature and urea on the thermodynamics and kinetics of folding and unfolding of hisactophilin. J Mol Biol. 2004;344(4):1089–107.
63. Kumar S, Tsai CJ, Nussinov R. Maximal stabilities of reversible two-state proteins. Biochemistry (Mosc). 2002;41(17):5359–74.
64. Ganesh C, Eswar N, Srivastava S, Ramakrishnan C, Varadarajan R. Prediction of the maximal stability temperature of monomeric globular proteins solely from amino acid sequence. FEBS Lett. 1999;454(1–2):31–6.
65. Dias CL, Ala-Nissila T, Wong-ekkabut J, Vattulainen I, Grant M, Karttunen M. The hydrophobic effect and its role in cold denaturation. Cryobiology. 2010;60(1):91–9.
66. Privalov PL. Stability of proteins. Adv Protein Chem. 1979;33:167–241.
67. Hilser VJ, García-Moreno EB, Oas TG, Kapp G, Whitten ST. A statistical thermodynamic model of the protein ensemble. Chem Rev. 2012;106(5):1545–58.
68. Eisenberg D, Nelson R, Sawaya MR, Balbirnie M, Sambashivan S, Ivanova MI, *et al.* The structural biology of protein aggregation diseases: fundamental questions and some answers. Acc Chem Res. 2006;39(9):568–75.
69. Cairoli S, Iametti S, Bonomi F. Reversible and irreversible modifications of beta-lactoglobulin upon exposure to heat. J Protein Chem. 1994;13(3):347–54.
70. Liu D, Ren D, Huang H, Dankberg J, Rosenfeld R, Cocco MJ, *et al.* Structure and stability changes of human IgG1 Fc as a consequence of methionine oxidation. Biochemistry. 2008;47(18):5088–100.
71. Tomizawa H, Yamada H, Imoto T. The mechanism of irreversible inactivation of lysozyme at pH 4 and 100 degrees C. Biochemistry. 1994;33(44):13032–7.
72. Bernacki JP, Murphy RM. Length-dependent aggregation of uninterrupted polyalanine. Biochemistry (Mosc): Peptides; 2011.
73. Banks DD, Hambly DM, Scavezze JL, Siska CC, Stackhouse NL, Gadgil HS. The effect of sucrose hydrolysis on the stability of protein therapeutics during accelerated formulation studies. J Pharm Sci. 2009;98(12):4501–10.
74. Kolhe P, Amend E, Singh SK. Impact of freezing on pH of buffered solutions and consequences for monoclonal antibody aggregation. Biotechnol Prog. 2010;26(3):727–33.
75. Harris RJ, Kabakoff B, Macchi FD, Shen FJ, Kwong M, Andya JD, *et al.* Identification of multiple sources of charge heterogeneity in a recombinant antibody. J Chromatogr B: Biomed Sci Appl. 2001;752(2):233–45.
76. Duggleby RG. Regression analysis of nonlinear Arrhenius plots: an empirical model and a computer program. Comput Biol Med. 1984;14(4):447–55.
77. Héberger K, Kemény S, Vidóczy T. On the errors of Arrhenius parameters and estimated rate constant values. Int J Chem Kinet. 1987;19(3):171–81.
78. Gierczak T, Talukdar RK, Herndon SC, Vaghjiani GL, Ravishankara AR. Rate coefficients for the reactions of hydroxyl radicals with methane and deuterated methanes. J Phys Chem A. 1997;101(17):3125–34.