DISSERTATION

PART I. FITTING PROTEIN AGGREGATION KINETIC DATA RELEVANT TO NEURODEGENERATIVE DISEASES WITH AN "OCKHAM'S RAZOR" MODEL EN ROUTE TO MEANINGFUL RATE CONSTANTS AND MECHANISTIC INSIGHTS. PART II. DIOXYGENASES: THE DEVELOPMENT OF NEW, AND THE REINVESTIGATION OF PRIOR, PRECATALYSTS

Submitted by

Aimee M. Morris

Department of Chemistry

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2009

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

COLORADO STATE UNIVERSITY

May 6, 2009

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY AIMEE M. MORRIS ENTITLED PART I. FITTING PROTEIN AGGREGATION KINETIC DATA RELEVANT TO NEURODEGENERATIVE DISEASES WITH AN "OCKHAM'S RAZOR" MODEL EN ROUTE TO MEANINGFUL RATE CONSTANTS AND MECHANISTIC INSIGHTS. PART II. DIOXYGENASES: THE DEVELOPMENT OF NEW, AND THE REINVESTIGATION OF PRIOR, PRECATALYSTS BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Oren P. Anderson

Anthony K. Rappé

Thomas Meersmann

Robert Woody

Advisor Richard G. Finke

Department Head Anthony K. Rappé

ABSTRACT OF DISSERTATION

PART I. FITTING PROTEIN AGGREGATION KINETIC DATA RELEVANT TO NEURODEGENERATIVE DISEASES WITH AN "OCKHAM'S RAZOR" MODEL EN ROUTE TO MEANINGFUL RATE CONSTANTS AND MECHANISTIC INSIGHTS. PART II. DIOXYGENASES: THE DEVELOPMENT OF NEW, AND THE REINVESTIGATION OF PRIOR, PRECATALYSTS

This dissertation is presented in two parts. Part I starts with a review of models that have been used to curve-fit or obtain rate constants for protein aggregation kinetic data. Following the review, the research presented in Part I is primarily focused on fitting protein aggregation literature relevant to neurodegenerative diseases using the Finke-Watzky (hereafter F-W) 2-step model of nucleation and autocatalytic growth. Part I includes: (i) the fits to the F-W model and resultant nucleation and growth rate constants of 14 representative data sets of amyloid-β, α-synuclein, and polyglutamine aggregation relevant to Alzheimer's, Parkinson's, and Huntington's diseases, respectively; (ii) the fits of 27 data sets of yeast and mammalian prion aggregation, along with the resultant rate constants and interpretation of factors that contribute to nucleation and growth of prion aggregates; and (iii) a re-examination of variable temperature and variable pH α-synuclein aggregation data in which the insights are elucidated that: (a) the processes of nucleation and growth are energetically similar, (b) the net charge of the

protein affects nucleation, and (c) the lag-time does not, as previously thought, correspond to the rate of nucleation.

Part II begins with a brief review of the importance of dioxygenases followed by an introduction to two important synthetic dioxygenases, the catechol dioxygenase $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ (where 3,5-DTBC = 3,5-di-*tert*-butylcatechol and 3,5-DBSQ = 3,5-di-*tert*-butylsemiquinone) and the claimed polyoxometalate dioxygenase, $[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]^{11-}$. The synthesis and characterization of a new dioxygenase, $V(3,6\text{-DBSQ})(3,6\text{-DTBC})_2$, along with the initial catalytic results with the $H_2(3,6\text{-DTBC})$, substrate are given. Next is a full report of the dioxygenase activity with $H_2(3,5\text{-DTBC})$ and $H_2(3,6\text{-DTBC})$ substrates of three d^0 metal precatalysts: $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$, $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$, and $[MoO(3,5\text{-DTBC})_2]_2$. The d^0 vanadium bound to a semiquinone ligand in both V-precatalysts appears to be an important component for obtaining dioxygenase products from the $H_2(3,5\text{-DTBC})$ and $H_2(3,6\text{-DTBC})$ substrates. Finally, Part II concludes with a reinvestigation a claimed dioxygenase, $[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]^{11-}$ (**1**). Three independent samples of **1** from two different laboratories, samples that also give the same catalysis results as previously reported, are all consistent with the composition of the parent, Ru-free polyoxometalate, $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ (**2**). Also, simple mixtures of **2** plus $[Ru(DMSO)_4Cl_2]$ is a ca. 2-fold more efficient catalyst than "**1**", placing in serious doubt a prior *Nature* paper detailing the claim that "**1**" is a Ru-based, all-inorganic dioxygenase.

Aimee M. Morris
Chemistry Department
Colorado State University
Fort Collins, CO 80523
Summer 2009

TABLE OF CONTENTS

I. INTRODUCTION....1

II. PART I INTRODUCTION. PROTEIN AGGREGATION KINETICS, MECHANISM, AND CURVE-FITTING: A REVIEW OF THE LITERATURE....5

Abstract....6
Introduction....7
Advantages and disadvantages of the physical methods used to monitor protein aggregation....11
Starting proteins, products, and detectable intermediates of protein aggregation....12
Thermochemistry of protein aggregation....18
Approaches to determine the kinetics and mechanism of protein aggregation....19
Three problem areas in the protein aggregation kinetic and mechanistic literature....76
Summary and conclusions....78
References....81
Supporting Information....94

III. FITTING NEUROLOGICAL PROTEIN AGGREGATION KINETIC DATA VIA A 2-STEP, MINIMAL / "OCKHAM'S RAZOR" MODEL: THE FINKE-WATZKY MECHANISM OF NUCLEATION FOLLOWED BY AUTOCATALYTIC SURFACE GROWTH....132

Abstract....133
Introduction....135
Materials and Methods....146
Results and Discussion....148
Summary and conclusions....165
References....168
Supporting Information....175

IV. FITTING YEAST AND MAMMALIAN PROTEIN AGGREGATION KINETIC DATA WITH THE FINKE-WATZKY 2-STEP MODEL OF NUCLEATION AND AUTOCATALYTIC GROWTH.....199

Abstract.....200
Introduction.....201
Materials and Methods.....205
Results.....206
Discussion.....223
Conclusions.....230
References.....233
Supporting Information.....238

V. α-SYNUCLEIN AGGREGATION VARIABLE TEMPERATURE AND VARIABLE pH KINETIC DATA: A RE-ANALYSIS USING THE FINKE-WATZKY 2-STEP MODEL OF NUCLEATION AND AUTOCATALYTIC GROWTH.....247

Abstract.....248
Introduction.....250
Experimental.....252
Results and Discussion.....253
Conclusions.....265
References.....267
Supporting Information.....270

VI. PART II INTRODUCTION: A BRIEF REVIEW OF THE IMPORTANCE OF DIOXYGENASES PLUS AN INTRODUCTION TO TWO SIGNIFICANT SYNTHETIC DIOXYGENASE SYSTEMS.....279

Introduction.....280
The Importance of Dioxygenases.....281
Summary.....285
References.....286

VII. SYNTHESIS AND CHARACTERIZATION OF $V^{V}(3,6\text{-DBSQ})(3,6\text{-DBCAT})_2$, A d^0 METAL COMPLEX WITH DIOXYGENASE CATALYTIC ACTIVITY.....288

Abstract.....289
Main Text.....290
References.....298
Supporting Information.....300

VIII. DIOXYGENASE CATALYSIS BY d^0 METAL-CATECHOLATE COMPLEXES CONTAINING VANADIUM AND MOLYBDENUM FOR H_2(3,5-DTBC) AND H_2(3,6-DTBC) SUBSTRATES.....313

Abstract.....314
Introduction.....316
Results and Discussion.....319
Conclusions.....335
Experimental.....337
References.....343
Supporting Information.....345

IX. RE-INVESTIGATION OF A Ru_2-INCORPORATED POLYOXOMETALATE DIOXYGENASE PRECATALYST, "$[WZnRu_2^{III}(H_2O)(OH)(ZnW_9O_{34})_2]^{11-}$": EVIDENCE FOR MARGINAL, ≤0.2 EQUIVALENTS OF Ru INCORPORATION PLUS FASTER CATALYSIS BY PHYSICAL MIXTURES OF $[Ru^{II}(DMSO)_4Cl_2]$ AND THE PARENT POLYOXOMETALATE $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$.....354

Abstract.....355
Introduction.....357
Results and Discussion.....361
Conclusions.....374
Experimental.....377
References.....385
Supporting Information.....389

X. SUMMARY.....418

APPENDICES

Appendix A. General Statement on "Journals-Format" Theses.....421
Appendix B. Research Proposal.....425

CHAPTER I

INTRODUCTION

This dissertation is presented in two parts. Part I of this dissertation has the theme of fitting protein aggregation kinetic data with a mechanism able to elucidate information on the nucleation and growth of the overall aggregation process. Part II contains the broad theme of investigating dioxygenase catalysts. This dissertation is written in the "journals-format" style (see Appendix A for a discussion of this type of dissertation). It is based on six separate publications (Chapters III-V and VII-IX) and a published literature review (Chapter II). Chapter II is written in the format of *Biochimica et Biophysica Acta* (Elsevier), Chapters III, IV, VI, VII, and IX are written in a format set by the American Chemical Society, and Chapter VIII is written in the format of *Journal of Molecular Catalysis A: Chemical* (Elsevier). Consistency of this dissertation as a single document is achieved by (i) this introduction, (ii) the use of bridging paragraphs at the beginning of each chapter, (iii) a second brief literature introduction to Part II (Chapter VI), and (iv) a final summary chapter. A concise overview of each chapter's contents is presented below.

Chapter II is a literature review of the different models that have been used to fit protein aggregation kinetic data and/or obtain rate constants for aggregation. Each different type of model that has been used to treat protein aggregation kinetic data is discussed along with illustrative examples of the experimental kinetic data that each model is able to fit.

Chapter III is an investigation of protein aggregation kinetic data relevant to Alzheimer's, Parkinson's, and Huntington's diseases. Specifically, we show that a broad range of kinetic data (14 representative data sets) from the literature can be fit with the Finke-Watzky (F-W) 2-step model consisting of nucleation and autocatalytic growth. As such, we are able to deconvolute and report for the first time the nucleation (k_1) and autocatalytic growth (k_2) rate constants for the protein aggregation systems examined therein.

Chapter IV extends the use of the F-W model to prion protein aggregation systems. We show that the 27 different kinetic data sets examined, including both yeast and mammalian prion systems, are well fit by the simple 2-step F-W nucleation and autocatalytic growth model. Moreover, we are able to gain insights from the resultant rate constants on what factors, such as the N-terminus vs the C-terminus, affect nucleation and growth.

Chapter V presents a study of the aggregation of α-synuclein, a protein hypothesized to be an underlying cause of Parkinson's disease. This study involved re-examination using the F-W model, of α-synuclein aggregation kinetic data that was previously analyzed by an empirical equation. From our re-examination we deduced the following insights that were previously unavailable: (i) that contrary to what is believed

in the field, the lag-time is not a good predictor of the nucleation rate, and (ii) that net charge is an important variable controlling the nucleation rate constant of α-synuclein aggregation.

Chapter VI is a brief literature introduction to Part II of this dissertation in which dioxygenase catalysis is introduced. In addition, two notable synthetic dioxygenase systems are discussed. These systems include the record lifetime catalytic-cycle resting state dimer $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ (where 3,5-DTBC = 3,5-di-*tert*-butylcatecholate and 3,5-DBSQ = 3,5-di-*tert*-butylsemiquinone) and the claimed all-inorganic polyoxometalate dioxygenase catalyst $[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]^{11-}$.

Chapter VII contains the synthesis and characterization by multiple physical methods of a new catechol dioxygenase precatalyst, $V(3,6\text{-DBSQ})(3,6\text{-DBCat})_2$ (where 3,6-DBSQ = 3,6-di-*tert*-butylsemiquinone and 3,6-DBCat = 3,6-di-*tert*-butylcatecholate). Included in this chapter are the initial catalytic results of $V(3,6\text{-DBSQ})(3,6\text{-DBCat})_2$ with the substrate $H_2(3,6\text{-DBCat})$.

Chapter VIII presents the dioxygenase catalytic abilities of the d^0 metal precatalysts $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$, $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$, and $[MoO(3,5\text{-DTBC})_2]_2$ with the substrates $H_2(3,5\text{-DTBC})$ and $H_2(3,6\text{-DTBC})$. Interestingly, both of the V- and semiquinone-containing precatalysts give similar yields of the intradiol and extradiol dioxygenase products for the substrates $H_2(3,5\text{-DTBC})$ and $H_2(3,6\text{-DTBC})$, while the Mo-precatalyst gives a majority of the less desired autoxidation product, benzoquinone. Product studies as well as time-dependent EPR studies suggest that the precatalysts $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ and $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$ feed into the

same catalytic cycle, but there appears to be different catalytic cycles depending on whether the H_2(3,5-DTBC) or H_2(3,6-DTBC) substrate is used.

Chapter IX is a reinvestigation of the composition of the putative dioxygenase catalyst $[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]^{11-}$, **1**. We show that three different samples of **1** from two different laboratories are all consistent with the composition of the parent, Ru-free polyoxometalate, $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ (**2**), with little to no incorporation of Ru from the $[Ru(DMSO)_4Cl_2]$ precursor. Moreover, the same catalysis as **1** is observed for **2** alone or $[Ru(DMSO)_4Cl_2]$ alone. More importantly, faster catalysis converting adamantane to 1-adamantanol and 2-adamantanone is observed with simple mixtures of **2** plus 0.13 equiv of $[Ru(DMSO)_4Cl_2]$, strongly suggesting that even if "**1**" exists, it is not a kinetically dominant catalyst in what is also not dioxygenase, but actually autoxidation, catalysis.

Chapter X presents a summary of the material presented in this dissertation.

CHAPTER II

PART I INTRODUCTION. PROTEIN AGGREGATION KINETICS, MECHANISM, AND CURVE-FITTING: A REVIEW OF THE LITERATURE

This dissertation chapter contains the manuscript of a review published in *Biochimica et Biophysica Acta: Proteins and Proteomics* **2009**, *1794*, 375-397. This is a literature review of the various approaches that have been used to analyze protein aggregation kinetic data. A focus of the review is on the models that have been used to try to fit protein aggregation kinetic data.

The following are contributions to this review from Aimee M. Morris: (i) compiling the original literature on which this review is based; (ii) writing the first draft version of the review; (iii) then working on subsequent versions of this review with Dr. Murielle A. Watzky and Prof. Richard G. Finke acting in advisory roles to bring the manuscript up its current, published form.

Part I Introduction: Protein Aggregation Kinetics, Mechanism, and Curve-Fitting: A Review of the Literature

Aimee M. Morris, Murielle A. Watzky and Richard G. Finke

Abstract

Protein aggregation is an important phenomenon that alternatively is part of the normal functioning of nature or, central to this review, has negative consequences via its hypothesized central role in neurodegenerative diseases. A key to controlling protein aggregation is understanding the mechanism(s) of protein aggregation. Kinetic studies, data curve-fitting, and analysis are, in turn, keys to rigorous mechanistic studies. The main goal of this review is to analyze and report on the primary literature contributions to protein aggregation kinetics, mechanism, and curve-fitting. Following a brief introduction, the multiple different physical methods that have been employed to follow protein aggregation are presented and briefly discussed. Next, key information on the starting proteins and especially the products, and any detectable intermediates, involved in protein aggregation are presented. This is followed by tabulation (in the Supporting Information) and discussion (in the main text), of the many approaches in the literature striving to determine the kinetics and mechanism of protein aggregation. It is found that these approaches can be broadly divided into three categories: (i) kinetic and

thermodynamic, (ii) empirical, and (iii) other approaches. The first two approaches are the main focus of the present contribution, their goal being curve-fitting the available kinetic data and obtaining quantitative rate constants characterizing the nucleation, growth, and any other parts of the overall aggregation process. The large literature of protein aggregation is distilled down to five classes of postulated mechanisms: i) the subsequent monomer addition mechanism, ii) the reversible association mechanism, iii) prion aggregation mechanisms, iv) an "Ockham's razor"/minimalistic model first presented in 1997 and known as the Finke-Watzky 2-step model, and v) quantitative structure activity relationship models. These five classes of mechanisms are reviewed in detail in historical order; where possible corresponding kinetic equations, and fits to aggregation data via the proposed mechanisms, are analyzed and discussed. The five classes of mechanisms are then analyzed and discussed in terms of their similarities and differences to one another. Also included is a brief discussion of selected empirical approaches used to investigate protein aggregation. Three problem areas in the protein aggregation kinetic and mechanistic studies area are identified, and a Summary and Conclusions section is provided en route to moving the field forward towards the still unachieved goal of unequivocal elucidation of the mechanism(s) of protein aggregation.

Introduction

The aggregation[1] of proteins such as amyloid-β, polyglutamine, α-synuclein, and prions has been suggested to be intimately associated with neurodegenerative disorders

[1] Herein we use the term "protein aggregation" to describe the process of protein monomers reacting to form fibrils and amorphous aggregates. The term "aggregation" has as its chemical dictionary definition: "A process that results in the formation of

such as Alzheimer's [1], Huntington's [2], Parkinson's [3], and prion [4] diseases, respectively. Aggregation is also a nuisance in industrial applications where it can interfere with the production and characterization of therapeutic polypeptides [5]. Naturally occurring, productive protein aggregation is also important in nature in cases such as the protein fibrillation reaction of $n(\text{G-actin}) \rightarrow (\text{F-actin})_n$, where G-actin is the globular, and F-actin the fibrillar form, of the protein actin.

For the purposes of this review, we will categorize protein aggregation into three classes: (i) *naturally occurring, productive aggregation* as in the $n(\text{G-actin}) \rightarrow (\text{F-actin})_n$ example mentioned. This reaction occurs throughout the human body, as well as in other organisms, and is necessary in controlling the mobility and shape of the cells [6]. Another example of naturally occurring protein aggregation includes the enzyme glutamate dehydrogenase [7,8,9,10]. Both actin and glutamate dehydrogenase function with aggregation of a protein in its native state. A second class of aggregation phenomenon can be classified as (ii) *unwanted aggregation in biology*. This class includes α-synuclein, amyloid β, polyglutamine, and prions as common examples of proteins that aggregate and are suspected to play a key role in the neurodegenerative diseases Parkinson's [3], Alzheimer's [1], Huntington's [2], and prion [4] diseases, respectively. This type of aggregation is generally believed to involve aggregation of the protein in a non-native state (*vide infra*). The final class of aggregation phenomenon is (iii) *unwanted aggregation in an industrial setting*. This class of aggregation usually

---

[groups of atoms or molecules that are held together in any way]." Alternatively, one could use "agglomeration" since it is defined as: "An indiscriminately formed cluster of particles" [McGraw-Hill Dictionary of Scientific and Technical Terms, 6th Ed., McGraw-Hill, New York, 2003]. In short, we use herein the term "protein aggregation" since it is the most common term in the literature and is acceptable under the definition given above.

produces amorphous aggregates and its control and understanding is important to the biotechnology industry for keeping proteins in a non-aggregated, bottleable, long-shelf-life form [11].

Because of its importance, the kinetics and mechanism of protein aggregation have been of interest for approximately fifty years [12]. Protein aggregation is, therefore, a topic that has been the subject of numerous other recent reviews, although from perspectives different than herein [13,14,15,16,17,18,19,20,21,22,23,24,25,26]. Of particular interest is the excellent and critical review on the detailed steps of protein aggregation recently published by C. J. Roberts [14], as well as a review on entropy-driven polymerization of proteins by M. A. Lauffer [26]. However, still missing in our opinion among the available reviews of the expansive protein aggregation literature is an analysis and review focusing on models that can fit kinetic data, give useful, quantitative rate constants, and ideally provide mechanistic insight. Key questions to be answered herein include: (i) How many distinct mechanisms actually exist in the literature for protein aggregation? (ii) What is the essence of each mechanism? (iii) Which mechanisms or models have been used to curve-fit kinetic data? (iv) Do any of these mechanisms have similarities to each other? Also, (v) which of the terms used in the sometimes confusing nomenclature[2] in the protein aggregation literature have the same

[2] The nomenclature in the protein agglomeration literature can be confusing since multiple terms are used for more or less the same phenomenon or meaning. Specifically, used more or less equivalently are the terms: induction period and lag phase to indicate the time before measurable aggregation occurs, and the terms elongation, aggregation, fibrillation, and polymerization all meaning what we will term herein as growth. Adding further confusion is the use of the term "heterogeneous nucleation" to mean, we infer, the dictionary definition of "heterogeneous", that is, nucleation of different, new polymeric aggregates on the surface of existing ("heterogeneous") aggregates or polymers. In the extensive, traditional nucleation theory literature, "heterogeneous nucleation" has been

meaning? These are some of the key questions we hope to answer herein. In short, *a main goal of the present contribution is to analyze and report the primary (in our opinion) literature contributions to protein aggregation kinetics, mechanism, and curve-fitting.*

In what follows we have tried to identify key papers in terms of the 5 main classes (*vide infra*) of mechanistic models in the literature, and to trace each (class of) mechanism back to its earliest origins. Our goal is to distill the literature to its essential components, again in our view. We apologize in advance to the authors of literature we were not able to cover in the space available or have somehow inadvertently missed.

We begin with a brief survey of the physical methods used to measure protein aggregation noting whether the methods are direct or indirect, in-situ or ex-situ, and whether the method is able to measure kinetics. Next, we discuss what is known about the starting proteins, products, and intermediates of protein aggregation—since knowing one's products and intermediates is key to rigorous mechanistic science. Third, we tabulate and discuss the main thermodynamic and kinetic based studies we have found that support what turns out to be the 5 main classes of suggested mechanisms of protein aggregation, all in a somewhat historical order. We also briefly discuss empirical approaches that have been used to fit protein aggregation kinetic data. We end with a discussion of what seems to be some of the common pitfalls in attacking the highly complex problem of the mechanism(s) of protein aggregation. We also list some of the

---

used to mean nucleation in two different phases. Another important definition for any nucleation and growth phenomenon such as protein aggregation is the "critical nucleus", and we use this term herein as used in the nucleation theory literature.

important unsolved problems and hence needed future research directions, and then conclude with a Summary and Conclusions section of the highlights of this review.

Advantages and Disadvantages of the Physical Methods Used to Monitor Protein Aggregation.

The kinetics and products of protein aggregation have been measured using at least 18 different analytical techniques, each having its own intrinsic advantages and disadvantages. Each technique is summarized in Table 1 and discussed in more detail in the Supporting Information. The interested reader is also referred to the recent, excellent review by S. E. Bondos that compares the various methods used to detect protein aggregation, with an emphasis on the concentration and volume ranges for each method, and organized by techniques that: (i) detect protein aggregates, (ii) screen buffers to improve protein stability, and (iii) characterize protein aggregates [27]. Indicated in Table 1 is if the physical methods listed there are used as *direct* or *indirect*[3] methods for the systems at hand, and if they can be used *in-situ* or involve *ex-situ* use and sample preparation. In protein aggregation, as with all science, the use of multiple, complementary, ideally direct, in-situ physical methods is of course preferred.

One notable study comparing six different physical methods (ANS fluorescence, chymotrypsin resistance, congo red binding, light scattering, SDS-PAGE, and sedimentation) has appeared [28]. Each of the different physical methods shows, for the

[3] We were unable to find a definition for direct vs. indirect physical methods in the literature. Therefore, we will use the term direct physical method to mean a method that measures a property that is directly affected by the aggregation process, while an indirect method measures a property that is only indirectly affected by the aggregation process.

first time (to our knowledge) the same kinetic rate constants within experimental error [29], making that study an especially notable one [28].

Table 1. Physical methods used in the literature to analyze protein aggregation.

Method	Direct/ Indirect[3]	In-situ/Ex-situ	Measure Kinetics?	Selected Reference(s)
Absorbance	Direct	Concentration dependent[a]	Yes	[30]
Atomic Force Microscopy	Direct	In-situ	Yes	[30]
Calorimetry	Direct	In-situ	Yes	[27,31]
Circular Dichroism	Direct	Concentration dependent[a]	Yes	[32]
Dyes	Indirect	Concentration dependent[a]	Yes	[30,33]
Electron Microscopy	Direct	Ex-situ	No	[27,30]
Electron Paramagnetic Resonance Spectroscopy	Indirect	Concentration dependent[a]	Yes	[34]
Flow Birefringence	Direct	In-situ	Yes	[35,36]
Fluorescence Spectroscopy with an intrinsic fluorophore	Direct	Concentration dependent[a]	Yes	[33]
Fluorescence Spectroscopy with an extrinsic fluorophore	Indirect	Concentration dependent[a]	Yes	[33]
Fourier Transform Infrared Spectroscopy	Direct	Concentration dependent[a]	Yes	[27]
Light Scattering	Direct	In-situ	Yes	[37,38,39]
Mass Spectrometry	Direct	Ex-situ	No	[27,40,41]
Nuclear Magnetic Resonance Spectroscopy	Direct	Concentration dependent[a]	Yes	[42,43]
Quartz Crystal Oscillator Measurements	Direct	Concentration dependent[a]	Yes	[44]
Turbidity	Direct	In-situ	Yes	[30,45]
Viscosity	Direct	In-situ	Yes	[46]
X-ray Diffraction	Direct	Concentration dependent[a]	No[b]	[30,47]

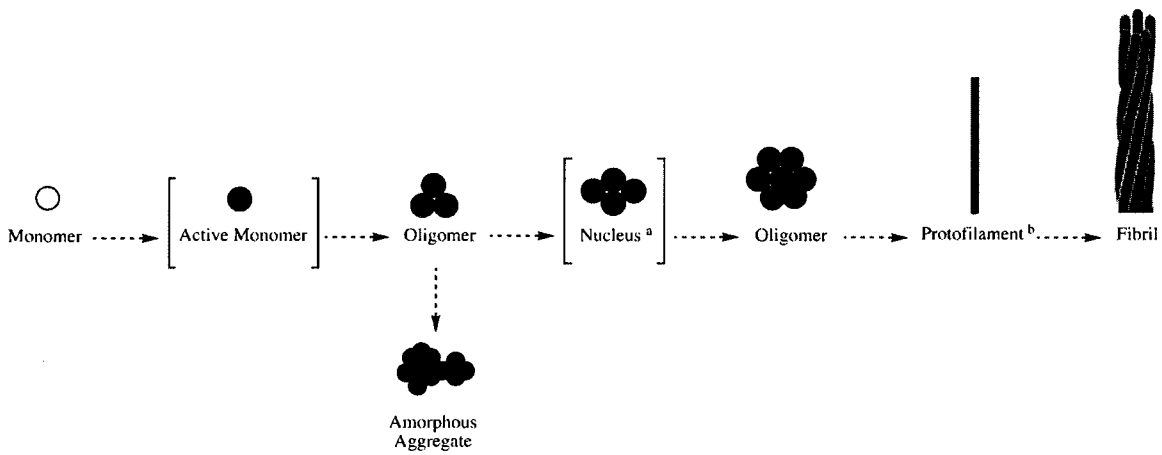
[a] By "concentration dependent" we mean that this method can be in-situ if the aggregation conditions are within the detection limits of the physical method.
[b] Or "yes" in principle if Synchrotron radiation is used.

Table 1 shows that many different methods for measuring protein aggregation exist, each with their own intrinsic advantages and disadvantages. Therefore, the use of as many multiple, complementary methods as possible when studying the kinetics and mechanism of protein aggregation is critical.

Starting Proteins, Products, and Detectable Intermediates of Protein Aggregation. It has long been known that "know your product(s)" is the first rule of

rigorous mechanistic science, since the steps in any proposed mechanism must add up to those observed products. We review, therefore, what is generally known vs. not known about the starting proteins, products, and any detectable intermediates of protein aggregation *when fibrils are formed*, as this category involves many important studies of (native and non-native) protein aggregation. While a main goal of this review is to examine protein aggregation mechanisms supported by kinetic data, presentation of a schematic representation of the overall pathway generally proposed for the aggregation (fibrillation) of the proteins is useful for the discussion that follows and is, therefore, shown in Scheme 1.

Scheme 1. Idealized schematic representation of a general, overall pathway for the formation of protein fibrils, including some possible intermediates.[a,b]



[a]There does seem to be a consensus that protein aggregation occurs via a mechanism that involves nucleation, therefore we included a nucleus formation step.
[b]There does not seem to be a consensus as to whether the protofilament results from an assembly of oligomers or from the addition of monomers, so we chose to represent protofilament formation in a generic way.

The starting reactant in protein aggregation is the monomeric form of the protein. However, in the cases of *unwanted aggregation*, there is an issue as to whether the

monomer begins in its native form, or in a denatured, more "active" form. Indeed, misfolded proteins are believed to be part of the first steps in the mechanism of unwanted protein aggregation [48]. For example, the natively unfolded protein α-synuclein [49] is found to have an increased ability both to misfold and aggregate under oxidative conditions, such as pathological conditions that result in the overproduction of oxidants [50,51]. Also, A. L. Fink and co-workers have shown evidence for the formation of partially folded intermediates in the aggregation of α-synuclein [52,53,54]. It is therefore hypothesized that in some cases the protein monomer undergoes a conformational change, leading to higher β-sheet character or to the exposure of "sticky" hydrophobic patches [52], that then gives the monomer a higher propensity to aggregate, causing it to become "active" [14].

The aggregation of the so-called *amyloidogenic proteins* eventually leads to the formation of insoluble fibers, called amyloid fibrils, as the final product. (In some instances, amorphous aggregates, a hypothesized *off-pathway product* [14,55], have also been observed in the aggregation process.) The morphologies of amyloid fibrils obtained from different proteins are remarkably similar [47]. They typically consist of several concentric protofilaments [56], which in turn are formed by (i) one or more β-pleated sheets [47,57] with each β-sheet consisting of (ii) polypeptide β-strands usually arranged in an anti-parallel configuration [58]. M. Sunde and co-workers reported Synchrotron X-ray diffraction studies on amyloid fibrils obtained from six different proteins [47], and found a "common core structure" among these proteins at least at the level of the protofilament, as represented in Figure 1. The protofilament is found to consist of a helical array of one or more β-sheets (four β-sheets are shown in Figure 1) twisted

around the protofilament main axis, with their constituent β-strands perpendicular to that axis [47].

Figure 1. Model of the structure of the common core protofilament in a generic amyloid fibril. This structure was determined using Synchrotron X-ray diffraction on amyloid fibrils of six different proteins. Reprinted from ref. [47]. Copyright (1997) with permission from Elsevier.

An example of perhaps one of the best-characterized protein fibrils comes from the classic sickle cell hemoglobin system. The product that results from the aggregation of sickle cell hemoglobin under solution conditions is a 21 nm diameter fibril consisting of 14 helical strands arranged in 7 twisted double strands (Figure 2) [59,60,61]. A 2.05 Å resolution structure of the helically twisted double strand of sickle cell hemoglobin, believed to be a building block, was solved in 1997 [62].

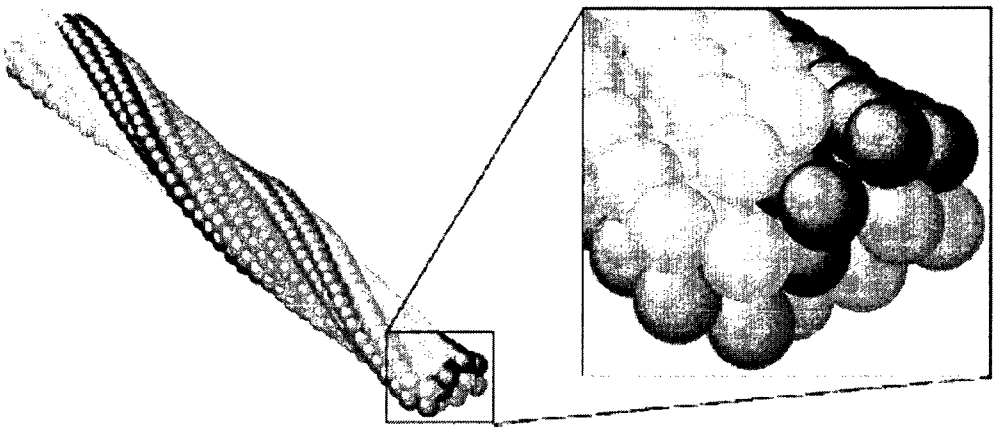


Figure 2. A schematic representation of the polymerized product of sickle cell hemoglobin. The fibril contains a total of 14 helical strands arranged in twisted pairs. Reprinted from ref. [63]. Copyright (2006) with permission from Elsevier.

Other proteins such as actin, collagen, and portions of prion proteins have also been well-characterized by X-ray diffraction [64,65,66,67,68,69,70]. However, most protein fibrils are hard to obtain in a crystalline form. Nevertheless, fibrils of amyloid β and α-synuclein among others have been extensively characterized by microscopy, solid-state NMR, and modeling studies as detailed in two recent reviews on the structures of amyloid fibrils [71,72]. These extensive characterization studies have lead to a fairly comprehensive understanding of the general structure of protein fibrils that result from protein aggregation, as well as to a more detailed understanding of the structures of specific fibrils originating from proteins such as actin, collagen, prions, amyloid β, and α-synuclein.

While there is no consensus yet as to whether all or some of the intermediate species of protein aggregation are "on" or "off" the amyloid fibril formation pathway [16,55,73,74], there is some evidence in the literature for the existence of toxic (*vide infra*) intermediate species. Such intermediates can be loosely classified into two

categories: (i) soluble oligomers, and (ii) insoluble oligomers or protofibrils. Protofibrils of amyloid-β and α-synuclein have been observed by Lansbury and co-workers using both electron microscopy (EM) and atomic force microscopy (AFM) [75,76,77,78]. Fink and co-workers have demonstrated the presence of (soluble) oligomers of α-synuclein using fluorescence studies in conjunction with various other physical methods [52,54,79], and of (insoluble) oligomers of light-chain immunoglobulin using AFM [80,81]. The morphological studies of insoluble oligomers or protofibrils have revealed that these intermediate species contain a ring or annular type structure. Recently, it was suggested that soluble *dimers* of amyloid β are responsible for the toxicity observed in Alzheimer's disease [82].

From their structural characterization, protofibrils of α-synuclein have been proposed to be toxic to neuronal dopamine cells by creating pores within the cell membrane (thus disturbing Ca^{2+} flow and balance [3]), or pores within the membrane of dopaminergic vesicles [83]. Several research groups have also looked at the toxicity of soluble oligomers of amyloid-β and proposed that toxicity could occur by way of (i) reacting with O_2 in the neuronal cell membrane to start a chain of radical reactions leading to the formation of lipid peroxidation products [84]; (ii) generating H_2O_2 within the neuronal cell [85]; (iii) inducing actin/cofilin rods in the neuronal cell [86]; or (iv) by impairing synapse structure and function [82]. In fact in a recent review, Dobson pointed out that smaller ordered aggregates (as compared to disordered or larger ordered aggregates) have a higher proportion on their surface of *hydrophobic* residues normally buried within the interior of the folded protein [87]. These small ordered aggregates are

more likely to interact with cellular components such as membranes and receptors and possibly cause disruptive interactions within the cell [87].

However, still needed is more information on the intermediate species and their putative role(s) in toxicity [20]. A possible way to gain information on the intermediate species was recently described in a review covering the technique of high hydrostatic pressure (HHR) for monitoring the aggregation of proteins [24]. HHR is postulated to stabilize and allow for the intermediate species to be isolated [24]. Also, time-resolved methods have been employed to elucidate structural information on the intermediate species of protein aggregation [34,88,89,90] and, therefore, insights into the mechanism of protein aggregation (We thank a referee for pointing this out and directing us to the key references).

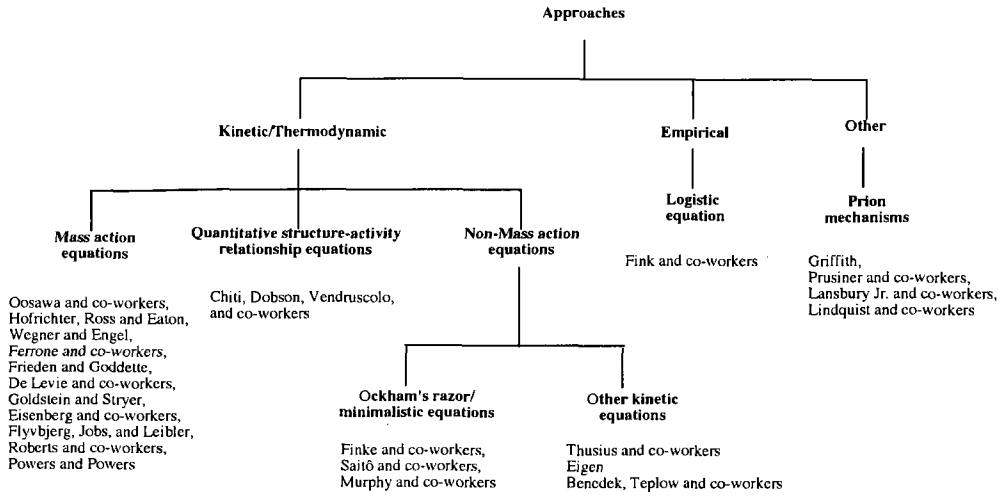
While there seems to be physical evidence for the starting proteins and final products of protein aggregation, further information is still needed on the (hypothesized) toxic intermediate species [83,84,85,86]. With the future development of new techniques and continued interest in this area, more information about intermediate species should become available and allow for further elucidation of the mechanisms of protein aggregation and toxicity [13,20].

Thermochemistry of Protein Aggregation. The aggregation of many biologically important proteins including, but not limited to, tobacco mosaic virus, tubulin, sickle cell hemoglobin, collagen, actin, myosin, flagellin, glutamate dehydrogenase, and α-chymotrypsin have been shown to exhibit a positive enthalpy *and entropy* [26,91]. We refer the readers to the scholarly work contained in references 26

and 91 for the quantitative $\Delta H°$ and $\Delta S°$ values. The positive enthalpy (i.e., endothermic nature) has been verified by calorimetry [26]. The high positive entropy is believed to result from the release of water molecules upon aggregation [26,91,92]. Therefore, protein aggregation is generally an entropy-driven process.

Approaches to Determine the Kinetics and Mechanism of Protein Aggregation. As Scheme 2 illustrates, many approaches exist in the literature for determining the kinetics and mechanism of protein aggregation. Kinetic, thermodynamic, empirical, or other approaches can provide useful information depending upon what one is trying to obtain from the analysis. Herein, we are most interested in being able to curve-fit kinetic data and extract useful information from that data—kinetics being a required part of reliable mechanistic studies. Therefore, the majority of what follows in this review is dedicated to discussing and analyzing the kinetic approaches that have been used in the literature for determining the mechanism of aggregation.

Scheme 2. Literature approaches to determining the kinetics and mechanism for protein aggregation.



Historical Review of the Kinetic and Thermodynamic Approaches to Protein Aggregation Mechanisms and Rate Constants. Kinetic and thermodynamic approaches toward determining the mechanism of protein aggregation began approximately 50 years ago with the work of Oosawa and co-workers [12]. Interest and research has continued to give what we have organized into *five classes of kinetic mechanisms*, as displayed in Table 2 from 55 primary references investigated (Table S1 of the Supporting Information). We have distilled down the available literature into what we believe to be the earliest original contributions to each of the different mechanisms proposed. Table 2 that follows is a condensed version of a fuller literature table, Table S1 of the Supporting Information, available to the interested reader (a different version of Table 2 was provided in a previous publication [93], but Table 2 displayed herein is more complete). Also provided in Table 2 are the equations, taken directly from the papers under discussion, that one in principle might use to fit experimental data. To avoid confusion,

the mathematical symbols and nomenclature original to each paper have been retained in Table 2 (i.e., no attempt was made at this time to express the equations in Table 2 in a common nomenclature, although doing so would simplify the area and is arguably a useful, future goal). Following Table 2 is a discussion of each of the five classes of kinetic mechanisms citing the work of those who have contributed to the understanding and practical applications of the mechanisms discussed. Where possible, we have also provided sample data fits to illustrate the practical applications of the mechanisms and corresponding equations obtained. Finally, we have applied a simplification to the subsequent monomer addition and prion mechanisms that will allow for simplified rate equations to be obtained for future applications.

Table 2. Historical summary of the kinetic, thermodynamic, and other models and associated equations for protein aggregation proposed in the literature.

Entry	Mechanism	Proposed by	Year	Protein System of Study	Equation Given
1	**Subsequent Monomer Addition I. Early Contributions**	Oosawa et al. [95]	1962	Actin	$\frac{dn_h}{dt} = [k_+\lambda(t) - k_-]\int[k_+'\lambda_3(t) - k_-'\lambda_{3h}(t)]dt + 3[k_+'\lambda_3(t) - k_-'\lambda_{3h}(t)] + k_-\lambda_{3h}(t)$ n_h = total concentration of monomers participating in helical polymers k_+ = forward rate constant for the attachment of monomers k_- = reverse rate constant for the detachment of monomers $k_+'\lambda_3$ = forward tranformation rate of ordinary trimers to nuclei of helical polymers $k_-'\lambda_{3h}$ = reverse transformation rate of nuclei of helical polymers to ordinary trimers Using assumptions and approximations, the above equation simplifies to: $\ln\frac{[1+(1-\lambda^p/\lambda_0{}^p)^{1/2}]}{[1-(1-\lambda^p/\lambda_0{}^p)^{1/2}]} = p(2/pk_+c)^{1/2}\lambda_0{}^{p/2}t$ p = number of monomers per turn in the helix λ = concentration of monomers λ_0 = initial concentration of monomers c = constant

2	Hofrichter et al. [97]	1974	Sickle Cell Hemoglobin	$\frac{dM_n}{dt} = \frac{[M]^{n-1} k_f^{n-2} (\sigma K - 1)(1 - K)}{k_b^{n-3} (\sigma - 1)}$ M_n = monomer of critical nucleus size M = monomer k_f = forward rate constant k_b = backward rate constant $K = \frac{k_f [M]}{k_b}$ $K < 1$ before nucleation $\sigma K > 1$ after nucleation
3	Wegner and Engel [101]	1975	Actin	$\frac{dc_2}{dt} = k_N c_1^2 - k_N' c_2 - k c_1 c_2 + k' c_3$ $\frac{dc_i}{dt} = k c_1 c_{i-1} - k' c_i - k c_1 c_i + k' c_{i+1}$ k_N = rate constant for the formation of the dimer k_N' = rate constant for the destruction of the dimer k = rate constant for the binding of protomers k' = rate constant for the dissociation of protomers c_i = concentration of i - mer with $\infty > i > 3$ Using assumptions and approximations, the above equation simplifies to: $\frac{dc_p}{dt} \approx \frac{k_N (k c_1 - k') c_1^2}{k'_N + k c_1 - k'}$ $\frac{dc_p^*}{dt} \approx (k c_1 - k') c_p$ $c_p = \Sigma c_i$ = concentration of polymers $c_p^* = \Sigma i c_i$ = concentration of monomers incorporated into polymers

4	Ferrone et al. [100]	1980	Sickle-Cell Hemoglobin	Homogeneous Nucleation : $\frac{dc_p}{dt} = K_N k_+ (\gamma c)^n + K_M \phi k_+ (c_0 - c)(\gamma c)^m$ " Heterogeneous Nucleation :" $-\frac{dc}{dt} = nK_N k_+ (\gamma c)^n + mK_M \phi k_+ (c_0 - c)(\gamma c)^m + (k_+ \gamma c - k_-)c_p$ K_N = equilibrium constant for homogeneous nucleation K_M = equilibrium constant for " heterogeneous nucleation" k_+ = rate of monomer addition k_- = rate of monomer removal γ = activity coefficient of the monomer c_0 = monomer concentration at time 0 c = monomer concentration at time t c_p = number concentration of polymers at time t ϕ = scaling factor for the number of effective nucleation sites n = critical nucleus size

5	De Levie et al. [105]	1983	"Living" polymers	Nucleation + Irreversible growth: $J = \frac{dc_n}{dt} + k_f C c_n \approx k_f C c_n$ $\frac{-dM}{dt} = (k_f C)^2 c_n t$ Using approximations at short reaction times: $-\Delta M = \frac{1}{2}(k_f C)^2 c_n t^2$ Nucleation + Reversible growth: $J = \frac{dc_n}{dt} + (k_f C - k_r) c_n \approx (k_f C - k_r) c_n$ $\frac{-dM}{dt} \approx (k_f C - k_r)^2 c_n t$ Using approximations at short reaction times: $-\Delta M = \frac{1}{2}(k_f C - k_r)^2 c_n t^2$ J = flux $-\Delta M$ = number of monomers consumed in polymer formation $\frac{-dM}{dt}$ = monomer uptake k_f = forward rate constant k_r = reverse rate constant C = monomer concentration c_n = critical nucleus concentration

6		Goldstein and Stryer [108]	1986	Not Specified	$\frac{dA_n}{dt} = k_+ A_1 (A_{n-1} - A_n) + k_- (A_{n+1} - A_n)$ when n < s $\frac{dA_n}{dt} = g_+ A_1 (A_{n-1} - A_n) + g_- (A_{n+1} - A_n)$ when n > s A_n = polymer of length n A_1 = monomer s = seed k_+ = forward rate constant prior to the formation of seed k_- = reverse rate constant prior to the formation of seed g_+ = forward rate constant after the formation of seed g_- = reverse rate constant after the formation of seed
7	**Reversible Association**	Eisenberg [7,114]	1970	Glutamate De-hydrogenase	$c = \Sigma\, c_i = \frac{c_1}{(1 - Kc_1)^2}$ $with\ K = K_i = \frac{c_{i+1}}{c_i c_1}$ $M_w^0 = M_1 \sqrt{1 + 4Kc}$ c = total concentration of the enzyme c_i = concentration of species P_i c_1 = concentration of monomer M_w^0 = weight - average molecular weight M_1 = molecular weight of monomer

8		Thusius [8]	1975	Glutamate De-hydrogenase	$\Delta A = \Delta A^0 e^{-\frac{t}{\tau}}$ $\frac{1}{\tau} = k_a \left[\left(\frac{1+4C_T}{K'} \right)^{\frac{1}{2}} - 1 \right] K + k_d$ A = total free association sites τ = relaxation time k_a = association rate constant k_d = dissociation rate constant C_T = total concentration $K = \frac{k_a}{k_d}$ K' = K times the molecular weight of the monomer
9	**Prion Aggregation Mechanisms**	Eigen[a] [115] Prusiner [4]	1996 1991	Prion Protein	$\frac{d[A]}{dt} = F_A - k_{-A}[A] - k_{AB}[A] - \frac{k_T[A]}{K_M + [A]}[B]$ $\frac{d[B]}{dt} = -(k_{BA} + k_{-B})[B] + k_{AB}[A] + \frac{k_T[A]}{K_M + [A]}[B]$ A = normal form of the host protein B = pathogenic form of A F_A = constant metabolic formation of A k_{-A} = metabolic decomposition of A k_{-B} = metabolic decomposition of B k_{AB} = noncatalytic conversion of A to B k_{BA} = noncatalytic conversion of B to A k_T = turnover number K_M = Michaelis - Menton constant

10	Eigen[a] [115] Lansbury [121]	1996 1993	Prion Protein	Rate of Nucleus Formation: $\frac{d[B_n]}{dt} \approx k_F'(\sigma q)^{n-2}[B]^2$ Growth Following Nucleus Formation: $\sum_{p=n} p[B_p] \approx k_F k_F'(\sigma q)^{n-2}[B]^3 \frac{t^2}{2}$ $\sigma q = \frac{k_F'[B]}{k_D'} < 1$ (prior to critical nucleus) $q = \frac{k_F[B]}{k_D} > 1$ (after critical nucleus) B = non - pathogenic monomeric protein with pre - prion conformation n = minimal nucleus size required for indefinite progression of aggregate formation i = size prior to critical nucleus formation p = size after critical nucleus formation k_F' = forward flux rate parameter prior to critical nucleus formation k_F = reverse flux rate parameter after the critical nucleus formation k_D' = reverse flux rate parameter prior to critical nucleus formation k_D = reverse flux rate parameter after the critical nucleus formation

11	**Subsequent Monomer Addition Mechanism II. Later Contributions**	Flyvbjerg et al. [126]	1996	Tubulin	$\frac{dc_1}{dt} = f_0 c^{n_0} - f_1 c^{n_1} c_1 + b_2 c_2 - d_1 c_1$ $\frac{dc_i}{dt} = f_{i-1} c^{n_{i-1}} c_{i-1} - f_i c^{n_i} c_i + b_i c_i + b_{i+1} c_{i+1} - d_i c_i$ $\frac{d\nu}{dt} = f_k c^{n_k} c_k$ $\frac{dM}{dt} = f_{k+1} c \nu$ c = monomer concentration c_i = number concentration of the ith relatively stable intermediate n_i = number of monomers added to form the (i+1)th intermediate k = number of intermediate assembly stages of the nucleus ν = number concentration of nuclei M = amount of mass polymerized f_i = forward rate constant b_i = backward rate constant d_i = disintegration rate constant

12	Benedek, Teplow et al. [127]	1996	Amyloid-β	$nL_f = \sqrt{\frac{mk_e c^*}{k_n}}$ $N_f = c_0 \sqrt{\frac{k_n}{mk_e c^*}}$ n : number of monomers (per unit length) in the fibril L_f : final length of the fibril m : number of monomers in a micelle c : concentration of protein c* : critical micelle concentration k_n : nucleation rate constant k_e : elongation rate constant N_f : number of fibrils c_0 : initial protein concentration	

13	Ferrone [130]	1999	Not Specified	$\frac{dc_p}{dt} = J * c *$ $\frac{d\Delta}{dt} = Jc_p$ $c* = K_{n*}c^{n*}$ $J* \approx J \approx k_+c$ Using approximations at short reaction times: $c_p = J * c * t$ $\Delta = \frac{1}{2}JJ * c * t^2$ Which can be rewritten as: $c_0 - c(t) = \frac{1}{2}(k_+c)^2 c * t^2$ c_p = concentration of polymers J = net elongation rate of polymers J* = rate of elongation of the nucleus Δ = concentration of polymerized monomers $\Delta(t) = c_0 - c(t)$ c = concentration of (free) monomer c_0 = initial concentration c* = concentration of nuclei k_+ = elongation rate constant K_{n*} = equilibrium constant for nucleation n* = critical nucleus size

14		Roberts et al. [137]	2007	Not Specified	$C_M = [N] + [I] + [U] + \sum_{i \geq 2} i\,[R_i] + \sum_{i=1}^{\delta} i \sum_{j=x}^{n^*-1} [A_j R_i]$ C_M : total monomer concentration "[ ]": concentration symbol N: native state monomer I: intermediate state monomer U: unfolded state monomer R_i : reversible oligomer of *i* monomers δ : number of monomers in each growth event n*: size at which condensation steps begin x: nucleus size A_j : aggregate composed of *j* monomers A_jR: reversibly associated A_j and R
15	**"Ockham's Razor"/ Minimalistic 2-Step Model**[b]	Finke et al. [29,93,107]	1997 2008	Amyloid β, α-Synuclein, Poly-glutamine, prions	$[A]_t = \frac{\frac{k_1}{k_2} + [A]_0}{1 + \frac{k_1}{k_2[A]_0} e^{(k_1 + k_2[A]_0)t}}$ or $[B]_t = [A]_0 - \frac{\frac{k_1}{k_2} + [A]_0}{1 + \frac{k_1}{k_2[A]_0} e^{(k_1 + k_2[A]_0)t}}$ A = monomeric protein B = aggregated, autocatalytic form of the protein k_1 = rate constant for nucleation k_2 = rate constant for growth

16		Saitô et al. [42]	2000	Human Calcitonin	$f = \frac{\rho\{\exp[(1+\rho)kt]-1\}}{\{1+\rho\exp[(1+\rho)kt]\}}$ f = fraction of calcitonin in the fibrillar form ρ = dimensionless value to describe the ratio of k_1 to k $k = k_2[A]_0$
17		Murphy et al. [150]	2006	Insulin	$A = \frac{k_S \cdot N_i(e^{t(k_S+k_f \cdot N_i)}-1)}{k_S \cdot e^{t(k_S+k_f \cdot N_i)}+k_f \cdot N_i}$ A = concentration of amyloidogenic species N_i = initial concentration of non-amyloidogenic species t = time k_S and k_f = rate constants
18	**Quantitative Structure-Activity Relationship Models**	Chiti, Dobson et al. [151]	2003	Amyloid proteins	$\ln(v_{mut}/v_{wt}) = A\Delta\text{Hydr.} + B(\Delta\Delta G_{coil-\alpha} - \Delta\Delta G_{\beta-coil}) + C\Delta\text{Charge}$ v_{mut} = aggregation rate of mutant v_{wt} = aggregaton rate of WT A,B,C = experimentally determined coefficients ΔHydr. = change in hydrophobicity resulting from the mutation $(\Delta\Delta G_{coil-\alpha} - \Delta\Delta G_{\beta-coil})$ = propensity to convert from α - helical to β - sheet structure ΔCharge = overall change in charge as a result of the mutation

19	Chiti, Dobson, et al. [152]	2004	Amyloid proteins	$\log(k) = \alpha_0 + \{\alpha_1 I^{\text{hydrophobicity}} + \alpha_2 I^{\text{patterns}} + \alpha_3 I^{\text{charge}}\} + \{\alpha_4 E^{\text{pH}} + \alpha_5 E^{\text{ionicstrength}} + \alpha_6 E^{\text{concentration}}$ k = aggregation rate I = intrinsic factors E = extrinsic factors hydrophobicity = normalized sum of hydrophobic contributions from each residue patterns = alternating behavior of hydrophobic and hydrophilic residues charge = absolute value of the net charge of the protein pH = pH of the solution ionic strength = ionic strength of the solution concentration = protein concentration

[a] These equations are not provided in the original reference but were later derived by Eigen. [b] This model is subsequently referred to as the Finke-Watzky (F-W) model.

The Subsequent-Monomer-Addition Mechanism I. Early Contributions. The first key paper probing the mechanism of protein aggregation dates back to 1959 by Oosawa and co-workers [12]. This paper investigated the aggregation of the native state G-actin protein to F-actin (*vide supra*). The observation was made that the aggregation process resembled a condensation reaction, as aggregation occurs only above a critical concentration [12]. The transformation of G-actin to F-actin as followed by flow birefringence, showed a "First phase characterized by a flat slope... followed by the second phase of a steep slope... expected in cooperative or autocatalytic phenomena" [94], Oosawa notes. The (positive) cooperativity was verified by observing that (i) the initial rate of transformation increased with G-actin concentration, and that (ii) the addition of F-actin produced an immediate transformation of G-actin to F-actin [94]. Connections were drawn to the early polymer literature, and kinetic analyses were performed [95,96]. The resulting rate equation for monomer uptake must be simplified several times with a number of assumptions (including irreversibility), before it can be approximately solved (see entry 1 in Table 2) [95]. In his treatment, Oosawa used a distinction between the concentration of (linear and helical) polymers ($\lambda_l+\lambda_h$) and the concentration of monomer in (linear and helical) polymers (n_l+n_h) [95]. An example of theoretical predictions (or simulations) using Oosawa's analytical expression is shown in Figure 3. Oosawa's expression, as well as modifications of that expression, have been routinely used over the years by other researchers [103].

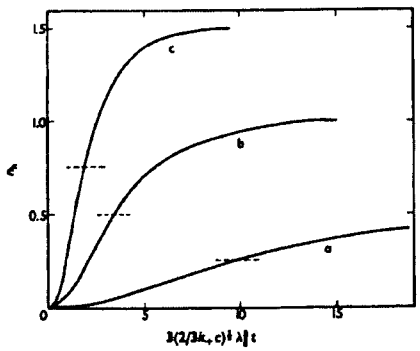
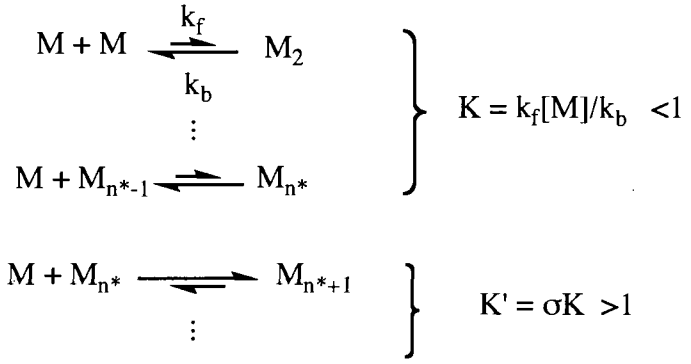


Figure 3. Oosawa's theoretical curves for actin polymerization, calculated using the equation shown in entry 1 of Table 2. Reprinted from ref. [95]. Copyright (1962) with permission from Elsevier.

In 1974, Hofrichter, Ross, and Eaton applied the subsequent monomer addition mechanism to the kinetics of sickle-cell hemoglobin gelation [97]. The authors distinguished nucleation from polymerization as shown in Scheme 3 where M is the monomer, M_n is a polymer of length n, k_f and k_b are the forward and backward rate constants respectively, and n* denotes the critical nucleus size (*vide infra*). The constant that describes the addition steps during nucleation is assumed to be less than one ($K=k_f[M]/k_b < 1$), but more than one during polymerization ($K'=\sigma K > 1$). That is, the addition steps are assumed to be thermodynamically unfavorable until a critical nucleus is formed (nucleation), but then thermodynamically favorable after nucleus formation (i.e., during polymerization) [97]. The authors introduced an expression for the rate of formation of the critical nucleus (see entry 2 in Table 2) [97,98], where the critical nucleus represents the least thermodynamically stable species in solution [99], that is, the minimum sized oligomer capable of initiating further growth [99]. Also, the authors noted, "The time course of the reaction suggests that this mechanism must involve either a nucleation or an autocatalytic process" [97].

Scheme 3. Hofrichter, Ross and Eaton's Nucleation and Polymerization Mechanism [97]. K indicates a constant with a favorable reverse reaction, while K'=σK indicates a favorable forward reaction.



In 1980, during further studies of sickle-cell hemoglobin gelation induced by photolysis, Ferrone et al. ruled out the Hofrichter et al. mechanism due to its inability to predict the shape of the early portion of the aggregation curve. In order to account for the observed "extreme autocatalysis" and strong concentration dependence of the 10^{th} time (the time required to complete one-tenth of the reaction), the authors proposed a new mechanism in which, in addition to homogeneous nucleation, "heterogeneous nucleation" also takes place on the surface of existing polymers [100]. The authors propose individual rate equations for homogeneous and heterogeneous nucleation solvable by numerical integration (see entry 4 of Table 2) [100]. The concept of "heterogeneous nucleation" was a novel and important contribution to the protein aggregation area [100]. An alternative view of "heterogeneous nucleation" might be *seeded autocatalytic growth.* (Experimentally, no one has shown whether the nuclei are forming on the surface of an existing aggregate or if growth is simply occurring on the

surface of that same aggregate. The distinction is whether a monomer vs. a critical nucleus is being added to the growing fibril.)

In 1975, Wegner and Engel applied the subsequent monomer addition model to the formation of F-actin, with the exception that the equilibrium constant for the formation of a dimer ($\sigma K = k_N/k_N'$) is different from that for each subsequent monomer addition step ($K = k/k'$) [101]. This is shown in Scheme 4, where M is the monomer, M_n is a polymer of length n, k_N and k_N' are the forward and reverse rate constants for the formation of a dimer, and k and k' are the forward and reverse rate constants for each subsequent monomer addition step [101]. The authors derived both an exact set of rate equations that must be solved by numerical integration, and an approximate (steady state) set of rate equations for the formation of F-actin (see entry 3 in Table 2). They follow an idea first introduced by Oosawa and co-workers [94] by distinguishing between the *concentration of polymers* ($c_p = \Sigma\, c_i$) and the *concentration of monomers incorporated into polymers* ($c_p^* = \Sigma\, i\, c_i$). Also, therein the authors defined the critical nucleus as being "The smallest aggregate for which [the rate of] elongation is faster than [the rate of] dissociation"; in terms of their model, the critical nucleus will be of size n=2 if $k[M]>k_N'$, and of size n=3 otherwise. They also note that while "Nucleation [is] extremely difficult... cooperativity [is] very high" [101].

Scheme 4. Wegner and Engel's Mechanism for Protein Aggregation [101].

$$M + M \underset{k'_N}{\overset{k_N}{\rightleftharpoons}} M_2 \quad \Big\} \quad \sigma K = k_N/k_{N'}$$

$$M + M_2 \underset{k'}{\overset{k}{\rightleftharpoons}} M_3$$

$$\vdots \qquad \Big\} \quad K = k/k'$$

$$M + M_j \underset{k'}{\overset{k}{\rightleftharpoons}} M_{j+1}$$

An example of Wegner and Engel's kinetic treatment is given in Figure 4, which shows actin polymerization curves (as measured by light scattering), along with the corresponding calculated kinetic curves. Other researchers have also used Wegner and Engel's polymerization kinetic treatments over the years [102,103].

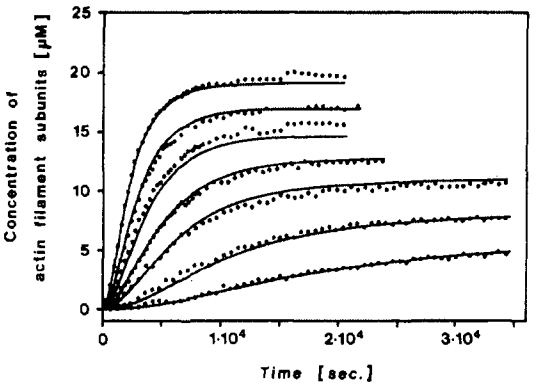


Figure 4. Wegner and co-workers's actin polymerization curves measured by light scattering (dotted lines), along with calculated kinetic curves using the equation shown in entry 3 of Table 2 (solid lines). Reprinted with permission from ref. [110]. Copyright (1982) American Chemical Society.

In 1983, Frieden and Goddette used the subsequent monomer addition model again for protein polymerization, with the exception that each monomer addition step had its own associated equilibrium constant. They also added a step that would occur prior to

all the other steps, that of an *activation* of the monomer, as shown in Scheme 5 where M' is the native monomer, M is the activated monomer, M_n is the polymer of length n, and K_n is the equilibrium constant [103]. The authors proposed that, in the case of actin, the activation step represented a ligand binding to a metal, followed by a conformational change that would accelerate the polymerization process [103]. The authors were able to build upon the work of Wegner and Savko in 1982 [110] to simplify their system of equations, and used computer simulation tools in their kinetic treatment. Based on their mechanistic analysis they note, "There is no simple measure (i.e., half-time, lag time) which would characterize the rate of polymerization... Thus, it is necessary to fit the full time course of polymerization" [103]. Our own experimental studies confirm this important directive [104,unpublished results].

Scheme 5. Frieden and Goddette's Mechanism with the Added Monomer Activation Step [103].

$$M' \overset{K'}{\rightleftharpoons} M$$

$$M + M \overset{K_1}{\rightleftharpoons} M_2$$

$$M + M_2 \overset{K_2}{\rightleftharpoons} M_3$$

$$\vdots$$

$$M + M_j \overset{K_j}{\rightleftharpoons} M_{j+1}$$

$$\vdots$$

Figure 5 shows polymerization curves of fluorescently labeled actin, along with the computer simulation curves obtained by Frieden and Goddette using an extension of Wegner and Engel's equation (entry 3 of Table 2).

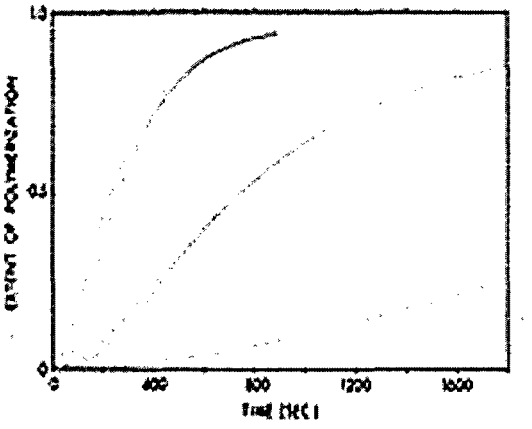


Figure 5. Fluorescently labeled actin polymerization curves (dotted lines), along with computer simulation curves (solid lines) from Frieden and Goddette. Figure 7b reprinted with permission from ref. [103]. Copyright (1983) American Chemical Society.

Also in 1983, De Levie et al. applied the subsequent monomer addition mechanism in a series of papers on the nucleation and linear growth of polymers [105,106], in which they used the assumptions that: (i) the active surface area (i.e., the end(s) of the polymer) is independent of particle size; (ii) there is an equilibrium constant for nucleation, and another for growth; (iii) the concentration of nuclei increases with time during the nucleation phase, but remains mostly constant during the growth phase; and (iv) the concentration of monomer is mainly constant at short times. The authors used Laplace transforms and a range of confluent hypergeometric functions to solve the sets of differential equations, both in the case of reversible and irreversible growth. Then using steady-state approximations, they obtained analytical expressions for a *flux in concentration* ($J = d/dt \Sigma c_i$) and for the *monomer uptake* ($-dM/dt = d/dt \Sigma i c_i$) (see entry

5 in Table 2). However, fits to the proposed equations were not provided [106]. The authors concluded that the sigmoidal ("S") shape of the aggregation curve could be obtained from the growth alone, but that nucleation is necessary to keep the concentration of nuclei at a steady-state value [106]. However, later work [107] shows that *both* nucleation and growth are crucial to obtaining sigmoidal-shaped curves.

In 1986, Goldstein and Stryer added a modification to the subsequent monomer addition mechanism in which, instead of defining the critical nucleus (whose size will depend on the relative rates of elongation and dissociation, and may ultimately depend on monomer concentration, see Hofrichter et al. [97] and Wegner and Engel [101]), they define instead a "seed" s, as "...the length s, where the kinetic constants change" [108]. Thus the equilibrium constant ($K_a = k_+/k_-$) is the same for each step prior to the formation of the seed and after the formation of the seed, the equilibrium constant ($K_b = g_+/g_-$) is the same for every subsequent step (see entry 6 of Table 2) [108]. Scheme 6 shows their proposed mechanism where M is the monomer, M_n is a polymer of length n, k_+ and k_- are the forward and reverse rate constants prior to the formation of the seed, and g_+ and g_- are the forward and reverse rate constants after the formation of the seed [108]. The authors aptly note that "The basic equations governing protein polymerization form an infinite interrelated set of differential equations that cannot be solved exactly, and simplifying assumptions must be used" [108]. Also provided in the original reference [108] are simulations of the equations shown in entry 6 of Table 2.

Scheme 6. Goldstein and Stryer's Mechanism for Protein Aggregation [108].

$$n < \text{seed:} \quad M + M_n \underset{k_-}{\overset{k_+}{\rightleftharpoons}} M_{n+1} \quad \Big\} \; K_a = k_+/k_-$$

$$m \geq \text{seed:} \quad M + M_m \underset{g_-}{\overset{g_+}{\rightleftharpoons}} M_{m+1} \quad \Big\} \; K_b = g_+/g_-$$

Authors like Oosawa and co-workers in 1975 [109] and Wegner et al. in 1982 [110], added additional steps of polymer fragmentation (Oosawa and Wegner), displayed with rate constant k_f, and of polymer association (Oosawa), shown with rate constant k_a, to their models of the subsequent monomer addition mechanism, as shown below in Scheme 7.

Scheme 7. Subsequent Monomer Addition Mechanism which Includes Polymer Fragmentation and Association.

$$M + M \overset{K_1}{\rightleftharpoons} M_2$$

$$\vdots$$

$$M + M_j \overset{K_j}{\rightleftharpoons} M_{j+1}$$

$$\vdots$$

$$M_j + M_k \overset{k_a}{\dashrightarrow} M_{j+k}$$

$$M_{m+n} \overset{k_f}{\dashrightarrow} M_m + M_n$$

The Reversible Association Mechanism. Another class of mechanisms involves studies looking at the aggregation of glutamate dehydrogenase, an allosteric enzyme linking the Krebs cycle and amino acid synthesis. Nucleation is usually not accounted for kinetically in these reversible association mechanistic postulates.

Building upon previous studies by Ts'o et al. in 1963 [111] and by Van Holde and Rossetti in 1967 [112] for purine association, and using Adams, Jr. and Lewis's 1968 study of β-lactoglobulin self-association [113], in 1970 Eisenberg suggested a reversible association mechanism for the aggregation of glutamate dehydrogenase. This mechanism is shown in Scheme 8 where M is the monomeric species, M_n is the polymerized species of length n, and K is the equilibrium constant that is assumed to be the same for each association step [114].

Scheme 8. Eisenberg's Reversible Association Mechanism [114].

$$M + M_n \overset{K}{\rightleftharpoons} M_{n+1}$$

The use of approximations at low enzyme concentration [113] lead to an estimate of average molecular weights found to be in good agreement with sedimentation experiments (see entry 7 in Table 2) [7,114]. An example of the concentration dependence of these molecular weights, along with calculated curves, is shown in Figure 6.

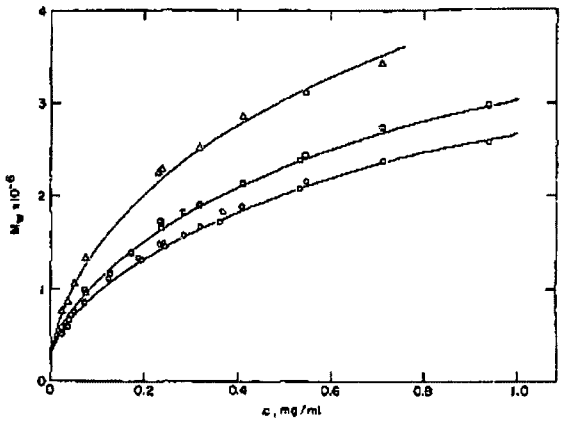


Figure 6. Average molecular weights of glutamate dehydrogenase measured by light scattering (points), along with calculated concentration dependence curves (line) using the equation shown in entry 7 of Table 2 from Eisenberg. Reprinted with permission from ref. [7]. Copyright (1971) American Chemical Society.

Thusius in 1975 and 1976 [8,9,10] challenged the reversible association mechanism put forth by Eisenberg (Scheme 8) [7] for the aggregation of glutamate dehydrogenase. Thusius claimed that Eisenberg's reversible association mechanism *alone* is not able to account for all the aggregation data in varying concentration ranges [8]. Thusius and co-workers proposed a variation of the reversible association mechanism that they termed the random association mechanism, Scheme 9, in which two (*monomeric* or *polymeric*) units of any size can associate together to form a larger polymer [8].

Scheme 9. Thusius' Random Association Mechanism [8].

$$M_i + M_j \underset{k_d}{\overset{k_a}{\rightleftharpoons}} M_{i+j}$$

$$i, j = 1, 2...\infty$$

Since Thusius' mechanism of reversible self-association occurs with "condensation occurring between all species without discrimination" [8], the authors simplified the mechanism by approximating it as a "two-state" mechanism, Scheme 10 [8]. In the "two-state" mechanism, A represents "free association sites", and B represents "bonds between units". The authors note that, "Although [their mechanism] involves more interactions than [the mechanism of sequential monomer addition only], it is the simplest kinetic model... in the sense that all elementary reactions are formally accounted for in a single transformation" [8].

Scheme 10. A Simplified Version of the Random Association Mechanism called the "Two-State" Mechanism [8].

$$2\,A \underset{k_d}{\overset{k_a}{\rightleftharpoons}} B$$

From the simplified "two-state" mechanism, the authors were able to write (and solve) a rate equation for the concentration of "free association sites" A (see entry 8 of Table 2) which they used to test their random association mechanism using light scattering aggregation data [8]. An example of the dependence of relaxation times on concentration, along with a calculated curve, is shown in Figure 7. The observable data are well fit by the two-state mechanism.

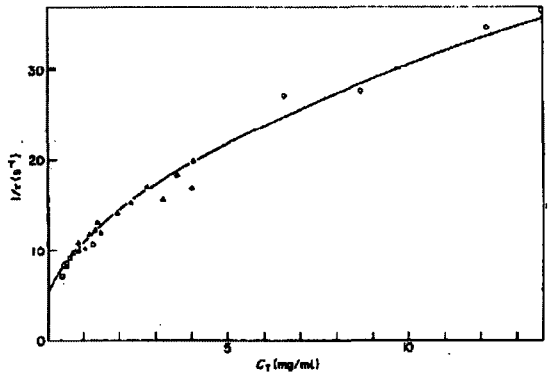
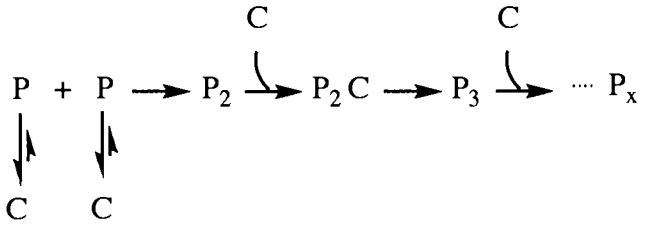


Figure 7. Thusius's glutamate dehydrogenase reciprocal relaxation times (points), and calculated concentration dependence curve (line) using the equation shown in entry 8 of Table 2. Reprinted from ref. [9]. Copyright (1975) with permission from Elsevier.

The Mechanisms of Prion Aggregation. The exact molecular mechanism for the transmission and propagation of prion diseases is generally believed to be intimately linked to the mechanism of aggregation of the prion protein. Many researchers in the area of prion aggregation have, therefore, focused on the molecular details of prion aggregation. Four different mechanisms have been proposed in the literature that provide alternative descriptions of how transmission and propagation of the prion disease may occur through a prion agent. We have chosen to express these mechanisms as chemical reactions rather than pictures or words (as given in some of the original references [4,28]), and have followed Eigen's [115] convention in the choice of arrows used to describe the mechanisms of Prusiner and Lansbury. In Schemes 11 to 15 below, C represents the cellular form of the prion protein (of predominantly α-helix secondary structure), and P represents the prion form of the protein (which is aggregation prone with a secondary structure of high β-sheet character). The non-prion (i.e., cellular) and prion forms of the prion protein, C and P respectively, are believed to be conformational isomers [4].

In 1967, a chemist with exceptional mathematical skills named Griffith proposed three possible theoretical mechanisms that could explain the self-replication of the prion disease agent scrapie [116]. These mechanisms made use of the new hypothesis that the scrapie agent could be a protein, and one is shown in Scheme 11. In Griffith's words, C would be the "stable" form of the protein, and P the "reactive" form. Griffith also proposed that *P be a different conformation of C.* In this mechanism (sometimes referred to as Template Assembly or TA [28]), the conformational change from C to P is thermodynamically unfavorable. Yet if the aggregation of P is thermodynamically favorable enough, it can act as a driving force to promote the overall conversion of two C monomers into a P_2 dimer. Further conversion of C monomers into an aggregated P form is then assisted by the presence of dimers and larger oligomers of P, which act as *templates* [116]. In this mechanism, the infectious species is an oligomer P_x (with $x \geq 2$), and propagation occurs through templated assembly and conversion of C monomers.

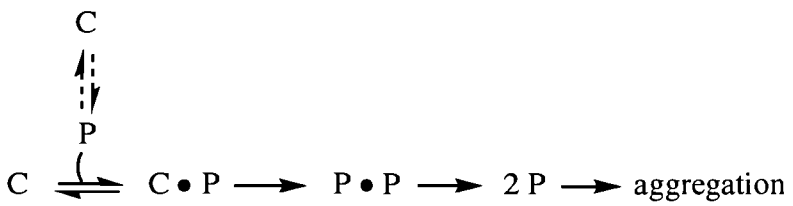
Scheme 11. Griffith's Mechanism of Scrapie Aggregation [116] Expressed in More Commonly Denoted P and C Terms.



In 1982, Prusiner presented enough experimental data to support the then controversial idea that the infectious scrapie agent is made of protein, and coined it prion

(for *pr*oteinaceous *i*nfectious *-on*) [117]. In 1991, he proposed a mechanism of prion transmission and propagation shown in Scheme 12 [4,118] (this mechanism is sometimes referred to as Monomer-Directed Conversion or MDC [28]). In this mechanism, the infectious species is a P monomer that is able to catalyze the otherwise slow conversion of C monomer into P monomer (through the formation of a C•P dimer); propagation thus occurs (we and others [93,115] note) by *autocatalysis* as P is both a reactant and a product of the reaction. Prusiner later acknowledged subsequent aggregation of the infectious form of the protein [119].

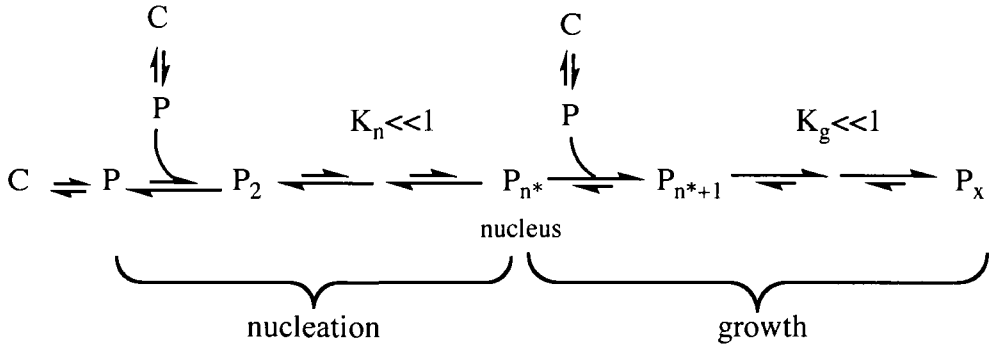
Scheme 12. Prusiner's Mechanism for Prion Aggregation [4] Expressed via Chemical Equations.



In 1996, Eigen published a kinetic analysis of Prusiner's mechanism (see entry 9 in Table 2) [115]. He also provided an alternative version of that mechanism by pointing out that the catalytic (infectious) agent could be either a P monomer (his "linear Prusiner mechanism") or a P_x oligomer (his "cooperative Prusiner mechanism") [115]. Researchers like M. Laurent have provided support for Eigen's "cooperative Prusiner mechanism" by pointing out that unrealistic rate constants would be needed for P to be the catalytic agent [120].

Following the work of Prusiner, in 1993 Lansbury proposed a mechanism of prion transmission and propagation (referred to as Nucleation-Dependent Polymerization, NP), which includes features from the mechanism of subsequent monomer addition discussed above [121,122,123,124]. This mechanism is shown in Scheme 13, where n* represents the critical nucleus size, and K_n and K_g are the equilibrium constants for nucleation and growth, respectively. Here C and P monomers are in equilibrium, and P monomers are slowly aggregating to form a critical nucleus (nucleation). Once the P_x oligomer is past the critical nucleus size ($x \geq n^*$), further aggregation becomes thermodynamically favorable and leads to growth. In this model, the aggregation of P through *nucleated polymerization* is the driving force that displaces the equilibrium between C and P (towards P) [121]. The infectious species is a seed P_x of size $x \geq n^*$, and propagation occurs through growth on the seed in this model.

Scheme 13. Lansbury's Mechanism for Prion Aggregation [121].

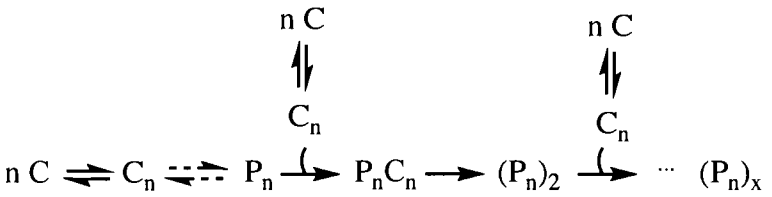


In 1996, Eigen also published a kinetic analysis of Lansbury's mechanism (see entry 10 in Table 2) [115]. There he points out that catalysis *per se* is not necessary in

this mechanism (and he defines both a *passive* and an *active* form of autocatalysis,[4] although these terms have not been extensively used in subsequent literature) [115]. Eigen further states that neither the Prusiner nor the Lansbury mechanisms can be ruled out as possible mechanisms of prion transmission and propagation [115].

While studying the aggregation of yeast prion proteins, Linquist proposed a new mechanism in 2000, which can be seen as a combination of the mechanisms of Griffiith and Lansbury, named Nucleated Conformational Conversion (NCC) and shown in Scheme 14 [28,125]. Here the formation of nuclei of C is followed by slow conversion to nuclei of P. Once nuclei of P are present, further assembly is proposed to occur rapidly. The formation of larger aggregates occurs by *nuclei of P* acting as *templates*, which combine with and convert nuclei of C [28]. In this mechanism, the infectious species is a P nucleus (P_n), and propagation occurs through templated assembly and conversion of C nuclei (C_n).

Scheme 14. Lindquist's Nucleated Conformational Conversion Mechanism [28] Expressed via Chemical Equations.



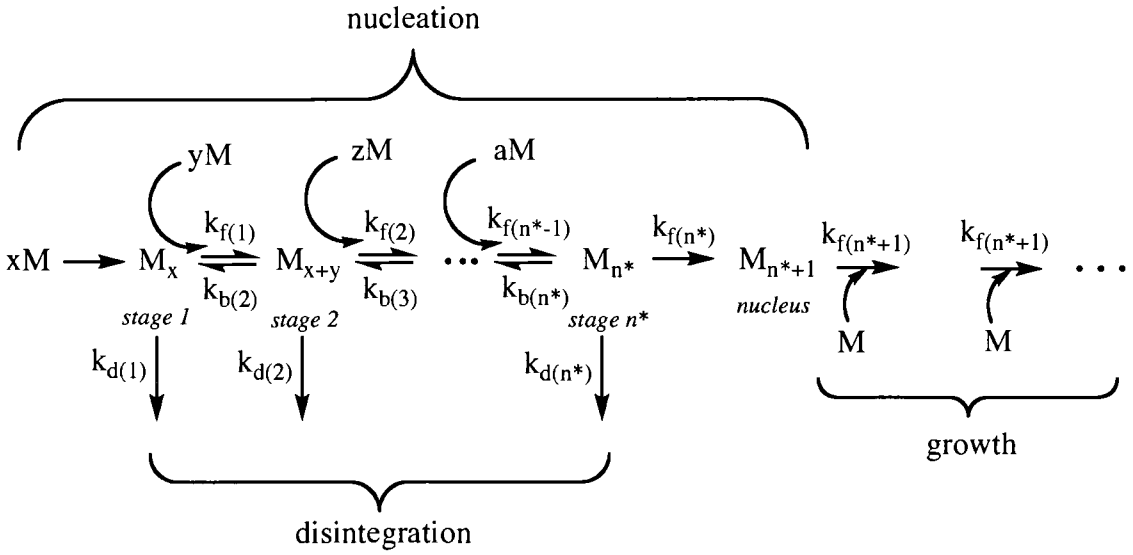
[4] According to Eigen: "The nucleus may be considered as a 'passive' autocatalyst...[S]uch a process is virtually indistinguishable from any 'active' form of catalysis, that is, a direct interaction of A-particles with some 'catalytic surface' of the B-aggregates." It is not clear whether the distinction of such putative forms of autocatalysis has any physical or chemical basis.

Attempts have been made in the prion literature to rule out some of the proposed prion mechanisms [28]. However, proper kinetic analyses are lacking [29] particularly because the mechanisms are usually displayed in words and pictures only (a form which is not amenable to rigorous kinetic and associated mathematical analyses).

The Subsequent Monomer Addition Mechanism II. Later Contributions. In 1996 Flyvbjerg, Jobs, and Leibler examined the self-assembly of microtubules from tubulin using a model described in Scheme 15 where M is the monomer, M_{n^*+1} is the nucleus of size (n*+1), and k_f, k_b, and k_d are the rate constants for the forward, backward, and disintegration steps, respectively [126].[5] The authors used the assumptions that (i) there is only one pathway for aggregation during nucleation, and (ii) during growth, every stage in the pathway is linked to the next stage by the addition of monomers only [126]. Note that while the steps are reversible up to the formation of the nucleus, they are irreversible thereafter. This model also accounts for the possible disintegration of the intermediate aggregate species during the nucleation phase. The associated rate equations for this model are shown in entry 11 of Table 2.

Scheme 15. Flyvbjerg, Jobs, and Leibler's Model for the Self-Assembly of Microtubules [126] Expressed in Terms of Chemical Equations.

[5] We thank a referee for directing us to this reference.



The authors applied a method of so-called "phenomenological scaling" to a series of kinetic aggregation curves (measured by turbidity at different tubulin concentrations), in which they essentially reduce the data to a common turbidity scale (A/A_∞, where A_∞ represents the maximum turbidity value) and a common time scale (t/t_0, where t_0 represents a "characteristic time" value determined during the scaling process), Figure 8. By applying their "phenomenological scaling" to variables present in the model's associated rate equations (see entry 11 of Table 2), such as time, monomer concentration, number concentration of monomers, and mass polymerized, the authors were able to find an approximate solution for their equation system [126]. Figure 8 also shows the fit to the proposed model, but which the authors note "was obtained by assuming scaling, a property that is only approximately satisfied by the experimental data" [126].

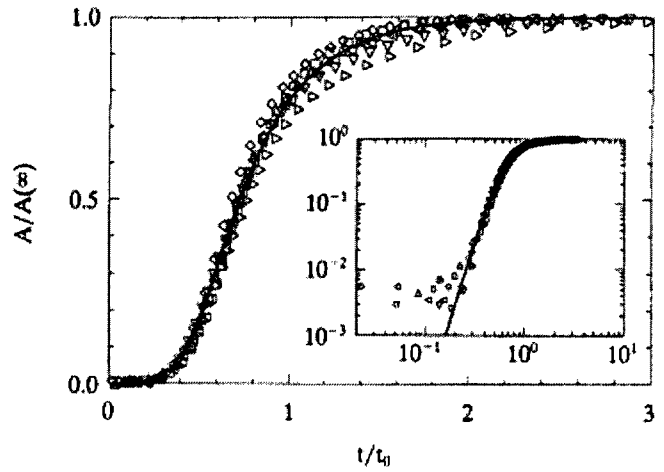
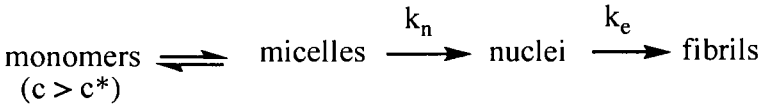


Figure 8. Flyvbjerg et al.'s microtubule assembly kinetics at various concentrations with the "phenomenological scaling" applied to the data on both the x- and y-axes. The fit is to the model described in Scheme 15. Note that the inset uses a log-log scale which tends to visually exaggerate the differences in values near zero. Ref [126]. Copyright (1996) National Academy of Sciences, USA.

Flyvbjerg et al. first used four rate constant parameters (from an empirical estimation of the number of assembly steps occurring during nucleation), but found that three of those four rate constants were equivalent, so that their model could be reduced to a two parameter model, which the authors state can be seen as a generalization of Oosawa's classical nucleation-polymerization model. The authors tested this model with turbidity vs. time data for tubulin aggregation at various concentrations, and reported average parameter values of $\Pi_{i=0}^{4} k_{f(i)} = (1.2\pm0.3)\times10^{3}$ cm^{15}/min^{5} and $k_{f(1)}=k_{f(2)}=k_{f(3)}=1.0\pm0.9$ cm^{3}/min. Interestingly, we show that the same data can also be fit with the (phenomenological) 2-step Finke-Watzky (F-W, *vide infra*) model of nucleation and autocatalytic growth (see Figure S1 of the Supporting Information). The averaged rate constants resulting from the F-W fits differ in units from the above results (with $k_1=(2\pm2)\times10^{-2}$ min^{-1} and $k_2=1.7\pm0.4$ cm/min), and thus are not directly comparable.

Also in 1996, Benedek, Teplow and co-workers looked at the fibrillation of amyloid-β, a peptide whose deposition is associated with Alzheimer's disease, under conditions where the protein is in high concentrations (0.1 mM range), but where the aggregation is slowed down in this case by acidic conditions, so that it may be followed (over time) by quasi-elastic light scattering (QLS) [127]. The authors proposed a model interpreted in Scheme 16, where c represents the protein concentration, c* the critical micelle concentration (CMC), k_n the rate constant for nucleation, and k_e the rate constant for elongation. In this model and at protein concentrations above the CMC ($c > c^*$), (i) the protein forms micelles, (ii) nucleation occurs within the micelles, and (iii) elongation takes place on the nuclei (by irreversible binding of monomers to the fibrils ends). For protein concentrations below the CMC ($c < c^*$), no nucleation occurs so that (seeded) growth may only take place on impurities [127]. Other researchers have since proposed that soluble oligomers of amyloid-β actually represent protein micelles [128].

Scheme 16. Word Interpretation of Benedek and Teplow's Mechanism for Protein Fibrillation [127].



The authors emphasized the need to deconvolute kinetic parameters for nucleation *and* growth, as illustrated by the point that an increase in nucleation (*relative* to growth) may both be seen as *promoting* fibrillation (by giving rise to a higher number of fibrils) or as *inhibiting* fibrillation (by leading to shorter fibrils), and vice versa for a decrease in

nucleation [127]. From their model the authors derived analytical expressions for quantitative aspects such as the number of fibrils and the final length of fibrils (see entry 12 in Table 2). The authors used these equations to extract rate constants from the observed concentration dependences of L_f, the final length of fibrils, and of dL/dt, the rate of change in fibril length (Figure 9).

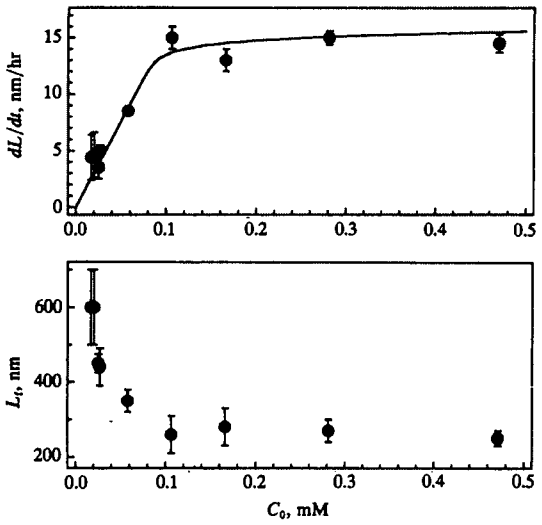


Figure 9. Benedek and Teplow's concentration dependence of the rate of change in fibril length (dL/dt) and the final fibril length (L_f) observed for amyloid-β with quasi-elastic light scattering. Ref. [127]. Copyright (1996) National Academy of Sciences, USA.

In 1997, the same team published a more detailed mathematical treatment where they relate the measured hydrodynamic radius to a fibril distribution [37]. The authors did so with the use of n^{th} order mathematical moments of fibril distribution: in their treatment, 0^{th} order corresponds to the *sum of fibrils of all sizes* (analogous to the quantity 'concentration of polymers' quantity used by Oosawa [95], *vide supra*), and 1^{st} order corresponds to the *sum of all protein in fibril form* (again, see the 'concentration of monomers in polymers' quantity used by Oosawa [95] *vide supra*) [37]. Figure 10

shows the evolution of the measured hydrodynamic radius, R_H, as a function of time, along with calculated curves. This work represents an early and detailed effort at expressing aggregate size as a function of time, a problem more recently investigated by M. M. Pallitto and R. M. Murphy [129].

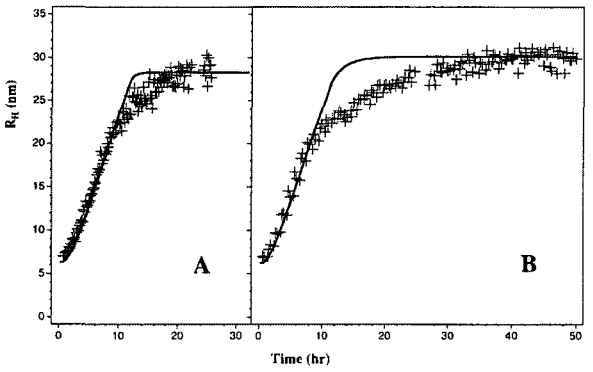


Figure 10. Amyloid-β aggregation followed by quasi-elastic light scattering: hydrodynamic radius (R_H) as a function of time, along with calculated curves from Benedek, Teplow, and co-workers. Ref. [37]. Copyright (1997) National Academy of Sciences, USA.

In 1999, Ferrone published a paper on the kinetics of protein aggregation in which he used a mechanism of subsequent monomer addition analogous to the one described in Scheme 3 above, where he defines the critical nucleus as the aggregate size after which "the association rate exceeds the dissociation rate for the first time" [130]. With assumptions similar to those made by De Levie in 1983 [106], but using a perturbation theory approach instead [131], the author derived an expression for the monomer concentration as a function of time (see entry 13 in Table 2) [130]. Ferrone's expression turns out to be mathematically identical to that obtained by De Levie in the case of irreversible growth [106] (see entry 5 in Table 2); for a detailed analysis of the equivalence between De Levie and Ferrone's expressions, see Scheme S1 of the

Supporting Information. The author also made use of the distinction employed by Oosawa [95] (*vide supra*) between the concentration of *polymers* and that of *monomers incorporated into polymers* [130].

Shown in Scheme 17 is our translation of Ferrone's equations [130] into chemical reactions, where M represents the protein monomer, M_{n^*} the nucleus of size n*, $M_{m>n^*}$ any polymer of size $m > n^*$, K_{n^*} is the nucleation equilibrium constant, and k_+ the rate constant for growth.

Scheme 17. Ferrone's Proposed Mechanism as Interpreted as Part of the Present Work.

$$n^*M \overset{K_{n^*}}{\rightleftharpoons} M_{n^*}$$

$$M + M_{n^*} \xrightarrow{k_+} M_{m>n^*}$$

$$M + M_{m>n^*} \xrightarrow{k_+} M_{m>n^*}$$

Using our interpreted scheme of Ferrone's mechanism along with the assumptions employed in his treatment [130], we have also attempted to express the corresponding rate equations in equation 1 below. The assumptions Ferrone used that we also employed in attaining Eqs. 1(a)-(c) are that: (i) all polymers can be counted by their ends *only* (so that they appear unchanged on both sides of the reaction, see the third step in Scheme 17); (ii) the concentration of nuclei is *small* (so that the monomer uptake occurs mostly through the growth of polymers, see eq. 1(c)); and (iii) initially, the concentration of monomers incorporated into polymers is *very small* compared to the initial monomer concentration (in the first 10-15% of the aggregation curve), so that there the concentration of free monomer can be assumed to be constant and equal to the initial

value (and the rate eqs. 1(a)-(c) can be integrated using a constant term $[M]=[M]_0$) [130]. The resulting equations are:

$$\begin{aligned} &(a) \quad [M_{n^*}] = K_{n^*}[M]^{n^*} \\ &(b) \quad \frac{d[M_{m>n^*}]}{dt} = k_+[M][M_{n^*}] \\ &(c) \quad -\frac{d[M]}{dt} = k_+[M][M_{m>n^*}] \end{aligned} \tag{1}$$

In 2002, Wetzel applied Ferrone's equation to the aggregation of polyglutamine [132,133], as shown in equation 2 below, where Δ is the concentration of polymerized monomer, c_0 is the initial concentration of monomer, c is the concentration of free monomer, k_+ is the second-order elongation rate constant, K_{n^*} is the nucleation equilibrium constant, and n* is the critical nucleus.

$$\Delta_{(t)} = c_0 - c_{(t)} = \frac{1}{2} k_+^2 K_{n^*} c^{(n^*+2)} t^2 \tag{2}$$

The use of Equation 2 is limited to the early portion of the kinetic curve [25], where the concentration of monomer may be taken as constant and equal to the initial value. In the early portion of the curve ($\leq$ 15% of the aggregation curve), a plot of $\{\Delta_{(t)}$ *vs.* $t^2\}$, as in Figure 11, gives a slope of $[\frac{1}{2} k_+^2 K_{n^*} c^{(n^*+2)}]$; in turn a plot of {ln(slope) *vs.* $\ln(c_0)$} gives $[(n^*+2)]$ as a slope and $[\ln(\frac{1}{2} k_+^2 K_{n^*})]$ as an intercept [134]. This formula has been used

to extract information about the critical nucleus size [132,133,134,135], but it should be noted that nucleation and growth are still convoluted in the intercept.

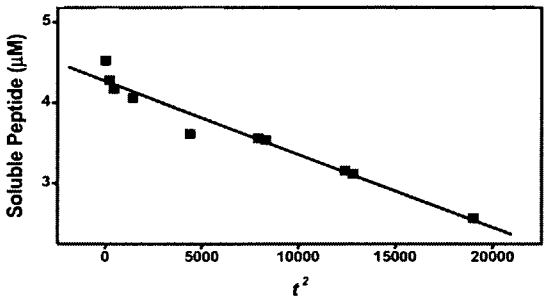
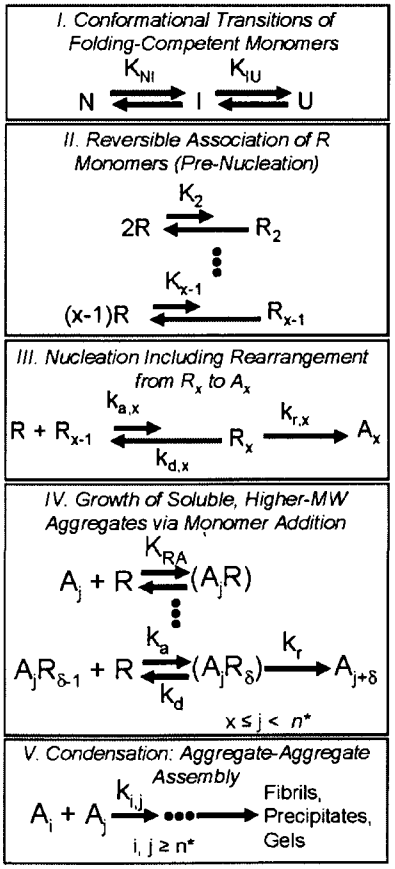


Figure 11. An example of Ferrone and Wetzel's soluble monomer concentration vs. t^2 plot for the early portion of the aggregation curve of polyglutamine fit using Eq. 2. Figure 2d of Figure 2a-d reproduced from Ref. [132]. Copyright (2002) National Academy of Sciences, USA.

In 2003, while exploring ways to predict a protein shelf-life, Roberts proposed a mathematical model for the irreversible aggregation of proteins in which the final product is amorphous aggregates [11]. In his approach, Roberts used an extended Lumry-Eyring model [136] to account for protein denaturation prior to aggregation. In 2007, Roberts adapted his model to the aggregation of proteins for which the final product of aggregation is ordered aggregates (i.e., fibrils) [137]. He proposed a mechanism reproduced in Scheme 18 (also see entry 14 in Table 2) [14,138]. In stage I (which corresponds to conformational changes involving the monomer), N represents a monomer in its native state, I in its intermediate state, and U in its unfolded state, while K_{NI} and K_{IU} are equilibrium constants for $N \rightleftharpoons I$ and $I \rightleftharpoons U$, respectively. In stage II (which is a "pre-nucleation" stage of reversible monomer association), R represents the reactive form of the monomer (N, I or U), R_i a reversible oligomer (composed of i monomers), and K_i is the equilibrium constant for $iR \rightleftharpoons R_i$. In stage III (which corresponds to nucleation

involving a rearrangement), A_x represents the aggregate nucleus, x the nucleus size, and $k_{a,x}$, $k_{d,x}$, and $k_{r,x}$ are the association, dissociation, and rearrangement rate constants, respectively, for the nucleation step. In stage IV (in which the soluble aggregates grow through monomer addition), A_j represents an aggregate composed of j monomers, A_jR the reversibly associated A_j and R, K_{RA} is the equilibrium constant, and δ is the number of monomers in each growth event. In stage V (where the condensation of aggregates leads to the formation of say, fibrils), n* represents the size at which condensation begins, and k_{ij} is the rate constant for the condensation of A_i+A_j.

Scheme 18. Roberts' Nucleated Polymerization Mechanism. Reprinted with permission from ref. [137]. Copyright (2007) American Chemical Society.



In this model, the protein aggregates are in a non-native (reactive) state, and Roberts has since proposed a version of this model for aggregation of protein in its native state (where stages I, III and V are omitted) [14]. The authors introduce and examine several kinetic regimes that are defined by the relative rates of nucleation and growth, and by the size at which the aggregates condense [137]. Simulations predict that these kinetic regimes should be differentiable experimentally by a combination of (i) apparent reaction order; (ii) dependence on the initial protein concentration; and (iii) aggregate size distribution [137]. Roberts also states that his model has the ability to encompass (most) previously proposed models, as these can be classified based on which stages they do include and where the rate limiting step occurs [14]. We agree that Roberts' model is able to account for many of the previous models for protein aggregation; however, application of this model to experimental protein aggregation data is not a simple task, but has been reported [139,140,141].

In 2006 Powers and Powers, while looking at the kinetics of subsequent monomer addition, proposed that three distinct regimes exist based on the total protein concentration: (i) low, (ii) medium and (iii) high [142]. The concentration "cut-offs" for these regimes are determined by the values of $K_{c(critical)}$, the equilibrium constant for monomer dissociation *subsequent* to the formation of a (structural [142]) nucleus and $K_{s(supercritical)}$, the equilibrium constant for monomer dissociation *prior* to the formation of that nucleus, see Figure 12.

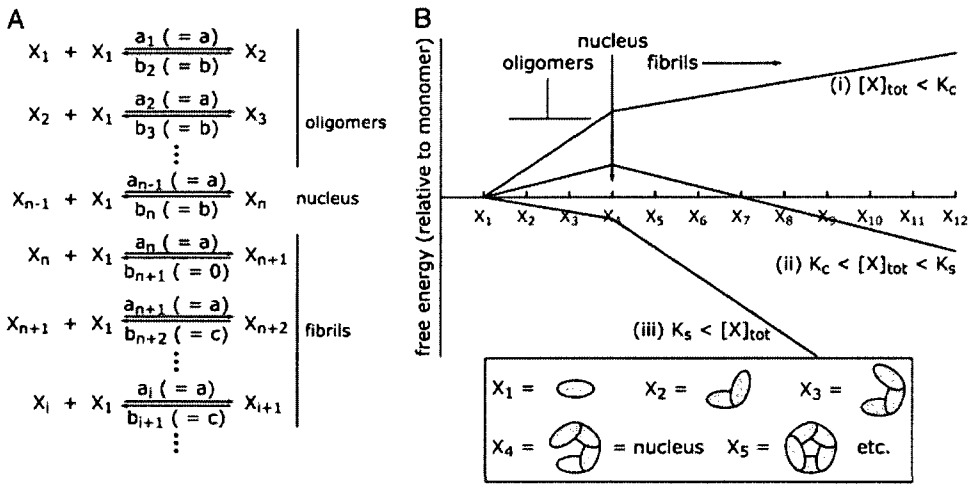


Figure 12. Three kinetic regimes proposed by Powers and Powers for the mechanism of subsequent monomer addition. Reprinted with permission from ref. [142]. Copyright (2006) Biophysical Society.

(i) In the low regime, defined for total protein concentrations lower than K_c, all the polymers (including oligomers, *structural* nucleus [142] and fibrils) are higher in energy than the monomer, and as a consequence no fibrillation occurs. (ii) In the medium regime, defined for protein concentrations larger than K_c but less than K_s, the highest energy species is the structural nucleus [142], and fibril formation occurs through a (classic) nucleated polymerization pathway. (iii) In the high regime though, defined for protein concentrations larger than K_s, the highest energy species is now the monomer, and the authors warn that (classic) models of nucleated polymerization may no longer apply [142]. The authors test their proposed mechanism with *simulations* using various concentrations. The simulations are consistent with the above three regime model, although no experimental kinetic data were used [142]. Recently, Powers and Powers have added off-pathway aggregation to their mechanism and tested this updated model again using simulations, but this time with more experimentally relevant concentrations [74].

An "Ockham's Razor"/Minimalistic 2-Step Model. This class of kinetic mechanisms for particle aggregation was first proposed in 1997 by Finke and Watzky for transition-metal nanocluster formation [107], but was recently shown to apply to a broad spectrum of aggregating proteins, including α-synuclein, amyloid-β, polyglutamine and prions, relevant in the neurodegenerative disorders of Parkinson's, Alzheimer's, Huntington's and prion diseases, respectively (entry 15 of Table 2) [29,93]. In Scheme 19 and in the case of protein aggregation, A represents a (precatalytic) form of the protein monomer while B represents any (catalytic) aggregated form of the protein past the critical nucleus size. Hence, all aggregates able to perform autocatalysis are treated as kinetically equivalent species in this minimalistic kinetic model. Importantly, the rate constants k_1 and k_2 correspond to nucleation and growth, respectively, so that this two-step model specifically and easily separates (average) nucleation from (average) growth. Also, as the species B is both a catalyst and a product in the growth step reaction, the latter step is the definition of autocatalysis (A + B → 2B being the elementary step that defines autocatalysis).[6]

---

[6] The Finke-Watzky (F-W) mechanism can for reasons of visual clarity be expressed in a more descriptive form as shown below, where A again represents the precatalytic form of the protein monomer, and now it is B_n that represents the catalytic form of the aggregated protein. The observed kinetics of the descriptive form would still obey the rate laws of the steps as written in Scheme 18, *during the initial phase of the reaction (i.e., as t→0)* where the concentration of monomer A is approximately constant (i.e., $k_{1\,(obs)} \approx k \times K^{n-1} \times [A]_0^{n-1} \approx$ constant, where K is the equilibrium constant for any (reversible) bimolecular association step prior to formation of the nucleus, k is the rate constant for the last (irreversible) bimolecular association step leading to formation of the nucleus, and $[A]_0$ is the initial monomer concentration). However, as t becomes larger, the rate laws and their solutions will deviate (e.g., in the simple case of bimolecular nucleation, when $\frac{-d[A]}{dt} = k[A]^2$ is approximated to $\frac{-d[A]}{dt} \approx k[A]_0[A]$, the solution to the first equation, $[A]_t = \frac{[A]_0}{1 + k[A]_0 t}$, is only equal to the solution of the second equation,

Scheme 19. The Finke-Watzky (F-W) Mechanism [107] that has been Applied to 41 Literature Protein Aggregation Data Sets [29,93].

$$A \xrightarrow{k_1} B$$

$$A + B \xrightarrow{k_2} 2B$$

As shown in Scheme 19, the F-W model consists of two simple *pseudoelementary* steps.[7] By definition a pseudoelementary step can be employed *kinetically* as an elementary step, and as such its rate law can be determined from the stoichiometry of the reaction. Within the F-W model, all the probably hundreds to thousands of actual steps occurring at the molecular level of the aggregation process can be combined into two pseudoelementary steps [107], as demonstrated by the fact that these two steps alone are able to account for a whole range of protein aggregation kinetic data [29,93]. The two pseudoelementary steps in the F-W mechanism represent respectively, typically slow nucleation and typically fast autocatalytic growth [107]. The rate constants, k_1 and k_2, are *average and pseudoelementary* rate constants for those nucleation and growth steps.

The F-W 2-step mechanism is an "Ockham's razor," [143] minimal kinetic model, *the first model in which quantitative rate constants for nucleation and growth can*

$[A]_t = [A]_0 \times e^{-k[A]_0 t}$, if t→0, as they both approximate through a Taylor series to $[A]_t = [A]_0(1 - k[A]_0 t)$.) We thank a referee for the specific and useful comments that lead us to clarify this footnote.

$$nA \longrightarrow B_n$$

$$A + B_n \longrightarrow B_{n+1}$$

[7] The concept of pseudoelementary steps was introduced in the 1970s by Noyes, who developed this concept using kinetic studies of complex oscillating reactions (*J. Am. Chem. Soc.* **1972**, *94*, 8649-8664.; *J. Chem. Phys.* **1974**, *60*, 1877-1884.; *Acc. Chem. Res.* **1977**, *10*, 214-221.).

be routinely and easily obtained using either of the analytic equations shown in entry 15 of Table 2 and corresponding to the model in Scheme 19 [107]. An example of a F-W fit previously obtained [93] for the aggregation of α-synuclein [55] is shown in Figure 13; it is apparent that this simple 2-step mechanism is capable of accounting for the aggregation data. Fits to the aggregation kinetic data from polyglutamine, amyloid-β, prion, and other α-synuclein aggregation have been published, with each of the 41 data sets analyzed showing good to excellent fits to the F-W 2-step model [29,93].

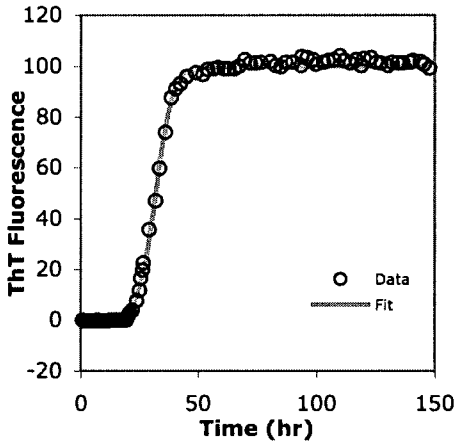


Figure 13. An example of aggregation data published by A. L. Fink [55] and fit to the F-W 2-step mechanism. In this case, aggregation of α-synuclein in the presence of macromolecular crowding was measured by fluorescence [55]. Data were digitized and fit to the F-W mechanism resulting in $k_1 = 4.0(8) \times 10^{-5}$ hr^{-1}, $k_2 = 4.0(1) \times 10^{-3}$ $\mu M^{-1} hr^{-1}$, and a coefficient of determination (R^2) of 0.999. Reprinted with permission from ref. [93]. Copyright (2008) American Chemical Society.

However, due primarily to its minimal complexity, the F-W mechanism does have limitations, including: (i) hundreds if not thousands of steps are condensed into two pseudo-elementary steps, (ii) the rate constants, k_1 and k_2, are average rate constants, (iii)

a higher kinetic order in [A] may be hidden kinetically in the nucleation step in particular, and (iv) all growing aggregates are hidden behind the descriptor 'B' which describes a catalytically active form of the "polymerized monomer", the latter concept having been originally introduced by Oosawa [95] and subsequently used by others [101,130]. Recently, we have found that the general descriptor 'B' can also hide processes such as fragmentation [29], which is a topic of future studies [unpublished results]. These limitations have been previously discussed in detail [29,93] for the interested reader.

In 2000, Saitô and co-workers published a paper analyzing the aggregation of human calcitonin with the mechanism shown in Scheme 20, where n_0 represents micelles of the same aggregation number, A_{n0} is the micelle of n_0 monomers, B_{n0} is the nucleus, and B_n and B_{n+1} are the elongated fibrils with n and n+1 molecules of the protein [42]. The concept of (protein) micelle formation [128] during the nucleation phase, invoked in this mechanism, is built upon the work of Benedek, Teplow and co-workers [127], see Scheme 16. Also published was a resultant integrated equation shown in entry 16 of Table 2 [42].

Scheme 20. Saitô and Co-workers'Mechanism for the Aggregation of Human Calcitonin [42].

$$n_0A \text{ (monomers)} \overset{K}{\rightleftharpoons} A_{n0} \text{ (micelle)}$$

$$A_{n0} \xrightarrow{k_1} B_{n0}$$

$$A + B_n \xrightarrow{k_2} B_{n+1}$$

Although not readily apparent, the integrated equation shown in entry 16 of Table 2 is mathematically identical to the integrated equation expressed in terms of $[B]_t$ for the F-W mechanism (entry 15 of Table 2); the analysis of this equivalence can be found in Scheme S2 of the Supporting Information in reference [93]. Sample fits using Saitô's fractional equation (entry 16 of Table 2) of human calcitonin aggregation data measured by amino acid labeled solid-state ^{13}C NMR and circular dichroism at two different concentrations of protein are shown in Figure 14.

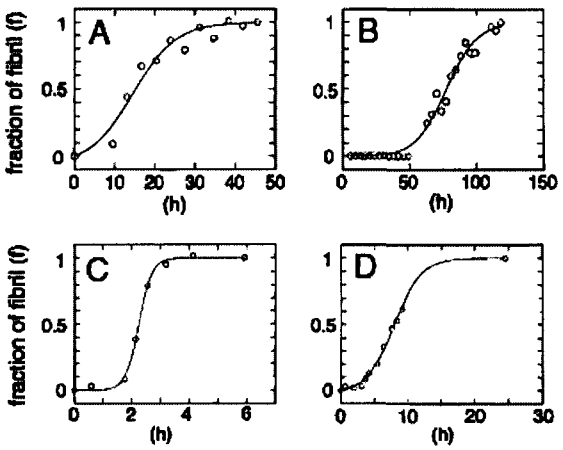


Figure 14. Saitô's fraction of fibril vs. time human calcitonin aggregation data measured using amino acid labeled solid-state ^{13}C-NMR (A and B) or circular dichroism (C and D, where C is the protein at 1.5 mg/mL and D is the protein at 0.2 mg/mL). The aggregation data was fit using Saitô's fractional equation shown in entry 16 of Table 2. Reprinted with permission from ref. [42]. Copyright (2000) Cold Spring Harbor Laboratory Press.

Although Saitô's 2000 mechanism is equivalent to the earlier F-W mechanism, Saitô was the first to use the equation shown in entry 16 of Table 2 to analyze the aggregation of proteins (Figure 14). Other research groups have subsequently used

Saitô's fractional form of the F-W equation to fit other protein aggregation kinetic data [144,145,146,147,148,149].[8]

Later, in 2006, Murphy and Gibson published a "two-step kinetic model of aggregation" for the protein insulin, along with its corresponding integrated rate equation (see entry 17 of Table 2) [150]. The integrated rate equation given by Murphy and co-workers in 2006 is equivalent to the integrated rate equation in the 1997 F-W and 2000 Saitô models.

Quantitative Structure-Activity Relationship Models. A phenomenological approach to protein aggregation that uses the protein physicochemical properties, was originally developed by Chiti et al., in which they looked at the effect of amino acid mutations on protein aggregation rates. Their approach is to correlate the observed aggregation rate changes to calculated changes in protein physicochemical properties such as hydrophobicity, charge, and propensity to convert from a α-helical to a β-sheet secondary structure [151]. The coefficients for the equation shown in entry 18 of Table 2 are determined by standard regression to fit the observed $\ln(\upsilon_{mut}/\upsilon_{wt})$ values that were obtained using experimental data from the literature (υ_{mut} and υ_{wt} represent the aggregation rates of the mutant and wild-type protein, respectively). It was shown that the hydrophobicity and hydrophobic/hydrophilic patterns (or the propensity to convert

[8] The mechanism displayed in Scheme 19 and corresponding equation in entry 16 of Table 2 are original works of Saitô and co-workers in 2000. Although Saitô's work has been properly referenced by some investigators [144,145], it has also been mis-cited by other investigators [146]. In one unfortunate case, other authors stated that they used a modification when in fact they used the same mathematics originally given by Saitô [146]. This illustrates a point in the problem areas section: that knowing the prior protein aggregation literature is necessary and key to avoid repetition and mis-citation.

from α-helix to β-sheet) contributed positively to the aggregation rate (A,B > 0), while charge contributed negatively (C < 0), see entry 18 of Table 2 [151]. As shown in Figure 15, the authors tested their model using 27 single-mutation data sets and compared their calculated $\ln(\upsilon_{mut}/\upsilon_{wt})$ values to the observed $\ln(\upsilon_{mut}/\upsilon_{wt})$ values; they found that their model predicted the observed values with a modest correlation coefficient (r) of 0.85.

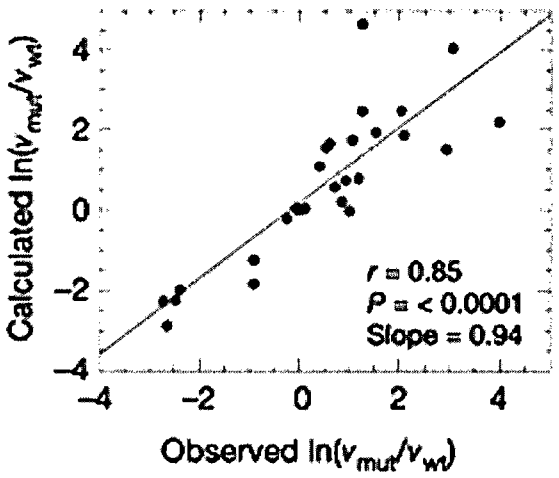


Figure 15. Dobson and co-workers' calculated and observed changes in the aggregation rate upon a single mutation of 27 proteins data sets including the proteins amylin, amyloid β, α-synuclein, and tau. Figure 2a of Figure 2a-b reprinted with permission from ref. [151]. Copyright (2003) Nature Publishing Group.

Later, Dobson and co-workers along with Vendruscolo expressed an absolute protein aggregation rate (k) in terms of a combination of physicochemical factors intrinsic and extrinsic to the protein, as shown in entry 19 of Table 2 [152]. This second equation in principle allows the rate of aggregation of the protein to be determined without the need for the mutant aggregation data. Here again the coefficients of the equation are determined by a standard regression of the observed log(k) values. The authors found here, too, that both hydrophobicity and propensity to convert from α-helix

to β-sheet (or hydrophobic/hydrophilic patterns) contributed positively to the aggregation rate (α_1, $\alpha_2 > 0$), while charge contributed negatively ($\alpha_3 < 0$), see entry 19 of Table 2 [152]. As shown in Figure 16 below, the authors tested their equation using 79 data sets; comparison of the calculated log(k) vs. the experimental log (k) showed a good correlation with r=0.92.

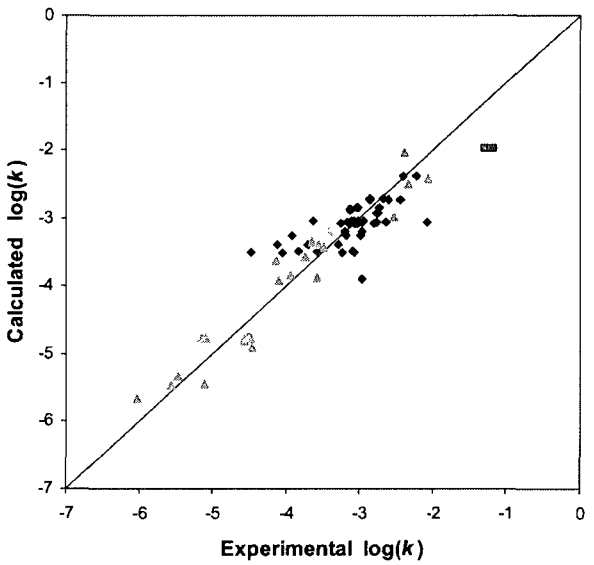


Figure 16. Calculated and experimentally observed changes in the absolute rate of aggregation (k) for 79 protein aggregation data sets including 59 data sets of WT and mutant human muscle acylphosphatase from DuBay et al. Reprinted from ref. [152]. Copyright (2004) with permission from Elsevier.

Although Dobson and co-workers' equations provide a powerful example of a phenomenological approach in which the rate of protein aggregation can be determined, the rate constants obtained from Dobson's equations still contain a *convolution of nucleation and growth* [unpublished results]. Restated, it would be of considerable interest to use Dobson and co-workers' approach, but on separate nucleation and growth rate constants such as k_1 and k_2 of the F-W 2-step model.

A Closer Look at the Five Classes of Mechanisms Using the Concept of Polymerized Monomer. The concept of polymerized monomer, first introduced by Oosawa [95] and subsequently used by others [101,130], can be used as an approximation tool to encompass *all monomer polymerized into aggregates of different sizes* (with the distinction, for example, of aggregates only past the critical nucleus), into one unique *polymerized monomer* species. This approximation can greatly reduce the complexity of the associated rate equations. For example, using the polymerized monomer approximation allows the mechanism of subsequent monomer addition to be rewritten as in Scheme 21, where M represents the protein monomer, P represents *polymerized monomer past the critical nucleus size*, and n represents the critical nucleus size.

Scheme 21. The Subsequent Monomer Addition Mechanism Incorporating the Polymerized Monomer Concept, with a Critical Nucleus Size of n.

$$nM \rightleftharpoons nP$$
$$M + P \rightleftharpoons 2P$$

In another example, the prion mechanisms of Griffith and Lansbury can be rewritten as in Scheme 22 below, where M_c is the cellular prion monomer, M_p is the infectious prion monomer, and P is now the *polymerized infectious prion monomer.* We did not include herein the prion mechanisms of Prusiner, as it did not originally involve aggregation steps, or Lindquist, as that mechanism is not easily simplified by the use of the polymerized monomer concept. Similarly, the reversible association mechanism can not be rewritten and simplified with P.

Scheme 22. (a) Griffith's and (b) Lansbury's Prion Mechanisms Rewritten with the Incorporation of the Polymerized Monomer Concept.

(a)

$$M_c \rightleftharpoons M_p$$
$$2M_p \longrightarrow 2P$$
$$M_c + P \longrightarrow M_c{\bullet}P \longrightarrow 2P$$

(b)

$$M_c \rightleftharpoons M_p$$
$$nM_p \rightleftharpoons nP$$
$$M_p + P \longrightarrow 2P$$

The F-W mechanism already employs the polymerized monomer concept. Rewriting this mechanism in terms of the nomenclature used above results in Scheme 23, where M represents the protein monomer, and P represents the *catalytically active polymerized monomer*.

Scheme 23. F-W mechanism Rewritten in Terms of Monomer, M, and Polymerized Monomer, P.

$$M \rightarrow P$$
$$M + P \rightarrow 2P$$

Using the polymerized monomer approximation with the subsequent monomer addition mechanism along with Griffith's and Lansbury's prion mechanisms allows for simplification. In the case of the subsequent monomer addition mechanism (Scheme 21), the polymerized monomer approximation simplifies this mechanism in such a way that it resembles the minimalistic F-W mechanism (Scheme 23).

Empirical Approaches to Determine Empirical Constants of Protein Aggregation. Empirically based equations are a different approach than those reported in the previous section (and shown in Table 2) to examine protein aggregation kinetic data.

Empirical equations typically provide a convenient and easy way to fit the kinetic data of protein aggregation, but they have a significant weakness: there is little or no physical meaning behind the typical empirical equation and, hence, its variables. As such, empirical approaches typically have not provided mechanistic insight into the aggregation process. However, primarily for the sake of completeness, some empirical methods are examined briefly below.

Logistic functions. Naiki and co-workers, who developed a first-order kinetic model for the growth only of amyloid-β fibrils (see Table S1 of the Supporting Information), also noticed that the sigmoidal aggregation curve obtained by incubation of an amyloid-β monomer, obeys a *logistic function* [153,154,155]. Cerny points out that logistic curves have indeed been shown to empirically fit a wide variety of sigmoidal data [156]. However, to our knowledge, no mechanistic insight has typically come out of fitting protein aggregation data with a logistic function.

Fink and co-workers have used a modified version [157] of Richards' function to fit sigmoidal aggregation curves observed in the fibrillation of proteins [52,55,158]. Richards' function is itself a generalized form of the logistic function [159], and is commonly used in botany to determine both the growth period length and the ultimate weight of plants [160]. The logistic equation used by Fink and co-workers is shown in Equation 3, where Y is the fluorescence intensity used to measure the aggregation, $(y_i + m_i x)$ is the initial slope during the lag phase, $(y_f + m_f x)$ is the final slope after growth has ended, x_0 is the time at 50% intensity, and $1/\tau$ is the length of time for the lag phase. The lag phase is calculated as $(x_0 \times 2\tau)$ and the apparent rate constant k_{app} is $(1/\tau)$ [52].

$$Y = (y_i + m_i x) + \frac{(y_f + m_f x)}{1 + e^{-\left(\frac{x - x_0}{\tau}\right)}} \quad (3)$$

The authors are careful to note that, "This expression is unrelated to the underlying molecular events, but provides a convenient method for comparison of the kinetics of fibrillation" [55]. Fink's empirical equation has been used to analyze numerous data sets thus allowing for comparisons to be made between related data sets [52,53,55,158]; an example is shown in Figure 17. Again, however, meaningful kinetic and mechanistic information is lacking. Recent work shows that Fink and co-workers' expert biophysical studies and resultant kinetic data can be analyzed by the F-W 2-step model [104].

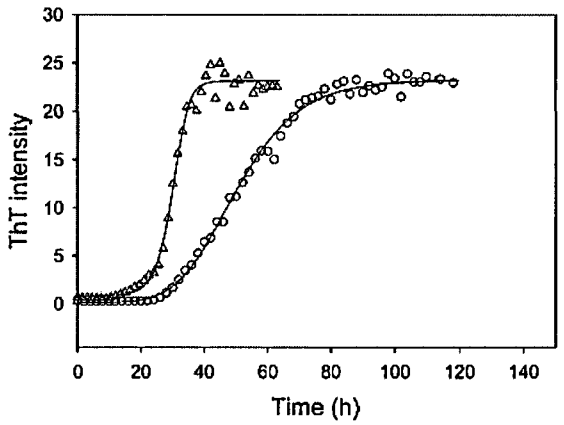


Figure 17. Aggregation of wild-type and mutant (Y39W) α-synuclein from Fink and co-workers measured by ThT fluorescence (points), along with corresponding fits using Eq. 3 (lines). Figure 1b of Figure 1a-b reprinted with permission from ref. [54]. Copyright (2006) American Chemical Society.

Recently, McRae and co-workers were able to express the empirical parameters in Fink's logistic equation in terms of kinetic variables from their model [161]. The

resulting correlations are complicated, and show that each of these empirical parameters is a convolution of variables such as initial monomer concentration, nucleation, elongation and dissociation rate constants, and average number of monomers per fibril [161]. Again, it would be of use to reanalyze this data in terms of models able to deconvolute average nucleation from average growth [107].

Three Problem Areas in the Protein Aggregation Kinetic and Mechanistic Literature. One must ask, why is such an important question, as 'what are the mechanism(s) of protein aggregation?', yet to be unequivocally answered despite the numerous contributions? The simple answer is that protein aggregation is a highly complex problem with complicated molecular level and kinetic details, along with associated complex mathematics. In hopes of moving this area forward, we list a few possible problem areas of importance, in our opinion, for future studies in protein aggregation.

Lack of a Full Appreciation of the Previous Literature. A complete understanding of the literature in any area is a very important and early requirement for the best and most efficient science—and especially in areas where a large literature exists such as in protein aggregation. However, and unfortunately, constructing this review makes apparent that not all papers published in the protein aggregation area have a clear and comprehensive understanding of the previous literature. Some mechanisms may be proposed only for a specific system of study, with little generalization to try to tie together a broad spectrum of protein aggregation literature. Relevant here is Platt's definition of good science as building upon others' ideas via the disproof of all possible

alternative hypotheses [162]. We hope that the present review will help alleviate this problem somewhat—we also offer our sincere apologies to any authors whose crucial literature we have inadvertently overlooked as we strove to distill the protein aggregation literature to its essential components.

Confused Nomenclature. One problem in the area of protein aggregation is that numerous terms are used for identical meanings. An example is with the multiple terms used for growth that include polymerization, elongation, fibrillation and maybe "heterogeneous nucleation"[2] [100]. Also, terms such as "extreme autocatalysis" due to the observation that "the first 10% of the reaction proceeds as the 10^{th} to 20^{th} power of time" [100] are confusing. Note that if one views "heterogeneous nucleation" as seeded autocatalytic growth, then terms such as "extreme catalysis" are taken care of by autocatalysis. In addition, it is not clear whether the descriptive terms "passive" vs. "active" autocatalysis [115] have any physical or chemical basis.[4] Also used alternately but with identical meaning (to our knowledge), are the terms "positive [kinetic] cooperativity," "positive feedback activation" [163], and "autocatalysis." Many other examples exist, so that we find a need for a clear, unified nomenclature in the area of protein aggregation. The early footnotes herein on nomenclature[1,2] are offered as an aid towards a more unified nomenclature in the protein aggregation area.

Using Word- or Picture-Only Mechanisms. At times, a word or picture-only mechanism may be useful to describe what would account to a very complex set of equations (that may not be tested or even testable due to their associated mathematical complexity). But, as such they do not lead to precise kinetic equations and corresponding differential equations (and then, ideally, their integrated versions), such kinetic equations

being required to test the mechanism vs. experimental kinetic data. Another serious issue arises when a word or picture-only mechanism is proposed along with equation(s) that cannot be derived from that (word or picture) mechanism: *a disconnect then results between the mechanism proposed and the equation(s) given.* This is turn leads to the use of at least rigorously undefined, if not the wrong, words for the physical phenomenon at hand. In rigorous mechanistic studies, balanced chemical equations define *both the rate constant* and the *word concepts* that one can use. This is a fundamental, central and important point for proceeding more efficiently in the area of protein aggregation kinetics and mechanism.

Summary and Conclusions

Despite the importance of the problem and the nearly 50 years of research aimed at determining the mechanism(s) and rate constant(s) for protein aggregation, many questions still remain. Frieden concisely summarizes the state of affairs in his recent review: "In spite of the extensive literature, however, the mechanism of [protein] aggregation is poorly understood" [16]. The focus of the present review has been to examine the protein aggregation literature from the perspective of trying to fit the available protein aggregation kinetic data to obtain quantitative, useful rate constants and mechanistic information on the aggregation process. The main contributions from this review are believed to be the following:

- We have briefly reviewed the methods used in the literature to follow protein aggregation, pointing out those methods that are direct vs. indirect, and in-situ vs. ex-situ. We have emphasized that the use of multiple techniques is advisable if

not necessary to avoid over-interpretation and to ensure that accurate kinetic data are collected.

- We then reviewed what is known about the starting proteins, products, and intermediates of protein aggregation, since knowing the products is a necessary prerequisite to reliable kinetic and mechanistic work. In the case of *amyloidogenic proteins*, the product of protein aggregation has been widely demonstrated to be aggregated protein fibrils. However, the pathway by which normal monomeric forms of the protein become fibrils and the intermediate species that are formed along the way, either on- or off-pathway to fibril formation is not yet known. Future studies are needed, since a consensus on the structure and toxicity of intermediate species remains elusive.
- Many approaches exist for determining the mechanism(s) and rate constant(s) of protein aggregation. These approaches can be broken broadly into three categories: (i) kinetic and thermodynamic, (ii) other, and (iii) empirical approaches as summarized in Scheme 2.
- We have focused on reporting, in historical order, the kinetic, thermodynamic, and other approaches that have been used in determining the kinetics and mechanism(s) of protein aggregation. We find that there are five main classes of mechanisms or approaches in the protein aggregation literature: i) subsequent monomer addition, ii) reversible association, iii) prion aggregation mechanisms, iv) the "Ockham's razor"/minimalistic 2-step (F-W) model, and v) quantitative structure-activity relationship models.

- Each of the five classes of kinetic and thermodynamic mechanisms was reviewed in detail, and where possible the fits obtained from the proposed mechanisms and corresponding equations were given to demonstrate how well those mechanisms/models were able to fit protein aggregation kinetic data.
- We applied the concept of "polymerized monomer" originally introduced by Oosawa [95], to the subsequent monomer addition mechanism along with Griffith's and Lansbury's prion mechanisms. This approximation method helps simplify the systems of associated rate equations so that they should become easier to solve.
- In what may prove to be an important part of this review, we expressed where possible the (four) mechanisms of protein aggregation shown in Schemes 21, 22a, 22b, and 23 in terms of common monomer (M) and polymerized monomer (P) terms so that they can be compared, contrasted, and serve as a stepping-stone for future work.
- We also briefly discussed empirical approaches that have been used in the literature. The empirical approaches and resultant equations provide good fits to the kinetic data, but the resultant fitting parameters lack physical meaning, greatly limiting their usefulness.

Acknowledgements

We thank Mr. Steve Hays for his graphic design of the schematic fibril pictured in Scheme 1. We also thank Professors Eric D. Ross and Jeffrey N. Agar for providing their

insightful comments on the manuscript. Finally, we gratefully acknowledge NSF grant #0611588 for partial support of this project.

---

[1] K. Blennow, M.J. de Leon, H. Zetterberg, Alzheimer's disease, Lancet 368 (2006) 387-403 and references therein.

[2] G.P. Bates, C. Benn, The polyglutamine diseases, in: G.P. Bates, P.S. Harper, L. Jones (Eds.), Huntington's Disease, Oxford University Press, Oxford, 2002; pp 429-472.

[3] W. Dauer, S. Przedborski, Parkinson's disease: mechanisms and models, Neuron 39 (2003) 889-909 and references therein.

[4] S.B. Prusiner, Molecular biology of prion diseases, Science 252 (1991) 1515-1522.

[5] L. Morozova-Roche, M. Malisauskas, A false paradise - Mixed blessings in the protein universe: the amyloid as a new challenge in drug development, Curr. Med. Chem. 14 (2007) 1221-1230.

[6] J.M. Berg, J. L. Tymoczko, L. Stryer, Actin is a polar, self-assembling, dynamic polymer, in: Biochemistry 5th edition, W. H. Freeman, New York, 2002, pp. 958-960.

[7] H. Eisenberg, Glutamate dehydrogenase: anatomy of a regulatory enzyme, Acc. Chem. Res. 4 (1971) 379-385.

[8] D. Thusius, Mechanism of bovine liver glutamate dehydrogenase self-assembly: II. Simulation of relaxation spectra for an open linear polymerization proceeding via a sequential addition of monomer units, J. Mol. Biol. 94 (1975) 367-383.

[9] D. Thusius, P. Dessen, J.-M. Jallon, Mechanism of bovine liver glutamate dehydrogenase self-association I. Kinetic evidence for a random association of polymer chains, J. Mol. Biol. 92 (1975) 413-432.

[10] M. Jullien, D. Thusius, Mechanism of bovine liver glutamate dehdrogenase self-assembly III. Characterization of the association-dissociation stoichiometry with quasi-elastic light scattering, J. Mol. Biol. 101 (1976) 397-416.

[11] C.J. Roberts, Kinetics of irreversible protein aggregation: analysis of extended Lumry-Eyring models and implications for predicting protein shelf life, J. Phys. Chem. B 107 (2003) 1194-1207.

[12] F. Oosawa, S. Asakura, K. Hotta, I. Nobuhisa, T. Ooi, G-F transformation of actin as a fibrous condensation, J. Polym. Sci. 37 (1959) 323-336.

[13] D.M. Walsh and D.J. Selkoe, Aβ oligomers. A decade of discovery, J. Neurochem. 101 (2007) 1172-1184.

[14] C.J. Roberts, Non-native protein aggregation kinetics, Biotechnol. Bioeng. 98 (2007) 927-938.

[15] R.M. Murphy, B.S. Kendrick, Protein misfolding and aggregation, Biotechnol. Prog. 23 (2007) 548-552.

[16] C. Frieden, Protein aggregation processes: in search of the mechanism, Protein Science 16 (2007) 2334-2344.

[17] E. van der Linden, P.Venema, Self-assembly and aggregation of proteins, Curr. Opin. Colloid and Interface Sci. 12 (2007) 158-165.

[18] J. Gsponer, M. Vendruscolo, Theoretical approaches to protein aggregation, Protein and Peptide Lett. 13 (2006) 287-293.

[19] S. Ohnishi, K. Takano, Amyloid fibrils from the viewpoint of protein folding, Cell. Mol. Life Sci. 61 (2004) 511-524.

[20] S.T. Ferreira, M.N.N. Vieira, F.G. De Felice, Soluble protein oligomers as emerging toxins in Alzheimer's and other amyloid diseases, Life 59 (2007) 332-345.

[21] E. Monsellier, F. Chiti, Prevention of amyloid-like aggregation as a driving force of protein evolution, EMBO Reports 8 (2007) 737-742.

[22] K. Trzesniewska, M. Brzyska, D. Elbaum, Neurodegenerative aspects of protein aggregation, Acta Neurobiol. Exp. 64 (2004) 41-52.

[23] R. Tycko, Insights into the amyloid folding problem from solid-state NMR, Biochemistry 42 (2003) 3151-3159.

[24] D. Foguel, J.L. Silva, New insights into the mechanisms of protein misfolding and aggregation in amyloidogenic diseases derived from pressure studies, Biochemistry 43 (2004) 11361-11370.

[25] R.M. Murphy, Kinetics of amyloid formation and membrane interaction with amyloidogenic proteins, Biochim. Biophys. Acta 1768 (2007) 1923-1934.

[26] M.A. Lauffer, Entropy-driven polymerization of proteins: tobacco mosaic virus protein and other proteins of biological importance, Developments in Biochemistry 30 (1978) 115-170.

[27] S.E. Bondos, Methods for measuring protein aggregation, Current Analytical Chemistry 2 (2006) 157-170.

[28] T.R. Serio, A.G. Cashikar, A.S. Kowal, G.J. Sawicki, J.J. Moslehi, L. Serpell, M.F. Arnsdorf, S.L. Lindquist, Nucleated conformational conversion and the replication of conformational information by a prion determinant, Science 289 (2000) 1317-1321.

[29] M.A. Watzky, A.M. Morris, E.D. Ross, R.G. Finke, Fitting yeast and mammalian prion aggregation kinetic data with the Finke-Watzky 2-step model of nucleation and autocatalytic growth, Biochemistry 47 (2008) 10790-10800.

[30] J.D. Harper, P.T. Lansbury, Jr., Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins, Annu. Rev. Biochem. 66 (1997) 385-407.

[31] C. LeBlond, J. Wang, R.D. Larsen, C.J. Orella, A.L. Forman, R.N. Landau, J. Laquidara, J.R. Sowa, Jr., D.G. Blackmond, Y.-K. Sun, Reaction calorimetry as an in-situ kinetic tool for characterizing complex reactions, Thermochimica Acta 289 (1996) 189-207.

[32] R.W. Woody, Theory of circular dichroism of proteins, in: G.D. Fasman, (Ed.), Theory of Circular Dichroism of Proteins in Circular Dichroism and the Conformational Analysis of Biomolecules, Plenum Press, New York, 1996, pp. 25-68.

[33] L.A. Munishkina, A.L. Fink, Fluorescence as a method to reveal structures and membrane-interactions of amyloidogenic proteins, Biochim. Biophys. Acta 1768 (2007) 1862-1885.

[34] K.M. Lundberg, C.J. Stenland, F.E. Cohen, S.B. Prusiner, G. L. Millhauser, Kinetics and mechanism of amyloid formation by the prion protein H1 peptide as determined by time-dependent ESR, Chemistry & Biology 4 (1997) 345-355.

[35] F. Oosawa, S. Asakura, K. Hotta, I. Nobuhisa, T. Ooi, G-F transformation of actin as a fibrous condensation, J. Polym. Sci. 37 (1959) 323-336.

[36] M. Kasai, F. Oosawa, Flow birefringence, Meth. Enzym. 26 (1972) 289-323.

[37] A. Lomakin, D.B. Teplow, D.A. Kirschner, G.B. Benedek, Kinetic theory of fibrillogenesis of amyloid β-protein, Proc. Natl. Acad. Sci. USA 94 (1997) 7942-7947.

[38] S.P.F.M. Roefs, K.G. De Kruif, A model for the denaturation and aggregation of beta-lactoglobulin, Eur. J. Biochem. 226 (1994) 883-889.

[39] S.F. Sun, Diffusion, Light scattering, in: Physical Chemistry of Macromolecules: Basic Principles and Issues, 2nd Ed., Wiley and Sons, New York, 2004, pp. 223-242.

[40] M. Moniatte, F.G. van der Goot, J.T. Buckley, F. Pattus, A. van Dorsselaer, Characterisation of the heptameric pore-forming complex of the *Aeromonas* toxin aerolysin using MALDI-TOF mass spectrometry, FEBS Lett. 384 (1996) 269-272.

[41] W.J. Henzel, J.T. Stults, Matrix-assisted laser desorption/ionization time-of-flight mass analysis of peptides, in: J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, P.T. Wingfield (Eds.), Current Protocols in Protein Science, Wiley and Sons, New York, 2004, Unit 16.2.

[42] M. Kamihira, A. Naito, S. Tuzi, A. Nosaka, H. Saitô, Conformational transitions and fibrillation mechanism of human calcitonin as studied by high-resolution solid-state ^{13}C NMR, Protein Sci. 9 (2000) 867-877.

[43] C.O. Fernandéz, W. Hoyer, M. Zweckstetter, E.A. Jares-Erijman, V. Subramaniam, C. Griesinger, T.M. Jovin, NMR of alpha-synuclein-polyamine complexes elucidates the mechanism and kinetics of induced aggregation, EMBO J. 23 (2004) 2039-2046.

[44] T.P.J. Knowles, W. Shu, G.L. Devlin, S. Meehan, S. Auer, C.M. Dobson, M.E. Welland, Kinetics and thermodynamics of amyloid formation from direct measurements of flucuations in fibril mass, Proc. Natl. Acad. Sci. USA 104 (2007) 10016-10021.

[45] B.J. Berne, Interpretation of the light scattering from long rods, J. Mol. Biol. 89 (1974) 755-758.

[46] S.E. Harding, The intrinsic viscosity of biological macromolecules. Progress in measurement, interpretation and application to structure in dilute solution, Prog. Biophys. Mol. Biol. 68 (1997) 207-262.

[47] M. Sunde, L.C. Serpell, M. Bartlam, P.E. Fraser, M.B. Pepys, C.C.F. Blake, Common core structure of amyloid fibrils by synchrotron X-ray diffraction, J. Mol. Biol. 273 (1997) 729-739.

[48] R. Wetzel, For protein misassembly, it's the "I" decade, Cell 86 (1996) 699-702.

[49] M. Sandal, F. Valle, I. Tessari, S. Mammi, E. Bergantino, F. Musiani, M. Brucale, L. Bubacco, B. Samori, Conformational equilibria in monomeric alpha-synuclein at the single-molecule level, PLoS Biol. 6 (2008) 99-108.

[50] H. Ischiropoulos, Oxidative modifications of alpha-synuclein, Ann. N.Y. Acad. Sci. 991 (2003) 93-100.

[51] B.I. Giasson, J.E. Duda, I.V.J. Murray, Q. Chen, J.M. Souza, H.I. Hurtig, H. Ischiropoulos, J.Q. Trojanowski, V. M.-Y. Lee, Oxidative damage linked to

neurodegeneration by selective alpha-synuclein nitration in synucleinopathy lesions, Science 290 (2000) 985-989.

[52] V.N. Uversky, J. Li, A.L. Fink, Evidence for a partially folded intermediate in α-synuclein fibril formation, J. Biol. Chem. 276 (2001) 10737-10744.

[53] V.N. Uversky, A.L. Fink, Conformational constraints for amyloid fibrillation: the importance of being unfolded, Biochim. Biophys. Acta 1698 (2004) 131-153.

[54] A. Dusa, J. Kaylor, S. Edridge, N. Bodner, D.-P. Hong, A.L. Fink, Characterization of oligomers during α-synuclein aggregation using intrinsic tryptophan fluorescence, Biochemistry 45 (2006) 2752-2760.

[55] A.L. Fink, The aggregation and fibrillation of α-synuclein, Acc. Chem. Res. 39 (2006) 628-634.

[56] M.F. Perutz, J.T. Finch, J. Berriman, A. Lesk, Amyloid fibers are water-filled nanotubes, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 5591-5595.

[57] S.B. Malinchik, H. Inouye, K.E. Szumowski, D.A. Kirschner, Biophys. J. 74 (1998) 537-545.

[58] L. Serpell, C.C.F. Blake, P.E. Fraser, Structural analysis of Alzheimer's beta(1-40) amyloid: Protofilament assembly of tubular fibrils, Biochemistry 39 (2000) 13269-13275.

[59] W.A. Eaton, J. Hofrichter, Sickle cell hemoglobin polymerization, in: C.B. Anfinsen, J.T. Edsall, F.M. Richards, D.S. Eisenberg (Eds.), Advances in Protein Chemistry, Academic Press, San Diego, 1990.

[60] G. Dykes, R.H. Crepeau, S.J. Edelstein, Three-dimensional reconstruction of the fibres of sickle cell haemoglobin, Nature 272 (1978) 506-510.

[61] G.W. Dykes, R.H. Crepeau, S.J. Edelstein, Three-dimensional reconstruction of the 14-filament fibers of hemoglobin S, J. Mol. Biol. 130 (1979) 451-472.

[62] D.L. Harrington, K. Adachi, W.E. Royer Jr., The high resolution crystal structure of deoxyhemoglobin S, J. Mol. Biol 272 (1997) 398-407.

[63] F.A. Ferrone, Nucleation: the connections between equilibrium and kinetic behavior, Methods Enzymol. 412 (2006) 285-299.

[64] C.C. Selby, R.S. Bear, The structure of actin-rich filaments of muscles according to X-ray diffraction, J. Biophysic. Biochem. Cytol. 2 (1956) 71-85.

[65] J. Hanson, J. Lowy, The structure of F-actin and of actin filaments isolated from muscle, J. Mol. Biol. 6 (1963) 46-60.

[66] J. Woodhead-Galloway, Structure of collagen fibril: some variations on a theme of tetragonally packed dimmers, Proc. R. Soc. Lond. B 209 (1980) 275-297.

[67] V. Ottani, D. Martini, M. Franchi, A. Ruggeri, M. Raspanti, Hierarchical structures of fibrillar collagens, Micron 33 (2002) 587-596.

[68] H. Wille, M.D. Michelitsch, V. Guénebaut, S. Supattapone, A. Serban, F.E. Cohen, D.A. Agard, S.B. Prusiner, Structural studies of the scrapie prion protein by electron crystallography, Proc. Natl. Acad. Sci. USA 99 (2002) 3563-3568.

[69] R. Nelson, M.R. Sawaya, M. Balbirnie, A.O. Madsen, C. Riekel, R. Grothe, D. Eisenberg, Structure of the cross-β spine of amyloid-like fibrils, Nature 435 (2005) 773-778.

[70] O.S. Makin, E. Atkins, P. Sikorski, J. Johansson, L.C. Serpell, Molecular basis for amyloid fibril formation and stability, Proc. Natl. Acad. Sci. USA 102 (2005) 315-320.

[71] O.S. Makin, L.C. Serpell, Structures for amyloid fibrils, FEBS J. 272 (2005) 5950-5961.

[72] X. Fernandez-Busquets, N.S. de Groot, D. Fernandez, S. Ventura, Recent structural and computational insights into conformational diseases, Curr. Med. Chem. 15 (2008) 1336-1349.

[73] W.S. Gosal, I.J. Morten, E.W. Hewitt, D.A. Smith, N.H. Thomson, S.E. Radford, Competing pathways determine fibril morphology in the self-assembly of beta(2)-microglobulin into amyloid, J. Mol. Biol. 351 (2005) 850-864.

[74] E.T. Powers, D.L. Powers, Mechanism of protein fibril formation: nucleated polymerization with competing off-pathway aggregation, Biophys. J. 94 (2008) 379-391.

[75] H.A. Lashuel, B.M. Petre, J. Wall, M. Simon, R.J. Nowak, T. Walz, P.T. Lansbury, Jr., α-Synuclein, especially the Parkinson's disease-associated mutants, forms pore-like annular and tubular protofibrils, J. Mol. Biol. 322 (2002) 1089-1102.

[76] H.A. Lashuel, D. Harley, B.M. Petre, T. Walz, P.T. Lansbury, Jr., Amyloid pores from pathogenic mutations, Nature 418 (2002) 291.

[77] T.T. Ding, S.J. Lee, J.-C. Rochet, P.T. Lansbury, Jr., Annular α-synuclein protofibrils are produced when spherical protofibrils are incubated in solution or bound to brain-derived membranes, Biochemistry 41 (2002) 10209-10217.

[78] B. Caughey, P.T. Lansbury, Jr., Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders, Annu. Rev. Neurosci. 26 (2003) 267-298.

[79] J. Kaylor, N. Bodner, S. Edridge, G. Yamin, D.P. Hong, A.L. Fink, Characterization of oligomeric intermediates in α-synuclein fibrillation: FRET studies of Y125W/Y133F/Y136F α-synuclein, J. Mol. Biol. 353 (2005) 357-372.

[80] M. Zhu, S. Han, F. Zhou, S.A. Carter, A.L. Fink, Annular oligomeric amyloid intermediates observed by in situ atomic force microscopy, J. Biol. Chem. 279 (2004) 24452-24459.

[81] R. Khurana, C. Ionescu-Zanetti, M. Pope, J. Li, L. Nielson, M. Ramirez-Alvarado, L. Regan, A.L. Fink, S.A. Carter, A general model for amyloid fibril assembly based on morphological studies using atomic force microscopy, Biophys. J. 85 (2003) 1135-1144.

[82] G.M. Shankar, S. Li, T.H. Mehta, A. Garcia-Munoz, N.E. Shepardson, I. Smith, F.M. Brett, M.A. Farrell, M.J. Rowan, C.A. Lemere, C.M. Regan, D.M. Walsh, B.L. Sabatini, D.J. Selkoe, Amyloid-β protein dimmers isolated directly from Alzheimer's brains impair synaptic plasticity and memory, Nature Medicine 14 (2008) 837-842.

[83] M.J.Volles, P.T. Lansbury, Jr., Zeroing in on the pathogenic form of alpha-synuclein and its mechanism of neurotoxicity in Parkinson's disease, Biochemistry 42 (2003) 7871-7878.

[84] D.A. Butterfield, A. Castegna, C.M. Lauderback, J. Drake, Evidence that amyloid beta-peptide-induced lipid peroxidation and its sequelae in Alzheimer's disease brain contribute to neuronal death, Neurobiology of Aging 23 (2002) 655-664.

[85] B.J. Tabner, S. Turnbull, N.J. Fullwood, M. German, D. Allsop, The production of hydrogen peroxide during early-stage protein aggregation: a common pathological mechanism in different neurodegenerative diseases?, Biochem. Soc. Trans. 33 (2005) 548-550.

[86] M.T. Maloney, J.R. Bamburg, Cofilin-mediated neurodegeneration in Alzheimer's disease and other amyloidopathies, Mol. Neurobiol. 35 (2007) 21-44.

[87] C.M. Dobson, Protein aggregation and its consequences for human disease, Protein Peptide Lett. 13 (2006) 219-227.

[88] F. Sokolowski, A.J. Modler, R. Masuch, D. Zirwer, M. Baier, G. Lutsch, D.A. Moss, K. Gast, D. Naumann, Formation of critical oligomers is a key event during conformational transition of recombinant Syrian hamster prion protein, J. Biol. Chem. 278 (2003) 40481-40492.

[89] J. Ollesch, E. Künnemann, R. Glockshuber, K. Gerwert, Prion protein α-to-β transition monitored by time-resolved Fourier transform infrared spectroscopy, Appl. Spectrosc. 61 (2007) 1025-1031.

[90] A. Perálvarez-Marín, A. Barth, A. Gräslund, Time-resolved infrared spectroscopy of pH-induced aggregation of the Alzheimer $A\beta_{1-28}$ peptide, J. Mol. Biol. 379 (2008) 589-596.

[91] M.A. Lauffer, in: A. Kleinzeller, G.F. Springer, H.G. Wittmann (Eds.), Entropy-driven processes in biology: polymerization of tobacco mosaic virus protein and similar reactions, Springer-Verlag, Berlin, 1975.

[92] S.S. Licht, C.C. Lawrence, J. Stubbe, Thermodynamic and kinetic studies on carbon-cobalt bond homolysisby ribonucleoside triphosphate reductase: the importance of entropy in catalysis, Biochemistry 38 (1999) 1234-1242.

[93] A.M. Morris, M.A. Watzky, J.N. Agar, R.G. Finke, Fitting neurological protein aggregation kinetic data via a 2-step minimal/"Ockham's razor" model: the Finke-Watzky mechanism of nucleation followed by autocatalytic surface growth, Biochemistry 47 (2008) 2413-2427.

[94] M. Kasai, S. Asakura, F. Oosawa, The cooperative nature of G-F transformation of actin, Biochim. Biophys. Acta 57 (1962) 22-31.

[95] F. Oosawa, M. Kasai, A theory of linear and helical aggregations of macromolecules, J. Mol. Biol. 4 (1962) 10-21.

[96] F. Oosawa, Size distribution of protein polymers, J. Theor. Biol. 27 (1970) 69-86.

[97] J. Hofrichter, P.D. Ross, W.A. Eaton, Kinetics and mechanism of deoxyhemoglobin S gelation: a new approach to understanding sickle cell disease, Proc. Natl. Acad. Sci. USA 71 (1974) 4864-4868.

[98] M. Saunders, P.D. Ross, A simple model of the reaction between polyadenylic acid and polyuridylic acid, Biochem. Biophys. Res. Comm. 3 (1960) 314-318.

[99] V.K. La Mer, Nucleation in phase transitions, Ind. Eng. Chem. 44 (1952) 1270-1277.

[100] F.A. Ferrone, J. Hofrichter, H.R. Sunshine, W.A. Eaton, Kinetic studies on photolysis-induced gelation of sickle cell hemoglobin suggest a new mechanism, Biophys. J. 32 (1980) 361-377.

[101] A. Wegner, J. Engel, Kinetics of the cooperative association of actin to actin filaments, Biophys. Chem. 3 (1975) 215-225.

[102] L.S. Tobacman, E.D. Korn, The kinetics of actin nucleation and polymerization, J. Biol. Chem. 258 (1983) 3207-3214.

[103] C. Frieden, D.W. Goddette, Polymerization of actin and actin-like systems: evaluation of the time course of polymerization in relation to the mechanism, Biochemistry 22 (1983) 5836-5843.

[104] A.M. Morris, R.G. Finke, α-Synuclein aggregation variable temperature and variable pH kinetic data: A re-analysis using the Finke-Watzky 2-step model of nucleation and autocatalytic growth, Biophys. Chem. 140 (2009) 9-15.

[105] M.P. Firestone, R. De Levie, S.K. Rangarajan, On one-dimensional nucleation and growth of "living" polymers I. Homogeneous nucleation, J. Theor. Biol. 104 (1983) 553-570.

[106] S.K. Rangarajan, R. De Levie, On one-dimensional nucleation and growth of "living" polymers II. Growth at constant monomer concentration, J. Theor. Biol. 104 (1983) 553-570.

[107] M.A. Watzky, R.G. Finke, Transition metal nanocluster formation kinetic and mechanistic studies. A new mechanism when hydrogen is the reductant: slow, continuous nucleation and fast autocatalytic surface growth, J. Am. Chem. Soc. 119 (1997) 10382-10400.

[108] R.F. Goldstein, L. Stryer, Cooperative polymerization reactions: analytical approximations, numerical examples, and experimental strategy, Biophys. J. 50 (1986) 583-599.

[109] F. Oosawa, S. Asakura, Kinetics of polymerization 4. Fragmentation and association of polymers, in: B. Horecker, N.O. Kaplan, J. Marmur, H.A. Scheraga (Eds.), Thermodynamics of the Polymerization of Protein, Academic Press, New York, 1975, p. 55.

[110] A. Wegner, P. Savko, Fragmentation of actin filaments, Biochemistry 21 (1982) 1909-1913.

[111] P.O.P. Ts'o, I.S. Melvin, A.C. Olson, Interaction and association of bases and nucleosides in aqueous solutions, J. Am. Chem. Soc. 85 (1963) 1289-1296.

[112] K.E. Van Holde, G.P. Rossetti, A sedimentation equilibrium study of the association of purine in aqueous solutions, Biochemistry 6 (1967) 2189-2194.

[113] E.T. Adams, Jr., M.S. Lewis, Sedimentation equilibrium in reacting systems. VI. Some applications to indefinite self-association. Studies with β-lactoglobulin A, Biochemistry 7 (1968) 1044-1053.

[114] E. Reisler, J. Pouyet, H. Eisenberg, Molecular weights, association, and frictional resistance of bovine liver glutamate dehydrogenase at low concentrations. Equilibrium and velocity sedimentation, light-scattering studies, and settling experiments with macroscopic models of the enzyme oligomer, Biochemistry 9 (1970) 3095-3102.

[115] M. Eigen, Prionics or the kinetic basis of prion diseases, Biophys. Chem. 63 (1996) A1-A18.

[116] J.S. Griffith, Self-replication and scrapie, Nature 215 (1967) 1043-1044.

[117] S.B. Prusiner, Novel proteinaceous infectious particles cause scrapie, Science 216 (1982) 136-144.

[118] F.E. Cohen, K.-M. Pan, Z. Huang, M. Baldwin, R.J. Fletterick, S.B. Prusiner, Structural clues to prion replication, Science 264 (1994) 530-531.

[119] I.V. Baskakov, G. Legname, M.A. Baldwin, S.B. Prusiner, F.E. Cohen, Pathway complexity of prion protein assembly into amyloid, J. Biol. Chem. 277 (2002) 21140-21148.

[120] M. Laurent, Autocatalytic processes in cooperative mechanisms of prion diseases, FEBS Lett. 407 (1997) 1-6.

[121] J.T. Jarrett, P.T. Lansbury, Jr., Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie?, Cell 73 (1993) 1055-1058.

[122] J.H. Come, P.E. Fraser, P.T. Lansbury, Jr., A kinetic model for amyloid formation in prion diseases: importance of seeding, Proc. Natl. Acad. Sci. USA 90 (1993) 5959-5963.

[123] P. T. Lansbury, Jr., B. Caughey, The chemistry of scrapie infection: implications of the 'ice 9' metaphor, Chemistry and Biology 2 (1995) 1-5.

[124] B. Caughey, D.A. Kocisko, G.J. Raymond, P.T. Lansbury, Jr., Aggregates of scrapie-associated prion protein induce the cell-free conversion of protease-sensitive prion protein to the protease-resistant state, Chemistry and Biology 2 (1995) 807-817.

[125] T. Scheibel, J. Bloom, S.L. Lindquist, The elongation of yeast prion fibers involves separable steps of association and conversion, Proc. Natl. Acad. Sci. USA 101 (2004) 2287-2292.

[126] H. Flyvbjerg, E. Jobs, S. Leibler, Kinetics of self-assembling microtubules: an "inverse problem" in biochemistry, Proc. Natl. Acad. Sci. USA 93 (1996) 5975-5979.

[127] A. Lomakin, D.S. Chung, G.B. Benedek, D.A. Kirschner, D.B. Teplow, On the nucleation and growth of amyloid β-protein fibrils: detection of nuclei and quantitation of rate constants, Proc. Natl. Acad. Sci. USA 93 (1996) 1125-1129.

[128] R. Kayed, E. Head, J.L. Thompson, T.M. McIntire, S.C. Milton, C.W. Cotman, C.G. Glabe, Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis, Science 300 (2003) 486-489.

[129] M.M. Pallitto, R.M. Murphy, A mathematical model of the kinetics of beta-amyloid fibril growth from the denatured state, Biophys. J. 81 (2001) 1805-1822.

[130] F.A. Ferrone, Analysis of protein aggregation kinetics, Methods Enzymol. 309 (1999) 256-274.

[131] M.F. Bishop, F.A. Ferrone, Kinetics of nucleation-controlled polymerization, Biophys. J. 46 (1984) 631-644.

[132] S. Chen, F.A. Ferrone, R. Wetzel, Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation, Proc. Natl. Acad. Sci. 99 (2002) 11884-11889.

[133] R. Wetzel, Kinetics and thermodynamics of amyloid fibril assembly, Acc. Chem. Res. 39 (2006) 671-679.

[134] T. Cellmer, R. Douma, A. Huebner, J. Prausnitz, H. Blanch, Kinetic studies of protein L aggregation and disaggregation, Biophys. Biochem. 125 (2007) 350-359.

[135] C.C. Lee, R.H. Walters, R.M. Murphy, Reconsidering the mechanism of polyglutamine peptide aggregation, Biochemistry 46 (2007) 12810-12820.

[136] R. Lumry, H. Eyring, Conformation changes of proteins, J. Phys. Chem. 58 (1954) 110-120.

[137] J.M. Andrews, C.J. Roberts, A Lumry-Eyring nucleated polymerization model of protein aggregation kinetics: 1. Aggregation with pre-equilibrated unfolding, J. Phys. Chem. B 111 (2007) 7897-7913.

[138] C.J. Roberts, The kinetics of nucleated polymerizations at high concentrations: amyloid fibril formation near and above the "supercritical concentration, in: R.M. Murphy, A.M. Tsai, (Eds), Misbehaving Proteins: Protein (Mis)Folding, Aggregation and Stability, Springer, New York, 2006, pp. 17-46.

[139] J.M. Andrews, C.J. Roberts, Non-native aggregation of alpha-chymotrypsinogen occurs through nucleation and growth with competing nucleus sizes and negative activation energies, Biochemistry 46 (2007) 7558-7571.

[140] W.F. Weiss IV, T.K. Hogdon, E.W. Kaler, A.M. Lenhoff, C.J. Roberts, Nonnative protein polymers: Structure, morphology, and relation to nucleation and growth, Biophys. J. 93 (2007) 4392-4403.

[141] J.M. Andrews, W.F. Weiss IV, C.J. Roberts, Nucleation, growth, and activation energies for seeded and unseeded aggregation of alpha-chymotrypsinogen A, Biochemistry 47 (2008) 2397-2403.

[142] E.T. Powers, D.L. Powers, The kinetics of nucleated polymerizations at high concentrations: amyloid fibril formation near and above the "supercritical concentration", Biophys. J. 91 (2006) 122-132.

[143] William of Ockham, 1285-1349, as cited by E.A. Moody in, The Encyclopedia of Philosophy, vol 7, McMillan, 1967.

[144] S.S.-S. Wang, Y.-T. Chen, P.-H. Chen, K.-N. Liu, A kinetic study on the aggregation behavior of beta-amyloid peptides in different initial solvent environments, Biochem. Eng. J. 29 (2006) 129-138.

[145] M.-S. Lin, L.-Y. Chen, H.-T. Tsai, S. S.-S. Wang, Y. Chang, A. Higuchi, W.-Y. Chen, Investigation of the mechanism of β-amyloid fibril formation by kinetic and thermodynamic analyses, Langmuir 24 (2008) 5802-5808.

[146] R. Sabaté, M. Gallardo, J. Estelrich, An autocatalytic reaction as a model for the kinetics of the aggregation of β-amyloid, Biopolymers 71 (2003) 190-195.

[147] R. Sabaté, J. Estelrich, Stimulatory and inhibitory effects of alkyl bromide surfactants on beta-amyloid fibrillogenesis, Langmuir 21 (2005) 6944-6949.

[148] R. Sabaté, M. Gallardo, J. Estelrich, Spontaneous incorporation of β-amyloid peptide into neutral liposomes, Colloids and Surfaces A 270-271 (2005) 13-17.

[149] R. Sabaté, U. Baxa, L. Benkemoun, N. Sánchez de Groot, B. Coulary-Salin, M.-L. Maddelein, L. Malato, S. Ventura, A.C. Steven, S.J. Saupe, Prion and non-prion amyloids of the HET-s prion forming domain, J. Mol. Biol. 370 (2007) 768-783.

[150] T.J. Gibson, R.M. Murphy, Inhibition of insulin fibrillogenesis with targeted peptides, Protein Sci. (2006) 1133-1141.

[151] F. Chiti, M. Stefani, N. Taddei, G. Ramponi, C.M. Dobson, Rationalization of the effects of mutations on peptide and protein aggregation rates, Nature 424 (2003) 805-808.

[152] K.F. DuBay, A.P. Pawar, F. Chiti, J. Zurdo, C.M. Dobson, M. Vendruscolo, Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains, J. Mol. Biol. 341 (2004) 1317-1326.

[153] H. Naiki, K. Nakakuki, First-order kinetic model of Alzheimer's β-amyloid fibril extension in vitro, Lab. Invest. 74 (1996) 374-383.

[154] K. Hasegawa, K. Ono, M. Yamada, H. Naiki, Kinetic modeling and determination of reaction constants of Alzheimer's β-amyloid fibril extension and dissociation using surface plasmon resonance, Biochemistry 41 (2002) 13489-13498.

[155] H. Naiki, F. Gejyo, Kinetic analysis of amyloid fibril formation, Methods Enzymol. 309 (1999) 305-318.

[156] L.C. Cerny, D.M. Stasiw, W. Zuk, The logistic curve for fitting sigmoidal data, Physiol. Chem. Phys. 13 (1981) 221-230.

[157] J.S. Pedersen, D. Dikov, J.L. Flink, H.A. Hjuler, G. Christiansen, D.E. Otzen, The changing face of glucagon fibrillation: structural polymorphism and conformational imprinting, J. Mol. Biol. 355 (2006) 501-523.

[158] L. Nielsen, R. Khurana, A. Coats, S. Frokjaer, J. Brange, S. Vyas, V.N. Uversky, A.L. Fink, Effect of environmental factors on the kinetics of insulin fibril formation: Elucidation of the molecular mechanism, Biochemistry 40 (2001) 6036-6046.

[159] Richards, F. J., A flexible growth function for empirical use, *J. Exp. Bot.* **1959**, *10*, 290-300.

[160] X.Y. Yin, J. Goudriaan, E.A. Lantinga, J. Vos, H.J. Spiertz, A flexible sigmoid function of determinate growth, Ann. Bot. 91 (2003) 361-371.

[161] C.-C. Lee, A. Nayak, A. Sethuraman, G. Belfort, G.J. McRae, A three-stage kinetic model of amyloid fibrillation, Biophys. J. 92 (2007) 3448-3458.

[162] J.R. Platt, Strong inference: certain systematic methods of scientific thinking may produce much more rapid progress than others, Science 146 (1964) 347-353.

[163] H. Kacser, J.R. Small, How many phenotypes from one genotype? The case of prion diseases, J. Theor. Biol. 182 (1996) 209-218.

Supporting Information for:

Protein Aggregation Kinetics, Mechanism, and Curve Fitting: A Review of the Literature

Aimee M. Morris, Murielle A. Watzky, and Richard G. Finke

Physical Methods: Additional Information. In Table 1 of the main text, 18 physical methods are listed. Herein, we present each of these physical methods along with sedimentation and size-exclusion chromatography in alphabetical order noting whether each method is direct or indirect,[1] in-situ or ex-situ, and discuss the advantages and disadvantages of each technique for studying protein aggregation.

Absorbance. Absorbance spectroscopy is a direct, easily employed and thus popular technique used to measure the aggregation of proteins. It can be in-situ if the concentration of the aggregating protein is within the detection limits of the absorbance experiment. If the monomer under study shows negligible absorbance in the region being measured, and if the soluble aggregates are the only absorbing species at the wavelength of interest, then this method is in principle easily applied under conditions where Beer's law is obeyed. Alternatively and under reverse absorbing conditions (i.e., where the monomer absorbs at the wavelength employed and the oligomeric species do not), absorbance spectroscopy can be used to follow the loss of monomer in the soluble fraction of an aggregating mixture (e.g., see the section on *Sedimentation*) [1]. While effectively a direct method, an underlying issue is that absorption spectroscopy is in general not able to distinguish intermediate fibrils of different sizes or shapes.

Atomic Force Microscopy (AFM). AFM is a direct, in-situ method for visualizing the surface morphology of fibrils being studied. AFM works by a flexible force-sensing cantilever which is scanned over the surface of the sample in a raster pattern to reveal

[1] We searched the literature for definitions of direct and indirect physical methods but were unable to find any. Therefore, we will use the term direct physical method to mean a method that measures a property that is directly affected by the aggregation process, while an indirect method measures a property that is only indirectly affected by the aggregation process.

atomic level resolution [2]. An advantage of AFM is the little to no sample preparation allowing for real time imaging, thereby allowing the measurement of the kinetics of aggregation to be obtained [1]. However, some disadvantages of AFM include issues of the uniformity of the sample being measured, the possibility of damage to the sample from the force being applied by the AFM measurement, the fact that this is only a surface morphology technique, and that AFM does not provide any information about the actual composition of the aggregate being measured.

Calorimetry. Calorimetry measurements obtained by detecting temperature changes vs. a reference cell can provide an in-situ, direct measurement of the thermodynamic changes accompanying protein aggregation [3]. Analysis of calorimetry data (e.g., from microtitration calorimetry) can in principle provide the full range of desired thermochemical information, the free energy, enthalpy, entropy, equilibrium constant, and heat capacity of the protein aggregation or other process being measured. Kinetic data is also possible from calorimetry, and in fact calorimetry is arguably one of the more universally applicable, in-situ methods of obtaining kinetic data [4]. In principle, it should be used more often and up-front where applicable in protein agglomeration studies.

Circular Dichroism (CD). CD is a direct, usually in-situ measurement and involves the differential absorption of left and right circularly polarized light [5]. One of the main advantages to this technique is the simple relationship between concentration and signal, as these are related by Beer's law. CD is also capable of giving information about the secondary structure of the protein [5], so that a conformational change, for example, can be followed over time. Another useful advantage of this technique is the

ability to differentiate between the formation of α-helices and β-sheets by their signature spectra [5]. The use of this method does of course require that the sample is optically active, although this is not a problem for proteins.

Dyes. The use of dyes for the detection of protein aggregation is an indirect method that is commonly employed due to the sensitivity of the technique [1]. It is usually an ex-situ methods as the concentrations of proteins used for these experiments must be increased in order to be detected by the dyes used. This technique involves the addition of a dye, with the two most common dyes being Thioflavin T (abbreviated ThT or TfT) and Congo Red (CR), to a solution of the sample under question. The result is a spectrum that shows an increase in the spectrum with the formation of more aggregates [6]. One of the primary advantages of this technique is the ability to measure the fluorescence vs. time to yield the aggregation kinetics. A disadvantage of this technique is that again different aggregate species cannot be differentiated from one another and all contribute to the fluorescence in a combined, averaged way. There also seems to be a discrepancy in what fluorescence is actually measuring as commented on by Munishkina and Fink for amyloid β and ThT, "It is tempting to speculate that ThT may bind between the beta-sheets of the fibril, however, no experimental data exist to support this assumption" [6]. For this reason, the use of dyes should be considered an indirect method, at least at present, and until needed control experiments comparing and calibrating this method to multiple other physical methods appear [7] for each dye used.

Electron Microscopy (EM). EM is an in principle powerful, albeit indirect, ex-situ, non-solution method for visualizing the surface morphology of the fibrils being studied. EM works by a focused beam of electrons being scanned across the surface of

the sample producing backscattered and secondary electrons as a way to image the sample [2]. EM techniques can prove useful for determination of the morphology of the sample, especially when used in conjunction with other, complementary physical methods. However, EM is a method fraught with issues regarding sample distribution uniformity during the grid preparation [1], so that careful microscopy necessarily involves looking at low to high resolution to ensure the homogeneity (or, likely more common, lack thereof) of the sample under visualization. Also, it is a "too sensitive" method with the ability to detect at atomic resolution and thus to see "a" species rather than "*the*" dominant species more accurately represented by Avogadro's number of molecules, for example—what we call the "Avogadro's number problem" of EM (i.e., imaging a small number of species far below the desired larger "Avogadro's number" level of statistics to avoid sampling errors). Another well-known issue is the possibility of the electron beam damaging the sample by either causing nucleation of the fibrils or changing the morphology of the fibrils already present in the sample [8]. As noted in one EM text book for transmission EM (TEM) specifically, "Certain materials are more susceptible than others, but in the end, you can damage virtually anything you put into the TEM" [8]. Thus, EM is a technique that should be used with caution and it is not an effective means for studying the kinetics of aggregation due to the long sample preparation times, its ex-situ nature, and the possibility of sampling errors.

Electron Paramagnetic Resonance (EPR) Spectroscopy. EPR has been used to monitor the kinetics of aggregation by covalently linking a spin label to the amyloid protein [9]. In this sense, the technique would be considered an indirect method for the detection of aggregates. Depending on the concentration of the sample being measured,

the technique could either be considered in-situ or ex-situ. In addition to being able to measure the kinetics of protein aggregation, EPR can also provide information on the structure of the monomeric and polymeric forms of the protein under study [9]. The disadvantages of this technique are that a spin label is usually required in order to obtain an EPR signal and higher than in vivo concentrations are usually required in order to obtain a reasonable signal.

Flow Birefringence. Flow birefringence is a hydrodynamic, bulk solution technique for determining rotational diffusion constants; this in principle direct, in-situ method involves measuring the equilibrium process of tumbling and re-orientation as a function of the flow. This in turn allows the measurement of the hydrodynamic radius [10]. Because an orientation of the protein is induced by the flow, while a re-orientation of the protein is induced by the tumbling, the degree of flow birefringence has been employed to monitor protein fibril formation [11]. The main disadvantage of this technique is it is not well understood exactly what species (i.e., monomer, oligomers, fibrils, or any combination of these species) cause the birefringence intensity and therefore, this technique cannot be used as a direct method to distinguish species in solution.

Fluorescence Spectroscopy with Intrinsic and Extrinsic Fluorophores. Fluorescence spectroscopy can be a direct, in-situ method to measure protein aggregation. It involves measuring the fluorescence intensity of either intrinsic or extrinsic (see *Dyes* above) fluorophores to gain information on conformation changes occurring in the protein [6]. This techniques is advantageous in that it is broadly applicable, requires pM-nM concentration, its sensitivity gives good signal to noise

ratios, and the kinetics of aggregation can be readily measured [6]. There are several subcatergories of fluorescence spectroscopy that are used to elucidate information on folding, membrane-protein interactions, and fibril formation, including but not limited to: time-resolved spectroscopy, Förster resonance energy transfer (FRET), fluorescence correlation spectroscopy, and fluorescence microscopy [6]. While fluorescence spectroscopy is a powerful technique for following structural changes, other complementary techniques should be used in conjunction with it to verify the interpretation of the observed fluorescence, especially if an extrinsic fluorophore or label is added.

Fourier Transform Infrared Spectroscopy (FTIR). FTIR is a direct, usually ex-situ method that can be used to detect the β-structure of proteins as well as the orientation of adjacent β-sheets to determine whether they are in a parallel or anti-parallel configuration [3]. Therefore, FTIR can give quantitative information about the larger aggregates being formed and information about the morphology of the fibrils. However, little use at least to date, has appeared of FTIR to detect intermediate oligomeric or amorphous protein aggregates.

Light Scattering. Light scattering can be considered a direct, in-situ method and has been extensively used to follow fibril size as a function of time [12]. Its extensive use is despite the lack of a simple, direct correlation between the light scattering signal and the amount of aggregated protein. If one assumes that all particles are of identical size, then the scattering intensity, R(Q), is dependent upon the particle concentration, C_p (g/l), the molar mass of the particles, M_p (g/mol), the particle form factor, P(Q), and the structure factor of the dispersion, $S(Q,C_p)$, according to the following equation:

$R(Q) \cong K \times C_p M_p P(Q) S(Q,C_p)$, where K is a constant dependent upon the instrumentation and difference in refractive index of particles and solvent [13]. This can also be thought of in a simplified way: the light scattering depends upon Brownian motion and the speed at which particles move, or the diffusion of the molecules [14]. Therefore, the molecular weight of a polymeric species can be determined by the Einstein-Stokes relation: $M = \frac{4\pi r^3 N_A}{3\upsilon}$, where M is the molecular weight of the polymer, r is the radius, N_A is Avogadro's number, and υ is the partial specific volume of the polymeric molecule [14]. This implies that diffusion is inversely proportional to the cubic root of the molecular weight [14], making correlations between the extent of light scattering measurements and fibril size problematic, especially for heterogeneous samples.

Mass Spectrometry. Mass spectrometry, in particular matrix-assisted laser desorption ionization time of flight (MALDI-TOF) mass spectrometry, is an approaching direct, ex-situ method that has been occassionally used to monitor the formation of protein complexes at specific time intervals [3]. Since MALDI-TOF has the ability to monitor species in the molecular weight range up to approximately 300,000 g/mol, both smaller oligomeric and larger, aggregated proteins can be detected [15]. One of the disadvantages of this method is that the use of MALDI-TOF typically requires the use of one of a number of buffer, salt, and co-solvent conditions. Hence, not every protein system may be analyzed using this technique [16] and, as with pretty much any of the physical methods that are not completely direct, multiple controls with different sample preparation methods and buffers, co-solvents, etc. are necessary to ensure the reliability of the observed results.

Nuclear Magnetic Resonance Spectroscopy (NMR). NMR is a powerful, direct, method that has been used in monitoring protein aggregation in both the solid-state and in solution. It is hard to use it strictly in-situ, however, since the concentrations needed for NMR (typically mM) are higher (and thus ex-situ) vs. typical concentrations of aggregating proteins (typically nM). Using solid state NMR, the amount of fibril that aggregates can be monitored and conformational changes to the fibril can be observed as well [17]. The use of solid-state NMR typically requires the use of labeled amino acid residues in order to obtain a reasonable signal [17]. Importantly, soluble intermediates will normally be missed by solid-state NMR (e.g., except when quantitative differences might imply their existence); however, solution phase NMR can be used to monitor the aggregation process of soluble species. One possible disadvantage is that at the higher concentrations required for solution NMR, proteins with a tendency to aggregate often become insoluble [18]. Due to the concentration and aggregation issues, typically, 2-D and polynuclear NMR techniques are required to monitor solution aggregation processes [18].

Quartz Crystal Oscillator Measurements. This direct, potentially in-situ technique was just described in 2007 for the measurement of kinetics of seeded aggregation of the protein insulin [19], despite quartz crystal oscillators being known since 1928 [20]. This technique employs an oscillating quartz crystal microbalance to detect the growing weight of fibrils on the surface of existing seeds. A control was conducted in the 2007 study showing no mass change is observed when the microbalance is in contact with the protein solution but no seed is present [19]. The mass change is taken as a direct measurement of aggregation [19], although this only strictly follows

after a determination of the complete composition of the deposited mass (something that is not routinely possible due to the small (ng to μg) changes of material typically detected). This technique has not yet been described for following the kinetics of unseeded growth. The main disadvantage of this otherwise sensitive method is that quartz crystal oscillator measurements detect a mass change but do not provide information about the heterogeneity or homogeneity of the sample. Perhaps the use of such microbalance methods in conjunction with mass spectrometry could alleviate this problem.

Sedimentation Studies. Sedimentation studies are used to separate soluble from insoluble proteins using centrifugation. In theory, the aggregated and thus heavier protein(s) should become separated from the monomeric protein [1]; the use of the absorbance spectrum of the soluble fraction allows one to follow the loss of the monomeric protein. However, there is evidence for the formation of soluble oligomeric intermediates that are in some cases hypothesized to be toxic [21], and this technique would not be able differentiate between monomeric and soluble oligomeric protein or between insoluble intermediates and fibrils. Sedimentation could, however, be used in conjunction with other techniques to distinguish between the species contained in the heterogeneous sample. Controls involving different centrifugation times, sample concentrations, and other conditions would seem to be typically needed, but are generally not reported, to show that the centrifugation process itself is not causing or changing any observed aggregation.

Size-Exclusion Chromatography. Size exclusion chromatography has been used as an (intrinsically ex-situ) method of separating aggregates of various sizes [22]. In

combination with one or more of the other technique(s) described above, size-separated aggregates can be detected, although it is not clear without suitable control studies (e.g., re-chromatographing the various size fractions obtained) if the separation process causes or just reports the separated sizes and. To date size exclusion chromatography appears to be most useful for separating out smaller protein oligomers subsequently detected by NMR or absorbance measurements, for example [3].

Turbidity. Turbidity directly measures aggregation in-situ by detecting all particles with a hydrodynamic radius greater than the wavelength of the incidence light [1]. The solution turbidity has been shown to be proportional to the number concentration (where the number concentration is equal to the number of entities of a mixture divided by the volume of the mixture) multiplied by the length of the aggregates in dilute solutions [23]; that is, turbidity is proportional to the total number of monomeric units in solution. Therefore, if the sample is homogeneous, then the turbidity is directly proportional to the concentration and the oligomer size can be determined [23]. Some drawbacks to this technique are: the inability to detect some intermediates if the hydrodynamic radii are smaller than the wavelength, and the inability to determine the size of the aggregates in a heterogeneous sample.

Viscosity. Viscosity is defined as the measure of resistance of a fluid to deform under shear stress [24]. As more aggregates form, the solution becomes more viscous, thereby giving a direct, in-situ measure of the degree of protein aggregation. The disadvantage of viscosity measurements is that the properties of the bulk solution are being examined rather then a direct monitoring of the species of interest present in

solution. In addition, differences between the "macroscopic" and "microscopic" viscosities are a long-standing problem [24].

X-Ray Diffraction (XRD). XRD is a direct, either in-situ or ex-situ technique (i.e., depending on how it is used) that can be used to examine fibril formation at the end of the aggregation process by examining the β-sheet structure and packing in crystalline fibril samples [3]. Due to the greater order intrinsic in crystalline samples the amount and precision of the structural information that results, XRD typically contains more information than other methods that do not examine crystalline samples. Also, the availability of synchrotron radiation sources for XRD is an advantage of this method. The main disadvantage of this method is that only crystalline samples can be examined. Therefore, many of the interesting intermediate and amorphous species cannot be examined by XRD.

In summary, at least 20 different physical methods have been used to detect protein aggregation. Greater use of direct, in-situ methods, especially methods that are capable of detecting intermediate fibril or other species, promises to be important in the future. Finally, since any physical method has its own intrinsic advantages and disadvantages, the use of multiple, complementary physical methods [7] will be an important part of future studies of protein aggregation.

Table S1. Key References from the Prior Literature of Protein Aggregation Mechanisms.

Entry	Ref.	Year	Title	Key Selected Contents
1	[11]	1959	G-F Transformation of Actin as a Fibrous Condensation	Oosawa and co-workers are pioneers in the field of protein polymerization with the recognition of actin polymerization of the general mechanism $n(\text{G-actin}) \overset{x\ Mg^{2+}}{\rightleftharpoons} (\text{F-actin})_n$, where G-actin is the globular and F-actin is the fibrous forms of actin. In 1959 they used the techniques of flow birefringence, light scattering, viscosity, and ultra-centrifugation as evidence for the fibrous condensation. In 1962, the aggregation of macromolecules with a focus on actin was proposed to aggregate by a helical rather than linear aggregation mechanism above a critical concentration of macromolecules. It was further proposed that F-actin is the helical aggregate of G-actin. The expression $\frac{dn_h}{dt} = [k_+\lambda(t) - k_-][\int [k_+'\lambda_3(t) - k_-'\lambda_{3h}(t)]dt + 3[k_+'\lambda_3(t) - k_-'\lambda_{3h}(t)] + k_-'\lambda_{3h}(t)$, is given for the rate of total monomers participating in helical polymers, where n_h is the total number concentration of monomers participating in helical polymers, k_+ and k_- are the forward and reverse rate constants for attachment and detachment of monomers, $\lambda(t)$ is the number concentration of monomers, $k_+'\lambda_3(t)$ and $k_-'\lambda_{3h}(t)$ are the forward and reverse transformation rates of ordinary trimers to nuclei of helical polymers. This expression can only be approximately solved with a number of assumptions, but is an early and important contribution to the mechanisms of protein aggregation.
	[25]	1962	A Theory of Linear and Helical Aggregations of Macromolecules	
2	[26]	1967	Self-replication and Scrapie	Griffith draws upon a mechanical model [27] to propose the mechanism shown on the right for prion replication where α' is the protein subunit stable conformation, α is the reactive conformation, α_2 is the dimeric form which the author describes as "an infective agent made of protein." Also noted in this paper is that in order for α to dimerize, there must be a catalytic influence from the α_2 molecules that are already present. This is an interesting, early, and unrecognized description of autocatalysis. This paper is also an early description of the infective agent for scrapie being a protein and this is recognized by many to be the first mechanism proposed for the infective agent made of protein, later termed prion. $\alpha' \rightleftharpoons \alpha$ $\alpha + \alpha \rightarrow \alpha_2$ $\alpha + \alpha_2 \rightarrow \alpha_3$ $\cdots$ $\alpha' + \alpha' \rightleftharpoons \alpha_2$ $\alpha' + \alpha_2 \rightarrow \alpha_3$ $\cdots$
3	[28]	1970	Molecular Weights, Association, and Frictional Resistance of Bovine Liver Glutamate Dehydrogenase at Low Concentrations. Equilibrium and Velocity Sedimentation, Light-	Eisenberg appears to have been the first to propose subsequent monomer addition as the mechanism for a glutamate dehydrogenase aggregation by drawing upon the purine stacking work of Ts'o and co-workers [30] and Van Holde and Rossetti [31], along with Adams and Lewis' work on lactoglobulin A [32]. In this case he examined the polymerization of glutamate dehydrogenase of the basic reaction, $n(\text{glutamate dehydrogenase}) \rightleftharpoons (\text{glutamate dehydrogenase})_n$. Of note is these authors' recognition of similarities between protein aggregation phenomena and the early polymerization literature. It was this recognition that allowed Eisenberg to proposed the mechanism shown on the left

			Scattering Studies, and Settling Experiments with Macroscopic Models of the Enzyme Oligomer	$P_i + P_1 \overset{K_i}{\rightleftharpoons} P_{i+1}$ this recognition that allowed Eisenberg to proposed the mechanism shown on the left for the polymerization of glutamate dehydrogenase in 1971, where P_i is the polymerized species of weight M_i, and K_i is the equilibrium constant for the reaction that is assumed to be the same for each monomer addition step.
	[29]	1971	Glutamate Dehydrogenase: Anatomy of a Regulatory Enzyme	
4	[33]	1974	Kinetics and Mechanism of Deoxyhemoglobin S Gelation: A New Approach to Understanding Sickle Cell Disease	The kinetics of deoxyhemoglobin S gelation were investigated by Hofrichter, Ross, and Eaton using temperature-jump and birefringence methods. From this data, the observation was made of a delay time prior to the onset of gelation and a sigmoidal curve. The authors also noted that the delay time had a dependence on approximately the 30th power of the concentration. We now know that this high power correlation can actually be accounted for by autocatalysis alone, which the authors did noted at the time may be part of the mechanism with the following statement, "the time course of the reaction further suggests that this mechanism must involve either a nucleation or an auto-catalytic process." They then proposed the mechanism shown on the right based on the work of McDonald on the condensation of water from supersaturated air [34]. This is again the subsequent monomer addition mechanism, this time applied deoxyhemoglobin S. The authors note that nucleation is a thermodynamically unfavorable series of bimolecular steps up to the critical nucleus, while polymerization is a thermodynamically favorable addition of monomer steps. From the proposed mechanism the following rate equation is given for the rate of formation of the critical nucleus, $\frac{dM_n}{dt} = \frac{[M]^{n-1} k_f^{n-2}(\sigma K - 1)(1-K)}{k_b^{n-8}(\sigma - 1)}$, where M is the monomer, k_f is the forward and k_b the backward rate constants, and $K = \frac{k_f[M]}{k_b}$, where K<1 and σK>1.
				Nucleation: $M + M \rightleftharpoons M_2$; $M_2 + M \rightleftharpoons M_3$; ⋮; $M_i + M \rightleftharpoons M_{i+1}$; ⋮; $M_{n-1} + M \rightleftharpoons M_n$ Polymerization: $M_n + M \rightleftharpoons M_{n+1}$; $M_{n+1} + M \rightleftharpoons M_{n+2}$; ⋮; $M_p + M \rightleftharpoons M_{p+1}$; ⋮
5	[35]	1975	Thermodynamics of the Polymerization of Protein	In 1975 Oosawa notes, "it has been shown that, generally, the polymerization of protein consists of nucleation and growth processes." Oosawa also maintains, as in entry 1, that the mechanism of aggregation occurs through the polymerization and depolymerization of monomers at the end of polymer, but the length distribution may be changed through fragmentation and association of polymers, a novel contribution at the time. He also gives an equation to account for the change in the

number of polymers with time due to fragmentation and association of polymers as, $\frac{dm}{dt} = k_+' c_h - \frac{k_-' m^2}{2}$, where m is the total number concentration of polymers, c_h is total concentration at the half polymerization time, $t_{1/2}$, and k_+' and k_-' are the rate constants for association and fragmentation, respectively. Also noted by these authors is that in the case of actin, fragmentation and association of polymers is likely to make a contribution to the overall polymerization process.

6 [36] 1975 Kinetics of the Cooperative Association of Actin to Actin Filaments

In 1975, Wegner and Engel proposed the subsequent monomer addition mechanism for actin fibril formation with the assumption that the rate constant for dimer formation would be different from other formation steps, but all subsequent rate constants would be equivalent. This proposed mechanism is shown on the left where A is an actin molecule, k_N and k_N' are the rate constants for the formation and destruction of the dimer, respectively, and k and k′ are the rate constants for the binding and dissociation of protomers in the elongation steps, respectively. The associatied differential equations are shown on the right where c_i represents the concentration of the i-mer with ∞ > i > 3. Their proposed mechanism was tested with experimental data, for the actin to actin filaments reaction, to test the ability of the mechanism to account for the experimental data. Visually poor to fair fits were observed.

$$2A \underset{k_N'}{\overset{k_N}{\rightleftharpoons}} A_2$$

$$A_2 + A \underset{k'}{\overset{k}{\rightleftharpoons}} A_3$$

$$\vdots$$

$$A_{i-1} + A \underset{k'}{\overset{k}{\rightleftharpoons}} A_i$$

$$\frac{dc_2}{dt} = k_N c_1^2 - k_N' c_2 - k c_1 c_2 + k' c_3$$

$$\frac{dc_i}{dt} = k c_1 c_{i-1} - k' c_i - k c_1 c_i + k' c_{i+1}$$

7 Mechanism of Bovine Liver Glutamate Dehydrogenase Self-Association

[37] 1975 I. Kinetic Evidence for a Random Association Mechanism

[38] 1975 II. Simulation of Relaxation Spectra for an Open Linear Polymerization Proceeding via a Sequential Addition of Monomer Units

[39] 1976 III. Characterization of the Association-Dissociation

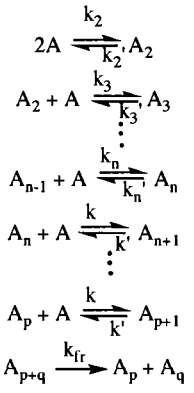
In 1975 and 1976, Eisenberg's subsequent monomer addition mechanism for glutamate dehydrogenase (entry 3) was questioned in a series of three papers. The basic driving force behind this series of papers is the notion that the subsequent monomer addition mechanism cannot account for all of the experimental data in varying concentration ranges. Due to this fact, the authors propose a new mechanism for the polymerization of glutamate dehydrogenase termed the "random association" mechanism shown to the right. The "random association" mechanism means that two polymers of any size can come together and that the rate constants for this process are the same regardless of the size of P_i and P_j. This mechanism was approximated by the authors to the "two-state" mechanism shown on the left, where A represents free association sites, B represents bonds between units, and k_a and k_d are the association and dissociation rate constants. From this simplified version, the mechanism was tested using previous aggregation data and also using quasi-elastic light scattering to measure to probe the kinetics of bovine liver glutamate dehydrogenase, although no aggregation kinetics along with a fit to the proposed mechanism are given. The authors noted that "although (their mechanism) involves more interactions than (the mechanism of sequential monomer addition only), it is the simplest kinetic model…in the sense that all elementary reactions are formally

$$P_i + P_j \rightleftharpoons P_{i+j} \qquad i, j = 1, 2, \ldots \infty$$

$$2A \underset{k_d}{\overset{k_a}{\rightleftharpoons}} B$$

			Stoichiometry with Quasi-elastic Light Scattering	accounted for in a single transformation."
8	[40]	1980	Kinetic Studies on Photolysis-Induced Gelation of Sickle Cell Hemoglobin Suggests New Mechanism	$M + M \rightleftharpoons M_2$ $M + M_2 \rightleftharpoons M_3$ . . . $M + M_{i-2} \rightleftharpoons M_{i-1}$ $M + M_{i-1} \rightleftharpoons M_i$ . . . $M + M_j \rightleftharpoons M_{j+1}$
	[41]	1985	Kinetics of Sickle Hemoglobin Polymerization	These authors propose (in 1980), an early word mechanism and then a pictorial mechanism (in 1985) for the polymerization of the polymerization of sickle hemoglobin, the basic stoichiometry of which is $n(hemoglobin - S) \rightarrow (hemoglobin - S)_n$. In 1980 the general mechanism for the polymerization process was given as shown on the left, essentially Eisenberg's monomer addition mechanism. This mechanism was examined in the classic and well studied sickle hemoglobin system; the proposed rate equations consist of double, so-called homogeneous and "heterogeneous," nucleation for the net polymerization process: $\frac{dc_p}{dt} = K_N k_+ (\gamma c)^n + K_M \phi k_+ (c_0 - c)(\gamma c)^m$ for the homogeneous nucleation and $-\frac{dc}{dt} = nK_N k_+ (\gamma c)^n + mK_M \phi k_+ (c_0 - c)(\gamma c)^m + (k_+ \gamma c - k_-)c_p$ for heterogeneous nucleation, where K_M and K_N are the equilibrium constants for "heterogeneous" and homogeneous nucleation respectively, k_+ and k_- describe the rates of monomer addition and removal, respectively, γ is the activity coefficient of the monomer, c_0 and c are the monomer concentrations at time zero and time t, and ϕ is a scaling factor for the number of effective sites for nucleation. The result, the authors note, is a five parameter model to describe the polymerization process. Also, the authors use (we infer) the term "heterogeneous nucleation" to mean the dictionary definition of heterogeneous meaning nucleation of "different" species, that is, new polymers nucleating on the surface of existing polymers. Traditionally, heterogeneous nucleation to indicate nucleation in a (typically solid) phase while the reactants are in a different (typically liquid) phase. The opinion that "heterogeneous nucleation" is actually seeded autocatalytic surface growth and other issues such as use of the term "extreme autocatalysis" are addressed in the main text.

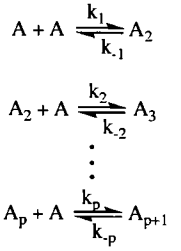
9 [42] 1982 Fragmentation of Actin Filaments



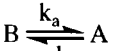
Wegner and Savko investigated the kinetics of actin polymerization by taking into account nucleation, elongation, and also fragmentation using the mechanism shown on the left where A is the actin monomer or subunit, n is the critical number of subunits, p > n, and k_{fr} is the fragmentation rate. These authors noted that "fragmentation depends strongly on the experimental conditions" and that fragmentation causes changes to the aggregation curve such as elongated lag phases and increased polymerization rates as observed in their light scattering kinetic experiments. With the assumptions that: i) longer fibrils comprise the bulk of the filaments and the concentration of short filaments is small, ii) the number of filaments is small compared to the number of monomers, iii) the concentration of nuclei is small and the monomers are mainly consumed by elongation of the filaments, and (iv) under the limiting case of $k_2' \ldots k_n' >> kc_1$ the equation on the right is given (where C is the total number concentration of filaments, c is filaments of length i, and the other variable are as defined above).

$$\frac{dC}{dt} = (kc_1 - k')c_1^n \prod_{i=2}^{i=n} \left(\frac{k_i}{k_i'}\right) + k_{fr}(c_{tot} - c_1)$$

10 [43] 1983 Polymerization of Actin and Actin-like Systems: Evaluation of the Time Course of Polymerization in Relation to the Mechanism



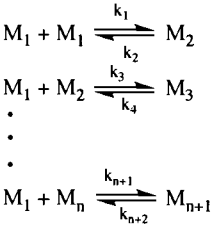
Frieden and Goddette examined the earlier mechanisms of Oosawa (see entry 1) and Wegner (see entry 6) using numerical integration techniques and the basic mechanism again of reversible subsequent monomer addition as shown to the right for the polymerization of actin. Their findings were that the Oosawa and Wegner equations were adequate for simple polymerization processes but not for more complex, realistic processes of aggregation. Thus the authors added an additional step into the subsequent monomer additon mechanism. Their additional step



desribed as is a step prior to polymerization and involves activation of the monomer. In the case of actin, this step represents ligand binding to a metal followed by a conformational change.

11 [44] 1983 On One-Dimensional Nucleation and Growth of "Living" Polymers I. Homogeneous Nucleation

[45] 1983 On One-Dimensional Nucleation and Growth of "Living" Polymers II.



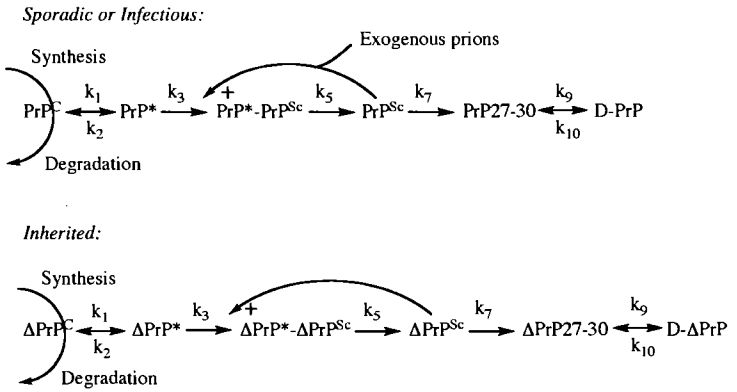
This classic set of papers investigates the one-dimensional homogeneous nucleation and growth/elongation proposing the subsequent monomer addition mechanism for "living" polymers also shown to the left with the following assumptions: i) there is only successive addition of monomers, ii) there is no geometric difference between any of the M_i, iii) $k_i = k'$ up to the critical nucleus and $k_1 = k$ past the critical nucleus, where k and k′ result from the two time constants of τ and τ', iv) a prior equilibrium exists up to the critical nucleus, and v) both reversible and irreversible growth are treated beyond the critical nucleus.

			Growth at Constant Monomer Concentration	Using the proposed general mechanism of successive monomer addition, four separate cases of initial rate, steady state, approaching equilibrium, and quasi-steady state are treated giving equations for nucleation for each of the four cases and an expression for reversible, $\frac{-dM_1}{dt} = (k_f C - k_r)^2 c_n t$, and irreversible, $-\frac{dM_1}{dt} = (k_f C)^2 c_n t$, growth are given where M_1 is the monomer, k_f and k_r are the forward and reverse rate constants, C is the monomer concentration, and c_n is the critical nucleus. These expressions for growth can be solved analytically using the assumptions of all forward and reverse rate constants being equal and if the monomer concentration is constant. Lacking herein is a test of the proposed mechanism and equations to experimental data although this is a more general treatment with good insights for one-dimensional growth.
12	[46]	1986	Cooperative Polymerization Reactions: Analytical Approximations, Numerical Examples, and Experimental Strategy	$A_1 + A_n \underset{k_-}{\overset{k_+}{\rightleftharpoons}} A_{n+1} \quad n < s$ $A_1 + A_n \underset{g_-}{\overset{g_+}{\rightleftharpoons}} A_{n+1} \quad n \geq s$ Goldstein and Stryer use the simple model shown to the left for stepwise addition and subtraction of monomers to one end of a polymer, where A_1 is the monomer, A_n is the polymer of length n, k_+ and k_- are the forward and reverse rate constants prior to the formation of the seed, s, g_+ and g_- are the rate constants after the formation of the seed, and the seed size is defined as the length s, where the kinetic rate constants change. In other words the first equation is meant to represent nucleation and the second, growth. The corresponding rate equations are then given as: $\frac{dA_n}{dt} = k_+ A_1 (A_{n-1} - A_n) + k_- (A_{n+1} - A_n)$ when n < s and $\frac{dA_n}{dt} = g_+ A_1 (A_{n-1} - A_n) + g_- (A_{n+1} - A_n)$ when n > s. The above rate equations are then transferred into a dimensionless, less easily understood form using the assumptions that i) the monomer concentration changes only by the addition or subtraction of monomer, ii) the polymer formation by seed production is irreversible, and that iii) the ratio of the concentration of the seed-minus-one-length to the (s-1) power of monomer concentration is essentially constant. The resulting equations are: $\frac{d\alpha_1}{d\tau} = C - C\alpha_1$ and $\frac{dC}{d\tau} = K\alpha_1^{s-1}(\alpha_1 - 1)$, where $\alpha_n = g_+/g_- * A_n$, $\tau = g_- t$, $C = \Sigma_{n-s+1}^{\infty} \alpha_n$, and "K is usually taken from equilibrium constants, but its exact value is immaterial for our discussion." A couple of important insights are apparent from the dimensionless form of the equation: first, the assumption of irreversible polymerization appears justified since the same simulated kinetic curves are observed with and without the approximation (except in the case of low concentration and long measurement time); second, the appearance of dimer, trimer, and tetramer species following the loss of monomers is apparent in the simulations. Lacking, however, are a test of the proposed mechanism and its corresponding dimensionless equations using experimental data. The lack of dimensional forms of the key equations have also probably contributed to the lack of use of these equations by others who might be more interested in the underlying physical picture and main

concepts behind protein aggregation.

13 [47] 1991 Molecular Biology of Prion Diseases

[48] 1994 Structural Clues to Prion Replication



Prusiner suggested the two mechanisms shown on the right for prion aggregation depending on whether the disease was sporadic or infectious vs. inherited. In this mechanism, PrP^C is the normal cellular prion, PrP* is a rare, partially unfolded monomer, PrP^{Sc} is an altered version of PrP^C that is pathogenic, PrP27-30 is a truncated form of PrP^{Sc} that has a high β-sheet content and polymerizes to form amyloids, and D-PrP is the denatured form of PrP^{Sc} and PrP27-30 that makes D-PrP protease sensitive and stops infectivity. The slow step in this mechanism is the conversion of the normal monomer to the partially unfolded monomeric state. This slow step differs from Lansbury's mechanism (see entry 14 below) in which nucleation is the slow step. While two mechanisms were given, there was not a rate equation written for either mechanism and therefore no kinetic data were tested with the suggested mechanisms.

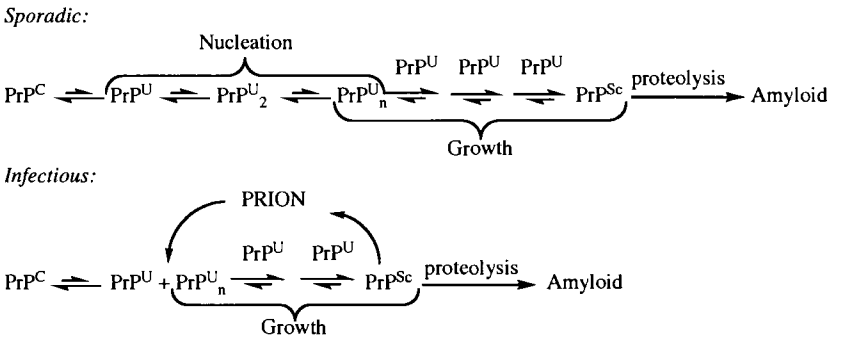
14 [49] 1993 Seeding "One-Dimensional Crystallization" of Amyloid: A Pathogenic Mechanism in Alzheimer's Disease and Scrapie?

[50] 1993 A Kinetic Model for Amyloid Formation in the Prion Diseases: Importance of Seeding

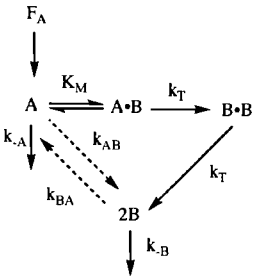
[51] 1994 Mechanism of Scrapie Replication

[52] 1995 The Chemistry of Scrapie

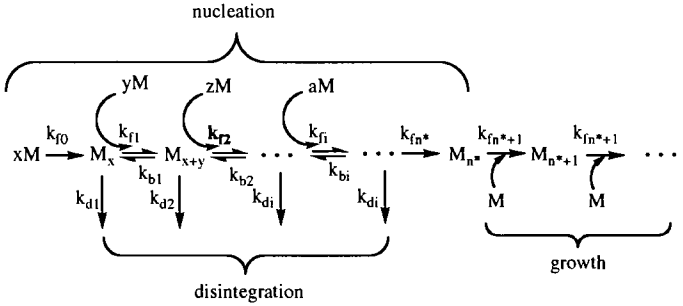
Lansbury proposed a nucleation-dependent polymerization mechanism for the self-propagating and transmittive prions that is shown on the right in two different forms, one corresponding to a sporadic form and the other an infectious form. In the proposed mechanism, PrP^C is the folded monomer form of the prion, PrP^U is the partially unfolded form of the monomer, and PrP^{SC} is protease resistant form of the prion aggregate. Lansbury's mechanism is consistent with the



			Infection: Implications of the 'ice 9' Metaphor	subsequent monomer addition mechanism first proposed in 1970 (see entry 3). In this mechanism, the nucleation is the slow step or the reaction and not the conformational change which is opposite of Prusiner's mechanism (see entry 13 above). The authors of these papers use a series of seeding experiments to explain the propagative and transmittive nature of prions. No integrated rate equation is given to test their mechanism with aggregative prion kinetic data, however.
	[53]	1995	Aggregates of Scrapie-Associated Prion Protein Induce the Cell-Free Conversion of Protease-Sensitive Prion Protein to the Protease-Resistant State	
15	[54]	1996	Prionics or the Kinetic Basis of Prion Diseases	This paper discusses the possibility of two different mechanisms for the formation of pathogenic prion species. The first of which is a condensed version of Prusiner's (see entry 13 above) linear autocatalysis mechanism shown on the right where F_A is the constant metabolic formation of A, A is the normal form of the host protein, B is the pathogenic form of A, K_M is the Michaelis constant, k_T is the turnover number, k_{-A} is the metabolic decomposition of A, k_{-B} is the metabolic decomposition of B, k_{AB} and k_{BA} are the noncatalytic conversion of A to B and B to A. From Prusiner's condensed mechanism the corresponding rate equation is obtained: $\frac{d[A]}{dt} = F_A - k_{-A}[A] - k_{AB}[A] - \frac{k_T[A]}{K_M + [A]}[B]$. Also discussed is Eigen's interpretation of Lansbury's (see entry 14 above) passive autocatalysis mechanism proposed for both the aggregation of prions and the plaque formation in Alzheimer's disease, $A \underset{k_{BA}}{\overset{k_{AB}}{\rightleftharpoons}} B \underset{k_D}{\overset{k_F[B]}{\rightleftharpoons}} B_2 \underset{k_D}{\overset{k_F[B]}{\rightleftharpoons}} \cdots \underset{k_D}{\overset{k_F[B]}{\rightleftharpoons}} B_n \underset{k_D}{\overset{k_F[B]}{\rightleftharpoons}} B_p$, where A is again the normal form of the host protein, B is the pathogenic monomeric form of A, B_2, B_n, and B_p are the aggregate forms of B, k_{AB} and k_{BA} are again the noncatalytic conversion of A to B and B to A, k_F is the formation rate constant, and k_D is the dissociation rate constant. Interestingly, the Lansbury mechanism does not necessarily require catalysis for the formation of aggregates. Neither mechanism can be ruled out in terms of the formation of aggregates of prion proteins and both represent valid possible mechanisms. Also, the author specifically states, "The aim of this paper is not a quantitative *description* of real data, but rather an *understanding* of the principles that control reality."



16 [55] 1996 Kinetics of Self-Assembling Microtubules: An "Inverse Problem" in Biochemistry

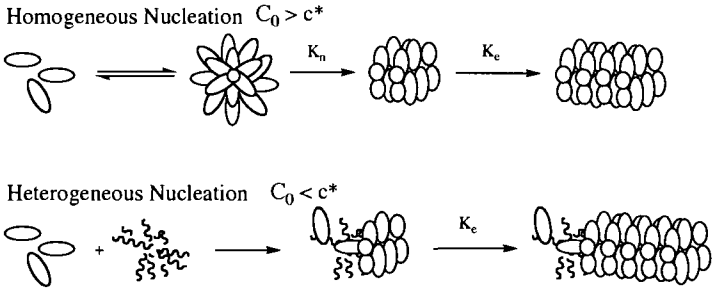


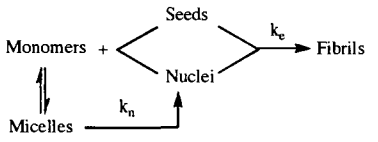
Flyvbjerg et al. develop a model based on the subsequent monomer addition mechanism as shown pictorially on the left where M is the monomer, M_{n^*} is the nucleus of size n, and k_f, k_b, and k_d are the rate constants for the forward, backward, and disintegration steps, respectively. Note that in this model the steps are reversible up to the formation of the nucleus and irreversible thereafter. This model also accounts for the possible disintegration of the intermediate aggregate species during the nucleation phase. Using the assumptions that (i) there is only one pathway for aggregation and (ii) every stage in the pathway is linked to the next stage by the addition of monomers only, the authors also developed rate equations that correspond to their model. However, the authors then applied so-called "phenomenological scaling" to the variables of time, monomer concentration, number concentration of monomers, and mass polymerized. The authors note that the scaling "only approximately satisfied by the experimental data". Four parameters were used to fit the scaled data but then it was assumed that three of the four rate constants obtained were equivalent. This assumption collapsed their four parameter model down to a two parameter model that the authors state can be seen as a generalization of Oosawa's classical nucleation-polymerization model. The proposed model appears to fit the "phenomenologically scaled" kinetic data well, but a couple of points remain troubling: (i) why use a four parameter model and collapse it down to a two parameter model? Wouldn't it be simpler to start with a two parameter model? In fact, the kinetic data without "phenomenological scaling can be fit with the 2-step Finke-Watzky (entry 38) model (Figure S1). More troubling, however, is (ii) the use of "phenomenological scaling". Is the "phenomenological scaling" necessary and justified, especially considering that the non-scaled data can be fit with a simple 2-step model? The authors admit that the scaling only approximately satisfies the experimental data while the entire model relies upon the scaling.

17 [56] 1996 On the Nucleation and Growth of Amyloid β-Protein Fibrils: Detection of Nuclei and Quantitation of Rate Constants

[12] 1997 Kinetic Theory of Fibrillogenesis of Amyloid β-Protein

Benedek and co-workers draw upon Ferrone's mechanism (see entry 8) to develop the pictorial kinetic model shown on the right for the fibrillogenesis of amyloid β that the authors state gives quantitative information on the nucleation, K_n, and elongation, K_e, rates. They use their model to obtain K_n and K_e for the aggregation of amyloid β as monitored by quasi-elastic light scattering using the following equations: $\frac{dL}{dt} = \frac{K_e C_0}{\lambda}$, $\frac{dL}{dt} = \frac{K_e c^*}{\lambda}$, and $\lambda L_f = \left(\frac{K_e c^*}{K_n}\right)^{\frac{1}{2}}$, where λ is the number of monomers per unit length of the fibril, L_f is the final length of the fibril, and C_0 is the initial concentration. The data examined, however, is admittedly on the later stages of the aggregation process. Also examined herein is the concentration dependence which showed that below the critical nucleus concentration, c*, the higher the concentration, the larger the fibrils that are produced but above c* the fibril size is independent of the concentration. What is not apparent is how the equations relate to the given pictorial model. In their later paper, the authors suggests a mathematical model for the fibrillogenesis of amyloid β that includes a nucleation step in which nuclei are produced either from micelles or seeds and irreversible elongation produced by the addition of monomers to fibril ends as shown pictorially to the left, and thenexpressed as a time derivative $N_p = ck_e N_{p-1} - ck_e N_p + k_n M\delta_{pn^*}$, where N_p is fibrils of size p, c is concentration of free amyloid β monomers, k_e is the coefficient of proportionality between attached and free c, k_n is the nucleation rate, M is the number concentration of micelles, and δ_{pn^*} was not defined but is presumed to be some sort of correction factor or scaling term related to the formation of a critical nucleus.



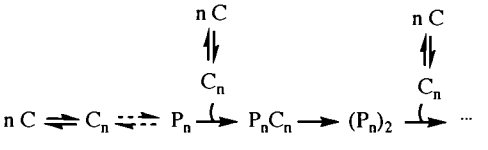


18 [57] 1996 First-Order Kinetic Model of Alzheimer's β-Amyloid Fibril Extension In Vitro

[58] 1999

A first-order kinetic model consisting of one-step, reversible, autocatalytic reaction, $[P] + [M] \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [P]$, where M is monomeric or dimeric amyloid β, P is the number concentration of seed fibrils, and k_{on} and k_{off} are the polymerization and depolymerization rate constants, respectively, was suggested for

	[59]	2002	Kinetic Analysis of Amyloid Fibril Formation Kinetic Modeling and Determination of Reaction Constants of Alzheimer's β-amyloid Fibril Extension and Dissociation Using Surface Plasmon Resonance	the fibrogenesis of β-amyloid. Also given is a derivitized form, $f'(t) = (k_{on}[M] - k_{off})[P]$, where f'(t) is the rate of amyloid β fibril extension, $k_{on}[M][P]$ is the rate of polymerization, and $k_{off}[P]$ is the rate of depolymerization. The authors' then rearrange the derivitized kinetic equation shown above to a form of the logistic equation. Lacking here is an initial step for the production of seed fibrils or the process of nucleation. In one of their later papers, the authors used the first-order kinetic model to analyze amyloid β aggregation data. Also obtained by the authors were some quantitative values; a critical monomer concentration of 20 nM and an equilibrium association constant of 5×10^7 M^{-1} for the amyloid β fibril extension were found.
19	[1]	1997	Models of Amyloid Seeding in Alzheimer's Disease and Scrapie: Mechanistic Truths and Physiological Consequences of the Time-Dependent Solubility of Amyloid Proteins	This review focuses on the aggregation of amyloid in Alzheimer's and prion in scrapie, both of which are believed to aggregate by a nucleation and growth polymerization in which seeding eliminates the lag phase. The simplest mechanism is shown as Lansbury's mechanism in entry 14 involving a series of entropically unfavorable association steps. Lacking is any mathematical equation for the treatment of kinetic data.
20	[60]	1999	Analysis of Protein Aggregation Kinetics	$A + A_n \rightleftharpoons A_{n+1}$ $n < n^*$ $A + A_n \rightleftharpoons A_{n+1}$ $n > n^*$ Ferrone worked with the subsequent monomer addition mechanism and equilibrium constants that differ significantly before (K<<1) and after (K>>1) nucleation as shown on the left. The equation below is obtained using perturbation theory along with the assumptions of i) the rate of formation of polymers is determined by the concentration of nuclei and the rate of elongation of the nuclei, ii) polymer addition all occurs by the same rate and does not depend on size, and iii) all species can be classified as either monomers or polymers: $\Delta = \frac{1}{2} JJ^* c^* t^2$, where Δ is the total concentration of monomers that have gone into polymers, J is the rate of polymer addition, J* is the rate of elongation of the nucleus, c* is the concentration of nuclei, and t is time. This equation is later applied to fit the initial portion of kinetic aggregation curves.
21	[61]	1999	α-Synuclein Fibrillogenesis is Nucleation-Dependent	These authors perform seeding kinetic experiments en-route to concluding that a "nucleation-dependent polymerization mechanism" consisting of an initial lag phase followed by a growth phase, is the kinetic mechanism by which α-synuclein aggregates. Lacking, however, is a detailed, balanced, proposed chemical equation and corresponding mechanism allowing for the quantitative determination of rate constants.

22	[17]	2000	Conformational Transitions and Fibrillation Mechanism of Human Calcitonin as Studied by High-Resolution Solid-State ^{13}C NMR	Saitô and co-workers suggest a mechanism for the fibrillation of the human calcitonin hormone with the three steps shown on the right where n_0 represents micellular intermediates of the same aggregation number, A_{n0} is the micellular aggregates made up of n_0 monomers, B_{n0} is the nucleus, and B_n and B_{n+1} are the elongated fibrils with n and n + 1 molecules of the protein. The corresponding analytical form of the mechanism is determined using the assumptions that: (i) the monomer and micelle states give the same measurable signal, (ii) the total monomer concentration is always much higher than the critical micelle concentration, and (ii) each B_n can act as a catalytic site for the conversion of A to B. The analytical equation is: $f = \frac{\rho\{\exp[(1+\rho)kt]-1\}}{\{1+\rho\exp[(1+\rho)kt]\}}$, where f is the fraction of calcitonin in the fibrillar form and ρ represents the dimensionless value to describe the ratio of k_1 to k (which is equal to $k_2*[A]_0$). The fractional integrated rate equation is then used to fit the calcitonin aggregation data under various concentration and pH conditions. n_0A (monomers) $\rightleftharpoons A_{n0}$ (micelle) $A_{n0} \xrightarrow{k_1} B_{n0}$ $A + B_n \xrightarrow{k_2} B_{n+1}$ Although not readily apparent, not recognized by the authors, and not demonstrated prior to our recent publication [62], Saitô and co-workers' above fractional form of the integrated equation for their 3-step mechanism turns out to be equivalent to the earlier, 1997 Finke-Watzky 2-step mechanism of nucleation followed by autocatalytic growth mechanism,[63] $[A]_t = \frac{\frac{k_1}{k_2}+[A]_0}{1+\frac{k_1}{k_2[A]_0}e^{(k_1+k_2[A]_0)t}}$, an important insight.
23	[7]	2000	Nucleated Conformational Conversion and the Replication of Conformational Information by a Prion Determinant	 Lindquist and co-workers suggested a fourth (molecular-level detail) mechanism for the transmission and propagation of prions using a yeast prion system as shown in equation form on the left where C represents the cellular and P the prion form of the protein. This new mechanism, which is a combination of Griffith's (entry 2) and Lansbury, Jr.'s (entry 14) mechanisms, is termed "nucleated conformational conversion." In addition to the proposed mechanism, the authors also provide a nice kinetic data analysis of prion aggregation using six different physical methods.
24	[64]	2001	Partially Folded Intermediates in α-Synuclein Fibrillation	Fink and co-workers give the following empirical formula for utilizing various slopes obtainable from the kinetic data: $Y = (y_i + m_i x) + \frac{(v_f + m_f x)}{1+e^{-\left(\frac{x-x_0}{\tau}\right)}}$, where Y is the fluorescence intensity used to measure the

aggregation, ($y_i + m_i x$) is the initial slope of the line during the lag phase, ($v_f + m_f x$) is the final slope of the line after growth has ended, χ_0 is the time at 50% intensity, and $1/\tau$ is the length of time for the lag phase. The authors comment that, "This expression is unrelated to the underlying molecular events, but provides a convenient method for comparison of the kinetics of fibrillation."

25 [65] 2002 Huntington's Disease Age-of-Onset Linked to Polyglutamine Aggregation Kinetics

[66] 2006 Kinetics and Theromodynamics of Amyloid Fibril Assembly

The authors of this paper and subsequent review suggest that the aggregation of polyglutamine, which is hypothesized to key in Huntington's disease, "has all the attributes of a nucleated growth polymerization reaction" [65]. They also suggest that the nucleation is a highly unfavorable event with a critical nucleus of a single protein. The initial portion of the kinetic data is analyzed with the following equation originally developed by Ferrone (see entry 20), $\Delta = \frac{1}{2} k_+^{\ 2} K_{n^*} c^{(n^*+2)} t^2$, where Δ is the concentration of monomers converted to polymers, k_+ is the forward elongation rate, K_{n^*} is the monomer nucleus equilibrium constant, c is the bulk concentration of monomers, n^* is the critical nucleus, and t is time. From this equation information on the convoluted rate constants, $k_+^{\ 2} K_n$, and the critical nucleus size was abstracted.

26 [67] 2003 Kinetics of Irreversible Protein Aggregation: Analysis of Extended Lumry-Eyring Models and Implications for Predicting Protein Shelf-Life

[68] 2007 Non-Native Aggregation of α-Chymotrypsinogen Occurs through Nucleation and Growth with Competing Nucleus Sizes and Negative Activation Energies

[69] 2007 A Lumry-Eyring Nucleated Polymerization Model of Protein Aggregation Kinetics: 1. Aggregation with Pre-Equilibrated Unfolding

In 2003, Roberts suggests a multi-step bimolecular nucleation and growth mechanism shown on the right for the irreversible aggregation of proteins via non-native conformations where the final product of aggregation is amorphous aggregates. Also given are mathematically sophisticated, mechanistically complex, integrated equations for four different scenarios shown on the left; the details of each scenario and the meanings of the parameters can be found in the original 2003 reference. In 2007, Roberts adapts his model to protein aggregation where the final product is ordered aggregates and gives a five part pictorial and equation mechanism supported by kinetics and thermodynamics consisting of reversible unfolding, prenucleus formation, nucleation, growth of soluble aggregates, and condensation. Also given is a simple rate equation for the loss of monomer, m, as $\frac{dm}{dt} = -k_{obs} m^{\delta}$, where δ is the number of monomers required to propagate growth and k_{obs} is a convolution of the nucleation and growth rate constant. The authors note that, "This [convoluted k_{obs}] is a common problem with all rate coefficients that are extracted from data such as monomer loss, aggregate mass conversion, overall dye binding to aggregate and turbidity." This statement is incorrect as other authors have been able to obtain rate constants that are not convoluted.

$$\text{Type IA}: r_0 \equiv -\frac{1}{M_0}\left(\frac{dM}{dt}\right)_{t\to 0} = 2k_{11}^{(app)}\left(1+\sqrt{\frac{\kappa_{ij}}{2\alpha_1}}\right)M_0$$

$$\text{Type IB}: r_0 = n^* k_{11}^{(app)} M_0$$

$$\text{Type II}: r_0 = k_1^{(app)}$$

$$\text{Type III}: r_0 = 2(1+\eta)k_{11}^{(app)} M_0 + \kappa_k^{(app)}$$

$$R_A + R_A \xrightarrow{k_{1,1}} A^{(2)}$$

$$A^{(2)} + R_A \xrightarrow{k_{1,2}} A^{(3)}$$

.

.

.

$$A^{(n^*-1)} + R_A \xrightarrow{k_{1,n^*-1}} A^{(n^*)}$$

$$A^{(2)} + A^{(2)} \xrightarrow{k_{2,2}} A^{(4)}$$

.

.

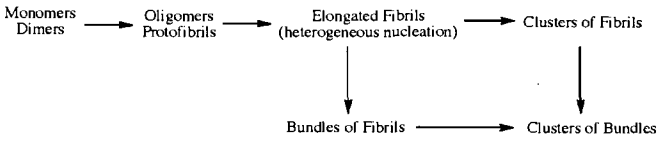
.

$$A^{(i)} + A^{(j)} \xrightarrow{k_{i,j}} A^{(i+j)}$$

27	[70]	2003	Rationalization of the Effects of mutations on Peptide and Protein Aggreation Rates	In 2003, Chiti, Dobson, and co-workers gave an equation to predict the difference in the aggregation rate of mutated vs. wild type proteins using hydrophobicity (ΔHydr.), overall charge (ΔCharge), and the propensity to convert from α-helix to β-sheet ($\Delta\Delta G_{coil-\alpha}+\Delta\Delta G_{\beta-coil}$). The equation is given as $\ln(v_{mut}/v_{wt}) = A\Delta\text{Hydr.} + B(\Delta\Delta G_{coil-\alpha} + \Delta\Delta G_{\beta-coil}) + C\Delta\text{Charge}$, where v_{mut} and v_{wt} are the rate of aggregation of mutant and wild type protein respectively, and A, B, and C are the experimentally determined dependencies of $\ln(v_{mut} / v_{wt})$ on the above three listed factors. Note that this equation predicts only the difference in the rates of a mutated vs. a wild type protein. In 2004, along with Vendruscolo, the authors were able to obtain values for the coefficients via standard regression fits and predict an absolute rate for a polypeptide chain to aggregate from previous literature for proteins of 8-127 residues, at pH 4.4-9.0, ionic strength 0.1-150 mM, and for peptide concentrations of 1×10^{-3}-2 mM, all using an expanded form of their previous equation. The expanded equation uses six properties to predict the aggregation rate: $\log(k) = \alpha_0 + \alpha_{hydr}I^{hydr} + \alpha_{pat}I^{pat} + \alpha_{ch}I^{ch} + \alpha_{pH}E^{pH} + \alpha_{ionic}E^{ionic} + \alpha_{conc}E^{conc}$, where k is the aggregation rate, α is the coefficient, I indicates an intrinsic, and E an extrinsic, property, hydr stands for hydrophobicity, pat for patterning, ch for charge, ionic for ionic strength, and conc for concentration of the polypeptide.
	[71]	2004	Prediction of the Absolute Aggregation Rates of Amyloidogenic Polypeptide Chains	
28	[72]	2003	Emerging Ideas on the Molecular Basis of Protein and Peptide Aggregation	This review focuses on the current (as of 2003) experimental and computational experiments, focusing mainly on the structural characteristics of proteins that have given insights into the mechanism of amyloid aggregation. They propose a generic nucleation and growth mechanism although the details may vary considerably from protein to protein. No actual balanced reaction equations or mechanisms are given.
29	[73]	2003	An Autocatalytic Reaction as a Model for the Kinetics of the Aggregation of β-Amyloid	These authors propose a qualitative 2-step mechanism for the autocatalytic aggregation of the protein, amyloid β (Aβ), $nM \xrightarrow{k_n} P_n$ and $M + P_n \xrightarrow{k_e} P_{n+1}$, where M is the monomeric Aβ, n is the number of Aβ molecules, and P_n is the nucleus of the fibril consisting of n molecules of Aβ. They also give an integrated rate equation of $f = \frac{\rho\{\exp[(1+\rho)kt]-1\}}{\{1+\rho\exp[(1+\rho)kt]\}}$, where f is the fraction of Aβ in the fibrillar form and ρ represents the dimensionless value to describe the ratio of k_n to the convoluted k consisting of both k_n and k_e. This exact integrated equation, however, was originally proposed in 2000 as shown above in entry 22, although credit was not given to the 2000 authors. Also in this article the units on the rate constants are not correct even though the 2000 publication that was copied gave the correct units for both k_n and k_e.

30	[18]	2004	NMR of α-Synuclein-Polyamine Complexes Elucidates the Mechanism and Kinetics of Induced Aggregation	Ligand Binding: $\text{ligand} + \text{monomer} \xleftrightarrow{K_{eq}} \text{liganded monomer}$ Nucleation: n (un)liganded monomers $\xleftrightarrow{K_{eq}}$ nucleation center Propogation: (un)liganded monomers + growth centers $\xrightarrow{k_{app}}$ aggregated state(s) Evolution: mechanical fragmentation of aggregates → more growth centers of polyamines consisting of an differential equation, $\alpha[t] = \frac{1 - e^{-k_{app}t}}{1 + ae^{-k_{app}t}}$, where α is the fractional monomer to aggregate conversion, k_{app} is a convolution of the rate constants, and a is a term related to the fractional amount of monomer active sites. Also included is a four-step word, but not balanced chemical equations, mechanism shown above. This word mechanism is not intuitively related to the proposed equation and does include the concept of autocatalysis. Again a problem here is the convolution of the nucleation and growth rate constants.	This paper suggests an "absolute minimal number of parameters" mechanism for the aggregation of α-synuclein in the presence
31	[74]	2004	Autocatalytic Self-Propagation of Misfolded Prion Protein	The authors of this paper suggest evidence in support of autocatalytic propagation of the misfolding of prion proteins supporting the "prion hypothesis" although, no balanced reaction mechanisms, or quantitative equations, not even a word mechanism, are suggested for the aggregation process.	
32	[75]	2005	Nucleation-Dependent Polymerization is an Essential Component of Amyloid-Mediated Neuronal Cell Death	These authors suggest that amyloid β aggregation occurs by a nucleation-dependent polymerization as a critical component in neuronal death; they further suggest this mechnaims will have large implications on suggested treatments for Alzheimer's disease. Lacking are any balanced equation (or even word) mechanism for the aggregation process as well as any equations capable of producing quantitative fits and rate constants.	
33	[76]	2005	Oxidative Metabolites Accelerate Alzheimer's Amyloidogenesis by a Two-Step Mechanism, Eliminating the Requirement for Nucleation	This paper provides evidence that oxidative metabolites accelerate the aggregation of amyloid β and suggest at least a 2-step, seeded-type mechanism for the acceleration process in which the first step modifies the protein and the second is the formation of spherical aggregates via a downhill polymerization. Neither step includes nucleation and neither step could explain the accelerative process of the metabolites. Lacking is a balanced reaction mechanism and quantitative equation for curve-fitting and the determination of rate constants.	
34	[77]	2006	Inhibition of Insulin Fibrillogenesis with Targeted Peptides	Gibson and Murphy apply a simple two step model shown on the right for insulin aggregation, where N is the nonamyloidogenic species, A is the amyloidogenic species, and k_s and k_f are the rate constants. From this two step model, the following integrated equation was used to fit experimental insulin aggregation kinetic data:	$N \xrightarrow{k_s} A$ $N + A \xrightarrow{k_f} 2A$

$A = \frac{k_s \cdot N_i (e^{t(k_s + k_f \cdot N_i)} - 1)}{k_s \cdot (e^{t(k_s + k_f \cdot N_i)} + k_f \cdot N_i}$. This equation is mathematically equivalent to both Saitô's 2000 equation for human calcitonin (entry 22) and the even earlier, 1997 F-W model (entry 38). As such, it shows good fits to the experimental data, although the data are rather noisy and limited.

35	[78]	2006	The Kinetics of Nucleated Polymerizations at High Concentrations: Amyloid Fibril Formation Near and Above the "Supercritical Concentration"	Powers and Powers suggest that three different concentration regimes of low, medium, and high concentrations exist within the subsequent monomer addition model. The scheme and equations shown in Figure 12 of the main text was proposed and test-case simulations were used to test the validity of the model. Later, the authors added an off-pathway aggregation step to their model and tested the updated model using simulations of test case values as well as experimentally obtained concentrations and parameters.
	[79]	2008	Mechanisms of Protein Fibril Formation: Nucleated Polymerization with Competing Off-Pathway Aggregation	
36	[80]	2007	Kinetics of Different Processes in Human Insulin Amyloid Formation	The general pattern of aggregation formation for insulin is suggested to be as shown on the right, eventually resulting in the precipitation of gelatin. This is the same general model that Ferrone suggested in 1985 for sickle hemoglobin (see entry 8). 
37	[19]	2007	Kinetics and Thermodynamics of Amyloid Formation from Direct Measurements of Fluctuations in Fibril Mass	This paper demonstrates the ability of quartz crystal oscillators to determine the quantitative kinetics and thermodynamics of amyloid protein aggregation. It should be noted that this is a new technique described to measure protein aggregation and as such differs from the other entries as it does not suggest any mechanisms or equations to fit protein aggregation data.

38	[62]	2008	Fitting Neurological Protein Aggregation Kinetic Data via a 2-Step, Minimal/ "Ockham's Razor" Model: The Finke-Watzky Mechanism of Nucleation Followed by Autocatalytic Surface Growth	$A \xrightarrow{k_1} B$ $A + B \xrightarrow{k_2} 2B$ This paper employs the Finke-Watzky (F-W) mechanism, originally developed in 1997 for transition-metal nanoclusters nucleation an autocatalytic growth, as shown on the left where A represents the unfolded or misfolded, monomeric form of the protein, B is the protein aggregate of all different sizes, k_1 is the nucleation rate constant, and k_2 is the growth rate constant. This 2-step, minimal, kinetic model was used to fit a broad range of protein aggregation data relevant to the neurodegenerative disorders including Parkinson's, Alzheimer's, and Huntington's, and prion diseases. The analytic form of the equation corresponding to the 2-step mechanism is shown on the right. All 41 representative kinetic aggregation data sets that were analyzed with the F-W 2-step model were fit with good to excellent coefficients of determination (R^2). From this minimalistic F-W model, quantitative rate constants corresponding separately to nucleation and growth were obtained in each of the 41 representative cases analyzed. $[B]_t = [A]_0 - \frac{\frac{k_1}{k_2} + [A]_0}{1 + \frac{k_1}{k_2[A]_0}\exp(k_1 + k_2[A]_0)t}$
	[81]	2008	Fitting Yeast and Mammalian Prion Aggregation Kinetic Data with the Finke-Watzky 2-Step Model of Nucleation and Autocatalytic Growth	

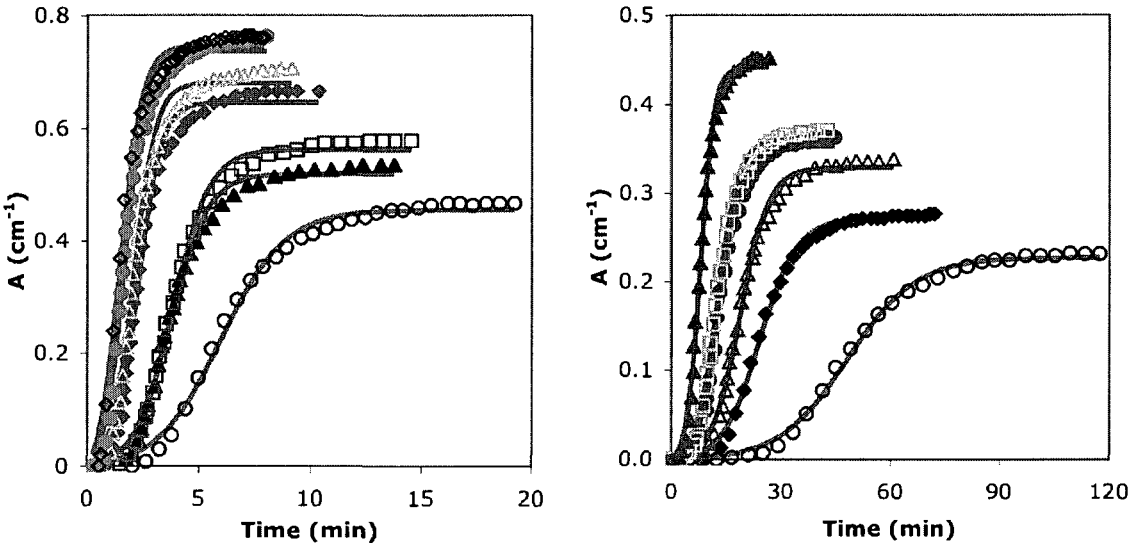


Figure S1. Digitized data (data points) of tubulin aggregation at various concentrations of tubulin from ref. [55]. This data was fit (lines) using the F-W model to illustrate that a 2-step model s capable of fitting these kinetic data sets and a four parameter model [55] is not required.

Scheme S1. The mathematical equivalence of De Levie's 1983 rate equation for irreversible growth and Ferrone's 1999 rate equation.

(I) ***De Levie's 1983 treatment*** [82]**:**

(a) De Levie's assumptions / approximations:

- The fibers grow from their ends only;
- The concentration of nuclei is constant during growth;
- The monomer concentration is mostly constant at short times.

(b) De Levie's rate equation (for irreversible growth):

(1) $\frac{-dM}{dt} = (k_f C)^2 c_n t$

$\frac{-dM}{dt}$ = monomer uptake

k_f = forward rate constant

C = monomer concentration

c_n = critical nucleus concentration

(c) De Levie's integrated rate equation (for irreversible growth) *using approximations at short reaction times*:

(2) $-\Delta M = \frac{1}{2}(k_f C)^2 c_n t^2$ - ΔM = monomer consumed during polymer formation

(II) ***Ferrone's 1999 treatment*** [83]**:**

(a) Ferrone's assumptions / approximations:

- The fibers grow from their ends only;
- The concentration of nuclei is constant during growth;
- The monomer uptake occurs mostly through growth (not nucleation);
- The monomer concentration is mostly constant at short times.

(b) Ferrone's rate equations:

(3) $\frac{d\Delta}{dt} = (k_+ c) c_p$ Δ = concentration of polymerized monomers

k_+ = elongation rate constant

c = concentration of (free) monomer

c_p = concentration of polymers

(4) $\frac{dc_p}{dt} = (k_+ c) c^*$ c* = concentration of nuclei

(c) Ferrone's integrated rate equations *using approximations at short reaction times*:

(5) $c_p = (k_+ c) c^* t$

(6) $\Delta = \frac{1}{2}(k_+ c)^2 c^* t^2$

(III) Equivalence between De Levie's and Ferrone's treatments:

It can be seen that De Levie's equation (3) and Ferrone's equation (6) above, are equivalent by making the following substitutions:

Ferrone's notation ⇒ De Levie's notation

$\Delta = -\Delta M$	(concentration of polymerized monomer)
$k_+ = k_f$	(elongation rate constant)
$c = C$	(monomer concentration)
$c^* = c_n$	(nuclei concentration)

[1] J.D. Harper, P.T. Lansbury, Jr., Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins, Annu. Rev. Biochem. 66 (1997) 385-407.

[2] D.A. Skoog, F.J. Holler, T.A. Nieman, Surface characterization by spectroscopy and microscopy, in: Principles of Instrumental Analysis, 5th Ed., Harcourt Brace, Philadelphia, 1998, pp. 535-562.

[3] S.E. Bondos, Methods for measuring protein aggregation, Current Analytical Chemistry 2 (2006) 157-170.

[4] C. LeBlond, J. Wang, R.D. Larsen, C.J. Orella, A.L. Forman, R.N. Landau, J. Laquidara, J.R. Sowa, Jr., D.G. Blackmond, Y.-K. Sun, Reaction calorimetry as an in-situ kinetic tool for characterizing complex reactions, Thermochimica Acta 289 (1996) 189-207.

[5] R.W. Woody, Theory of circular dichroism of proteins, in: G.D. Fasman, (Ed.), Theory of Circular Dichroism of Proteins in Circular Dichroism and the Conformational Analysis of Biomolecules, Plenum Press, New York, 1996, pp. 25-68.

[6] L.A. Munishkina, A.L. Fink, Fluorescence as a method to reveal structures and membrane-interactions of amyloidogenic proteins, Biochim. Biophys. Acta 1768 (2007) 1862-1885.

[7] T.R. Serio, A.G. Cashikar, A.S. Kowal, G.J. Sawicki, J.J. Moslehi, L. Serpell, M.F. Arnsdorf, S.L. Lindquist, Nucleated conformational conversion and the replication of conformational information by a prion determinant, Science 289 (2000) 1317-1321.

[8] D.B. Williams, C.B. Carter, Inelastic scattering and beam damage, in: Transmission Electron Microscopy, Plenum Press, New York, 1996, Chapter 4 pp. 49-65.

[9] K.M. Lundberg, C.J. Stenland, F.E. Cohen, S.B. Prusiner, and G.L. Millhauser, Kinetics and mechanism of amyloid formation by the prion protein H1 peptide as determined by time-dependent ESR, Chemistry & Biology 4 (1997) 345-355.

[10] M. Kasai, F. Oosawa, Flow birefringence, Meth. Enzym. 26 (1972) 289-323.

[11] F. Oosawa, S. Asakura, K. Hotta, I. Nobuhisa, T. Ooi, G-F transformation of actin as a fibrous condensation, J. Polym. Sci. 37 (1959) 323-336.

[12] A. Lomakin, D.B. Teplow, D.A. Kirschner, G.B. Benedek, Kinetic theory of fibrillogenesis of amyloid β-protein, Proc. Natl. Acad. Sci. USA 94 (1997) 7942-7947.

[13] S.P.F.M. Roefs, K.G. De Kruif, A model for the denaturation and aggregation of beta-lactoglobulin, Eur. J. Biochem. 226 (1994) 883-889.

[14] S.F. Sun, Diffusion, Light scattering, in: Physical Chemistry of Macromolecules: Basic Principles and Issues, 2nd Ed., Wiley and Sons, New York, 2004, pp. 223-242.

[15] M. Moniatte, F.G. van der Goot, J.T. Buckley, F. Pattus, A. van Dorsselaer, Characterisation of the heptameric pore-forming complex of the *Aeromonas* toxin aerolysin using MALDI-TOF mass spectrometry, FEBS Lett. 384 (1996) 269-272.

[16] W.J. Henzel, J.T. Stults, Matrix-assisted laser desorption/ionization time-of-flight mass analysis of peptides, in: J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, P.T. Wingfield (Eds.), Current Protocols in Protein Science, Wiley and Sons, New York, 2004, Unit 16.2.

[17] M. Kamihira, A. Naito, S. Tuzi, A. Nosaka, H. Saitô, Conformational transitions and fibrillation mechanism of human calcitonin as studied by high-resolution solid-state ^{13}C NMR, Protein Sci. 9 (2000) 867-877.

[18] C.O. Fernandéz, W. Hoyer, M. Zweckstetter, E.A. Jares-Erijman, V. Subramaniam, C. Griesinger, T.M. Jovin, NMR of alpha-synuclein-polyamine complexes elucidates the mechanism and kinetics of induced aggregation, EMBO J. 23 (2004) 2039-2046.

[19] T.P.J. Knowles, W. Shu, G.L. Devlin, S. Meehan, S. Auer, C.M. Dobson, M.E. Welland, Kinetics and thermodynamics of amyloid formation from direct measurements of flucuations in fibril mass, Proc. Natl. Acad. Sci. USA 104 (2007) 10016-10021.

[20] R.E. Hitchcock, Piezoelectric frequency control, Elec. J. 25 (1928) 503-503.

[21] (a) S.T. Ferreira, M.N.N. Vieira, F.G. De Felice, Soluble protein oligomers as emerging toxins in Alzheimer's and other amyloid diseases, IUBMB Life 59 (2007) 332-345 and references therein. (b) D.M. Walsh, D.J. Selkoe, Aβ oligomers—a decade of discovery, J. Neurochem. 101 (2007) 1172-1184. (c) J.D. Harper, S.S. Wong, C.M. Lieber, P.T. Lansbury, Jr., Assembly of Aβ amyloid protofibrils. An in vitro model for a possible early event in Alzheimer's disease, Biochemistry 38 (1999) 8972-8980. (d) M.J. Volles, P.T. Lansbury, Jr., Zeroing in on the pathogenic form of α-synuclein and its mechanism of neurotoxicity in Parkinson's disease, Biochemistry 42 (2003) 7871-7878. (e) B.A. Chromy, R.J. Nowak, M.P. Lambert, K.L. Viola, L. Chang, P.T. Velasco, B.W. Jones, S.J. Fernandez, P.N. Lacor, P. Horowitz, C.E. Finch, G.A. Krafft, W.L. Klein, Self-assembly of Aβ(1-42) into globular neurotoxins, Biochemistry 42 (2003) 12749-12760. (f) J. Kaylor, N. Bodner, S. Edridge, G. Yamin, D.-P. Hong, A.L. Fink, Characterization of oligomeric intermediates in α-synuclein fibrillation. FRET studies of Y125W/Y133F/Y136F α-synuclein, J. Mol. Biol. 353 (2005) 357-372. (g) B.J. Tabner, O.M.A. El-Agnaf, S. Turnbull, M.J. German, K.E. Paleologou, Y. Hayashi, L.J. Cooper, N.J. Fullwood, D. Allsop, Hydrogen peroxide is generated during the very early stages of

aggregation of the amyloid peptides implicated in Alzheimer's disease and familial British dementia, J. Biol. Chem. 280 (2005) 35789-35792. (h) O.M.A. El-Agnaf, S.A. Salem, K.E. Paleologou, M.D. Curran, M.J. Gibson, J.A. Court, M.G. Schlossmacher, D. Allsop, Detection of oligomeric forms of α-synuclein protein in human plasma as a potential biomarker for Parkinson's disease, FASEB J. 20 (2006) 419-425.

[22] K. Baussay, C. Le Bon, T. Nicolai, D. Durand, J.P. Busnel, Influence of the ionic strength on the heat-induced aggregation of the globular protein beta-lactoglobulin at pH 7, Int. J. Biol. Macromol. 34 (2004) 21-28.

[23] B.J. Berne, Interpretation of the light scattering from long rods, J. Mol. Biol. 89 (1974) 755-758.

[24] S.E. Harding, The intrinsic viscosity of biological macromolecules. Progress in measurement, interpretation and application to structure in dilute solution, Prog. Biophys. Mol. Biol. 68 (1997) 207-262.

[25] F. Oosawa, M. Kasai, A theory of linear and helical aggregations of macromolecules, J. Mol. Biol. 4 (1962) 10-21.

[26] J.S. Griffith, Self-replication and scrapie, Nature 215 (1967) 1043-1044.

[27] L.S. Penrose, R. Penrose, Self-reproducing analogue, Nature 179 (1957) 1183-1183.

[28] E. Reisler, J. Pouyet, H. Eisenberg, Molecular weights, association, and frictional resistance of bovine liver glutamate dehydrogenase at low concentrations. Equilibrium and velocity sedimentation, light-scattering studies, and settling experiments with macroscopic models of the enzyme oligomer, Biochemistry 9 (1970) 3095-3102.

[29] H. Eisenberg, Glutamate dehydrogenase: anatomy of a regulatory enzyme, Acc. Chem. Res. 4 (1971) 379-385.

[30] P.O.P. Ts'o, I.S. Melvin, A.C. Olson, Interaction and association of bases and nucleosides in aqueous solutions, J. Am. Chem. Soc. 85 (1963) 1289-1296.

[31] K.E. Van Holde, G.P. Rossetti, A sedimentation equilibrium study of the association of purine in aqueous solutions, Biochemistry 6 (1967) 2189-2194.

[32] E.T. Adams, Jr., M.S. Lewis, Sedimentation equilibrium in reacting systems. VI. Some applications to indefinite self-association. Studies with β-lactoglobulin A, Biochemistry 7 (1968) 1044-1053.

[33] J. Hofrichter, P.D. Ross, W.A. Eaton, Kinetics and mechanism of deoxyhemoglobin S gelation: a new approach to understanding sickle cell disease, Proc. Natl. Acad. Sci. USA 71 (1974) 4864-4868.

[34] J.E. McDonald, Homogeneous nucleation of vapor condensation. 2. Kinetic aspects, Amer. J. Phys. 31 (1963) 31-41.

[35] F. Oosawa, S. Asakura, Kinetics of polymerization 4. Fragmentation and association of polymers, in: B. Horecker, N.O. Kaplan, J. Marmur, H.A. Scheraga (Eds.), Thermodynamics of the Polymerization of Protein, Academic Press, New York, 1975, p. 55.

[36] A. Wegner, J. Engel, Kinetics of the cooperative association of actin to actin filaments, Biophys. Chem. 3 (1975) 215-225.

[37] D. Thusius, Mechanism of bovine liver glutamate dehydrogenase self-assembly: II. Simulation of relaxation spectra for an open linear polymerization proceeding via a sequential addition of monomer units, J. Mol. Biol. 94 (1975) 367-383.

[38] D. Thusius, P. Dessen, J.-M. Jallon, Mechanism of bovine liver glutamate dehydrogenase self-association I. Kinetic evidence for a random association of polymer chains, J. Mol. Biol. 92 (1975) 413-432.

[39] M. Jullien, D. Thusius, Mechanism of bovine liver glutamate dehdrogenase self-assembly III. Characterization of the association-dissociation stoichiometry with quasi-elastic light scattering, J. Mol. Biol. 101 (1976) 397-416.

[40] F.A. Ferrone, J. Hofrichter, H.R. Sunshine, W.A. Eaton, Kinetic studies on photolysis-induced gelation of sickle cell hemoglobin suggest a new mechanism, Biophys. J. 32 (1980) 361-377.

[41] F.A. Ferrone, J. Hofrichter, W.A. Eaton, Kinetics of Sickle Hemoglobin Polymerization, J. Mol. Biol. 183 (1985) 611-631.

[42] A. Wegner, P. Savko, Fragmentation of actin filaments, Biochemistry 21 (1982) 1909-1913.

[43] C. Frieden, D.W. Goddette, Polymerization of actin and actin-like systems: evaluation of the time course of polymerization in relation to the mechanism, Biochemistry 22 (1983) 5836-5843.

[44] M.P. Firestone, R. De Levie, S.K. Rangarajan, On one-dimensional nucleation and growth of "living" polymers I. Homogeneous nucleation, J. Theor. Biol. 104 (1983) 553-570.

[45] S.K. Rangarajan, R. De Levie, On one-dimensional nucleation and growth of "living" polymers II. Growth at constant monomer concentration, J. Theor. Biol. 104 (1983) 553-570.

[46] R.F. Goldstein, L. Stryer, Cooperative polymerization reactions: analytical approximations, numerical examples, and experimental strategy, Biophys. J. 50 (1986) 583-599.

[47] S.B. Prusiner, Molecular biology of prion diseases, Science 252 (1991) 1515-1522.

[48] F.E. Cohen, K.-M. Pan, Z. Huang, M. Baldwin, R.J. Fletterick, S.B. Prusiner, Structural clues to prion replication, Science 264 (1994) 530-531.

[49] J.T. Jarrett, P.T. Lansbury, Jr., Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie?, Cell 73 (1993) 1055-1058.

[50] J.H. Come, P.E. Fraser, P.T. Lansbury, Jr., A kinetic model for amyloid formation in prion diseases: importance of seeding, Proc. Natl. Acad. Sci. USA 90 (1993) 5959-5963.

[51] P.T. Lansbury, Mechanism of scrapie replication, Science 265 (1994) 1510-1510.

[52] P. T. Lansbury, Jr., B. Caughey, The chemistry of scrapie infection: implications of the 'ice 9' metaphor, Chemistry and Biology 2 (1995) 1-5.

[53] B. Caughey, D.A. Kocisko, G.J. Raymond, P.T. Lansbury, Jr., Aggregates of scrapie-associated prion protein induce the cell-free conversion of protease-sensitive prion protein to the protease-resistant state, Chemistry and Biology 2 (1995) 807-817.

[54] M. Eigen, Prionics or the kinetic basis of prion diseases, Biophys. Chem. 63 (1996) A1-A18.

[55] H. Flyvbjerg, E. Jobs, S. Leibler, Kinetics of self-assembling microtubules: an "inverse problem" in biochemistry, Proc. Natl. Acad. Sci. USA 93 (1996) 5975-5979.

[56] A. Lomakin, D.S. Chung, G.B. Benedek, D.A. Kirschner, D.B. Teplow, On the nucleation and growth of amyloid β-protein fibrils: detection of nuclei and quantitation of rate constants, Proc. Natl. Acad. Sci. USA 93 (1996) 1125-1129.

[57] H. Naiki, K. Nakakuki, First-order kinetic model of Alzheimer's β-amyloid fibril extension in vitro, Lab. Invest. 74 (1996) 374-383.

[58] H. Naiki, F. Gejyo, Kinetic analysis of amyloid fibril formation, Methods Enzymol. 309 (1999) 305-318.

[59] K. Hasegawa, K. Ono, M. Yamada, H. Naiki, Kinetic modeling and determination of reaction constants of Alzheimer's β-amyloid fibril extension and dissociation using surface plasmon resonance, Biochemistry 41 (2002) 13489-13498.

[60] F.A. Ferrone, Analysis of protein aggregation kinetics, Methods Enzymol. 309 (1999) 256-274.

[61] S.J. Wood, J. Wypych, S. Steavenson, J.-C. Louis, M. Citron, A.L. Biere, α-Synuclein fibrillogenesis is nucleation-dependent, J. Biol. Chem. 274 (1999) 19509-19512.

[62] A.M. Morris, M.A. Watzky, J.N. Agar, R.G. Finke, Fitting neurological protein aggregation kinetic data via a 2-step minimal/"Ockham's razor" model: the Finke-Watzky mechanism of nucleation followed by autocatalytic surface growth, Biochemistry 47 (2008) 2413-2427.

[63] M.A. Watzky, R.G. Finke, Transition metal nanocluster formation kinetic and mechanistic studies. A new mechanism when hydrogen is the reductant: slow, continuous nucleation and fast autocatalytic surface growth, J. Am. Chem. Soc. 119 (1997) 10382-10400.

[64] V.N. Uversky, J. Li, A.L. Fink, Evidence for a partially folded intermediate in α-synuclein fibril formation, J. Biol. Chem. 276 (2001) 10737-10744.

[65] S. Chen, F.A. Ferrone, R. Wetzel, Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation, Proc. Natl. Acad. Sci. 99 (2002) 11884-11889.

[66] R. Wetzel, Kinetics and thermodynamics of amyloid fibril assembly, Acc. Chem. Res. 39 (2006) 671-679.

[67] C.J. Roberts, Kinetics of irreversible protein aggregation: analysis of extended Lumry-Eyring models and implications for predicting protein shelf life, J. Phys. Chem. B 107 (2003) 1194-1207.

[68] J.M. Andrews, C.J. Roberts, Non-native aggregation of α-chymotrypsinogen occurs through nucleation and growth with competing nucleus sizes and negative activation energies, Biochemistry 46 (2007) 7558-7571.

[69] J.M. Andrews, C.J. Roberts, A Lumry-Eyring nucleated polymerization model of protein aggregation kinetics: 1. Aggregation with pre-equilibrated unfolding, J. Phys. Chem. B 111 (2007) 7897-7913.

[70] F. Chiti, M. Stefani, N. Taddei, G. Ramponi, C.M. Dobson, Rationalization of the effects of mutations on peptide and protein aggregation rates, Nature 424 (2003) 805-808.

[71] K.F. DuBay, A.P. Pawar, F. Chiti, J. Zurdo, C.M. Dobson, M. Vendruscolo, Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains, J. Mol. Biol. 341 (2004) 1317-1326.

[72] D. Thirumalai, D.K. Klimov, R.I. Dima, Emerging ideas on the molecular basis of protein and peptide aggregation, Curr. Opin. Struct. Biol. 13 (2003) 146-159.

[73] R. Sabaté, M. Gallardo, J. Estelrich, An autocatalytic reaction as a model for the kinetics of the aggregation of β-amyloid, Biopolymers 71 (2003) 190-195.

[74] J. Bieschke, P. Weber, N. Sarafoff, M. Beekes, A. Giese, H. Kretzschmar, Autocatalytic self-propagation of misfolded prion protein, Proc. Natl. Acad. Sci. 101 (2004) 12207-12211.

[75] M. Wogulis, S. Wright, D. Cunningham, T. Chilcote, K. Powell, R.E. Rydel, Nucleation-dependent polymerization is an essential component of amyloid-mediated neuronal cell death, J. Neurosci. 25 (2005) 1071-1080.

[76] J. Bieschke, Q. Zhang, E.T. Powers, R.A. Lerner, J.W. Kelly, Oxidative metabolites accelerate Alzheimer's amyloidogenesis by a two-step mechanism, eliminating the requirement for nucleation, Biochemistry 44 (2005) 4977-4983.

[77] T.J. Gibson, R.M. Murphy, Inhibition of insulin fibrillogenesis with targeted peptides, Protein Sci. (2006) 1133-1141.

[78] E.T. Powers, D.L. Powers, The kinetics of nucleated polymerizations at high concentrations: amyloid fibril formation near and above the "supercritical concentration", Biophys. J. 91 (2006) 122-132.

[79] E.T. Powers, D.L. Powers, Mechanism of protein fibril formation: nucleated polymerization with competing off-pathway aggregation, Biophys. J. 94 (2008) 379-391.

[80] M. Mauro, E.F. Craparo, A. Podestá, D. Bulone, R. Carrotta, V. Martorana, G. Tiana, P.L. San Biagio, Kinetics of different processes in human insulin amyloid formation, J. Mol. Biol. 366 (2007) 258-274.

[81] M.A. Watzky, A.M. Morris, E.D. Ross, R.G. Finke, Fitting yeast and mammalian prion aggregation kinetic data with the Finke-Watzky 2-step model of nucleation and autocatalytic growth, Biochemistry, submitted.

[82] M.P. Firestone, R. De Levie, S.K. Rangarajan, On one-dimensional nucleation and growth of "living" polymers I. Homogeneous nucleation, J. Theor. Biol. 104 (1983) 553-570.

[83] F.A. Ferrone, Analysis of protein aggregation kinetics, Methods Enzymol. 309 (1999) 256-274.

CHAPTER III

FITTING NEUROLOGICAL PROTEIN AGGREGATION KINETIC DATA VIA A 2-STEP, MINIMAL/"OCKHAM'S RAZOR" MODEL: THE FINKE-WATZKY MECHANISM OF NUCLEATION FOLLOWED BY AUTOCATALYTIC SURFACE GROWTH

This dissertation chapter contains the manuscript of a full article published in *Biochemistry* **2008**, *47*, 2413-2427. This chapter presents the fits of 14 representative protein aggregation kinetic data sets relevant to neurodegenerative disorders to the Finke-Watzky (F-W) 2-step mechanism of nucleation and autocatalytic growth originally developed for transition-metal nanoclusters. The F-W mechanism is able to account for each data set analyzed and to deconvolute quantitative rate constants for the separate processes of nucleation and autocatalytic growth.

The manuscript was prepared by Aimee M. Morris with the helpful insights, critical analysis, and mathematical equivalence derivation by Dr. Murielle A. Watzky, along with the assistance and editing of Prof. Richard G. Finke and also helpful comments from Prof. Jeffrey N. Agar.

Fitting Neurological Protein Aggregation Kinetic Data via a 2-Step, Minimal / "Ockham's Razor" Model: the Finke-Watzky Mechanism of Nucleation Followed by Autocatalytic Surface Growth

Aimee M. Morris, Murielle A. Watzky, Jeffrey N. Agar, and Richard G. Finke

Abstract

The aggregation of proteins has been hypothesized to be an underlying cause of many neurological disorders including Alzheimer's, Parkinson's, and Huntington's diseases; protein aggregation is also important to normal life function in cases such as G to F-actin, regulation of glutamate dehydrogenase, and microtubule and flagella formation. For this reason, the underlying mechanism of protein aggregation, and accompanying kinetic models for protein nucleation and growth (growth also being called elongation, polymerization, or fibrillation in the literature), have been investigated for more than 50 years. As a way to concisely present the key prior literature in the protein aggregation area, Table 1 in the main text summarizes 23 papers by 10 groups of authors that provide five basic classes of mechanisms for protein aggregation over the period from 1959 to 2007. However, and despite this major prior effort, still lacking are both (i) anything approaching a consensus mechanism (or mechanisms), and (ii) a

generally useful, and thus widely used, simplest / "Ockham's razor" kinetic model and associated equations that can be routinely employed to analyze a broader range of protein aggregation kinetic data. Herein we demonstrate that the 1997 Finke-Watzky (F-W) 2-step mechanism of slow continuous nucleation, $A \rightarrow B$ (rate constant k_1), followed by typically fast, autocatalytic surface growth, $A + B \rightarrow 2B$ (rate constant k_2), is able to quantitatively account for the kinetic curves from all 14 representative data sets of neurological protein aggregation found by a literature search (the prion literature was largely excluded for the purposes of this study in order provide some limit to the resultant literature that was covered). The F-W model is able to deconvolute the desired nucleation, k_1, and growth, k_2, rate constants from those 14 data sets obtained by four different physical methods, for three different proteins, and in nine different labs. The fits are generally good, and in many cases excellent, with R^2 values ≥ 0.98 in all cases. As such, this contribution is the current record of the widest set of protein aggregation data best fit by what is also the simplest model offered to date. Also provided is the mathematical connection between the 1997 F-W 2-step mechanism and Saitô's 3-step mechanism proposed by Saitô and coworkers. In particular, the kinetic equation for the 2000 3-step mechanism is shown to be mathematically identical to the earlier 1997 F-W 2-step mechanism and under the 3 simplifying assumptions Saitô and coworkers used to derive their kinetic equation. A list of the 3 main caveats / limitations of the F-W kinetic model is provided, followed by the main conclusions from this study as well as some needed future experiments.

Introduction

Protein aggregation is known to be intimately associated with many different types of neurological disorders, including Alzheimer's, Parkinson's and Huntington's diseases (1,2,3). Understanding the underlying mechanism(s) of protein aggregation is, therefore, central to controlling and otherwise rationally dealing with this important aspect of the etiology of these neurological disorders. Indeed, protein aggregation has been postulated to be the underlying cause of such diseases (4,5,6,7,8,9,10,11,12,13,14), although cause and effect are still not unequivocally demonstrated and, therefore, remain controversial. Naturally occurring, non-disease state protein aggregation is also a part of normal life function in cases such as G to F-actin conversion (15,16,17,18), regulation of glutamate dehydrogenase (19,20,21,22), and microtubule (16,18) and flagella (18) formation. Hence, understanding especially the simplest description of the mechanism(s) of protein aggregation is critically important for both the aforementioned disease as well as non-disease states. For the above reasons, the basic mechanism(s) of protein aggregation has been the focus of considerable prior effort in at least 30 laboratories over a period of approximately 50 years (23).

Protein aggregation studies are generally acknowledged (16,18,24,25,26,27) to have begun in earnest with Oosawa's 1959 and 1962 studies of the G-actin to F-actin polymerization where G stands for the globular and F the fibril form of actin (15,16). In order to proceed rationally and in as best informed mode as possible, we have recently constructed a critical review of the prior protein aggregation literature (23). We find 5 basic classes of mechanisms for protein aggregation over the period from 1959 to 2007 (23). Those studies are listed in Table 1 below, Table 1 being the first of two key tables

of prior work that is discussed in our forthcoming critical review (23). The essence of Table 1 directly relevant to the present contribution is as follow: (i) the most cited, "complete" mechanism is the subsequent monomer addition mechanism, Scheme 1 (16,17,28,29); and (ii) the resulting mathematics and kinetic equations corresponding to the complete mechanism are, however, not routinely used to fit experimental data (25,26,28,29). Alternatively, when fits are possible (following numerous assumptions and approximations (16,18,27)), the resultant equation obfuscates verification (or refutation) of whether nucleation and then growth are present and deconvoluteable (growth also being called polymerization, elongation, or "heterogeneous nucleation" in some treatments[1]). The deconvolution of nucleation from growth is key since these are central concepts in the broader literature of aggregation phenomenon in Nature (23,30,31,32,33). In addition our review of the literature (23) reveals that: (iii) nothing approaching a consensus mechanism—or even an accepted primary model for fitting data—has emerged; and especially important for the present work (iv) a simplest

---

[1] The nomenclature in the protein aggregation literature is varied and can be confusing, with multiple terms being used for (we infer) more or less the same intended meaning. Specifically, for growth or agglomeration one also sees the terms "elongation, aggregation, fibrillation, and polymerization". For the nucleation period one sees the phenomenological terms of "induction period" or "lag phase". Adding considerable confusion is the term "heterogeneous nucleation" used, we surmise, in the dictionary definition / meaning of heterogeneous as nucleation of "different" species, for example, new protein polymers on the surface of existing polymers. We would simply call this autocatalytic growth. The confusion here arises since, traditionally *in the nucleation literature*, "heterogeneous nucleation" means nucleation occurring within a different phase (such as nucleation on a solid surface or phase despite the reactants being soluble in solution). Herein we will use the terms (homogeneous) nucleation and (autocatalytic surface) growth as defined rigorously before and also herein by the kinetic equations and respective k_1 and k_2 rate constants in Scheme 2—except in Table 1 where we retain the terms used by the original authors so as to properly represent and reproduce the extant literature.

possible, "Ockham's razor" model (34) that can be used to fit experimental data has not appeared, even though mechanistic scientists know that an Ockham's razor approach to mechanism is an essential part of rigorous mechanistic science. Finally, also sorely lacking is: (vi) a minimal model that can routinely and successfully be used for the badly needed deconvolution of nucleation from growth[1] for a broad range of kinetic data, even if that minimal model is, by necessity, phenomenological and unsophisticated relative to a more "complete" mechanism.

Scheme 1. The Subsequent Monomer Addition Mechanism Proposed for the Aggregation of Proteins.

$$P_1 + P_1 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} P_2$$

$$P_1 + P_2 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} P_3$$

$$P_1 + P_i \underset{k_{-i}}{\overset{k_i}{\rightleftharpoons}} P_{i+1}$$

Herein we report the following results and insights: (i) that the 1997 Finke-Watzky (hereafter F-W) 2-step mechanism (31) of slow, continuous nucleation followed by typically fast autocatalytic surface growth,[2] Scheme 2 and Figure 1 with its associated rate law, equation 1, and analytic equation, equation 2, is able to fit all of the 14 representative protein aggregation data sets we have been able to mine out of the (non-prion) neurological protein aggregation and related literature, (ii) that those in general

---

[2] It is Noteworthy that the A + B → 2B step is the kinetic definition of autocatalysis, that is, a reaction in which a product (B) is also a reactant; this is why the curves "turn on" with almost step-function like appearance in some cases with the middle part of the reaction going faster and faster (i.e., and until the reaction runs out of A, so that the reaction then slows down).

excellent, often seemingly best-to-date fits demonstrate the broad applicability of the F-W mechanism to amyloid, α-synuclein, and polyglutamine disease proteins; and (iii) that the F-W mechanism is both the simplest ("Ockham's razor" (34)) model yet reported, as well as perhaps the simplest conceivable model (see Scheme 2), for protein aggregation that is also able to fit a wide body of protein aggregation data. In addition, we report (iv) that the F-W mechanism achieves the key result of deconvoluting the crucial nucleation step from the growth step,[1] and we also show herein (iv) that the 3-step mechanism proposed in 2000 by Saitô and co-workers (35) (entry 10 of Table 1) yields, after the necessary simplification and assumptions those authors made to make the mathematics manageable, a kinetic equation that we demonstrate herein is mathematically identical to the simpler, 2-step F-W mechanism published 3 years earlier (31).

Scheme 2. The F-W Mechanism (31) for Nucleation and Growth Phenomenon.[a]

$$A \xrightarrow{k_1} B$$
$$A + B \xrightarrow{k_2} 2B$$

[a] This mechanism was first worked out for transition-metal nanocluster nucleation and growth (31,32) (for a more expanded list of references on nanoclusters, please refer to the Supporting Information), but this work as well as another in progress (33) show that it also applies to other larger particle and aggregate formation phenomenon in Nature. Note that (i) the resultant kinetic curves are generally sigmoidal (see Figure 1); (ii) the A → B and A + B → 2B steps are typically composites of what may be hundreds to even thousands of actual elementary steps (31,32) (e.g., those back in Scheme 1). Hence, the resultant rate constants, k1 and k2, are necessarily average values for all the underlying steps.

$$\frac{-d[A]}{dt} = k_1[A] + k_2[A][B] \tag{1}$$

$$[A]_t = \frac{\frac{k_1}{k_2} + [A]_0}{1 + \frac{k_1}{k_2[A]_0}\exp(k_1 + k_2[A]_0)t} \tag{2}$$

Finally, equation 3 will prove to be most convenient form for analyzing all but one of the literature data sets reexamined herein. Equation 3 is readily derivable from equation 2 in the case where one has a clean reaction, A → B, with associated clean mass balance, $[A]_t = [A]_0 - [B]_t$. Substitution for $[A]_t$ into equation 2 and solving for $[B]_t$ yield eq. 3.

$$[B]_t = [A]_0 - \frac{\frac{k_1}{k_2} + [A]_0}{1 + \frac{k_1}{k_2[A]_0}\exp(k_1 + k_2[A]_0)t} \qquad (3)$$

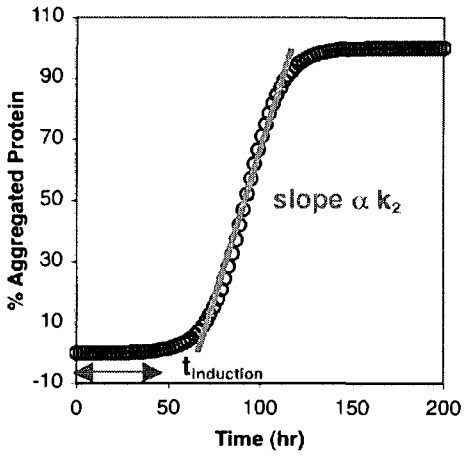


Figure 1. A typical kinetic curve seen for the F-W, 2-step mechanism in Scheme 1 with representative rate constants of $k_1 = 1 \times 10^{-5}$ hr^{-1}, $k_2 = 1 \times 10^{-3}$ $mM^{-1}hr^{-1}$ (i.e., rate constants chosen for the sake of illustration but which are also within the range of values we will find for the protein agglomeration examples which follow). Two useful facts about this example sigmoidal curve is, as shown in the curve, k_1 (units of hr^{-1}) is proportional to the inverse of the induction period time, $t_{induction}$ (units hr^1); that is, k_1 gives quantitative information about the nucleation period (also called the induction period or lag phase[1]). In addition, $k_2*[A]_0$ (units hr^{-1}) is proportional to the normalized slope of the line following the induction period, $+d[B]/dt \times 1/[A]_0$ (units hr^{-1}) (31, 54).

Table 1. Key Lead References from the Prior (Primarily Non-Prion) Literature of Protein Aggregation Mechanisms.

Entry	Reference	Year	Title	Key Selected Contents
1	(15)	1959	G-F Transformation of Actin as a Fibrous Condensation	Oosawa and coworkers are pioneers in the field of protein polymerization with the recognition of actin polymerization of the general mechanism $n(\text{G-actin}) \overset{x\ Mg^{2+}}{\rightleftharpoons} (\text{F-actin})_n$, where G-actin is the globular and F-actin is the fibrous forms of actin. In 1959 they used the techniques of flow birefringence, light scattering, viscosity, and ultra-centrifugation as evidence for the fibrous condensation. In 1962, the aggregation of macromolecules with a focus on actin was proposed to aggregate by a helical rather than linear aggregation mechanism above a critical concentration of macromolecules. It was further proposed that F-actin is the helical aggregate of G-actin. The expression $\frac{dn_h}{dt} = [k_+\lambda(t) - k_-]\int [k_+'\lambda_3(t) - k_-'\lambda_{3h}(t)]dt + 3[k_+'\lambda_3(t) - k_-'\lambda_{3h}(t)] + k_-\lambda_{3h}(t)$, is given for the rate of total monomers participating in helical polymers, where n_h is the total number concentration of monomers participating in helical polymers, k_+ and k_- are the forward and reverse rate constants for attachment and detachment of monomers, $\lambda(t)$ is the number concentration of monomers, $k_+'\lambda_3(t)$ and $k_-'\lambda_{3h}(t)$ are the forward and reverse transformation rates of ordinary trimers to nuclei of helical polymers. This expression can only be approximately solved with a number of assumptions, but is an early and important contribution to the mechanisms of protein aggregation.
	(16)	1962	A Theory of Linear and Helical Aggregations of Macromolecules	
2	(19)	1971	Glutamate Dehydrogenase: Anatomy of a Regulatory Enzyme	Eisenberg was, to the best of our knowledge, the first to propose subsequent monomer addition as the (complete) mechanism for a protein aggregation system. In this case he examined the polymerization of glutamate dehydrogenase of the basic reaction shown on the right. Of note is his recognition of similarities between protein aggregation phenomena and early polymerization literature. It was this recognition that allowed Eisenberg to propose in 1971 the mechanism shown to the left for the polymerization of glutamate dehydrogenase, where P_i is the polymerized species of weight M_i, and K_i is the equilibrium constant for the reaction, K_i being assumed to be the same for each monomer-addition step. $n(\text{glutamate dehydrogenase}) \rightleftharpoons (\text{glutamate dehydrogenase})_n$ $P_i + P_1 \overset{K_i}{\rightleftharpoons} P_{i+1}$
3	(20)	1975	Mechanism of Bovine Liver Glutamate Dehydrogenase Self-Association	In 1975 and 1976, Eisenberg's subsequent monomer addition mechanism for glutamate dehydrogenase (entry 2) was questioned in a series of three papers. The basic driving force behind this series of papers is that the subsequent monomer addition mechanism cannot account for all of the experimental data in varying concentration ranges. Due to this fact, the

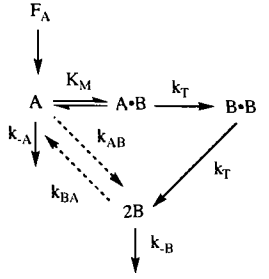
	(21)	1975	I. Kinetic Evidence for a Random Association Mechanism	authors propose a new mechanism for the polymerization of glutamate dehydrogenase termed the "random association" mechanism shown to the right. The "random association" mechanism means that two polymers of any size can come together and that the rate constants for this process are the same regardless of the size of P_i and P_j. This random aggregation mechanism was tested using previous aggregation data and also using quasi-elastic light scattering to measure to probe the kinetics of bovine liver glutamate dehydrogenase although, no aggregation kinetics along with a fit to the proposed mechanism are given. $P_i + P_j \rightleftharpoons P_{i+j}$ $i, j = 1, 2, \ldots \infty$
	(22)	1976	II. Simulation of Relaxation Spectra for an Open Linear Polymerization Proceeding via a Sequential Addition of Monomer Units	
			III. Characterization of the Association-Dissociation Stoichiometry with Quasi-elastic Light Scattering	
4	(17)	1975	Kinetics of the Cooperative Association of Actin to Actin Filaments	In 1975, Wegner and Engel proposed the subsequent monomer addition mechanism for actin fibril formation with the assumption that the rate constant for dimer formation would be different, but all subsequent rate constants would be equivalent. This proposed mechanism is shown on the left where A is an actin molecule, k_N and k_N' are the rate constants for the formation and destruction of the dimer, and k and k′ are the rate constants for the binding and dissociation of protomers in the so-called elongation steps. The associated differential equations are shown on the right where c_i represents the concentration of the i-mer with $\infty > i > 3$. Also determined in this 1975 work is a critical nucleus size of three or four. Their proposed mechanism was tested with experimental data, for the actin to actin filaments aggregation reaction, to demonstrate the mechanism's ability to account for the experimental data, and visually poor to fair fits were observed. $2A \underset{k_N'}{\overset{k_N}{\rightleftharpoons}} A_2$ $A_2 + A \underset{k'}{\overset{k}{\rightleftharpoons}} A_3$ $\vdots$ $A_{i-1} + A \underset{k'}{\overset{k}{\rightleftharpoons}} A_i$ $\frac{dc_2}{dt} = k_N c_1^2 - k_N' c_2 - k c_1 c_2 + k' c_3$ $\frac{dc_i}{dt} = k c_1 c_{i-1} - k' c_i - k c_1 c_i + k' c_{i+1}$

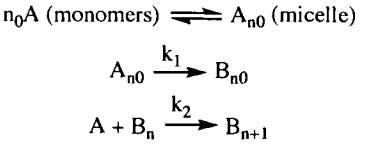
5	(28) (29)	1980 1985	Kinetic Studies on Photolysis-Induced Gelation of Sickle Cell Hemoglobin Suggests New Mechanism Kinetics of Sickle Hemoglobin Polymerization	These authors propose an early word mechanism in 1985, for the polymerization of sickle hemoglobin, the basic stoichiometry of which is $n(hemoglobin-S) \rightarrow (hemoglobin-S)_n$. In 1980 the general mechanism for the polymerization process was given as shown on the left, essentially Eisenberg's monomer addition mechanism. This mechanism was examined in the classic and well studied sickle hemoglobin system; the proposed rate equations consist of double, so-called homogeneous and "heterogeneous," nucleation for the net polymerization process: $M + M \rightleftharpoons M_2$ $M + M_2 \rightleftharpoons M_3$ $M + M_{i-2} \rightleftharpoons M_{i-1}$ $M + M_{i-1} \rightleftharpoons M_i$ $M + M_j \rightleftharpoons M_{j+1}$ $\frac{dc_p}{dt} = K_N k_+ (\gamma c)^n + K_M \phi k_+ (c_0 - c)(\gamma c)^m$ for the former and $-\frac{dc}{dt} = nK_N k_+ (\gamma c)^n + mK_M \phi k_+ (c_0 - c)(\gamma c)^m + (k_+ \gamma c - k_-)c_p$ for the latter, where KM and KN are the equilibrium constants for "heterogeneous" and homogeneous nucleation respectively, k_+ and k_- describe the rates of monomer addition and removal, γ is the activity coefficient of the monomer, c_0 and c are the monomer concentrations at time zero and time t, and ϕ is a scaling factor for the number of effective sites for nucleation. The result, the authors note, is a five parameter model to describe the polymerization process. Also, the authors use (we infer) the term "heterogeneous nucleation" in the dictionary sense of heterogeneous meaning the nucleation of "different" species, that is, new polymers growing on the surface of existing polymers. Elsewhere we note that heterogeneous nucleation has been traditionally used in the nucleation literature to indicate nucleation in a (typically solid) phase while the reactants are in a different (typically liquid) phase (23). The opinion that "heterogeneous nucleation" is actually seeded autocatalytic surface growth, and other issues such as the term "extreme autocatalysis" are also addressed in our review elsewhere (23).
6	(18)	1983	Polymerization of Actin and Actin-like Systems: Evaluation of the Time Course of Polymerization in Relation to the Mechanism	Frieden and Goddette examined the earlier mechanisms of Oosawa (entry 1) and Wegner (entry 4) using numerical integration techniques; again, the basic mechanism of reversible subsequent monomer as shown on the right is employed in this case for the polymerization of actin. Their findings are that the Oosawa and Wegner equations are adequate for simple polymerization processes, but not for the more complex, more realistic processes of aggregation. Thus the authors added an additional step into the subsequent monomer addition mechanism. $A + A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} A_2$ $A_2 + A \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} A_3$ $A_p + A \underset{k_{-p}}{\overset{k_p}{\rightleftharpoons}} A_{p+1}$ $B \underset{k_{-a}}{\overset{k_a}{\rightleftharpoons}} A$ Their additional step described as is a step prior to polymerization and involves activation of the monomer. In the case of actin, this reversible step represents ligand binding to a metal followed by a protein conformational change.

7	(25) (26)	1983 1983	On One-Dimensional Nucleation and Growth of "Living" Polymers I. Homogeneous Nucleation On One-Dimensional Nucleation and Growth of "Living" Polymers II. Growth at Constant Monomer Concentration	$M_1 + M_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} M_2$ $M_1 + M_2 \underset{k_4}{\overset{k_3}{\rightleftharpoons}} M_3$ $\vdots$ $M_1 + M_n \underset{k_{n+2}}{\overset{k_{n+1}}{\rightleftharpoons}} M_{n+1}$	This classic set of papers investigates the one-dimensional homogeneous nucleation and growth or elongation, and proposes the subsequent monomer addition mechanism for "living" polymers shown to the left. The following assumptions are also employed: i) there is only successive addition of monomers, ii) there is no geometric difference between any of the M_i, iii) $k_i=k'$ up to the critical nucleus and $k_1=k$ past the critical nucleus, where k and k′ result from the two time constants of τ and τ', iv) a prior equilibrium exists up to the critical nucleus, and v) both reversible and irreversible growth are treated beyond the critical nucleus. Using the proposed general mechanism of successive monomer addition, four separate cases of initial rate, steady state, approaching equilibrium, and quasi-steady state are treated giving non-integrable equations for nucleation for each of the four cases and an expression for reversible growth, $\frac{-dM_1}{dt} = (k_f C - k_r)^2 c_n t$, where M_1 is the monomer, k_f and k_r are the forward and reverse rate constants, C is the monomer concentration, and c_n is the critical nucleus. This expression for growth comes from employing the assumptions that all forward and reverse rate constants are equal and if the monomer concentration is constant. Lacking in this 1983 work is a test of the proposed mechanism and equations by fitting experimental data, although this is a more general treatment with good insights for one-dimensional growth, in our opinion.
8	(27)	1986	Cooperative Polymerization Reactions: Analytical Approximations, Numerical Examples, and Experimental Strategy	$A_1 + A_n \underset{k_-}{\overset{k_+}{\rightleftharpoons}} A_{n+1}$ $\quad n < s$ $A_1 + A_n \underset{g_-}{\overset{g_+}{\rightleftharpoons}} A_{n+1}$ $\quad n \geq s$	Goldstein and Stryer use the simple model shown to the left for stepwise addition and subtraction of monomers to one end of a polymer, where A_1 is the monomer, A_n is the polymer of length n, k_+ and k_- are the forward and reverse rate constants prior to the formation of the seed, s, g_+ and g_- are the rate constants after the formation of the seed, and the seed size is defined as the length s, where the kinetic rate constants change. In other words, the first equation represents nucleation and the second, growth. The corresponding rate equations are then given as: $\frac{dA_n}{dt} = k_+ A_1 (A_{n-1} - A_n) + k_- (A_{n+1} - A_n)$ when $n < s$ and $\frac{dA_n}{dt} = g_+ A_1 (A_{n-1} - A_n) + g_- (A_{n+1} - A_n)$ when $n > s$. The above rate equations are then transferred into a dimensionless, more mathematically convenient but less physically obvious form using the assumptions of i) monomer concentration changes only by the addition or subtraction of monomer, ii) polymer formation by seed production is irreversible, and iii) the ratio of the concentration of the seed-minus-one-length to the (s-1) power of monomer concentration is essentially constant,

yielding: $\frac{d\alpha_1}{d\tau} = C - C\alpha_1$ and $\frac{dC}{d\tau} = K\alpha_1^{s-1}(\alpha_1 - 1)$, where $\alpha_n = g_+/g_-{}^*A_n$, $\tau = g_-t$, $C = \Sigma_{n-s+1}^{\infty}\alpha_n$, and "K is usually taken from equilibrum constants, but its exact value is immaterial for our discussion." A couple of important insights are apparent from the dimensionless form of the equation: first, the assumption of irreversible polymerization is justified since the same kinetic curves are observed with and without the approximation, except in the case of low concentration and long measurement time; second, the appearance of dimer, trimer, and tetramer species following the loss of monomers is apparent from simulations. Lacking, however, are a test of the proposed mechanism and its corresponding dimensionless equations using experimental data. The lack of dimensional forms of the key equations have also probably contributed to the lack of use of these equations by others who might be more interested in the underlying physical picture and main concepts behind protein aggregation.

9	(59)	1996	Prionics or the Kinetic Basis of Prion Diseases	This paper discusses the possibility of two different mechanisms for the formation of pathogenic prion species. The first of which is a condensed verion of Prusiner's (60,61) so-called "linear autocatalysis" mechanism shown on the right where F_A is the constant metabolic formation of A, A is the normal form of the host protein, B is the pathogenic form of A, K_M is the Michaelis constant, k_T is the turnover number, k_{-A} is the metabolic decomposition of A, k_{-B} is the metabolic decomposition of B, k_{AB} and k_{BA} are the noncatalytic conversion of A to B and B to A. From Prusiner's condensed mechanism the corresponding rate equation is obtained: $\frac{d[A]}{dt} = F_A - k_{-A}[A] - k_{AB}[A] - \frac{k_T[A]}{K_M + [A]}[B]$. Also discussed is Lansbury's (62,63,64,65,66) "passive autocatalysis mechanism" proposed for both the aggregation of prions and the plaque formation in Alzheimer's disease, $A \underset{k_{BA}}{\overset{k_{AB}}{\rightleftharpoons}} B \underset{k_D}{\overset{k_F[B]}{\rightleftharpoons}} B_2 \underset{k_D}{\overset{k_F[B]}{\rightleftharpoons}} \cdots \underset{k_D}{\overset{k_F[B]}{\rightleftharpoons}} B_n \underset{k_D}{\overset{k_F[B]}{\rightleftharpoons}} B_p$, where A is again the normal form of the host protein, B is the pathogenic monomeric form of A, B_2, B_n, and B_p are the aggregate forms of B, kAB and kBA are again the noncatalytic conversion of A to B and B to A, k_F is the formation rate constant, and k_D is the dissociation rate constant. Neither mechanism can be ruled out in terms of the formation of aggregates of prion proteins, and both remain as possible valid mechanisms. Also, the authors specifically state, "The aim of this paper is not a quantitative description of real data, but rather an understanding of the principles that control reality." Accordingly, this classic work has apparently not been used to treat raw kinetic data.



10	(35)	2000	Conformational Transitions and Fibrillation Mechanism of Human Calcitonin as Studied by High-Resolution Solid-State 13C NMR	Saitô and coworkers suggest a mechanism for the fibrillation of the human calcitonin hormone with these three steps shown on the right where n_0 represents micellular intermediates of the same aggregation number, A_{n0} is the micellular aggregates made up of n_0 monomers of the protein, B_{n0} is the nucleus, and B_n and B_{n+1} are the elongated fibrils with n and n + 1 molecules of the protein. The corresponding analytical form of the mechanism was obtained using the three assumptions of: monomer and micelle states give the same measurable signal, the total monomer concentration is always much higher than the critical micelle concentration, and each B_n can act as a catalytic site for the conversion of A to B is represented by the following equation $f = \frac{\rho\{\exp[(1+\rho)kt]-1\}}{\{1+\rho\exp[(1+\rho)kt]\}}$, where f is the fraction of calcitonin in the fibrillar form and ρ represents the dimensionless value to describe the ratio of k_1 to the convoluted k which is equal to $k_2*[A]_0$. The fractional integrated rate equation is then used to fit the calcitonin aggregation data under various concentration and pH conditions. Although not readily apparent, not recognized by the authors, and not heretofore demonstrated, herein we show that Saitô and coworkers' above fractional form of the integrated equation for their 3-step mechanims turns out to be equivalent to the earlier, 1997 Finke-Watzky 2-step mechanism of nucleation followed by autocatalytic growth mechanism (31), $[A]_t = \frac{\frac{k_1}{k_2}+[A]_0}{1+\frac{k_1}{k_2[A]_0}e^{(k_1+k_2[A]_0)t}}$, an important insight of the present contribution. (The mathematical derivation proving this equivalence is provided in the Supporting Information.) 

Materials and Methods

Selection of Data Sets for Analysis. The literature was searched, using both SciFinder Scholar and the Web of Science, for the protein aggregation studies relevant to the neurodegenerative disorders, Alzheimer's, Parkinson's, and Huntington's diseases; for this first study prion related diseases were deliberately, albeit arbitrarily, omitted from the literature search in order to keep the number of papers examined to ≤100. We are constructing a separate review and kinetic analysis of the prion protein aggregation literature and will report on that separately in due course. The searches were then narrowed to articles containing usable protein aggregation vs time kinetic data resulting in a representative 14 data sets as presented below.

Data Analysis and Curve Fitting. Data were extracted (digitized) from published kinetic curves using Engauge Digitizer 2.12. In 13 of the 14 data sets, the data were displayed as the formation of product (B) versus time (t), and these digitized data sets were fit by the analytical equation shown in equation 3 using the nonlinear least-squares curve-fitting program in OriginLab Corporation's Origin version 7.0. In one case (see Figure 6a), the kinetic data were reported as the loss of monomeric protein (A) versus time (t); in this case, the kinetic data were fit using equation 2 and, again, Origin version 7.0. The derivation of the analytical equation 3 is given in the Supporting Information.

In the fits presented in Figures 2-12 and summarized in Table 2, we have by necessity assumed that all forms of kinetic measurement used previously in the literature are directly proportional to the percent of aggregated protein. This is the same assumption underlying virtually all the protein agglomeration kinetic studies to date, but is an assumption that we note merits reinvestigation in future studies. For example,

changes in protein conformation, without aggregation, may also influence fluorescence, circular dichroism (CD), absorption, and dynamic light scattering (DLS), even though many papers in the relevant literature do not account for these contributions. Furthermore, for these first studies we have again by necessity assumed that the data measured by fluorescence, CD, and absorption are ideal: that is, it is assumed that there is no absorption for the monomer at the wavelength measured, and no deviation from Beer's law, so that the absorption present is due to the aggregated protein. This appears to be true for the cases where the spectra are shown (4,6,7,13), but in some cases the monomer-alone spectrum is not provided to the reader (5,8,9,10), so that we were unable to verify (or refute) this, again at present necessary, assumption. The resultant signal intensity vs. time may, therefore, not necessarily be a direct measure of the aggregated protein3 (36,37), but for the purposes of the initial treatments herein we assume that it is. It should also be noted that in the literature cases presented herein, the protein generally

---

[3] In the case of dynamic light scattering, there is not a simple correlation between the observed light scattering and the amount of aggregated protein; instead, the relationship is more complex. Assuming particles all of identical size, the following equation shows that the scattering intensity, R(Q), is dependent upon the particle concentration, C_p (g/l), the molar mass of the particles, M_p (g/mol), the particle form factor, P(Q), and the structure factor of the dispersion, S(Q,C_p),that is: $R(Q) \cong K \times C_p M_p P(Q) S(Q,C_p)$ (where K is a constant dependent upon the instrumentation and difference in refractive index of particles and solvent). This can also be thought of in a more simplified way: dynamic light scattering measurements depend upon Brownian motion and the speed at which particles move (the diffusion of the molecules), and therefore the molecular weight of the polymer species can be determined by the Einstein-Stokes relation: $M = \frac{4\pi r^3 N_A}{3\upsilon}$, where M is the molecular weight of the polymer, r is the radius, N_A is Avogadro's number, and υ is the partial specific volume of the polymer molecule. This implies that diffusion is inversely proportional to the cubic root of the molecular weight rather than the concentration in solution. This in turn means that light scattering measurements are not directly proportional to the concentration of the solution being measured.

begins in an unfolded or misfolded state, either natively or by denaturation, prior to measurements of the kinetics of protein aggregation.

Results and Discussion

To start, the literature was searched for protein agglomeration kinetic data sets as detailed in the Experimental section. Both Scifinder Scholar and Web of Science were employed to search for kinetic aggregation data relevant to the neurodegenerative disorders Parkinson's disease, Alzheimer's disease, and Huntington's disease. A total of 14 representative kinetic data sets were found (4,5,6,7,8,9,10,11,12,13,14) encompassing four main categories of proteins: (i) amyloid-β, the peptide that is hypothesized to be central to Alzheimer's disease (1), (ii) α-synuclein, the peptide hypothesized to be key in Parkinson's disease (2), (iii) polyglutamine, the peptide hypothesized to be crucial in the etiology of Huntington's disease (3), and (iv) other amyloidogenic proteins not belonging to the previous three categories.

Each kinetic data set was in turn analyzed by the F-W 2-step mechanism given in Scheme 1. The results are presented in Figures 2-4, 5a-c, 6a-b, and 7-12 that follow, with Table 2 summarizing the nucleation, k_1, and growth, k_2, rate constants obtained from the kinetic analyses. The fit shown in Figure 12 overestimates the second half of the data; for this reason the data was also analyzed by the 3- and 4-step, expanded versions of the F-W mechanism which include bimolecular agglomerative growth, $B + B \rightarrow C$ (rate constant k_3) and a novel, only recently discovered autocatalytic agglomeration step, $B + C \rightarrow 1.5C$ (rate constant, k_4) (32). For further information on these fits the interested reader is referred to the Supporting Information (see Figure S7 and the section titled

Alternative Mechanisms Considered: the 3- and 4-Step Mechanisms). Interestingly, the 2-, 3-, and 4-step fits all give similar residual values. On the basis of Ockham's razor, the fit to the simpler 2-step mechanism was chosen for presentation in Figure 12 and since at present there is no compelling evidence to expand the mechanism to the 3- or 4-step variants of the original 1997 F-W mechanism.

Fitting Amyloid β Aggregation Literature Data to the 2-Step F-W Mechanism. The aggregation of proteins to form β-sheet rich amyloid fibrils has been hypothesized to be intimately involved in more than 20 human diseases (13,38). One very important amyloidogenic protein is amyloid β, which forms plaques hypothesized to be a crucial component of Alzheimer's disease (39). Shown in Figures 2-6 are the digitized data from the literature on aggregation of amyloid β protein along with the curve-fits to the F-W 2-step mechanism (31) of slow, continuous nucleation plus subsequent autocatalytic surface growth. In all cases the fits are very good if not excellent.

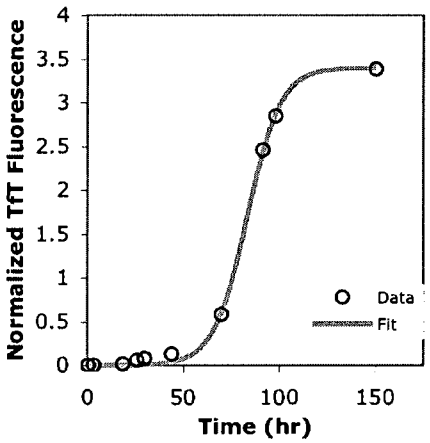


Figure 2. Digitized data of Kelly and coworkers' (4) amyloid β peptide aggregation measured (and normalized by the authors) by TfT fluorescence and fit to the F-W 2-step mechanism. Amyloid β peptide aggregation has been hypothesized to be central in the pathology of Alzheimer's disease (1). The fit values for this dataset are $k_1 = 8(3) \times 10^{-6}$ hr^{-1} and $k_2 = 3.4(1) \times 10^{-2}$ $\mu M^{-1} \bullet hr^{-1}$.

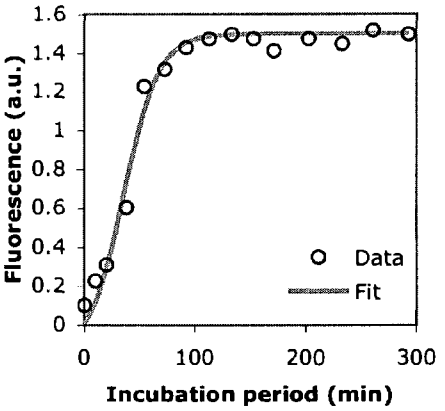


Figure 3. Digitized data of Vestergaard and coworkers' (5) aggregation of amyloid β (Aβ-42) peptide measured by ThT fluorescence and fit to the F-W 2-step mechanism. Again aggregation of this peptide is believed to be key in Alzheimer's disease (39). Although there is considerably less sigmoidal character to this data set (e.g. vs. Figure 1), the fit is still quite good. The resultant $k_1 = 6(2) \times 10^{-3}$ min^{-1} and $k_2 = 3.8(8) \times 10^{-2}$ $\mu M^{-1} \bullet min^{-1}$.

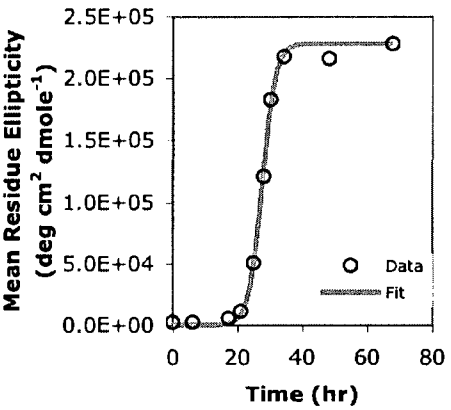


Figure 4. Data of Lynn and coworkers' (6) amyloid β peptide aggregation measured by circular dichroism (CD) digitized, and fit to the F-W 2-step mechanism. This data set and fit provide a third example of the excellent fit by the F-W mechanism to protein aggregation related to the etiology of Alzheimer's disease (39). The observed $k_1 = 6(7) \times 10^{-7}$ hr^{-1} and $k_2 = 2.2(2) \times 10^{-6}$ $mM^{-1} \bullet hr^{-1}$.

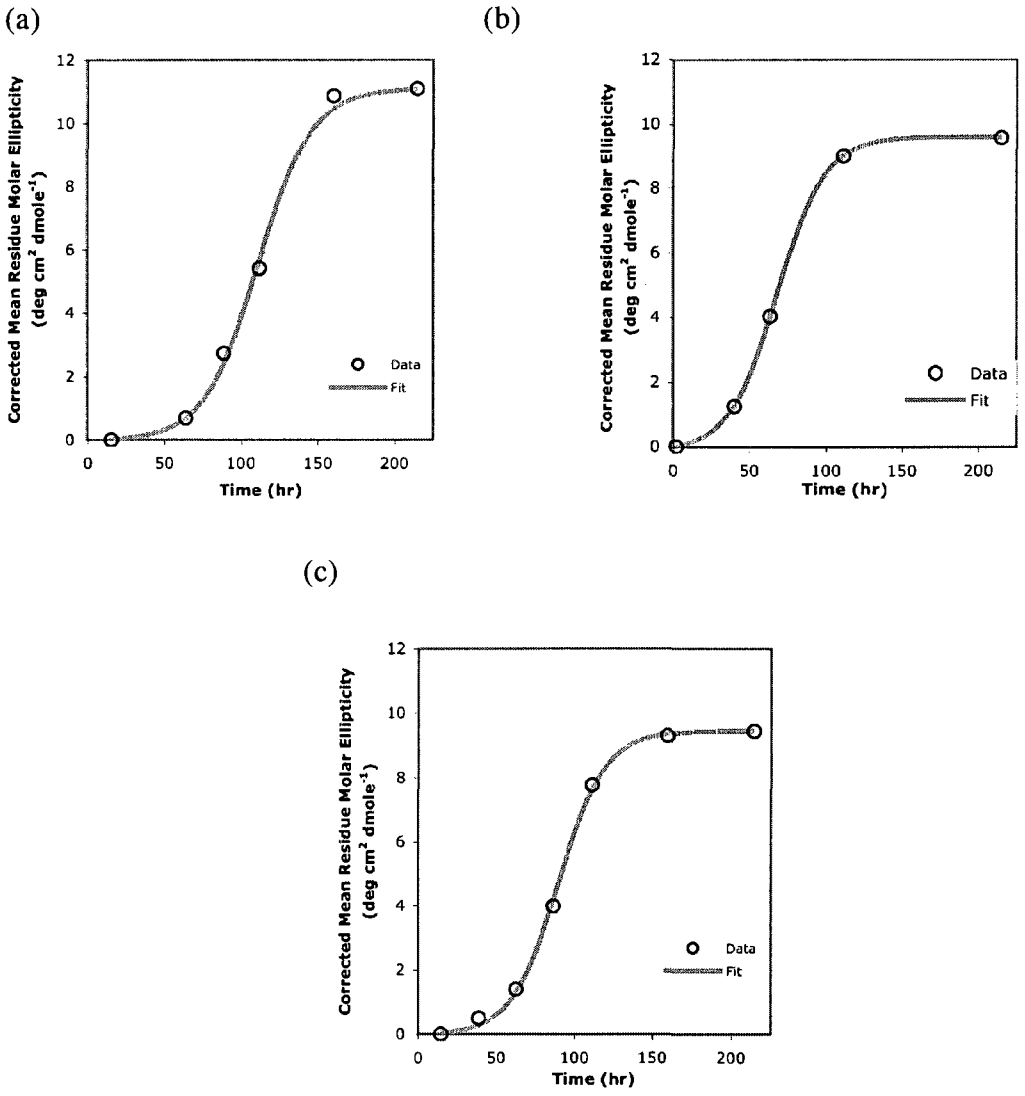


Figure 5. Digitized data of Lynn and coworkers' (7) amyloid β peptide measured as a function of added Zn2+ and fit to the F-W 2-step mechanism. The authors speculate that Zn2+ binding sites exist along the β sheets of amyloid β peptides; therefore, they measured the aggregation as a function of added $[Zn^{2+}]$ and observed a reduced nucleation time with increased $[Zn^{2+}]$. The data in this figure are corrected for the negative values of ellipticity that were reported in the original reference by normalizing the data to have a minimum value of zero (7). By making the assumption that the negative ellipticity values, resulting from the difference in absorption of left vs right circularly polarized, would remain constant throughout a blank sample, a constant value was added to each data point shown so as to render positive all the corrected values. (For the raw, uncorrected data and fits please see Figure S1 of the Supporting Information.) Part (a) is for 1mM of the peptide alone with $k_1 = 1.0(5) \times 10^{-4}$ hr^{-1} and $k_2 = 5.1(5) \times 10^{-3}$ $mM^{-1} \cdot hr^{-1}$, (b) is for $[Zn^{2+}]$:[peptide] of 0.2 with $k_1 = 8.3(3) \times 10^{-4}$ hr^{-1} and $k_2 = 6.55(9) \times 10^{-3}$ $mM^{-1} \cdot hr^{-1}$, and (c) represents a $[Zn^{2+}]$:[peptide] of 0.4 with $k_1 = 1.6(4) \times 10^{-4}$ hr^{-1} and

k_2= 7.1(4) × 10^{-3} $mM^{-1}•hr^{-1}$. A plot of the resultant k_1 and k_2 values vs. [Zn^{2+}], provided as Figure S2 of the Supporting Information, shows little to no effect on the [Zn^{2+}] to the k_1 and 2 values in this specific case.

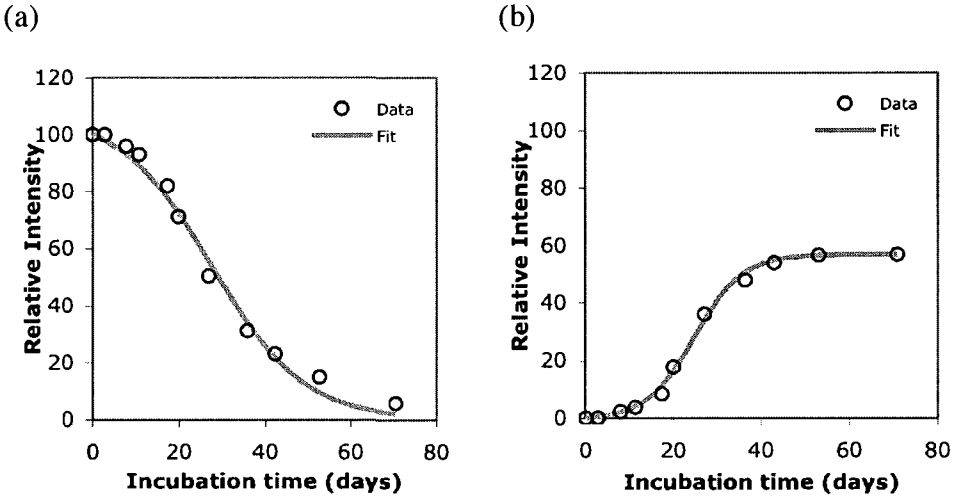


Figure 6. Digitized data of Rebuffat and coworkers' (8) amyloid β peptide aggregation measured by reversed-phase HPLC with absorbance detection plus the fits to the F-W 2-step mechanism. Figure 6 (a) is for Aβ1-16, the first 16 peptides of amyloid β (k_1= 7(1) × 10^{-3} $days^{-1}$, k_2= 9(4) × 10^{-4} $mM^{-1}•days^{-1}$) and (b) Aβ5-16, peptides 5-16 of amyloid β, (k_1= 2.1(5) × 10^{-3} $days^{-1}$ and k_2= 3.1(2) × 10^{-3} $mM^{-1}•days^{-1}$).

Fitting α-Synuclein Aggregation Literature Data to the 2-Step F-W Mechanism.

The aggregation of α-synuclein has been hypothesized to be a crucial component in Parkinson's disease (2). Figures 7 and 8 illustrate the fits of the F-W mechanism to the aggregation of α-synuclein. Figure 7 in particular is a striking demonstration of how well the simple 2-step F-W mechanism fits the α-synuclein aggregation—that is, the nucleation and autocatalytic growth—kinetic data.

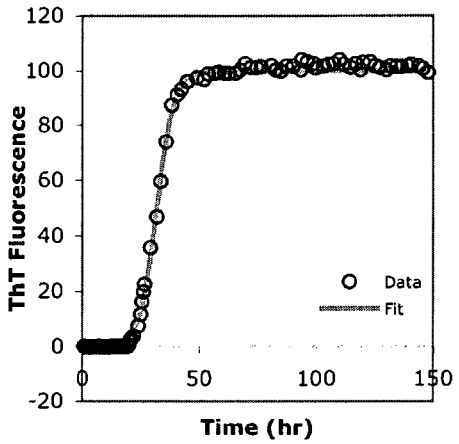


Figure 7. Digitized data of Fink's (9) α-synuclein aggregation, in the presence of macromolecular crowding induced by 100 mg/mL poly(ethylene glycol)-3350 and monitored by ThT fluorescence, and curve-fit to the F-W 2-step mechanism, $k_1 = 4.0(8) \times 10^{-5}$ hr^{-1}, $k_2 = 2.77(7) \times 10^{-3}$ μM^{-1}•hr^{-1}. Fink notes about α-synuclein aggregation, "The in vitro kinetic studies of α-synuclein fibril formation show an initial lag phase followed by an exponential growth phase and a final plateau..." (9). These are, again, the kinetic signatures of the F-W 2-step mechanism.

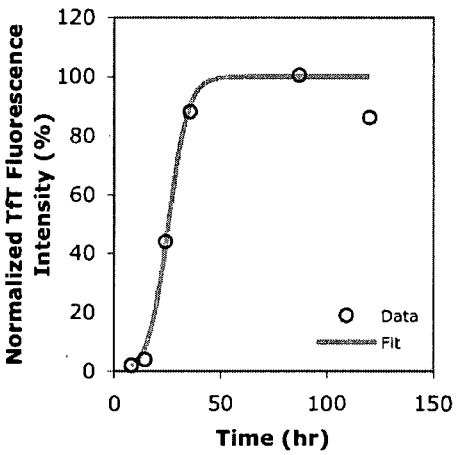


Figure 8. Digitized data of Sode and coworkers' (10) α-synuclein aggregation measured by TfT fluorescence assay analysis and fit to the F-W 2-step mechanism. The resultant $k_1 = 1(1) \times 10^{-3}$ hr^{-1} and $k_2 = 2.2(6) \times 10^{-3}$ μmol^{-1}•hr^{-1}.

There is also some kinetic data in the literature that looks at the aggregation of α-synuclein as a function of added seed protein. The interested reader is referred to Figures S9-S10 of the Supporting Information for the digitized data and corresponding fits.

Fitting Polyglutamine Aggregation Literature Data to the 2-Step F-W Mechanism. Aggregation of polyglutamine has been hypothesized to be intimately involved in the neurodegenerative disorder Huntington's disease (3). Figures 9 and 10 show the correlation between the aggregation of polyglutamine data and the fit to the F-W 2-step mechanism demonstrating a third example of protein aggregation kinetic data that can also be fit and deconvoluted into its nucleation (k_1) and autocatalytic surface growth (k_2) components by the 2-step F-W mechanism.

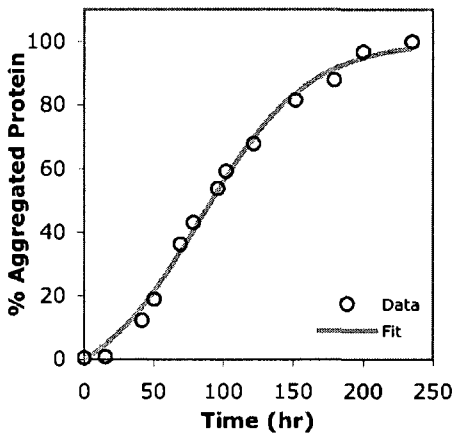


Figure 9. Digitized data of Wetzel and coworkers' (11) aggregation of polyglutamine monitored by light scattering and fit to the F-W 2-step mechanism with resultant rate constants, $k_1 = 2.6(3) \times 10^{-3}$ hr^{-1} and $k_2 = 2.3(2) \times 10^{-4}$ μM^{-1}•hr^{-1}.

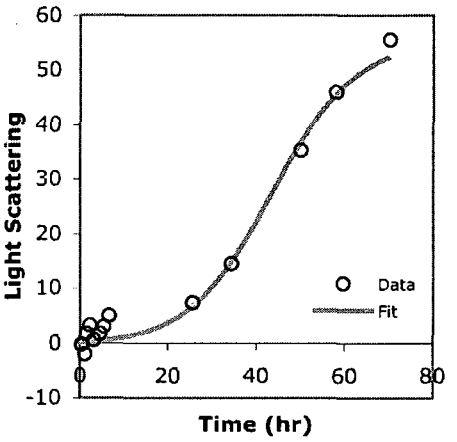


Figure 10. Digitized data of Wetzel's (12) polyglutamine aggregation without seeding monitored by light scattering and fit to the F-W 2-step mechanism. The fit values are k_1= 1.0(4) × 10^{-3} hr^{-1} and k_2= 1.9(2) × 10^{-3} $\mu M^{-1} \bullet hr^{-1}$.

Wetzel has also examined the aggregation kinetics of polyglutamine in the presence of a 5% (w/w) seed (12). This digitized and fit data set is contained in Figure S8 of the Supporting Information for the interested reader.

Fitting Other Amyloidogenic Proteins Aggregation Literature Data to the 2-Step F-W Mechanism. Figures 11 and 12 show, respectively, the fits of the F-W mechanism to the aggregation of β2-microglobulin and human lysozyme proteins.

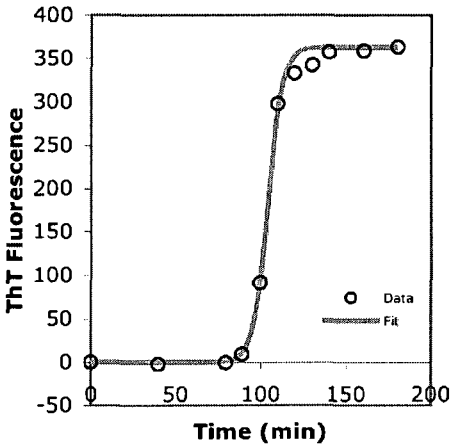


Figure 11. Digitized data of Goto and coworkers' (13) ultrasonication-induced aggregation of β2-microglobulin monitored by ThT fluorescence and fit (solid line) to the F-W 2-step mechanism. β2-Microglobulin is an amyloid protein that causes amyloidosis

resulting in the deposition of amyloid fibrils in the synovia of the carpal tunnel in patients who have been receiving dialysis for more than 10 years (67). Note the generally excellent fit to this highly sigmoidal, approaching step-function-like kinetic curve. The fit values are $k_1 = 1(3) \times 10^{-11}$ min^{-1} and $k_2 = 6(1) \times 10^{-4}$ min^{-1}•μM^{-1}.

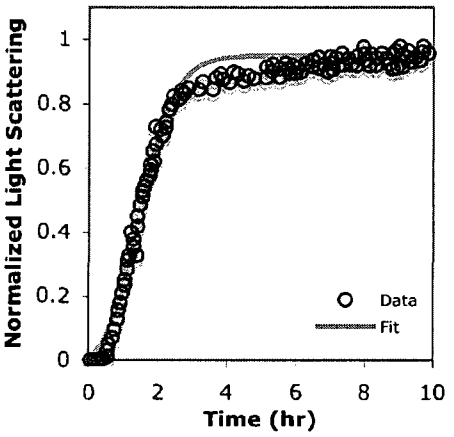


Figure 12. Digitized data of Dobson and coworkers' (14) in vitro D67H human lysozyme variant used to study the folding and misfolding of the protein, the latter leading to aggregate formation and amyloid disease, again with fits to the F-W 2-step mechanism. In 1993 the D67H lysozyme variant was discovered to be associated with a hereditary form of nonneuropathic systemic amyloidosis that causes amyloid deposits to form in various organs of the body (68). The aggregation data was measured by dynamic light scattering and normalized by the authors. The resultant $k_1 = 1.1(1) \times 10^{-1}$ hr^{-1}, $k_2 = 1.9(1)$ μM^{-1}•hr^{-1}.

In summation to this point, all 14 data sets found by our literature search fit the F-W 2-step mechanism with an R2 value ≥ 0.98 (Table 2). This includes data obtained by four different methods, for three different proteins, by nine different labs and spanning over four publication years. The fits are generally good, and in many cases excellent (e.g. Figures 2, 4, 6, 7, 9, and 11). As such, this contribution is the current record of the widest set of protein aggregation data best fit by the also simplest kinetic model to date. The implied conclusion is, therefore, that at least phenomenologically speaking, the F-W 2-step mechanism of nucleation and autocatalytic growth is a previously unrecognized,

"Ockham's razor" (34) / minimalistic model for deconvoluting protein agglomeration nucleation, k_1, from (autocatalytic) growth, k_2.

Table 2. Rate constants and coefficient of determination (R^2) values from fitting 14 literature protein aggregation kinetic data sets to the F-W 2-step mechanism, Scheme 2, *vide supra*. The plots of those fits are shown in Figures 2-12.[a]

Entry	Reference	System	Distinguishing Experimental Details	Data Collection Method	k_1[a]	k_2[a,b]	R^2 [c]
1	Kelly (4)	Amyloid β		Fluorescence	$8(3) \times 10^{-6}$ hr^{-1}	$2.3(1) \times 10^{-3}$ $\mu M^{-1}hr^{-1}$	0.9989
2	Vestergaard (5)	Amyloid β		Fluorescence	$6(2) \times 10^{-3}$ min^{-1}	$7(2) \times 10^{-4}$ $\mu M^{-1}min^{-1}$	0.9830
3	Lynn (6)	Amyloid β		CD	$6(7) \times 10^{-7}$ hr^{-1}	$3.0(3) \times 10^{-1}$ $mM^{-1}hr^{-1}$	0.9971
4	Lynn (7)	Amyloid β	1mM Peptide[d]	CD	$1.0(5) \times 10^{-4}$ hr^{-1}	$5.7(6) \times 10^{-2}$ $mM^{-1}hr^{-1}$	0.9973
5			[Zn]:[Peptide] = 0.2[d]		$8.3(3) \times 10^{-4}$ hr^{-1}	$6.29(9) \times 10^{-2}$ $mM^{-1}hr^{-1}$	1.0000
6			[Zn]:[Peptide] = 0.4[d]		$1.6(4) \times 10^{-4}$ hr^{-1}	$6.7(4) \times 10^{-2}$ $mM^{-1}hr^{-1}$	0.9988
7	Rebuffat (8)	Amyloid β	Aβ1-16	Absorbance	$7(1) \times 10^{-3}$ $days^{-1}$	$1.8(8) \times 10^{-3}$ $mM^{-1}days^{-1}$	0.9904
8			Aβ5-16		$2.1(5) \times 10^{-3}$ $days^{-1}$	$3.5(2) \times 10^{-3}$ $mM^{-1}days^{-1}$	0.9964
9	Fink (9)	α-synuclein		Fluorescence	$4.0(8) \times 10^{-5}$ hr^{-1}	$4.0(1) \times 10^{-3}$ $\mu M^{-1}hr^{-1}$	0.9986
10	Sode (10)	α-synuclein		Fluorescence	$1(1) \times 10^{-3}$ hr^{-1}	$1.1(3)$ $\mu M^{-1}hr^{-1}$	0.9778
11	Wetzel (11)	Polyglutamine		DLS	$2.6(3) \times 10^{-3}$ hr^{-1}	$2.3(2) \times 10^{-4}$ $\mu M^{-1}hr^{-1}$	0.9941
12	Wetzel (12)	Polyglutamine		DLS	$1.0(4) \times 10^{-3}$ hr^{-1}	$1.3(1) \times 10^{-2}$ $\mu M^{-1}hr^{-1}$	0.9877
13	Goto (13)	β-microglobulin		Fluorescence	$1(3) \times 10^{-11}$ min^{-1}	$9(1) \times 10^{-3}$ $\mu M^{-1}min^{-1}$	0.9968
14	Dobson (14)	D67H lysozyme		DLS	$1.1(1) \times 10^{-1}$ hr^{-1}	$1.41(7) \times 10^{-1}$ $\mu M^{-1}hr^{-1}$	0.9846

[a] The error bars are determined by the square root of the reduced $\chi 2$ based on a modified Levenberg-Marquardt algorithm. [b] The concentration units on k_2 were determined by multiplying by the maximum intensity and dividing by the initial starting concentration of the respective protein being measured. [c] R^2 is the coefficient of determination that gives information about the goodness of fit of the model applied (the closer this value is to 1.0, the more precise the fit is to the data). [d] These data points were corrected for the negative ellipticity values that were reported (*vide infra*).

The Relationship between a generalized form of the 1997 F-W mechanism and the 2000 Saitô and co-workers (35) Mechanism. Upon studying the protein aggregation literature and constructing Table 1 (*vide supra*), we encountered Saitô and co-worker's 2000 paper and 3-step mechanism, Scheme 3 (35). After some reflection, we in turn wondered if the 2-step F-W mechanism (Scheme 2) and Saitô and co-worker's 3-step mechanism might not be more closely similar than initially meets the eye (i.e., and after the three assumptions and simplification made by Saitô and co-workers, listed in Table 1, modifications those authors found necessary to simplify the mathematics into an equation usable for fitting experimental data). Indeed, in the Supporting Information we demonstrate the mathematical identity of the F-W and Saitô analytic equations. This means that the 3-step mechanism proposed by Saitô and co-workers in 2000 reduces to the 2-step, 1997 F-W mechanism.

Scheme 3. Saitô and coworkers' (35) mechanism for the aggregation of human calcitonin.

$$n_0A \text{ (monomers)} \overset{K_{eq}''}{\rightleftharpoons} A_{n0} \text{ (micelle)}$$

$$A_{n0} \xrightarrow{k_1''} B_{n0}$$

$$A + B_n \xrightarrow{k_2''} B_{n+1}$$

In addition, one can also see intuitively how the two mechanisms shown in Schemes 2 and 3 can lead to the same overall kinetic expression. For the case of a fast prior equilibrium that lies predominantly back to the left (towards A) in the Saitô et al. mechanism, their k_{obs}'' will equal $k_1''K_{eq}''[A]_{n-1}$ which in term will equal k_1 of the F-W mechanism, that is, $k_{obs}'' = k_1''K_{eq}''[A]_{n-1} = k_1$ (where we have substituted a simple "n" for

Saitô's "n0"(35)). That is, Saitô and co-worker's postulated pre-equilibrium of nA monomers to a A_n micellular-like aggregate—which others would just call soluble protein oligomers (40,41)—can be hidden within the k_1 of the F-W mechanism.

The mathematical identity of the F-W 2-step mechanism with the Saitô et al. 3-step mechanism (i.e., and using the assumptions Saitô et al. necessarily made to yield an equation that would allow them to analyze their protein aggregation kinetic data) means that, therefore, the Saitô et al. 3-step mechanism is, in the final analysis, a reinvention of the 2-step mechanism. It should, therefore, be replaced by the 2-step F-W mechanism. This comment is not intended to downgrade in any way the independent derivation and contribution by Professor Saitô and his co-workers; indeed, the fractional conversion ("f") form of the equation given in Table 1 is a convenient form for analyzing data (with units of reciprocal time). Moreover, it is entirely possible that the 3-step mechanism in Scheme 3 may well eventually prove to be a more accurate description of the underlying, true protein process in at least some cases.

Relevant here is a brief discussion of the original $A \rightarrow B$, $A + B \rightarrow 2B$ and more generalized form, $nA \rightarrow B_n$, $A + B_n \rightarrow B_{n+1}$ (Scheme 4), of the F-W 2-step mechanism. In Schemes S2 and S3 of the Supporting Information we show the mathematical relationship between these two. At the first level of approximation, they are equivalent with the generalized form of the equation introducing a statistical factor into k_1' and k_2'; that is, k_1 and k_2 of the original F-W mechanism are proportional, respectively, to $(n+1)k_1'[A]_{n-1}$ and $(n+1)k_2'$ of the generalized F-W mechanism under the conditions and assumptions of the derivation in the Supporting Information. In both k_2 and k_2' there is also a previously introduced (31) "scaling factor" of $(1+\chi_{growth})/2$ for spherical growth, as

further defined and discussed elsewhere (31) due to the increasing surface area of B onto which aggregation is taken to be increasingly likely as the aggregate grows. More recent (42) and upcoming (43) papers discuss a treatment of the scaling factor that is continuous with time and, therefore, allows a continuous correction to the changing fraction of B on the particle or aggregates's surface, although the specific treatment given is for the case of transition-metal nanoparticles (42,43)—one, however, that should be readily extendable by analogy to protein aggregation.

Scheme 4. The generalized form of the F-W mechanism.

$$nA \xrightarrow{k_1'} B_n$$

$$A + B_n \xrightarrow{k_2'} B_{n+1}$$

Does the F-W 2-step mechanism make physical sense for protein aggregation?

The good to excellent fits obtained for a wide range of protein aggregation data (Figures 2-12) argue that the F-W mechanism must be portraying at least some of the key features of the more complete, elementary step process of protein agglomeration. Nevertheless, the good fits still beg the question of "does the F-W mechanism make physical sense for protein aggregation?" A closely related question is what is a deeper physical description of "A" and "B" in the 2-step mechanism, $A \rightarrow B$, $A + B \rightarrow 2B$?

In their simplest forms, A is just the predominant form of the unfolded or misfolded, ready-to-aggregate protein in solution while B is more complex, B representing the protein aggregates, with all different sizes of B at a given time, t, being lumped into one average "B". Physically, B appears to be the (growing) surface area of

the aggregated protein, that is, the surface area that is primarily involved in, and leads to, aggregation as $A + B_n \rightarrow B_{n+1}$ grows to higher n. This is consistent with the description by Fink et al. in the case of α-synuclein where he describes "a partially folded intermediate is anticipated to have contiguous hydrophobic patches on its surface, which are likely to foster self-association and hence potentially fibrillation" (44)—an apt description of our "B". The broad descriptor "B" therefore covers a range anywhere from a hypothesized toxic intermediate species (45,46,47,48,49,50,51,52) all the way up to the hypothesized non-toxic fibril (45). This makes it rather apparent that much more effort needs to go into demonstrating by direct spectroscopic means what "B" is more precisely and as a function of time.

The Emerging Case for Protofibrils as the Toxic Intermediate(s). A consensus appears to be developing in the literature that intermediate protofibrils and/or soluble oligomers are the toxic species in neurodegenerative diseases (45,46,47,48,49,50,51,52). (For more references on intermediate protofibril and soluble oligomer species, please refer to the Supporting Information.) A very important implication of the fits to the F-W 2-step mechanism provided herein is that the soluble monomer-to-protofibril-to-final fibril size vs. time is in principle obtainable from the deconvoluted k_1 and k_2 and knowledge of the beginning protein concentration, A_0, so long as a final fibril geometry and final size are also known. The treatment is expected to be directly analogous to that we have developed for nanocluster size vs. time (53), but with the caveat that the shape and geometry treatment will of course be different than the treatment used for spherical nanoparticles (53). The needed equations are under development (54). In short, an implication of the fits herein deconvoluting nucleation, k_1, from surface autocatalytic

growth, k_2, is that protofibril size vs. time will be predictable (again, if the final geometry and average size can be firmly established). This in turn promises to greatly expedite the rational synthesis of fibrils of known size and, hence, their detailed study and all the insights those studies promise to provide.

Advantages and Especially Caveats / Limitations of the F-W 2-Step Kinetic Model. The F-W 2-step mechanism appears to fit a broad range of neurological protein aggregation data. The simplicity and quality of the fits, the deconvolution of the nucleation, k_1, from the growth, k_2, and the implications for fibril size vs. time studies are all extremely valuable components of the F-W kinetic model. Another significant advantage that is that the use of the 2-step F-W mechanism allows kinetic fits en route to the correct deduction of mechanism in cases where fewer data of lower precision are available (55). The above positives noted, the possible limitations of the F-W model also need to be clearly noted. Some main limitations apparent at present are:

(1) The F-W 2-step mechanism is obviously a highly condensed, oversimplified, phenomenological model of the real protein agglomeration that often consists of probably hundreds if not thousands of steps. The ability of such a 2-step model to account in at least an average way for the multi-step reaction is, in that sense, remarkable even if a significant oversimplification.

(2) The resultant k_1 and k_2 values are, therefore, averages over all of the true underlying steps. As such, important kinetic and mechanistic information must be hidden in the "average" pseudo-elementary steps (31) of $A \rightarrow B$, $A + B \rightarrow 2B$ or even the more generalized $nA \rightarrow B_n$, $A + B_n \rightarrow B_{n+1}$. Put another way, the F-W model assumes that k_1 and k_2 are independent of aggregate size (i.e., that for the nth step a given $k_n = k_{n+1} = k_{n+2}$

and so on). This is probably never exactly true, and in some cases may hide a very important increase or decrease in the rate of for example growth with aggregate particle size. Another issue here is that, as the kinetic equations in Scheme S2 (or just thinking about the following) make apparent that, during the nucleation process, [A] is approximately constant. Therefore, a higher kinetic order in [A], that is $k_1'[A]_n$, is easily hidden kinetically and can appear as an apparent $k_{1(apparent)}[A]_1$ dependence, where in fact $k_1 \propto k_1'[A]_{n-1}$. This situation is something we already have experimental evidence for in the case of transition-metal nanocluster nucleation and growth (43), and is something that we expect will be more common. This is actually an advantage when working up kinetic data, as fits to the simple, apparently first order nucleation step, $A \rightarrow B$, are seen—an advantage so long as the user is aware that further plotting of the pseudo-first-order $k_{1(apparent)}$ vs. [A] are needed to test for a higher order $nA \rightarrow B_n$ component to the reaction.

(3) The fact that all sizes of the growing aggregate are hidden behind the general descriptor "B" is a significant weakness, given the growing evidence that smaller, intermediate size fibrils may be the more toxic species (45,46,47,48,49,50,51,52). This weakness is mitigated somewhat by the fact that k_1 and k_2 can be used, along with a knowledge of the final geometry and final (average) size of the aggregates to obtain an (average) size vs. time of the aggregates (53) or fibrils. Nevertheless, a better experimental handle on the sizes of "B" vs. time is a very important avenue of future research.

The above are the main caveats and limitations of the F-W 2-step model that are apparent to us, at least at present. Other disadvantages (as well as other possible advantages?) will undoubtedly become apparent with time. The situation is perhaps best

stated by Eigen in his analysis of prion disease kinetics (59): "The model chosen must be simple enough to yield lucid results. On the other hand, it must be complex enough to comprise the essential influences that characterize nucleated aggregation."

Summary and Conclusions

The main insights from this contribution are the following:

(i) Fourteen protein aggregation data sets from primarily three different proteins (amyloid β, α-synuclein, and polyglutamine), four different monitoring methods, and nine individual research labs were fit to the F-W 2-step mechanism/model. The resultant fits separate and quantitate for the first time the nucleation and growth quantitative rate constants for the representative 14 data sets and under the assumptions noted in the Experimental section (e.g., assuming as in the prior literature that the signal intensity is directly proportional to the amount of aggregated protein).

(ii) The F-W 2-step mechanism of $A \rightarrow B$ and $A + B \rightarrow 2B$ with the generalized form, $nA \rightarrow B_n$ and $A + B_n \rightarrow B_{n+1}$, appears, therefore, to provide the simplest possible, "Ockham's razor" model for protein aggregation across a fairly wide body of data. Its most important feature is its ability to separate and quantitate nucleation from growth in a clear, well-defined, easily applied way.

(iii) The main limitations of the F-W kinetic model that are apparent at present were also noted, specifically the fact that is an oversimplified, phenomenological model of the true, multi-step process, that it provides average k_1 and k_2 values that may hide important changes in k_1 and k_2 with growing aggregate size, and that B is a catch-all for

the growing aggregate which, again, will hide any important differences in the growing aggregate and protofibrils as a function of their size.

(iv) Even given its limitations, since the F-W mechanism gives separate rate constants for nucleation and growth, therapeutic opportunities are apparent. For instance, quantitative studies aimed at understanding what factors start vs. stop nucleation (k_1) should in turn yield opportunities for better treatments—namely inhibiting the nucleation step and thus stopping aggregation from even beginning. The importance of mechanistic understanding to therapeutic approaches should not be underestimated. An interesting historical case backing this assertion is sickle cell hemoglobin polymerization where Eaton and Hofrichter noted: "Finally, kinetic studies have played a central role in understanding the pathophysiology of sickle cell disease and the design of strategies for therapy" (56).

(v) The deconvolution herein of protein aggregation into its nucleation, k_1, and autocatalytic growth, k_2, components also promises to allow more detailed studies and insights into established phenomenological linear-free-energy relationships correlating protein aggregation rates (typically after the "lag" phase; i.e., really during the k_2 growth phase) with calculated indices for hydrophobicity, charge, and intrinsic propensity to form β-sheets or α-helices (57) as well as extrinsic factors such as pH, ionic strength, and protein concentration (57). As Dobson, Chiti, and Vendruscolo (57) have noted, correlations for the crucial "lag" (really nucleation, k_1) phase need to be done, and that now is possible—that is, more rigorous correlations of both k_1 and k_2 values with the above intrinsic as well as extrinsic properties are now also possible.

(vi) In addition, a host of other, more detailed and more quantitative studies of protein agglomeration—that is, nucleation (k_1) and growth (k_2)—now become possible, specifically any and all desired studies looking at the factors that influence k_1 vs those that influence k_2.

(vii) Finally, an interesting, hypothesis results from this and our other (31,32,33) work—indeed, the most far-reaching hypothesis to result from this work in combination with our prior work on the F-W mechanism in other areas of science (31,32,33)—is that the F-W mechanism appears to be a (if not the) more general, simplest / Ockham's razor, phenomenological kinetic model for first-order phase transitions in solution (first order phase transitions being defined as "a change in state of aggregation of a system accompanied by a discontinuous change in enthalpy, entropy, and volume at a single temperature and pressure" (58)). We have shown previously that this mechanism fits nanocluster agglomeration (31,32) (see the Supporting Information for a more expanded list of nanocluster references), we have shown herein that the 2-step mechanism fits the completely different case of protein aggregation, and in other studies in progress we show that the F-W model can account for at least some solid-state kinetic data as well (33). The evidence to date suggests, then, that the F-W 2-step mechanism is the more general, Ockham's razor kinetic model or "mechanism" for first-order phase transitions in solution. We view and offer this as a hypothesis, one able to connecting a broad spectrum of previously largely unconnected literature (23,31,32,33).

Acknowledgements. This project grew out of a seminar given by R. Finke at Brandeis in November of 2005; the kind invitation from the faculty of the Brandeis Department of Chemistry, and the stimulating time at Brandeis, are a pleasure to acknowledge.

Supporting Information Available: Derivation of the analytic equation of the F-W 2-step mechanism; Derivation of Saitô and coworkers 2-step mechanism for the aggregation of calcitonin and its equivalence to the F-W 2-step mechanism; comparison of the original and generalized form of the F-W 2-step mechanism; examination of the scaling factors for the rate constants, k_1 and k_2; additional references for nanocluster formation data accounted for with the F-W mechanism; additional references for fibril intermediate species; numerical integration of kinetic data using MacKinetics, effect of added $[Zn^{2+}]$ on the rate constants of amyloid β aggregation; alternative mechanisms considered: the 3- and 4-step mechanisms; fits of amyloid β aggregation data to the 3- and 4-step mechanisms; comparison of rate constants obtained from the 2-, 3-, and 4-step mechanisms, derivation of analytic equations for autocatalysis alone and the 2-step mechanism when $[B]_0 \neq 0$; fits of polyglutamine seeded data for autocatalysis alone, the 2-step mechanism, and the 2-step mechanism with $[B]_0 \neq 0$, along with a table of resultant rate constants; fits of α-synuclein seeded data for the 2-step and 2-step with $[B]_0 \neq 0$ mechanisms with a table of resultant rate constants; correlations of the k_1 and k_2 rate constants with the $[B]_0$.

References

(1) Blennow, K., de Leon, M. J., Zetterberg, H. (2006) Alzheimer's disease, Lancet 368, 387-403.

(2) Dauer, W., and Przedborski, S. (2003) Parkinson's disease: mechanisms and models, Neuron 39, 889-909.

(3) Bates, G. P., and Benn, C. (2002) The polyglutamine diseases, in Huntington's Disease (Bates, G. P.; Harper, P. S.; Jones, L., Eds.) 3rd ed., pp 429-472, Oxford University Press, Oxford.

(4) Bieschke, J., Zhang, Q., Powers, E. T., Lerner, R. A., and Kelly, J. W. (2005) Oxidative metabolites accelerate Alzheimer's amyloidogenesis by a two-step mechanism, eliminating the requirement for nucleation, Biochemistry 44, 4977-4983.

(5) Vestergaard, M., Kerman, K., Saito, M., Nagatani, N., Takamura, Y., and Tamiya, E. (2005) A rapid label-free electrochemical detection and kinetic study of Alzheimer's amyloid beta aggregation, J. Am. Chem. Soc. 127, 11892-11893.

(6) Lu, K., Jacob, J., Thiyagarajan, P., Conticello, V. P., and Lynn, D. G. (2003) Exploiting amyloid fibril lamination for nanotube self-assembly, J. Am. Chem. Soc. 125, 6391-6393.

(7) Dong, J., Shokes, J. E., Scott, R. A., and Lynn, D. G. (2006) Modulating amyloid self-assembly and fibril morphology with Zn(II), J. Am. Chem. Soc. 128, 3540-3542.

(8) Zirah, S., Kozin, S., Mazur, A. K., Blond, A., Cheminant, M., Ségalas-Milazzo, I., Debey, P., and Rebuffat, S. (2006) Structural changes of region 1-16 of the Alzheimer disease amyloid β peptide upon zinc binding and in vitro aging, J. Biol. Chem. 281, 2151-2161.

(9) Fink, A. L. (2006) The aggregation and fibrillation of α-synuclein, Acc. Chem. Res. 39, 628-634.

(10) Sode, K., Usuzaka, E., Kobayashi, N., and Ochiai, S. (2005) Engineered α-synuclein prevents wild type and familial Parkin variant fibril formation, Biochem. Biophys. Res. Comm. 335, 432-436.

(11) Chen, S., Ferrone, F. A., and Wetzel, R. (2002) Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation, Proc. Natl. Acad. Sci. 99, 11884-11889.

(12) Wetzel, R. (2006) Kinetics and thermodynamics of amyloid fibril assembly, Acc. Chem. Res. 39, 671-679.

(13) Ban, T., Yamaguchi, K., and Goto, Y. (2006) Direct observation of amyloid fibril growth, propagation, and adaptation, Acc. Chem. Res. 39, 663-670.

(14) Dumoulin, M., Kumita, J. R., and Dobson, C. M. (2006) Normal and aberrant biological self-assembly: Insights from studies of human lysozyme and its amyloidogenic variants, Acc. Chem. Res. 39, 603-610.

(15) Oosawa, F., Asakura, S., Hotta, K., Imai, N., and Ooi, T. (1959) G-F transformation of actin as a fibrous condensation, J. Polym. Sci. 37, 323-336.

(16) Oosawa, F., and Kasai, M. (1962) A theory of linear and helical aggregations of macromolecules, J. Mol. Biol. 4, 10-21.

(17) Wegner, A., and Engel, J. (1975) Kinetics of the cooperative association of actin to actin filaments, Biophys. Chem. 3, 215-225.

(18) Frieden, C., and Goddette, D. W. (1983) Polymerization of actin and actin-like systems: evaluation of the time course of polymerization in relation to the mechanism, Biochemistry 22, 5836-5843.

(19) Eisenberg, H. (1971) Glutamate dehydrogenase: anatomy of a regulatory enzyme, Acc. Chem. Res. 4, 379-385.

(20) Thusius, D., Dessen, P., and Jallon, J.-M. (1975) Mechanism of bovine liver glutamate dehydrogenase self-association, J. Mol. Biol. 92, 413-432.

(21) Thusius, D. (1975) Mechanism of bovine liver glutamate dehydrogenase self-assembly: II. Simulation of relaxation spectra for an open linear polymerization proceeding via a sequential addition of monomer units, J. Mol. Biol. 94, 367-383.

(22) Jullien, M., and Thusius, D. (1976) Mechanism of bovine liver glutamate dehydrogenase self-assembly: III. Characterization of the association-dissociation stoichiometry with quasi-elastic light scattering, J. Mol. Biol. 101, 397-416.

(23) Morris, A. M., Watzky, M. A., and Finke, R. G., "A critical review of the literature and future directions towards achieving the goal of determining the mechanism of protein aggregation", in progress.

(24) Bishop, M. F., and Ferrone, F. A. (1984) Kinetics of nucleation-controlled polymerization, Biophys. J. 46, 631-644.

(25) Firestone, M. P., De Levie, R., and Rangarajan, S. K. (1983) On one-dimesional nucleation and growth of "living" polymers I. Homogeneous nucleation, J. Theor. Biol. 104, 535-552.

(26) Rangarajan, S. K., and De Levie, R. (1983) On one-dimensional nucleation and growth of "living" polymers II. Growth at constant monomer concentration, J. Theor. Biol. 104, 553-570.

(27) Goldstein, R. F., and Stryer, L. (1986) Cooperative polymerization reactions: analytical approximations, numerical examples, and experimental strategy, Biophys. J. 50, 583-599.

(28) Ferrone, F. A., Hofrichter, J., Sunshine, H. R., and Eaton, W. A. (1980) Kinetic studies on photolysis-induced gelation of sickle-cell hemoglobin suggest a new mechanism, Biophys. J. 32, 361-377.

(29) Ferrone, F. A., Hofrichter, J., and Eaton, W. A. (1985) Kinetics of sickle hemoglobin polymerization II. A double nucleation mechanism, J. Mol. Biol. 183, 611-631.

(30) LaMer, V. K., and Dinegar, R. H. (1950) Theory, production, and mechanism of formation of monodispersed hydrosols, J. Am. Chem. Soc. 72, 4847-4854.

(31) Watzky, M. A., and Finke, R. G. (1997) Transition metal nanocluster formation kinetic and mechanistic studies. A new mechanism when hydrogen is the reductant: slow, continuous nucleation and fast autocatalytic surface growth, J. Am. Chem. Soc. 119, 10382-10400.

(32) Besson, C., Finney, E. E., and Finke, R. G. (2005) A mechanism for transition-metal nanoparticle self-assembly, J. Am. Chem. Soc. 127, 8179-8184.

(33) Finney, E. E., and Finke, R. G., "Solid-state kinetics and the Avrami and subsequent kinetic models revisited: the Finke-Watzky 2-step nucleation and autocatalytic growth mechanism as the Ockham's razor, phenomenological mechanism for 1st-order phase transitions," in preparation.

(34) William of Ockham, 1285-1349, as cited in E. A. Moody (1967) The Encyclopedia of Philosophy, vol 7, MacMillan, New York.

(35) Kamihira, M., Naito, A., Tuzi, S., Nosaka, A. Y., and Saitô, H. (2000) Conformational transitions and fibrillation mechanism of human calcitonin as studied by high-resolution solid-state 13C NMR, Protein Science 9, 867-877.

(36) Roefs, S. P. F. M., and De Kruif, K. G. (1994) A model for the denaturation and aggregation of β-lactoglobulin, Eur. J. Biochem. 226, 883-889.

(37) Sun, S. F. Diffusion. Physical Chemistry of Macromolecules: Basic Principles and Issues, 2nd Ed.; Wiley and Sons: New York, 2004; 223-242.

(38) Westermark, P. (2005) Aspects on human amyloid forms and their fibril polypeptides, FEBS 272, 5942-5949.

(39) Golde, T. E., Dickson, D., and Hutton, M. (2006) Filling in the gaps in the Aβ cascade hypothesis of Alzheimer's disease, Current Alzheimer's Research 3, 493-504 and references therein.

(40) Kayed, R., Head, E., Thompson, J.L., McIntire, T.M., Milton, S.C., Cotman, C.W., and Glabe, C.G. (2003) Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis, Science 300, 486-489.

(41) Lomakin, A., Teplow, D.B., Kirschner, D.A., and Benedek, G.B. (1997) Kinetic theory of the fibrillogenesis of the amyloid-β protein, Proc. Natl. Acad. Sci. USA 94, 7942-7947.

(42) Schmidt, A. F., and Smirnov, V. V. (2005) Concept of "magic" number clusters as a new approach to the interpretation of unusual kinetics of the Heck reaction with aryl bromides, Top. Catal. 32, 71-75.

(43) Watzky, M.A., Finney, E. E., and Finke, R. G., Transition-metal nanocluster size vs formation time and the catalytically effective nucleus number: a mechanism-based treatment, J. Am. Chem. Soc. 130 (2008) 11959-11969.

(44) Uversky, V. N., Li, J., and Fink, A. L. (2001) Evidence for a partially folded intermediate in α-synuclein fibril formation, J. Biol. Chem. 276, 10737-10744.

(45) Ferreira, S. T., Vieira, M. N. N., and De Felice, F. G. (2007) Soluble protein oligomers as emerging toxins in Alzheimer's and other amyloid diseases, IUBMB Life 59, 332-345.

(46) Walsh, D. M., and Selkoe, D. J. (2007) Aβ oligomers—A decade of discovery, J. Neurochem. 101, 1172-1184.

(47) Harper, J. D., Wong, S. S., Lieber, C. M., and Lansbury, P. T., Jr., (1999) Assembly of Aβ amyloid protofibrils. An in vitro model for a possible early event in Alzheimer's disease, Biochemistry 38, 8972-8980.

(48) Volles, M. J., and Lansbury, P. T., Jr., (2003) Zeroing in on the pathogenic form of α-synuclein and its mechanism of neurotoxicity in Parkinson's disease, Biochemistry 42, 7871-7878.

(49) Chromy, B. A., Nowak, R. J., Lambert, M. P., Viola, K. L., Chang, L., Velasco, P. T., Jones, B. W., Fernandez, S. J., Lacor, P. N., Horowitz, P., Finch, C. E., Krafft, G. A., and Klein, W. L. (2003) Self-assembly of Aβ(1-42) into globular neurotoxins, Biochemistry 42, 12749-12760.

(50) Kaylor, J., Bodner, N., Edridge, S., Yamin, G., Hong, D.-P., and Fink, A. L. (2005) Characterization of oligomeric intermediates in α-synuclein fibrillation. FRET studies of Y125W/Y133F/Y136F α-synuclein, J. Mol. Biol. 353, 357-372.

(51) Tabner, B. J., El-Agnaf, O. M. A., Turnbull, S., German, M. J., Paleologou, K. E., Hayashi, Y., Cooper, L. J., Fullwood, N. J., and Allsop, D. (2005) Hydrogen peroxide is generated during the very early stages of aggregation of the amyloid peptides implicated in Alzheimer's disease and familial British dementia, J. Biol. Chem. 280, 35789-35792.

(52) El-Agnaf, O. M. A., Salem, S. A., Paleologou, K. E., Curran, M. D., Gibson, M. J., Court, J. A., Schlossmacher, M. G., and Allsop, D. (2006) Detection of oligomeric forms

of α-synuclein protein in human plasma as a potential biomarker for Parkinson's disease, FASEB J. 20, 419-425.

(53) Watzky, M.A., Finney, E. E., and Finke, R. G., Transition-metal nanocluster size vs formation time and the catalytically effective nucleus number: a mechanism-based treatment, J. Am. Chem. Soc. 130 (2008) 11959-11969.

(54) Watzky, M. A., Morris, A. M., and Finke, R. G. Studies in progress.

(55) Smith, S. E., Sasaki, J. M., Bergman, R. G., Mondloch, J. E., and Finke, R. G. "Platinum catalyzed Sn-Ph and Sn-Me transfer to Cp*(PMe3)IrCl2: evidence for an autocatalytic reaction pathway with an unusual preference for Sn-Me transfer," J. Am. Chem. Soc., accepted.

(56) Eaton, W. A., and Hofrichter, J. (1990) Sickle cell hemoglobin polymerization, in Advances in Protein Chemistry Vol. 40, p 159, Academic Press, San Diego.

(57) DuBay, K. F., Pawar, A. P., Chiti, F., Zurdo, J., Dobson, C. M., and Venruscolo, M. (2004) Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains, J. Mol. Biol. 341, 1317-1326.

(58) (2003) McGraw-Hill Dictionary of Scientific and Technical Terms 6th ed., p 807, McGraw-Hill, New York.

(59) Eigen, M. (1996) Prionics or the kinetic basis of prion disease, Biophys. Chem. 63, A1-A18.

(60) Prusiner, S. B. (1991) Molecular biology of prion diseases, Science 252, 1515-1522.

(61) Cohen, F. E., Pan, K.-M., Huang, Z., Baldwin, M., Fletterick, R. J., and Prusiner, S. B. (1994) Structural clues to prion replication, Science 264, 530-531.

(62) Jarrett, J. T., and Lansbury, P. T., Jr., (1993) Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and Scrapie?, Cell 73, 1055-1058.

(63) Come, J. H., Fraser, P. E., and Lansbury, P. T., Jr., (1993) A kinetic model for amyloid formation in the prion diseases: importance of seeding, Proc. Natl. Acad. Sci. USA 90, 5959-5963.

(64) Lansbury, P. T. (1994) Mechanism of scrapie replication, Science 265, 1510.

(65) Lansbury, P. T., Jr., and Caughey, B. (1995) The chemistry of scrapie infection: implications of the 'ice 9' metaphor, Chemistry and Biology 2, 1-5.

(66) Caughey, B., Kocisko, D. A., Raymond, G. J., and Lansbury, P. T., Jr., (1995) Aggregates of scrapie-associated prion protein induce the cell-free conversion of protease-sensitive prion protein to the protease-resistant state, Chem. Biol. 2, 807-817.

(67) Yamamoto, S., and Gejyo, F. (2005) Historical background and clinical treatment of dialysis-related amyloidosis, Biochim. Biophys. Acta 1753, 4-10.

(68) Pepys, M. B., Hawkins, P. N., Booth, D. R., Vigushin, D. M., Tennent, G. A., Soutar, A. K., Totty, N., Nguyen, O., Blake, C. C. F., Terry, C. J., Feest, T. G., Zalin, A. M., and Hsuan, J. J. (1993) Human lysozyme gene mutations cause hereditary systemic amyloidosis, Nature 362, 553-557.

Supporting Information for:

Fitting Neurological Protein Aggregation Kinetic Data via a 2-Step, Minimal / "Ockham's Razor" Model: the Finke-Watzky Mechanism of Nucleation Followed by Autocatalytic Surface Growth

Aimee M. Morris, Murielle A. Watzky, Jeffrey N. Agar, and Richard G. Finke

Scheme S1. The derivation of the analytic equation for the F-W mechanism. This derivation in the original reference (1) contains typesetting errors that were not caught at the time of publication and should read:

$$-\frac{d[A]}{dt} = k_1[A] + k_2[A]([A]_0 - [A])$$

$$\int_0^t \frac{-d[A]}{[A](k_1 + k_2([A]_0 - [A]))} = \int_0^t dt$$

$$\frac{1}{k_1 + k_2[A]_0} * \ln\left[\frac{k_1 + k_2([A]_0 - [A])}{[A]}\right]_0^t = t$$

$$\ln\left[\frac{(k_1 + k_2([A]_0 - [A]))[A]_0}{k_1[A]}\right] = (k_1 + k_2[A]_0)t$$

Scheme S2. Derivation showing the equivalence of the Finke-Watzky (F-W) 2-step mechanism to Saitô and coworkers' (2) 3-step mechanism for the aggregation of calcitonin.

Saitô's equation:

$$f = \frac{\rho(e^{(1+\rho)kt} - 1)}{1 + \rho e^{(1+\rho)kt}}$$

By definition:

$$\rho = \frac{k_1}{k}, \; k = k_2 a, \; a = [A]_0,$$

Substitution gives:

$$f = \frac{\frac{k_1}{k_2[A]_0}\left(e^{\left(1+\frac{k_1}{k_2[A]_0}\right)k_2[A]_0 t} - 1\right)}{1 + \frac{k_1}{k_2[A]_0} e^{\left(1+\frac{k_1}{k_2[A]_0}\right)k_2[A]_0 t}}$$

Simplifying:

$$f = \frac{\frac{k_1}{k_2[A]_0}\left(e^{(k_1+k_2[A]_0)t} - 1\right)}{1 + \frac{k_1}{k_2[A]_0} e^{(k_1+k_2[A]_0)t}}$$

By definition:

$$[A]_t = [A]_0(1 - f) \Rightarrow f = 1 - \frac{[A]_t}{[A]_0}$$

Substituting for f gives:

$$1 - \frac{[A]_t}{[A]_0} = \frac{\frac{k_1}{k_2[A]_0}\left(e^{(k_1+k_2[A]_0)t} - 1\right)}{1 + \frac{k_1}{k_2[A]_0} e^{(k_1+k_2[A]_0)t}}$$

Solving for $[A]_t$:

$$[A]_t = [A]_0\left(1 - \frac{\frac{k_1}{k_2[A]_0}\left(e^{(k_1+k_2[A]_0)t} - 1\right)}{1 + \frac{k_1}{k_2[A]_0} e^{(k_1+k_2[A]_0)t}}\right)$$

Simplification gives:

$$[A]_t = \frac{\frac{k_1}{k_2} + [A]_0}{1 + \frac{k_1}{k_2[A]_0} e^{(k_1+k_2[A]_0)t}}$$

This last equation is the Finke - Watzky analytic equation corresponding to the 2 - step mechanism.

Scheme S3. Comparison of the Original and Generalized Form of the F-W 2-Step Mechanism.

The original form of the equation is

$$A \xrightarrow{k_1} B$$

$$A + B \xrightarrow{k_2} 2B$$

where the sum reaction is

$$2A \rightarrow 2B$$

which gives the rate equation:

$$\frac{-d[A]}{dt} = k_1[A] + k_2[A][B]$$

Whereas the generalized form of the 2-step mechanism is given by

$$nA \xrightarrow{k_1'} B_n$$

$$A + B_n \xrightarrow{k_2'} B_{n+1}$$

where the sum reaction is

$$(n+1)A \rightarrow B_{n+1}$$

which gives the rate equation:

$$\frac{\frac{-d[A]}{dt}}{(n+1)} = k_1'[A]^n + k_2'[A][B_n]$$

At constant [A] (during nucleation), and with the zeroth-order approximation that one can equate [B] with [B_n] (*i.e.*, by treating B as one species), then k_1 vs k_1' and k_2 vs k_2' are related by

$$k_1 \approx (n+1)k_1'[A]^{n-1}$$

$$k_2 \approx (n+1)k_2'$$

Hence, there is just a statistical factor difference of (n+1) between the k_2 and k_2' rate constants and a similar (n+1) factor, as well as a multiplicative $[A]^{n-1}$ concentration factor, between the k_1 and k_1' rate constants.

Scheme S4. Examination of the Scaling Factors for the Rate Constants, k_1 and k_2.

Beginning with the generalized form of the 2-step mechanism

$$nA \xrightarrow{k_1'} B_n$$

$$A + B_n \xrightarrow{k_2'} B_{n+1}$$

where the sum reaction is

$$(n+1)A \rightarrow B_{n+1}$$

There are several factors suggested to increase the aggregation but for argument sake, if you assume only one factor contributes to aggregation and apply Fink's (3) definition of B to be a partially folded intermediate which has aggregative (i.e., hydrophobic, charged or other (4)) patches for increased fibrillation, then the following can be obtained:

$$nA \xrightarrow{k_1'} B_{n(\text{aggregative patches})}$$

$$A + B_{n(\text{aggregative patches})} \xrightarrow{k_2'} B_{n+1(\text{aggregative patches})}$$

In a zeroth-order treatment that mirrors our earlier one for transition-metal nanoclusters (1), we define $\chi_{(\text{nucleation or growth})}$ as the fraction of increased area of aggregative patches gained in a reaction step of nucleation or growth:

$$\chi_{(\text{nucleation or growth})} = \frac{\text{increase in area of aggregative patches}}{\text{total area of aggregative patches}}$$

The generalized form can be rewritten then for protein aggregation as:

$$A \xrightarrow{k_1'} B_{\chi_{\text{nucleation}}}$$

$$A + B_{\chi_{\text{nucleation}}} \xrightarrow{k_2'} B_{(\chi_{\text{nucleation}} + \chi_{\text{growth}})}$$

where the sum reaction is

$$2A \rightarrow B_{(\chi_{\text{nucleation}} + \chi_{\text{growth}})}$$

If you assume that B is needed in order to form aggregative patches then (as before (1)) a reasonable assumption is that $\chi_{nucleation} \approx 1$ which gives:

$$2A \rightarrow B_{(1+\chi_{growth})}$$

or rewritten

$$2A \rightarrow (1+\chi_{growth})B$$

since there are $(1 + \chi_{growth})$ number of aggregative patches in this given example. From this we obtain:

$$-\frac{1}{2}\frac{d[A]}{dt} = \frac{1}{1+\chi_{growth}}\frac{d[B]}{dt}$$

$$\frac{d[B]}{dt} = -\frac{1+\chi_{growth}}{2}\frac{d[A]}{dt}$$

$$[B]_t = \frac{1+\chi_{growth}}{2}([A]_0 - [A]_t)$$

so the rate equation becomes

$$-\frac{d[A]}{dt} = k_1[A] + k_2[A][B]$$

$$-\frac{d[A]}{dt} = k_1[A] + k_2\frac{1+\chi_{growth}}{2}[A]([A]_0 - [A]_t)$$

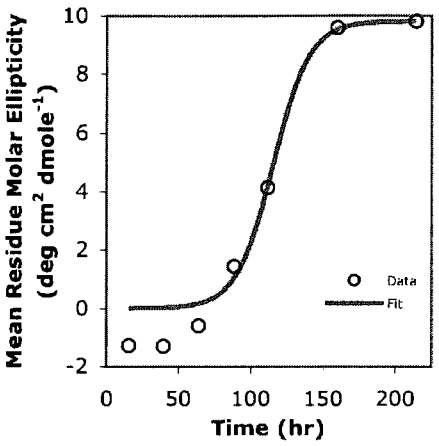
Therefore, the value of k_2 has a scaling factor of $(1 + \chi_{growth})/2$.

For an expanded list of nanocluster formation references in which the F-W mechanism has been demonstrated to account for the kinetic data, see references 1,5,6,7,8,9.

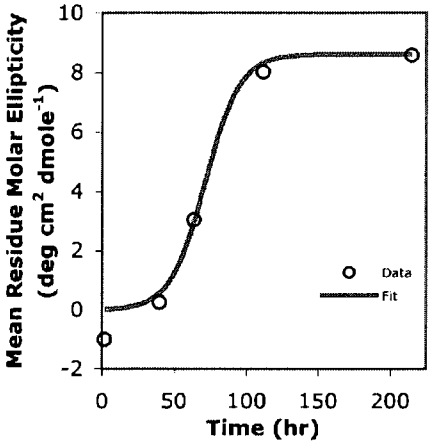
For other selected key references on fibril intermediates which includes soluble oligomers and protofibrils for which there was not space in the main text, please refer to references 10,11,12,13,14 below.

Figure S1 below shown the data for the uncorrected amyloid β aggregation from Lynn and coworkers' 2006 report (15) along with the fits to the F-W 2-step mechanism. These data were corrected in Figure 5 of the main text using the assumption that, if a blank of the sample were run, then the ellipticity values observed would have been negative throughout the measurement. The raw data is also provided below in Table S1, along with the rate constants and correlations for the fits from both the uncorrected and corrected data, so that the interested reader can compare the fits for themselves.

(a)



(b)



(c)

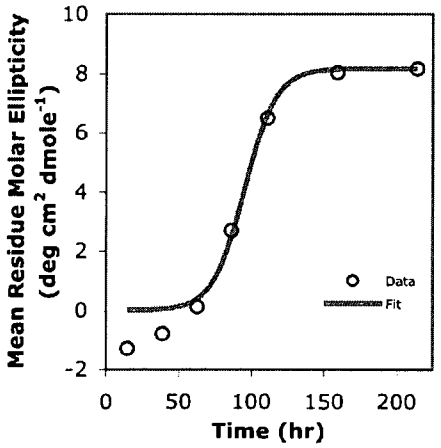


Figure S1. Digitized data of Lynn and coworkers' (15) amyloid β peptide measured as a function of added Zn^{2+} and fit to the F-W 2-step mechanism. The authors speculate that Zn^{2+} binding sites exist along the β sheets of amyloid β peptides; therefore, they measured the aggregation as a function of added [Zn^{2+}] and observed a reduced nucleation time with increased [Zn^{2+}]. Figure (a) is for 1mM of the peptide alone with $k_1 = 1(3) \times 10^{-5}$ hr^{-1} and $k_2 = 8(4) \times 10^{-3}$ $mM^{-1} \cdot hr^{-1}$, (b) is for [Zn^{2+}]:[peptide] of 0.2 with $k_1 = 2(4) \times 10^{-4}$ hr^{-1} and $k_2 = 1.0(4) \times 10^{-3}$ $mM^{-1} \cdot hr^{-1}$, and (c) represents a [Zn^{2+}]:[peptide] of 0.4 with $k_1 = 1(3) \times 10^{-5}$ hr^{-1} and $k_2 = 1.1(2) \times 10^{-2}$ $mM^{-1} \cdot hr^{-1}$. The induction periods are less well fit in this data set vs the fits shown in Figure 5 of the main text.

Table S1. Comparison of the rate constants and coefficient of determination for the uncorrected and corrected data from Lynn's (15) amyloid β aggregation system measured by circular dichroism.

	Data Corrected?	k_1 (hr^{-1})	k_2 (mM^{-1} hr^{-1})	R^2
1 mM peptide	No	$1(3) \times 10^{-5}$	$8(4) \times 10^{-3}$	0.9707
	Yes	$1.0(5) \times 10^{-4}$	$5.1(5) \times 10^{-3}$	0.9973
[Zn]:[peptide]=0.2	No	$2(4) \times 10^{-4}$	$1.0(4) \times 10^{-3}$	0.9839
	Yes	$8.3(3) \times 10^{-4}$	$6.55(9) \times 10^{-3}$	1.0000
[Zn]:[peptide]=0.4	No	$1(3) \times 10^{-5}$	$1.1(3) \times 10^{-2}$	0.9762
	Yes	$1.6(4) \times 10^{-4}$	$7.1(4) \times 10^{-3}$	0.9988

Effect of Added [Zn^{2+}] to the Rate Constants of Amyloid β Aggregation. From the fits shown in Figure 5a-c of the main text, the effects on k_1 and k_2 of 0, 0.2, and 0.4 mM added Zn^{2+} were examined, Figure S2. The results in Figure S2 show that there is no clean trend in the k_1 values, and little effect within what is probably the true experimental error on the k_2 values, at least within the range tested of 0-0.4 mM added Zn^{2+}.

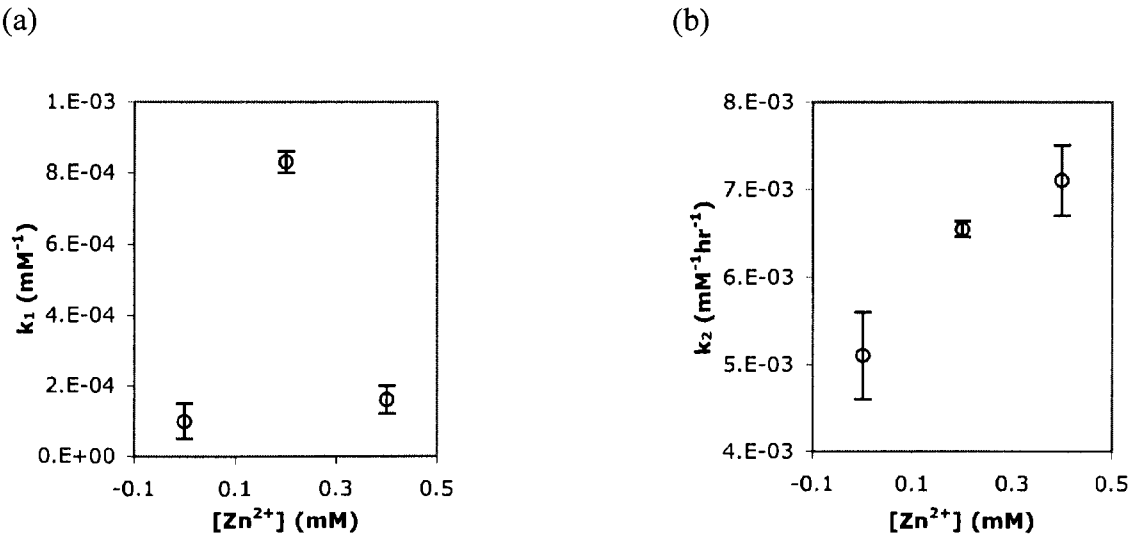


Figure S2. The effect of added Zn^{2+} of Lynn's (15) amyloid β aggregation on the (a) k_1 and (b) k_2 values obtained from the corrected data in Figure 5 of the main text.

Data Analysis and Curve Fitting Following the Loss of [A]. Data were extracted from published kinetic curves using Engauge Digitizer 2.12. In the cases presented below, the y-axis was inverted vs. that presented in the main text. That is, instead of displaying the data as it is reported in the literature with the formation of product [B] vs. time, the data is displayed and fit as the loss of reactant [A] vs. time. These two display modes are of course equivalent so long a mass balance, $[A]_o = [A]_t + [B]_t$ are obeyed throughout the reaction, see equations 2 and 3 of the main text.

Numerical integration curve fitting for the 2-, 3-, and 4-step mechanisms was accomplished with MacKinetics (version 0.9.1b) as detailed in our previous publications (6,9) in order to compare the residual values from the fits to each of the three mechanisms. The MacKinetics freeware used for numerical integration was downloaded at http://members.dca.net/leipold/mk/advert.html. The availability and support of MacKinetics has become problematic in the time since we first obtained MacKinetics, however. Hence, we have listed elsewhere (9) the available numerical integration curve-fitting software we can locate, and also recommend interested readers to take a look at the freeware Copasi available at http://www.copasi.org. Table S2 contains the rate constants obtained for each of the kinetic data sets fit in Figures S4-S7.

(a) (b)

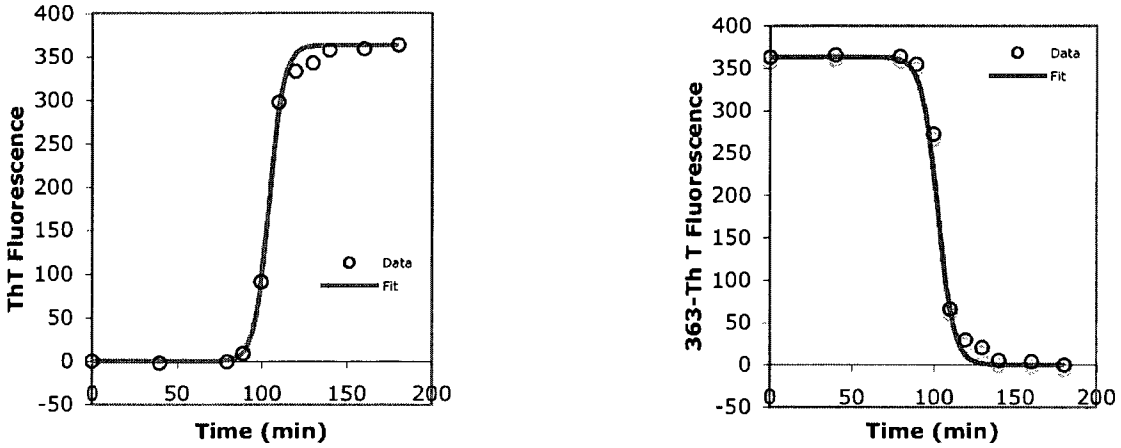


Figure S3. An example of plotting the aggregation data as (a) the formation of [B] vs. time, and (b) the loss of [A] vs. time. In both cases the rate constants obtained are identical with $k_1=1(3) \times 10^{-11}$ and $k_2=6(1) \times 10^{-4}$ as required since the data were interconverted via subtraction of the maximal intensity from each data point. As expected, the R^2 values are also equivalent, 0.9968 and $R^2=0.9968$ for the curves (a) and (b), respectively.

Alternative Mechanisms Considered: the 3- and 4-Step Mechanisms. The 3- and 4-step mechanisms (6,8,9) of nanocluster formation and subsequent agglomeration are shown in Schemes S5 and S6. The 4-step mechanism is a relatively rare example of a mechanism with two autocatalytic steps as discussed further elsewhere (8,9).

Scheme S5. The 3-step mechanism for transition-metal nanocluster formation and bimolecular agglomeration (6).

$$A \xrightarrow{k_1} B$$

$$A + B \xrightarrow{k_2} 2B$$

$$B + B \xrightarrow{k_3} C$$

Scheme S6. The 4-step mechanism for nanocluster formation that involves two autocatalytic steps (8).

$$A \xrightarrow{k_1} B$$

$$A + B \xrightarrow{k_2} 2B$$

$$B + B \xrightarrow{k_3} C$$

$$B + C \xrightarrow{k_4} 1.5C$$

Since we only disprove in science in general and in mechanism in particular, it was important to see if the 3- or 4-step mechanism might be able to provide superior fits to kinetic data that were less then very well fit by the 2-step F-W mechanism. Figures S4-S7 below reveal that three of the data sets originally in Figures 2-12 of the main text can also be fit by the 3- or 4-step mechanisms. However, by comparison of the residual values for the 2-, 3-, and 4-step mechanisms, Table S2, and in line with a conditional exclusion by Ockham's razor (16) of the more complicated mechanisms in the absence of compelling evidence for those more complex mechanisms, the simpler 2-step F-W model was used to analyze the data below and as shown in Figures 2-12 of the main text.

(a) (b)

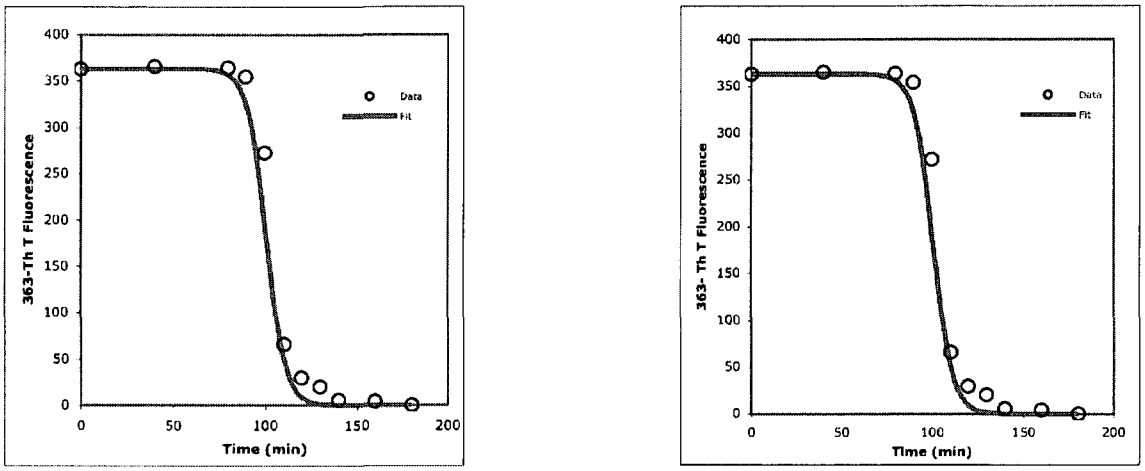


Figure S4. Goto and coworkers' (17) digitized data for the ultrasonification-induced aggregation of β2-microglobulin monitored by ThT fluorescence and fit to the (a) 3-step and (b) 4-step mechanisms developed by the Finke group.

(a) (b)

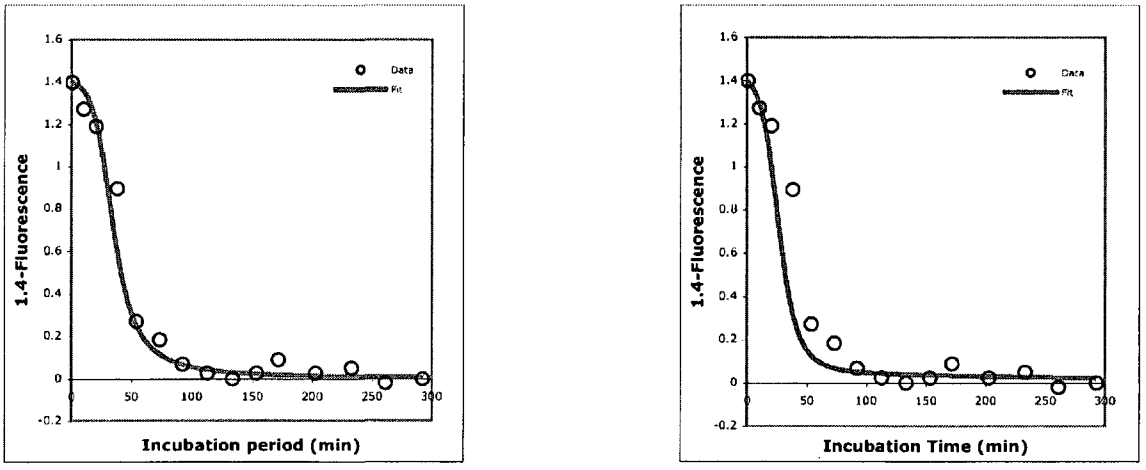


Figure S5. Vestergaard et al's (18) amyloid β (Aβ-40) peptide measured ThT fluorescence and fit to the (a) 3-step and (b) 4-step mechanisms.

(a) (b)

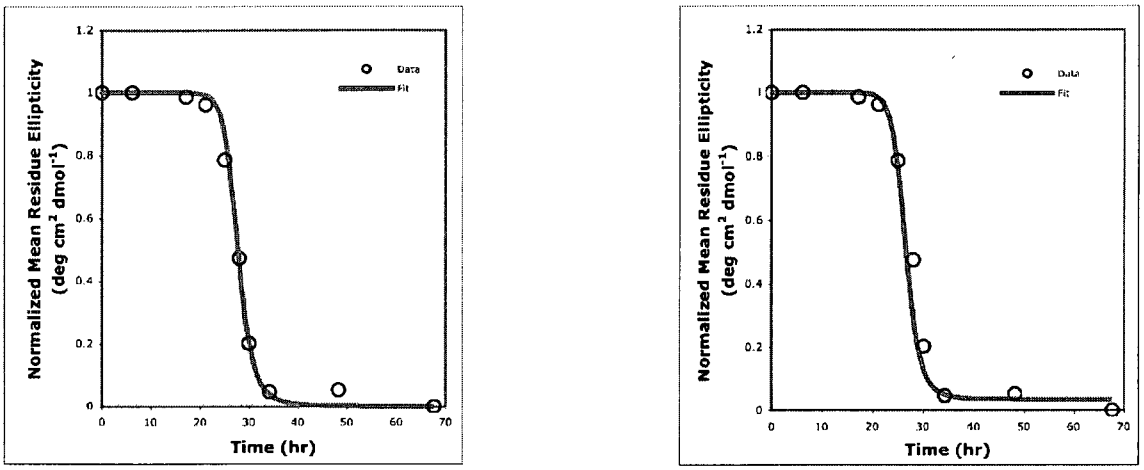


Figure S6. Lynn and coworkers' (19) amyloid β peptide measured by circular dichroism (CD) and fit to the (a) 3-step and (b) 4-step mechanisms.

(a) (b)

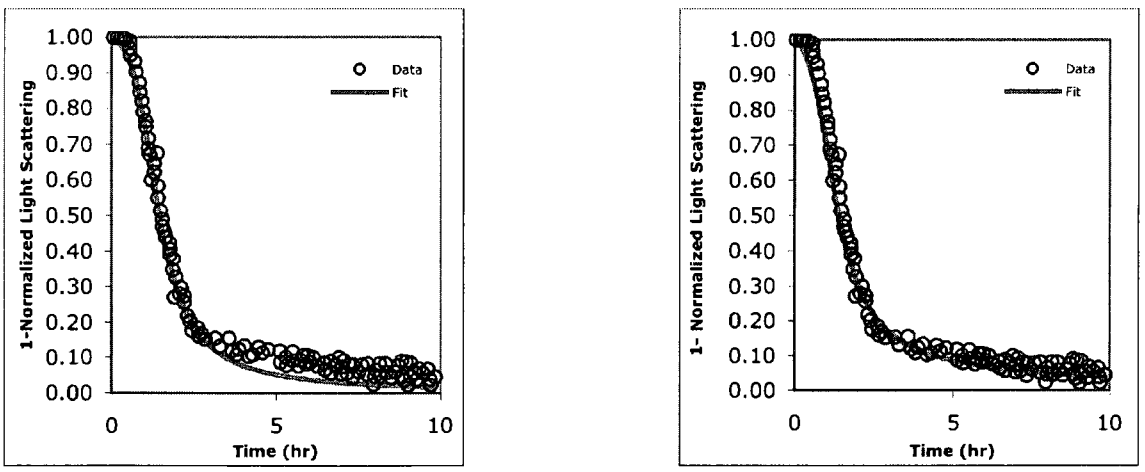


Figure S7. Dobson and coworkers' (20) in vitro D67H human lysozyme variant fit to the (a) 3-step and (b) 4-step mechanisms.

Figure S7 is better fit to the 3- and 4-step mechanisms for the latter half of the data. However, what is not easily observed here, but is important in choosing the best fit, is the undercutting of the first half of the data seen in the 3- and 4-step fits. This is a case

where the 2-step mechanism (Figure 12 of the main text) fits better for the first half of the data and the 3- and 4-step mechanisms fit somewhat better for the second half of the data.

While the 3- and 4-step mechanism appear to fit the second half of the data better (especially in the case of Figure S7; see Figure 12 of the main text for the 2-step fit), the residual values from each of the 2-, 3-, and 4-step fits show little difference, entry 4 of Table S2. Moreover, the the k_1 and k_2 values for the 2-, 3-, and 4-step mechanisms are similar in each case (entries 1-4 of Table S2). Hence, again we chose the simplest, Ockham's razor (16) 2-step mechanism to fit the data reported in the main text.

Table S2. Comparison of residual values obtained from the 2-, 3-, and 4-step mechanistic fits to the data presented in Figures S4-S7.

Entry	Reference	System		2-Step Fit	3-Step Fit	4-Step Fit
1	Goto (17)	β-microglobulin	k_1	2.5×10^{-10}	7.7×10^{-10}	1.0×10^{-9}
			k_2	5.5×10^{-4}	5.3×10^{-4}	5.2×10^{-4}
			k_3		8.3×10^{-10}	1.5×10^{-9}
			k_4			1.0×10^{-10}
			Residual	0.03497	0.05691	0.05771
2	Vestergaard (18)	Amyloid-β	k_1	7.0×10^{-3}	1.3×10^{-3}	3.3×10^{-3}
			k_2	3.4×10^{-2}	1.0×10^{-1}	1.0×10^{-1}
			k_3		4.9×10^{-2}	2.3×10^{-2}
			k_4			8.3×10^{-2}
			Residual	0.06901	0.05689	0.1116
3	Lynn (19)	Amyloid-β	k_1	7.5×10^{-10}	1.1×10^{-9}	1.1×10^{-9}
			k_2	7.5×10^{-1}	7.3×10^{-1}	7.6×10^{-1}
			k_3		1.1×10^{-1}	1.1×10^{-1}
			k_4			8.2×10^{-1}
			Residual	0.02117	0.02735	0.08707
4	Dobson (20)	D67H lysozyme	k_1	1.1×10^{-1}	4.9×10^{-2}	8.5×10^{-2}
			k_2	1.8	3.2	3.0
			k_3		2.4	2.7
			k_4			1.1
			Residual	0.08148	0.07403	0.11009

Fitting Seeded Protein Aggregation Data. Some of the protein aggregation kinetic data that we found in the literature relevant to neurodegenerative diseases is for

seeded protein agglomeration. Those data were digitized with Engauge Digitizer 2.12 and fit using Origin 7.0. Each data set was then fit (i) as a control to the original F-W 2-step mechanism, which assumes an initial concentration of B=0, via equations (1) or (2) of the main text, and then also (ii) to the analytical equation shown below in equation (S1) or (S2), depending upon the way the data was reported in the literature, S1 and S2 being the equations for the F-W 2-step mechanism, but with $[B]_0 \neq 0$ (i.e., for the seeded case). The data sets were also fit (iii) by a hypothetical mechanism of autocatalysis alone, namely just the A + B → 2B step, equations (S3) and (S4). The derivation of equations (S1) and (S2) is shown below in Scheme S7. The derivation of equations (S3) and (S4) can be found in several textbooks, for example reference (21).

$$[A]_t = \frac{(k_1 + k_2([A]_0 + [B]_0))[A]_0}{k_2[A]_0 + k_1 e^{(k_1 + k_2([A]_0 + [B]_0))t}} \tag{S1}$$

$$[B]_t = ([A]_0 - [B]_0) - \frac{(k_1 + k_2([A]_0 + [B]_0))[A]_0}{k_2[A]_0 + k_1 e^{(k_1 + k_2([A]_0 + [B]_0))t}} \tag{S2}$$

$$[A]_t = ([A]_0 + [B]_0) - \frac{[A]_0 + [B]_0}{1 + \frac{[A]_0}{[B]_0} e^{-k([A]_0 + [B]_0)t}} \tag{S3}$$

$$[B]_t = \frac{[A]_0 + [B]_0}{1 + \frac{[A]_0}{[B]_0} e^{-k([A]_0 + [B]_0)t}} \tag{S4}$$

Scheme S7. Derivation of the 2-step mechanism with $[B]_0 \neq 0$.

$$-\frac{d[A]}{dt} = k_1[A] + k_2[A][B]$$

By mass balance and assuming that $[B]_0 \neq 0$:

$$[B] = [B]_0 + [A]_0 - [A]$$

$$-\frac{d[A]}{dt} = k_1[A] + k_2[A]([B]_0 + [A]_0 - [A])$$

$$\int_0^t \frac{-d[A]}{k_1[A] + k_2[A]([B]_0 + [A]_0 - [A])} = \int dt$$

$$\frac{1}{k_1 + k_2([A]_0 + [B]_0)} \ln\left(\frac{k_1 + k_2([A]_0 + [B]_0 - [A])}{[A]}\right)\Big]_0^t = t$$

$$\ln\left(\frac{(k_1 + k_2([A]_0 + [B]_0 - [A]))[A]}{k_1[A]}\right) = (k_1 + k_2([A]_0 + [B]_0))t$$

$$\frac{(k_1 + k_2([A]_0 + [B]_0 - [A]))[A]}{k_1[A]} = e^{(k_1 + k_2([A]_0 + [B]_0))t}$$

$$[A]_t = \frac{(k_1 + k_2([A]_0 + [B]_0))[A]_0}{k_2[A]_0 + k_1 e^{(k_1 + k_2([A]_0 + [B]_0))t}}$$

Which by mass balance of:

$$[A]_t = [A]_0 - [B]_0 - [B]_t$$

can also be expressed as:

$$[B]_t = [A]_0 - [B]_0 - \frac{(k_1 + k_2([A]_0 + [B]_0))[A]_0}{k_2[A]_0 + k_1 e^{(k_1 + k_2([A]_0 + [B]_0))t}}$$

The data and fits shown below in Figure S8 correspond to the data found in Figure 10 of the main text for the unseeded case. What is observed in Figure S8 is that each of the proposed mechanisms seems to fit the seeded aggregation kinetic data, but by examination of the R^2 values in Table S2, the 2-step mechanism with $[B]_0 \neq 0$ provides the best overall fit.

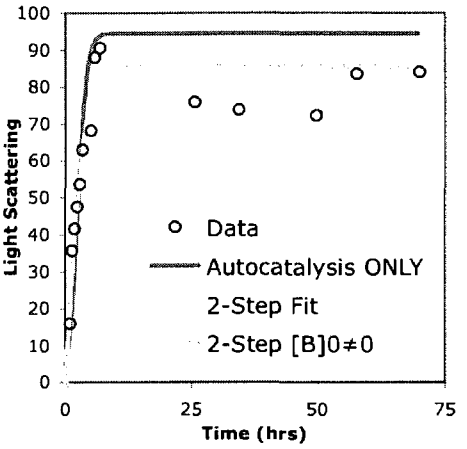


Figure S8. Wetzel's (22) seeded protein aggregation kinetic data measured by dynamic light scattering and fit to autocatalysis only (pink), the F-W 2-step mechanism (yellow), and the 2-step mechanism with $[B]_0 \neq 0$ (teal).

Table S3. Rate constants and R^2 values obtained for the data in shown in Figure S8

Autocatalysis only	$k=1.2(1) \times 10^{-2}\ hr^{-1}$	$R^2=0.75718$
Classic 2-step mechanism	$k_1=2.3(9) \times 10^{-1}\ hr^{-1}$	$R^2=0.89385$
	$k_2=3(3) \times 10^{-3}\ \mu M^{-1}hr^{-1}$	
2-step mechanism with $[B]_0 \neq 0$	$k_1=2.9(8) \times 10^{-1}\ hr^{-1}$	$R^2=0.92885$
	$k_2=3(3) \times 10^{-3}\ \mu M^{-1}hr^{-1}$	

A second set of seeded protein aggregation kinetic data was also fit. Note, in the original publication (23) the absorbance data does not reach zero. Therefore, it was assumed the that residual absorbance observed is some sort of background absorbance and, hence, that the data should be corrected by subtracting the difference between the last leveled off data points and zero in the cases of 0.01, 0.1, 1, and 10% seeded cases. (The data could not be corrected in the case of 0.001% seeded because there is no final leveling off point present, i.e., the growth phase is not completely shown in the original paper (23).) Both the raw data and corrected data were analyzed by both the 2-step mechanism and the 2-step mechanism with $[B]_0 \neq 0$; the results can be seen in Table S4. The corrected data gives the better fit and can be seen in Figure S9.

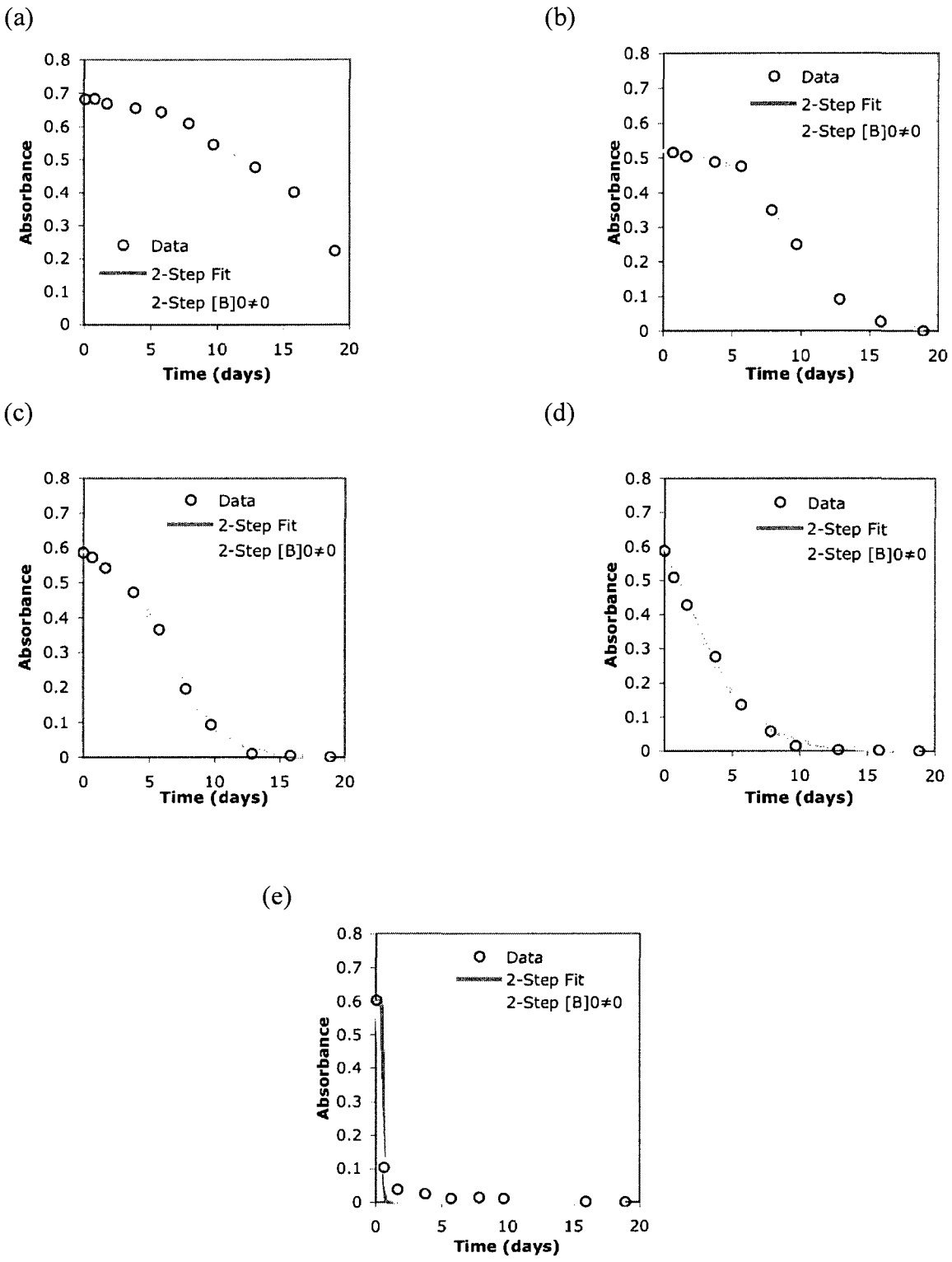


Figure S9. Biere's (23) (a) 0.001%, (b) 0.01%, (c) 0.1%, (d) 1%, and (e) 10% seeded α-synuclein data after correction for the non-zero residual absorbance and curve-fit to the F-W 2-step mechanism (as a control) and then to the 2-step mechanism with $[B]_0 \neq 0$.

Both the 2-step mechanism and the 2-step mechanism with $[B]_0 \neq 0$ give the same rate constants within experimental error for the seeded data with $\leq 0.1\%$ seed. However, the baseline-corrected data with a seed concentration $> 0.1\%$ is better fit by the 2-step mechanism with $[B]_0 \neq 0$. (When the uncorrected data were analyzed, negative rate constants were obtained for both the cases of 1% and 10% seeded shown in Table S4.) The 2-step mechanism with $[B]_0 \neq 0$ is not able to account for the data at higher seed concentrations; our belief is that the mechanism may change to a presently unknown mechanism at seed concentrations at or above certain concentration levels.

Table S4. The rate constants, k_1 and k_2, obtained from the uncorrected data as originally published (23) and the corrected data shown in Figure S9.

% Seeded	Fit	Corrected?	k_1 ($days^{-1}$)	k_2 ($\mu M^{-1}days^{-1}$)	R^2
0.001	2-step	No	$8(2) \times 10^{-3}$	$2.8(3) \times 10^{-1}$	0.98115
	2-step + $[B]_0 \neq 0$	No	$8(2) \times 10^{-3}$	$2.8(3) \times 10^{-1}$	0.98115
0.01	2-step	No	$1.6(4) \times 10^{-2}$	$3.0(5) \times 10^{-1}$	0.97507
		Yes	$4.8(9) \times 10^{-3}$	$9.2(5) \times 10^{-1}$	0.99808
	2-step + $[B]_0 \neq 0$	No	$1.6(4) \times 10^{-2}$	$3.0(5) \times 10^{-1}$	0.97508
		Yes	$4.8(9) \times 10^{-3}$	$9.2(5) \times 10^{-1}$	0.99808
0.1	2-step	No	$5(1) \times 10^{-2}$	$2.7(8) \times 10^{-1}$	0.97505
		Yes	$2.1(3) \times 10^{-2}$	$8.1(5) \times 10^{-1}$	0.99849
	2-step + $[B]_0 \neq 0$	No	$5(1) \times 10^{-2}$	$2.7(8) \times 10^{-1}$	0.97514
		Yes	$2.1(3) \times 10^{-2}$	$8.1(5) \times 10^{-1}$	0.99843
1	2-step	No	$2.2(3) \times 10^{-1}$	$-1.8(9) \times 10^{-1}$	0.98231
		Yes	$1.6(2) \times 10^{-1}$	$4(1) \times 10^{-1}$	0.99590
	2-step + $[B]_0 \neq 0$	No	$3.6(5) \times 10^{-1}$	$-5.7(8) \times 10^{-1}$	0.95212
		Yes	$1.6(2) \times 10^{-1}$	$4(1) \times 10^{-1}$	0.99538
10	2-step	No	3.0(8)	-5(1)	0.94538
		Yes	$2(20{,}000) \times 10^{-6}$ [a]	50(300,000) [a]	0.99177
	2-step + $[B]_0 \neq$	No	9.9(7)	-15(1)	0.99105
		Yes	1.0(3)	6(2)	0.99196

[a]These resulting errors that were 100-1000 times larger than the actual rate constants indicates the inability of the mechanism to account for the data.

The rate constants obtained from the fits to the corrected absorbance data (as displayed in the five graphs in Figure S9) were plotted as a function of the added seed in Figure S10. The k_1 dependence on the concentration of B_0 is first-order (i.e., if we did not need the logarithmic scale in x-axis of Figure S10a to display the data in the space allowed, then a linear k_1 vs concentration plot would be apparent). The k_2 dependence on added seed does not show any discernable dependence of the concentration of added seed, B_0.

(a) (b)

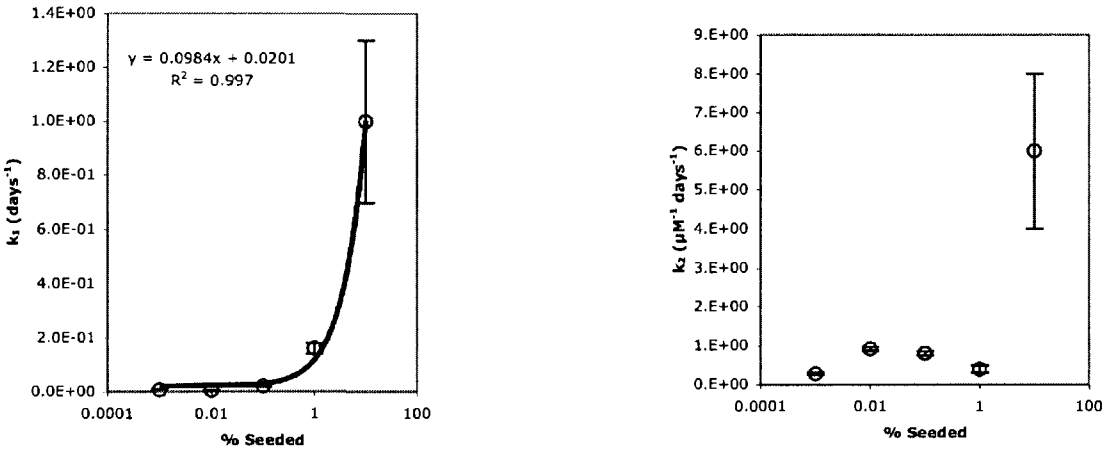


Figure S10. (a) k_1 and (b) k_2 correlations with the concentration of added seed for the corrected α-synuclein aggregation data from ref (23). Note the logarithmic x-axis in each graph.

Our current working hypothesis as far as the seeded kinetic data sets go is that the 2-step mechanism is able to account for the data up to a certain level of seed concentration after which it appears that the aggregation mechanism *may* be changing to

some at present, unknown pathway. Obviously, further studies of seeded protein aggregation will be needed before definitive statements about the underlying mechanism(s) can be made.

References for the Mathematical Derivations:

[1] Watzky, M. A., and Finke, R. G. (1997) Transition metal nanocluster formation kinetic and mechanistic studies. A new mechanism when hydrogen is the reductant: slow, continuous nucleation and fast autocatalytic surface growth, *J. Am. Chem. Soc. 119*, 10382-10400.

[2] Kamihira, M., Naito, A., Tuzi, S., Nosaka, A. Y., and Saitô, H. (2000) Conformational transitions and fibrillation mechanism of human calcitonin as studied by high-resolution solid-state ^{13}C NMR, *Protein Science 9*, 867-877.

[3] Uversky, V. N., Li, J., and Fink, A. L. (2001) Evidence for a partially folded intermediate in α-synuclein fibril formation, *J. Biol. Chem. 276*, 10737-10744.

[4] DuBay, K. F., Pawar, A. P., Chiti, F., Zurdo, J., Dobson, C. M., and Venruscolo, M. (2004) Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains, *J. Mol. Biol. 341*, 1317-1326.

References for Nanocluster Studies:

[5] Aiken III, J. D., and Finke, R. G. (1998) Nanocluster Formation Synthetic, Kinetic, and Mechanistic Studies. The Detection of, and Then Methods To Avoid, Hydrogen Mass-Transfer Limitations in the Synthesis of Polyoxoanion- and Tetrabutylammonium-Stabilized, Near-Monodisperse 40 ± 6 Å Rh(0) Nanoclusters, *J. Am. Chem. Soc. 120*, 9545-9554.

[6] Widegren, J. A.; Aiken III, J. D., Özkar, S., and Finke, R. G. (2001) Additional Investigations of a New Kinetic Method To Follow Transition-Metal Nanocluster Formation, Including the Discovery of Heterolytic Hydrogen Activation in Nanocluster Nucleation Reactions, *Chem. Mater. 13*, 312-324.

[7] Hornstein, B. J., and Finke, R. G. (2004) Transition-Metal Nanocluster Kinetic and Mechanistic Studies Emphasizing Nanocluster Agglomeration: Demonstration of a Kinetic Method That Allows Monitoring of All Three Phases of Nanocluster Formation and Aging, *Chem. Mater. 16*, 139-150. See also correction Hornstein, B. J., and Finke, R. G. (2004) *Chem. Mater. 16*, 3972-3972.

[8] Besson, C., Finney, E. E., and Finke, R. G. (2005) A Mechanism for Transition-Metal Nanoparticle Self-Assembly, *J. Am. Chem. Soc. 127*, 8179-8184.

[9] Besson, C., Finney, E. E., and Finke, R. G. (2005) Nanocluster Nucleation, Growth, and Then Agglomeration Kinetic and Mechanistic Studies: A More General, Four-Step Mechanism Involving Double Autocatalysis, *Chem. Mater. 17*, 4925-4938.

References for Protein Intermediate Protofibrils/Soluble Oligomers:

[10] Walsh, D.M., Lomakin, A., Benedek, G.B., Condron, M.M., and Teplow, D.B. (1997) Amyloid β-protein fibrillogenesis: Detection of a protofibrillar intermediate, *J. Biol. Chem. 272,* 22364-22372.

[11]Walsh, D.M., Hartley, D.M., Kusumoto, Y., Fezoui, Y., Condron, M.M., Lomakin, A., Benedek, G.B., Selkoe, D.J., and Teplow, D.B. (1999) Amyloid β-protein fibrillogenesis: Structure and biological activity of protofibrillar intermediates, *J. Biol. Chem. 274,* 25945-25952.

[12] Harper, J.D., Wong, S.S., Lieber, C.M., and Lansbury Jr., P.T. (1997) Observation of metastable Aβ amyloid protofibrils by atomic force microscopy, *Chem. Biol. 4,* 199-125.

[13] Volles, M.J., Lee, S.-J., Rochet, J.-C., Shtilerman, M.D., Ding, T.T., Kessler, J.C., and Lansbury Jr., P.T. (2001) Vesicle permeabilization by protofibrillar α-synuclein. Implications for the pathogenesis and treatment of Parkinson's disease, *Biochem. 40,* 7812-7819.

[14] Lambert, M.P. Barlow, A.K., Chromy, B.A., Edwards, C., Freed, R., Liosatos, M., Morgan, T.E., Rozovsky, I., Trommer, B., Viola, K.L., Wals, P., Zhang, C., Finch, C.E., Krafft, G.A., and Klein, W.L. (1998) Diffusible, nonfibrillar ligands derived from Aβ(1-42) are potent central nervous system neurotoxins, *Proc. Natl. Acad. Sci. USA 95,* 6448-6453.

Other References for the Supporting Information:

[15] Dong, J., Shokes, J. E., Scott, R. A., and Lynn, D. G. (2006) Modulating amyloid self-assembly and fibril morphology with Zn(II), *J. Am. Chem. Soc. 128*, 3540-3542.

[16] William of Ockham, 1285-1349, as cited by E. A. Moody (1967) in *The Encyclopedia of Philosophy*, vol 7, MacMillan, New York.

[17] Ban, T., Yamaguchi, K., and Goto, Y. (2006) Direct observation of amyloid fibril growth, propagation, and adaptation, *Acc. Chem. Res. 39*, 663-670.

[18] Vestergaard, M., Kerman, K., Saito, M., Nagatani, N., Takamura, Y., and Tamiya, E. (2005) A rapid label-free electrochemical detection and kinetic study of Alzheimer's amyloid beta aggregation, *J. Am. Chem. Soc. 127*, 11892-11893.

[19] Lu, K., Jacob, J., Thiyagarajan, P., Conticello, V. P., and Lynn, D. G. (2003) Exploiting amyloid fibril lamination for nanotube self-assembly, *J. Am. Chem. Soc. 125,* 6391-6393.

[20] Dumoulin, M., Kumita, J. R., and Dobson, C. M. (2006) Normal and aberrant biological self-assembly: Insights from studies of human lysozyme and its amyloidogenic variants, *Acc. Chem. Res. 39*, 603-610.

[21] Steinfeld, J. I.; Francisco, J. S.; Hase, W. L. (1989) Autocatalysis and oscillating reactions, in *Chemical Kinetics and Dynamics,* pp 182-183, Prentice-Hall, New Jersey.

[22] Wetzel, R. (2006) Kinetics and thermodynamics of amyloid fibril assembly, *Acc. Chem. Res. 39*, 671-679.

[23] Wood, S. J.; Wypych, J.; Steavenson, S.; Louis, J.-C.; Citron, M.; Biere, A. L. (1999) α-Synuclein fibrillogenesis is nucleation-dependent, *J. Biol. Chem. 274*, 19509-19512.

CHAPTER IV

FITTING YEAST AND MAMMALIAN PRION AGGREGATION KINETIC DATA WITH THE FINKE-WATZKY 2-STEP MODEL OF NUCLEATION AND AUTOCATALYTIC GROWTH

This dissertation chapter contains the manuscript of a full article published in *Biochemistry* **2008**, *47*, 10790-10800. This chapter presents the fits of 27 representative yeast and mammalian prion aggregation kinetic data sets relevant to prion diseases to the Finke-Watzky (F-W) 2-step model of nucleation and autocatalytic growth. The F-W model is able to account for each data set analyzed and elucidate quantitative, separate rate constants for the processes of nucleation and autocatalytic growth. In addition, the quantitative rate constants were used to analyze the data and suggest hypotheses for factors that may affect nucleation more than growth or vice versa.

The data were analyzed and manuscript prepared jointly by Aimee M. Morris and Dr. Murielle A. Watzky with the helpful insights, discussions, key references, and editing by Profs. Eric D. Ross and Richard G. Finke.

Fitting Yeast and Mammalian Prion Aggregation Kinetic Data with the Finke-Watzky 2-Step Model of Nucleation and Autocatalytic Growth

Murielle A. Watzky, Aimee M. Morris, Eric D. Ross, and Richard G. Finke

Abstract

Recently we reported 14 amyloid protein aggregation kinetic data sets that were fit using the "Ockham's razor"/minimalistic Finke-Watzky (F-W) 2-step model of slow nucleation (A→B, rate constant k_1) and fast autocatalytic growth (A+B→2B, rate constant k_2), yielding quantitative rate constants for nucleation (k_1) and growth (k_2), where A is the monomeric protein and B is the polymeric protein (*Biochemistry* **2008**, *47*, 2413-2427.). Herein we apply the F-W model to 27 representative prion aggregation kinetic data sets obtained from the literature. Each prion data set was successfully fit with the F-W model, including three different yeast prion proteins (Sup35p, Ure2p and Rnq1p) as well as mouse and human prions. These fits yield the first quantitative rate constants for the steps of nucleation and growth in prion aggregation. Examination of a Sup35p system shows that the same rate constants are obtained for nucleation and for growth within experimental error, regardless of which of six physical methods was used, a unique set of important control experiments in the protein aggregation literature. Also

provided herein are analyses of several factors influencing the aggregation of prions such as glutamine/asparagine rich regions and the number of oligopeptide repeats in the prion domain. Where possible, verification or refutation of previous correlations to glutamine/asparagine regions, or number of repeat sequences, in literature aggregation kinetics is given in light of the quantitative rate constants obtained herein for nucleation and growth during prion aggregation. The F-W model is then contrasted to four prior proposed mechanisms that address the molecular picture of prion transmission and propagation. Key limitations of the F-W model are listed in order to avoid overinterpretation of the data being analyzed, limitations that derive ultimately from the model's simplicity. Finally, possible avenues of future research are suggested.

Introduction

The mammalian prion diseases (or transmissible spongiform encephalopathies) are a subset of neurodegenerative disorders that include bovine spongiform encephalopathy (or "mad cow" disease), scrapie in sheep, chronic wasting disease in deer and elk, and the human kuru and Creutzfeldt-Jakob diseases (1). It is generally recognized that the infectious prion agent is composed of protein (thus the term *prion* for *pr*oteinaceous *i*nfectious -*on*) (2). However, in the case of mammalian prions it has not been irrefutably shown whether the infectious agent is composed of only protein (3,4). The prion agent in mammalian diseases is usually referred to as PrP^{Sc} (for prion protein, scrapie) and the normal form of the protein as PrP^{C} (for prion protein, cellular).[1] The two

[1] Abbreviations: ANS=ANS fluorescence; *CA=Candida albicans*; CHY=chymotrypsin resistance; CR=congo red; F-W=Finke-Watzky 2-step protein aggregation kinetic model; LS=light scattering; MDC=monomer directed conversion; NCC=nucleated

forms of the protein are believed to be conformational isomers (2), with the PrP^{Sc} form believed (i) to be aggregation-prone with a secondary structure of high β-sheet character (5) and (ii) to have the ability to convert PrP^{C} (of predominantly α-helix secondary structure) into PrP^{Sc} (1).

The prions found in yeast are now being widely used as models of the mammalian prion diseases, as (i) they appear to propagate in a similar manner as their mammalian equivalent and, importantly, (ii) unlike their mammalian equivalent, they are not contagious to humans (6). In the yeast species *Saccharomyces cerevisae* (*SC*), three different prions have been identified, the names of which are derived from the characteristic (change in) phenotype that accompanies them: [URE3], [*PSI*+], and [*PIN*+] (7,8). The "encoding" or "determinant" proteins whose propagating prion form is associated with these prion diseases are Ure2p, Sup35p, and Rnq1p, respectively (12). Hence, overexpression of the protein Ure2p, Sup35p, or Rnq1p will cause de novo appearance in *SC* yeast of the prion phenotype [URE3] (9), [*PSI*+] (10), or [*PIN*+] (11), respectively. For the yeast prions, it has been demonstrated that the infectious prion agent is composed only of protein, as amyloid fibrils formed in vitro by purified proteins are found to be infectious when transformed into yeast (12,13,14,15).

A common feature in many neurodegenerative disorders is the presence of insoluble protein aggregate fibrils (called amyloid fibrils) in the affected area of the brain. In prion diseases, it is the prion form of the protein that aggregates to form fibrils. Fibrils are typically composed of the protein in a β-sheet secondary structure, with all β-sheets

---

conformational conversion; NP=nucleated polymerization; PrD=prion domain; PrP^{C}=prion protein, cellular; PrP^{Sc}=prion protein, scrapie; *SC*=*Saccharomyces cerevisae*; SDS=SDS-PAGE; Sed=sedimentation; TA=templated assembly.

aligned perpendicular to the axis of the fibril. The final product is an insoluble fibril comprised of β-sheets; however, it is unclear if these insoluble fibrils are toxic (16). It is currently believed that smaller, soluble oligomeric intermediates in the protein aggregation process are toxic to the neuronal cells, but the structure of these intermediates is still under investigation (17,18). It is also not clear whether the oligomeric species that are toxic lead to the formation of ordered fibrils ("on-pathway"), or instead to the formation of disordered aggregates ("off-pathway") (19,20,21,22).

The exact (molecular) mechanism of transmission and propagation of prion diseases is still controversial, but generally believed to be intimately linked to prion aggregation (23,24). Overall, four different mechanisms have been proposed in the literature for how transmission and propagation of the prion disease may occur through an agent made (only) of protein, see Table 1 below. The first three mechanisms (I, II, and III) were originally developed for mammalian prion diseases (25,26,27,28,29) but have been widely cited in yeast prion studies. The last mechanism (IV) was proposed in a study of the yeast prion Sup35p (23).

Table 1. Four proposed literature mechanisms of transmission and propagation of the prion agent (23,30).

Entry	Reference	Year	Author	Mechanism [a]
I	(29)	1967	J. S. Griffith	Template Assembly (TA)
II	(25,26)	1991	S. B. Prusiner	Monomer Directed Conversion (MDC)
III	(27,28)	1993	P. T. Lansbury, Jr.	Nucleated Polymerization (NP)
IV	(23)	2000	S. L. Lindquist	Nucleated Conformational Conversion (NCC)

[a] The mechanistic nomenclature comes from reference 23.

These four mechanisms are often displayed as word and picture-only mechanisms that lack the numerically or analytically integrable balanced chemical equations corresponding to the proposed mechanism. This is not to say that these needed equations do not (or can not) exist, only that they are not routinely provided nor, therefore, used to fit prion aggregation kinetic data. Of course, without connecting specific balanced reactions to their corresponding differential/integrated kinetic equations, one cannot verify the postulated (word or picture only) mechanism (30). In short, it remains crucial to gain a better understanding of the mechanism(s) of protein aggregation. Having a minimalistic kinetic model that can account for the available prion aggregation data is a needed, key first step towards that better understanding.

Recently we were able to fit a broad range of amyloid protein aggregation (14 representative kinetic data sets) using the minimalistic/"Ockham's razor" Finke-Watzky (F-W) 2-step model of nucleation and autocatalytic growth (31) (Scheme 1). With those results, we were able to obtain separate quantitative rate constants for nucleation and growth of amyloid protein aggregation for those 14 data sets for the first time (31). However, due to the breadth of the protein aggregation literature, we previously focused our attention on the proteins of amyloid β, α-synuclein, and polyglutamine (31). We purposely, but arbitrarily, left out prion proteins in that first study (31).

Scheme 1. The minimalistic F-W 2-step model of slow nucleation followed by fast autocatalytic growth (42).[a]

$$A \xrightarrow{k_1} B$$

$$A + B \xrightarrow{k_2} 2B$$

[a] A is the monomeric form of the protein and B is the (auto)catalytic, polymeric form of the protein.

Herein, we apply the 2-step F-W model to the aggregation of prion proteins with a focus on the *SC* yeast prion proteins Ure2p, Sup35p, and Rnq1p. We report 24 representative fits of Ure2p, Sup35p, and Rnq1p aggregation kinetics to the F-W model. We also examine 3 representative examples of mammalian prion aggregation kinetics using the F-W model, again affording quantitative (albeit average, vide infra) rate constants for nucleation and growth. In all cases examined, the F-W model is able to fit the prion aggregation kinetic data. To our knowledge, this is the first nonempirical (30) equation that has been used to fit prion aggregation kinetic data and afford quantitative, separate (average) rate constants for both nucleation (k_1) and growth (k_2).

Materials and Methods

Selection of Data Sets for Analysis. The literature was searched, using both Scifinder Scholar and the Web of Science, for yeast prion as well as mammalian prion aggregation kinetics. For the yeast prion references, we chose representative data sets to fit from various research groups, under various conditions, and with the wild type and mutants of Ure2p, Sup35p, and Rnq1p proteins. For the mammalian prion aggregation kinetics three representative data sets were chosen (from a small number of data sets).

Data Analysis and Curve Fitting. Data were extracted (digitized) from the published kinetic curves cited herein using Engauge Digitizer 2.12. All data were fit using the integrated analytical equation (eq. 1) corresponding to the F-W 2-step model using Origin 7.0, as detailed previously (31). In eq. 1, $[B]_t$ represents the concentration of aggregated protein at time t (vide infra), $[A]_0$ represents the initial protein

concentration, and k_1 and k_2 are the aforementioned rate constants for nucleation and growth, respectively.

$$[B]_t = [A]_0 - \frac{\frac{k_1}{k_2} + [A]_0}{1 + \frac{k_1}{k_2[A]_0}\exp(k_1 + k_2[A]_0)t} \quad (1)$$

In each of the fits presented herein and as previously detailed (31), we have by necessity assumed (as generally done in the literature) that (i) all forms of kinetic measurement are directly proportional to the percent of aggregated protein (i.e., that the physical method used is a direct measure of the aggregated protein), and (ii) the data are ideal; that is, only the aggregated protein is responsible for the observed signal. Furthermore, where noted in the cases that follow, the published data sets were normalized to an initial minimum of zero.

Results

In what follows, we present fits of 24 representative yeast prion aggregation kinetic data sets, including the yeast prions Sup35p, Ure2p, and Rnq1p, as well as three representative mammalian (both mouse and human) prion aggregation kinetic data sets. Each data set was fit using the F-W model; the results are displayed in Figures 1-10 and summarized in Table 3.

Fitting Yeast Prion Aggregation Kinetic Data. Figures 1-8 show representative kinetic data sets for yeast prion aggregation fit using the F-W model from the three yeast prion systems Sup35p, Ure2p, and Rnq1p. To start, we looked at Sup35p aggregation

kinetic data obtained by Lindquist and co-workers using six different physical methods (23). Displayed below in Figure 1 are the F-W fits to each set of aggregation kinetic data measured by ANS fluorescence, chymotrypsin resistance, Congo Red binding, light scattering, SDS-PAGE, and sedimentation. Interestingly, and as displayed in Table 2, each of these physical methods gives the same nucleation (k_1) and growth (k_2) rate constants within experimental error. However, one issue with the data sets shown in Figure 1 is that the spectra appear to be somewhat noisy, resulting in relatively large error bars in the rate constants and in lower than desired coefficients of determination (R^2, Table 2). Nevertheless, Lindquist and co-workers are the first (to the best of our knowledge) to perform this important control of examining protein aggregation kinetic data with at least six physical methods. Our analysis herein is the first to deconvolute that data into k_1 and k_2 values showing that both rate constants are the same within experimental error with all methods used, at least for the Sup35p system.

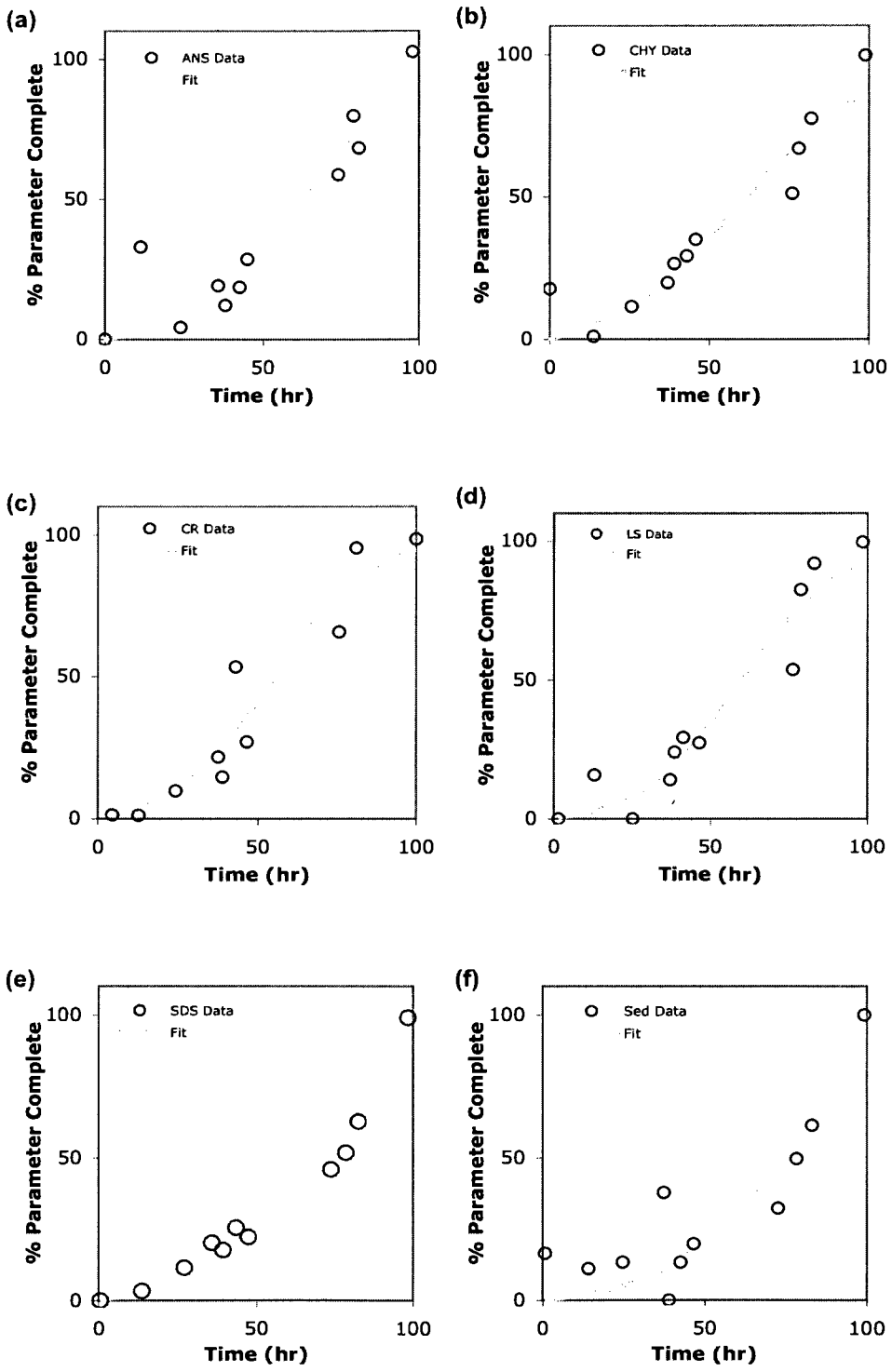


Figure 1. Aggregation of Sup35p measured by Lindquist and co-workers (23) using 6 different physical methods. Specifically, Sup35p aggregation was measured by (a) ANS fluorescence, (b) chymotrypsin resistance, (c) Congo Red binding, (d) light scattering, (e) SDS-PAGE, and (f) sedimentation methods. Each data set was digitized and fit using the

F-W model. Table 2 shows the k_1 and k_2 rate constants obtained for parts a-f. The important finding is that the k_1 values, as well as the k_2 values, are all equal to one another within experimental error for the six techniques used.

Table 2. Rate constants and coefficients of determination (R^2) obtained from the fits shown in Figure 1 measuring the aggregation of Sup35p by the six listed methods.

Physical Method	k_1 (hr^{-1})	k_2 ($\mu M^{-1} hr^{-1}$)	R^2 [a]
ANS fluorescence (ANS)	$1(1) \times 10^{-3}$	$1.2(2) \times 10^{-2}$	0.88158
Chymotrypsin resistance (CHY)	$3(1) \times 10^{-3}$	$1.0(2) \times 10^{-2}$	0.91165
Congo Red (CR)	$2(1) \times 10^{-3}$	$1.2(4) \times 10^{-2}$	0.91684
Light scattering (LS)	$1.4(8) \times 10^{-3}$	$1.2(2) \times 10^{-2}$	0.92883
SDS-PAGE (SDS)	$1.7(8) \times 10^{-3}$	$0.9(2) \times 10^{-2}$	0.93344
Sedimentation (Sed)	$1(1) \times 10^{-3}$	$1.0(4) \times 10^{-2}$	0.76354

[a] The coefficient of determination R^2 is a measure of the quality of the regression model; the closer the value of R^2 is to 1, the more precise the fit is to the data.

Another Sup35p data set we examined comes from Weissman and co-workers (32). Those authors looked at whether the prion function of Sup35p is conserved in distantly related yeasts, by expressing foreign prion domains (PrD) of Sup35p in *SC* (32). The authors found that the prion function of Sup35p was conserved among these yeasts. The ability to form amyloid fibrils in vitro was conserved as illustrated in Figure 2, which includes fits by the F-W model (32).[2] From the F-W fits in Figure 2, prion aggregation of Sup35p from the yeast species *SC* and *Candida Albicans* (*CA*) display (i) different nucleation kinetics, as shown by k_1 values that appear to differ by 5 orders of magnitude

[2] The authors also found that a species barrier inhibited prion induction by Sup35p from different species. That is, "foreign" PrDs could not efficiently seed prion formation by "wild-type" Sup35p (but "foreign" PrDs did behave as prions themselves). We chose to not fit cross-seeding experiments in this data set, as seeding introduces 'B' into the reaction, which then is not be properly modeled by eq.1. (We have, however, provided an equation elsewhere (31) that can in principle deal with seeded data sets.) To our knowledge, no data set treated herein involves any type of seeding.

(note, however, that there are large error bars in these k_1 values[3]); but (ii) similar growth kinetics, as demonstrated by k_2 values that are the same order of magnitude.

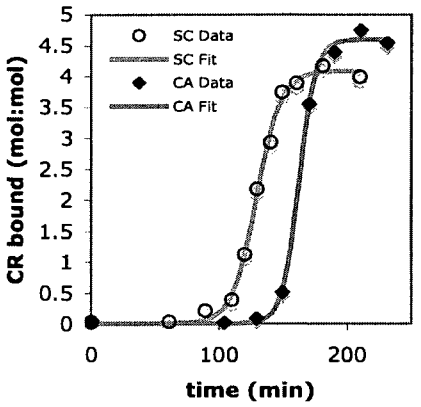


Figure 2. Weissman and co-workers' aggregation of Sup35p PrDM from different yeast species, measured by Congo Red binding at [PrD]= 2.5 μM with continuous rotation (32). The data sets were digitized, normalized to an initial minimum of zero, and then fit to the F-W model. Resultant rate constants for *Saccharomyces cerevisae* (*SC*): $k_1=1(1) \times 10^{-7}$ min^{-1} ($k_1=6(6) \times 10^{-6}$ h^{-1}), $k_2= 4.2(4) \times 10^{-2}$ $\mu M^{-1}.min^{-1}$ ($k_2=2.5(2)$ μM^{-1} h^{-1}), $R^2=0.997$; for *Candida albicans* (*CA*): $k_1=2(9) \times 10^{-12}$ min^{-1} ($k_1=1(5) \times 10^{-10}$ h^{-1}), $k_2= 6(1) \times 10^{-2}$ μM^{-1} min^{-1} ($k_2=3.6(6)$ μM^{-1} h^{-1}), $R^2=0.998$. The relatively large error bars in k_1 result from the lack of data around the initial, concave portion of the sigmoidal curve.

The third literature study we examined is an interesting paper by Weissman and co-workers in which they used a genetic screening technique to identify mutations that cause defects in the amyloid formation of Sup35p (33). Visual examination of the aggregation curves shown in Figure 3 below for WT Sup35p and a Q15R mutant (where Q at position 15 is substituted by R) reveals that the induction period is lengthened, qualitatively, on going from the WT to the mutant. Using the F-W model, we were able

[3] Ideal data sets would have more data points around the concave portion of the sigmoidal curve, in order to lessen error bars from the fit, especially the error in k_1.

to more quantitatively dissect the data. Specifically, we find that the mutation (i) does affect nucleation kinetics, decreasing the observed k_1 value by one order of magnitude on going from the WT to the mutant (although here, again, large error bars in k_1 values are present[3]), while (ii) the growth kinetics are less affected, with the k_2 value for the mutant being approximately one-half of that for the WT.

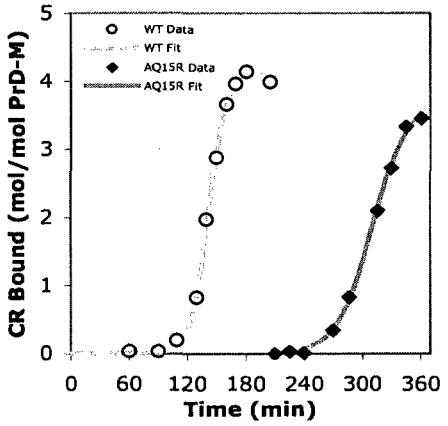


Figure 3. Weissman and co-workers' Sup35p and mutant AQ15R digitized aggregation data measured by Congo Red binding (33). The digitized data were normalized to an initial minimum of zero and then fit using the F-W model. The resultant rate constants for the WT are: $k_1=2(2) \times 10^{-8}$ min^{-1} ($k_1=1(1) \times 10^{-6}$ h^{-1}), $k_2=4.5(3) \times 10^{-2}$ μM^{-1} min^{-1} ($k_2=2.7(2)$ μM^{-1} h^{-1}), $R^2=0.999$; and AQ15R: $k_1=2(2) \times 10^{-9}$ min^{-1} ($k_1=1(1) \times 10^{-7}$ h^{-1}), $k_2=2.1(1) \times 10^{-2}$ μM^{-1} min^{-1} ($k_2=1.26(6)$ μM^{-1} h^{-1}), $R^2=0.999$. Here again, the relatively large error bars in k_1 result from the lack of data around the initial, concave portion of the sigmoidal curve.

In a different study in 1999, Lindquist and co-workers examined the importance of oligopeptide repeat expansions in Sup35p (34). In order to investigate this idea, they created the variants RΔ2-5 and R2E2 in which RΔ2-5 differs from the WT in that it has the last four repeats deleted, while R2E2 has two additional copies of the second repeat. Via the F-W fits to their aggregation data shown in Figure 4a, we now can look at the

effect of repeat expansion or deletion on nucleation and growth, separately. The fits reveal that no statistically significant trend due to the effect of repeat sequences on nucleation is observed (Figure 4b). (Although the k_1 values appear to be increased by one order of magnitude both for the deletion and repeat expansion variants as compared to WT,[4] the large error bars render the results statistically equivalent.[3]) On the other hand, the repeat sequences do appear to affect the growth kinetics, as shown by the increase in k_2 values with an increasing number of repeat sequences (Figure 4c). These observations are consistent with Osherovich et al.'s findings of residues 1-39 of Sup35p driving in vivo prion formation, and residues 40-114 driving prion propagation (35). These are valuable insights on the factors affecting nucleation vs. growth that were previously buried within this interesting 1999 study (34).

[4] We have observed that the (inverse of the) induction period is a good indicator of nucleation kinetics only when the sigmoidal aggregation curve is not "rounded" (i.e., when the initial portion of the curve is flat and there is a clean-cut transition to the linear slope portion). When the sigmoidal curve is "rounded" (i.e., when the initial portion of the curve is not flat and the transition to the linear slope portion shows a wide concave angle), we find that both nucleation and growth kinetics can be convoluted in the induction period. Additional studies on this issue are in progress.

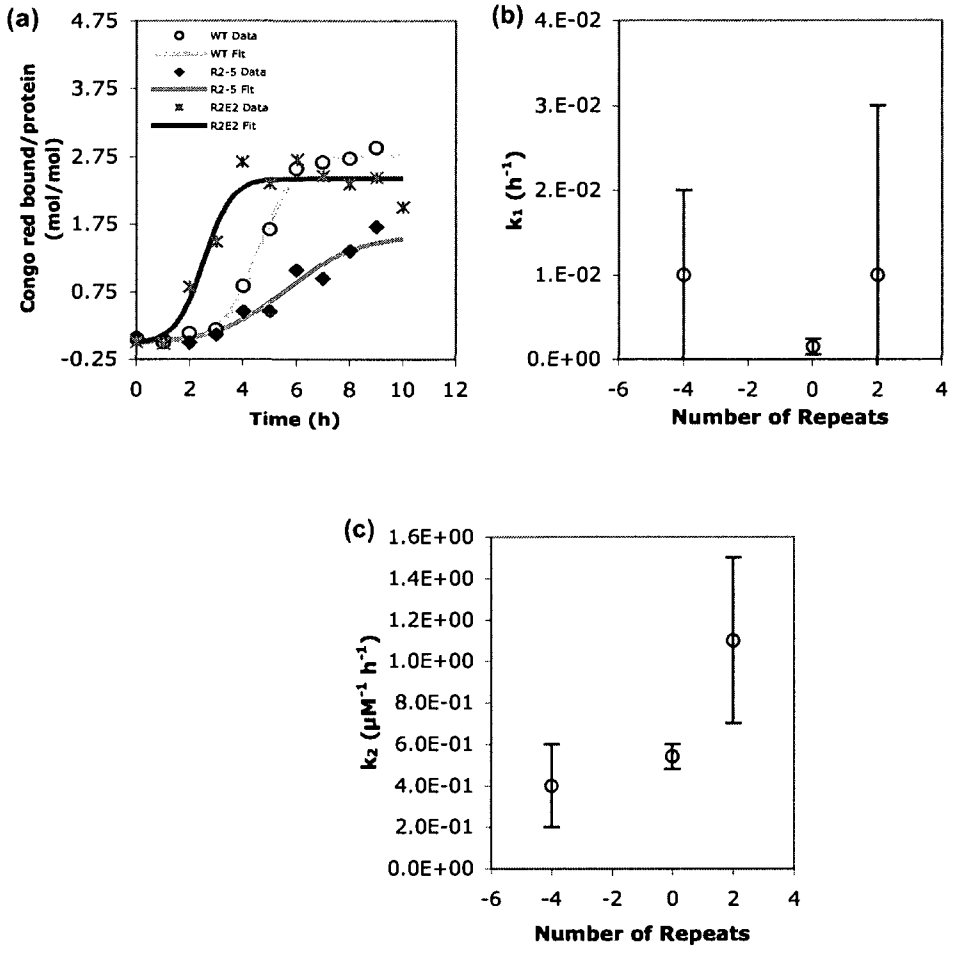


Figure 4. (a) Aggregation data for WT Sup35p along with mutants, RΔ2-5 and R2E2, measured by Congo Red binding from Lindquist and co-workers (34). The RΔ2-5 mutant contains four less repeat sequences than the WT, while R2E2 has two more repeat sequences (34). The data were digitized, normalized to an initial minimum of zero, and fit by the F-W model. The following rate constants were obtained: WT, k_1=1.5(9) × 10^{-3} h^{-1} and k_2=5.4(6) × 10^{-1} μM^{-1} h^{-1}, R^2=0.995; RΔ2-5, k_1=1(1) × 10^{-2} h^{-1} and k_2=4(2) × 10^{-1} μM^{-1} h^{-1}, R^2=0.945; R2E2, k_1=1(2) × 10^{-2} h^{-1} and k_2=1.1(4) μM^{-1} h^{-1}, R^2=0.951. (b), (c) Correlations between the difference in number of repeat sequences (set at zero for WT), and the k_1, k_2 values obtained from the data fits in Figure 4a.

The next interesting study that caught our eye was the investigation by Perrett and co-workers of the role of specific glutamine/asparagine (Gln/Asn) rich regions, contained in residues 1-14 and 42-90, of the prion domain of Ure2p (36). We re-examined the aggregation data shown in Figure 5 for both WT and mutant Ure2p, using the F-W model

to obtain quantitative k_1 and k_2 rate constants. For the mutant 15Ure2 (Figure 5a), which refers to Ure2p with residues 1-14 (a Gln/Asn rich region) removed, the nucleation kinetics are strongly affected, as shown by a k_1 value decreased by three orders of magnitude as compared to WT (where this difference is beyond the fitting error). On the other hand, the mutation has little effect on growth kinetics, with k_2 values the same within experimental error. Consistent with these observations is the previous work addressing nucleation and showing that deletion of residues 2-14 of Ure2p dramatically reduces the in vivo frequency with which new prions are detected (37). For the mutant Δ15-42Ure2 (Figure 5b), which refers to Ure2p with residues 15-42 (a region less enriched in Gln/Asn) removed, the authors note a strong effect on the intensity of the fluorescence signal and hypothesize that the less enriched Gln/Asn region which was removed might be a binding site for ThT (36). As a consequence, it is not clear at present whether the change in the mutant aggregation curve (and its F-W fit) actually represents different aggregation kinetics, or if it is some type of experimental artifact. The deconvolution of k_1 and k_2 is, however, useful in directing needed future control experiments.

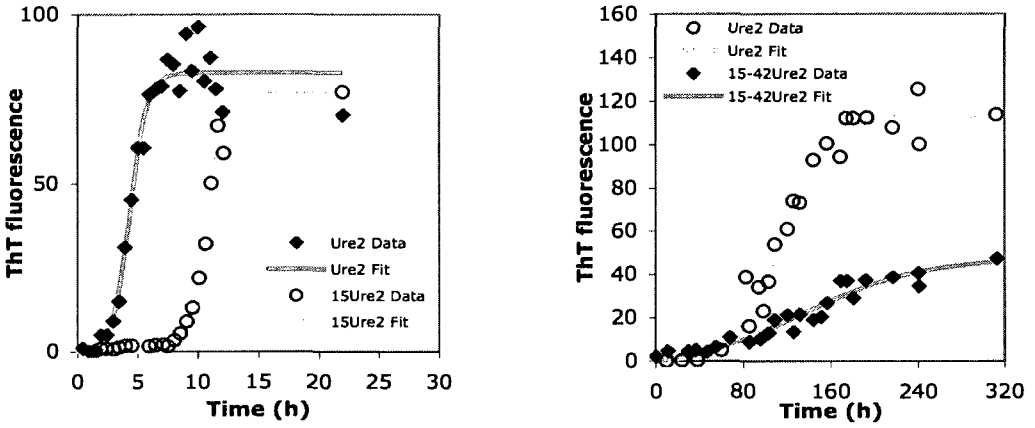


Figure 5. (a) Perrett and co-workers' 40 μM Ure2p and mutant 15Ure2 aggregation data measured by fluorescence, at pH = 7.5 and T = 25°C, with shaking (36). The data were digitized, normalized to an initial minimum of zero, and then fit using the F-W model. Rate constants for Ure2p are $k_1=3(2) \times 10^{-3}$ h^{-1} and $k_2=4.1(7) \times 10^{-2}$ μM^{-1} h^{-1}, $R^2=0.975$. Rate constants for 15Ure2 are $k_1= 1(1) \times 10^{-6}$ h^{-1} and $k_2=3.3(4) \times 10^{-2}$ μM^{-1} h^{-1}, $R^2=0.999$. (b) Digitized data of Perrett and co-workers' 40 μM Ure2p and mutant Δ15-42Ure2 aggregation data measured by fluorescence, at pH = 9 and T = 4°C, without shaking (36). The data were fit to the F-W model. Resultant rate constants for Ure2p: $k_1=2(1) \times 10^{-4}$ h^{-1}, $k_2=1.3(2) \times 10^{-3}$ μM^{-1} h^{-1}, $R^2=0.975$; and Δ15-42Ure2: $k_1=1.4(4) \times 10^{-3}$ h^{-1}, $k_2=4(1) \times 10^{-4}$ μM^{-1} h^{-1}, $R^2=0.931$.

Perrett and co-workers also looked at the stability of folding intermediates of Ure2p, and studied the effect of experimental variables on its aggregation (38). The effect of varying Ure2p concentration is shown in Figure 6, along with fits to the F-W model (38). [Unfortunately, the full range of concentrations could not be digitized or fit, since at lower concentrations the aggregation curves (in their published form) were convoluted in the early portion of the curve, that early portion being crucial in determining the nucleation kinetics (38).] In Figure 7, we show four representative examples of the effect of pH and temperature on the aggregation of Ure2p, along with fits by the F-W model (38). In this case, all the published curves could be digitized. However, some of the curves differed significantly enough from the average curve

obtained under the same conditions (as the results summarized in Table 4 of ref. 38 indicate), that we chose not to attempt to draw trends (38). The main conclusions from the data and fits in Figure 7, then, are (i) the F-W model fits the data well, but (ii) more precise data would be useful for drawing correlations for different concentrations of prion protein, varying temperatures, and varying pH values.

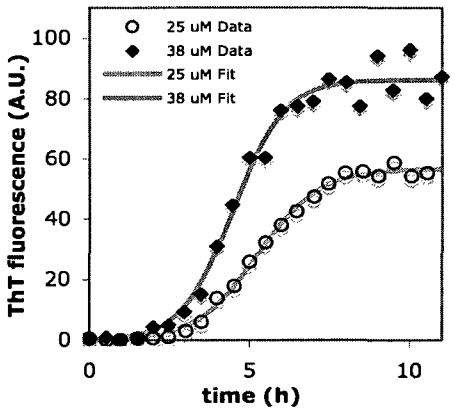


Figure 6. Perrett and co-workers' Ure2p aggregation measured by ThT fluorescence for different [Ure2p] concentrations at pH = 7.5 and T = 25°C, in 50 mM Tris buffer with shaking (38). The data was digitized, normalized to an initial minimum of zero, and then fit by the F-W model. Resultant rate constants for [Ure2p]=25 μM: k_1=4.2(9) × 10^{-3} h^{-1}, k_2=4.2(2) × 10^{-2} $\mu M^{-1}.h^{-1}$, R^2=0.997; [Ure2p]=38 μM: k_1=5(2) × 10^{-3} h^{-1}, k_2=3.3(3) × 10^{-2} $\mu M^{-1}.h^{-1}$, R^2=0.987.

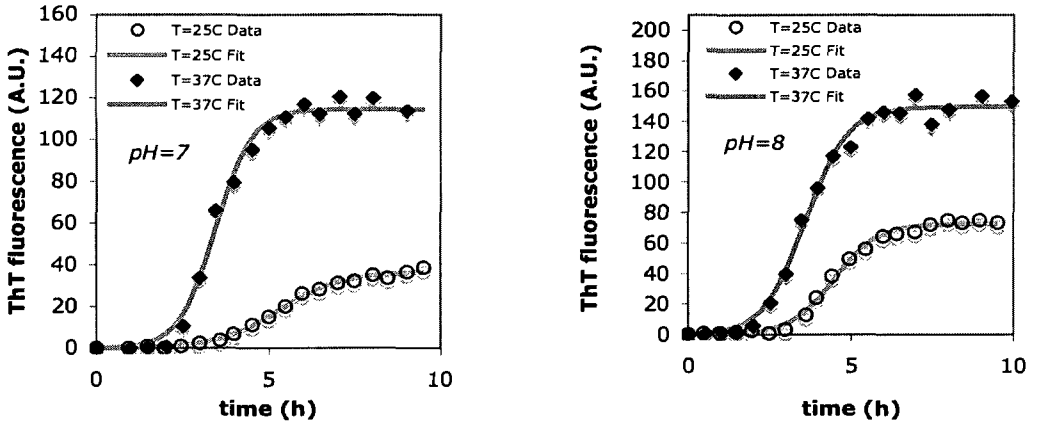


Figure 7. (a) Perrett and co-workers' Ure2p aggregation measured by ThT fluorescence for different temperatures at pH = 7 and [Ure2p]= 30μM, in 50 mM phosphate buffer with shaking (38). The data was digitized, normalized to an initial minimum of zero and then fit to the F-W model. Resultant rate constants for T = 25 °C: k_1= 3.2(9) × 10^{-3} h^{-1}, k_2=3.7(3) × 10^{-2} $\mu M^{-1}.h^{-1}$, R^2=0.996; for T = 37 °C: k_1= 3(1) × 10^{-3} h^{-1}, k_2=6.3(7) × 10^{-2} $\mu M^{-1}.h^{-1}$, R^2=0.992. (b) Perrett and co-workers' Ure2p aggregation measured by ThT fluorescence for different temperatures at pH = 8 and [Ure2p]= 30μM, in 50 mM phosphate buffer with shaking (38). The data was digitized, normalized to an initial minimum of zero and then fit to the F-W model. Resultant rate constants for T = 25 °C: k_1=1.8(7) × 10^{-3} h^{-1}, k_2=5.0(4) × 10^{-2} $\mu M^{-1}.h^{-1}$, R^2=0.995; for T = 37 °C : k_1=6(2) × 10^{-3} h^{-1}, k_2=5.1(5) × 10^{-2} $\mu M^{-1}.h^{-1}$, R^2=0.993.

In a study aimed at demonstrating a protein-only mechanism of infection for [*PIN*+], Liebman and co-workers used a Glu/Asn rich recombinant fragment of Rnq1p (residues 132-405) that polymerizes into fibrils (12). The authors studied the effect of agitation, concentration, and temperature on the Rnq1p-(132-405) aggregation curves measured by continuous ThT fluorescence. We show one representative example of Rnq1p-(132-405) aggregation in Figure 8, fit by the F-W model. [Unfortunately, most of the published curves could not be readily digitized (12).] The primary conclusion from Figure 8 is that the observed data are well fit by the F-W 2-step kinetic model.

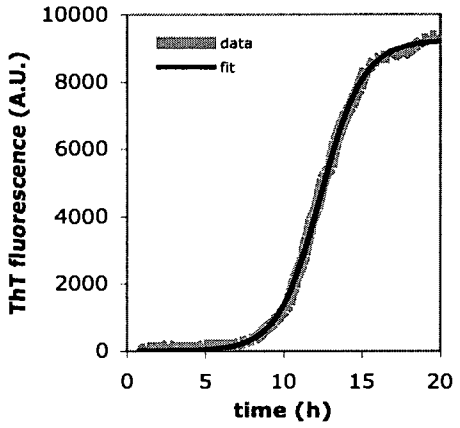


Figure 8. Liebman and co-workers' aggregation of Rnq1p-(132-405) measured by continuous ThT fluorescence at T = 27°C and [Rnq1p-(132-405)]=110 μM, with agitation (12). The data was digitized and fit to the F-W model. Resultant rate constants: k_1=8.0(9) × 10^{-5} h^{-1}, k_2=6.7(1) × 10^{-3} $\mu M^{-1}.h^{-1}$, R^2=0.999.

Fitting Mammalian Prion Aggregation Kinetic Data. Figures 9 and 10 show representative data sets of mammalian prion aggregation curves fit using the F-W model, for both mouse and human prion systems. Cohen and co-workers (39) were interested in the mechanism of conversion of the (monomeric, α-helical) cellular form of the prion protein, PrP^C, into its (polymeric, β-sheet) prion form, PrP^{Sc}; hence, they collected kinetic data for the aggregation of recombinant mouse prion protein (MoPrP). The aggregation curve is shown in Figure 9, illustrating the good fit of the F-W nucleation and growth model.

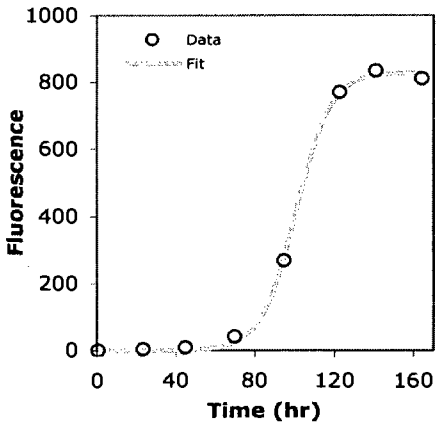


Figure 9. Digitized data for the assembly of 40 μM MoPrP from Cohen and co-workers normalized to an initial minimum of zero and then fit using the F-W model (39). The resultant rates constants are $k_1=1(1) \times 10^{-6}$ hr^{-1} and $k_2=2.9(2) \times 10^{-3}$ μM^{-1} hr^{-1} with $R^2=0.999$.

Next, we looked at Baskakov and co-workers' study of the aggregation of recombinant human prion protein (HuPrP) under different experimental conditions, including various prion protein concentrations (40). Two aggregation curves illustrating the effect of HuPrP concentration are shown in Figure 10, along with fits by the F-W model. Here again, the full range of concentrations could not be digitized or fit by us, since at lower concentrations the published aggregation curves were convoluted in the early portion of the curve (40). The fits that were possible are quite good, indicating that it would be of interest to investigate other concentrations of HuPrP.

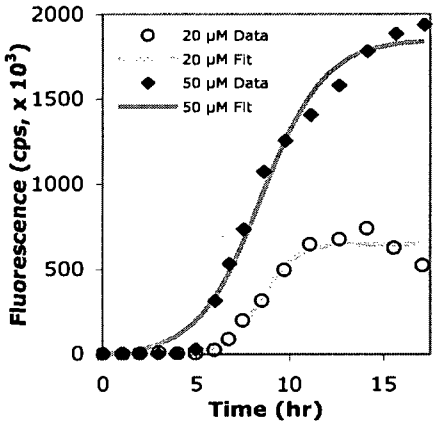


Figure 10. Baskakov and co-workers' kinetics of the human recombinant prion protein (Hu rPrP 90-231) digitized data from varying concentrations normalized to an initial minimum of zero and then fit to the F-W model (40). The following rate constants and R^2 values are obtained for each concentration: 20 μM, $k_1=1(2) \times 10^{-4}$ hr^{-1}, $k_2=5(1) \times 10^{1}$ μM^{-1} hr^{-1}, $R^2=0.977$; 50 μM, $k_1=4(2) \times 10^{-3}$ hr^{-1}, $k_2=1.2(2) \times 10^{1}$ μM^{-1} hr^{-1}, $R^2=0.988$.

Addressing the Issue of the Effects of Agitation. To address the issue of agitation (and resulting fragmentation) on prion aggregation, we searched the literature for prion kinetic data sets of a system under the same experimental conditions, with and without agitation. Only one unseeded data set was found that could be readily digitized; this data set and its analysis can be found in the Supporting Information. Unfortunately, the likely presence of several possible experimental issues in this otherwise important data set kept us from being able to properly address the important issue of fragmentation. In the future, we plan to obtain our own data sets to address this issue via the F-W model (41).

Table 3 summarizes the 24 representative cases of yeast prion and 3 representative cases of mammalian prion aggregation kinetics that were successfully curve-fit using the F-W model (Figures 1-10). Of course, any time one goes to the trouble to collect and tabulate this amount of data, the resultant table merits scrutiny for any insights or needed future experiments suggested by the data and the cross

comparisons. A couple of such comparisons in Table 3 caught our eye; the reader is encouraged to find their own, within the caveats noted below of the issues involved in comparing different systems and different experimental conditions. (i) A comparison of data for ostensibly similar systems (e.g., entries 3 and 9) yields k_1 values differing by $\sim 10^3$ and k_2 values differing by $\sim 10^2$. Hence, repeats of identical systems under identical (as well as a broader range) of conditions, and in multiple laboratories, remain to be done to determine what k_1 and k_2 values (and what level of reproducibility) results. Especially important is that (ii) considerable caution needs to be exercised in any such cross comparisons, even for identical proteins since rate constant can be sensitive to the precise experimental conditions just like any other physical constant of a system. Indeed, Table 3 is just the first such table of its kind, so there is a need for many more k_1 and k_2 values obtained under carefully controlled and reported conditions.

Table 3. Rate constants and coefficients of determination (R^2) from fitting 27 prion aggregation kinetic data sets to the F-W 2-step model.

Entry	Reference	System	Distinguishing experimental details	Data collection method	k_1 [a] (h^{-1})	k_2 [a,b] ($\mu M^{-1} h^{-1}$)	R^2 [c]
1	Lindquist (23)	Sup35p		ANS	$1(1) \times 10^{-3}$	$1.2(2) \times 10^{-2}$	0.882
2				CHY	$3(1) \times 10^{-3}$	$1.0(2) \times 10^{-2}$	0.912
3				CR	$2(1) \times 10^{-3}$	$1.2(4) \times 10^{-2}$	0.917
4				LS	$1.4(8) \times 10^{-3}$	$1.2(2) \times 10^{-2}$	0.929
5				SDS	$1.7(8) \times 10^{-3}$	$9(2) \times 10^{-3}$	0.933
6				Sed	$1(1) \times 10^{-3}$	$1.0(4) \times 10^{-2}$	0.764
7	Weissman (32)	Sup35p	From SC	CR	$6(6) \times 10^{-6}$	2.5(2)	0.997
8			From CA	CR	$1(5) \times 10^{-10}$	3.6(6)	0.998
9	Weissman (33)	Sup35p	WT	CR	$1(1) \times 10^{-6}$	2.7(2)	0.999
10			AQ15R	CR	$1(1) \times 10^{-7}$	1.26(6)	0.999
11	Lindquist (34)	Sup35p	WT	CR	$1.5(9) \times 10^{-3}$	$5.4(6) \times 10^{-1}$	0.995
12			RΔ2-5	CR	$1(1) \times 10^{-2}$	$4(2) \times 10^{-1}$	0.946
13			R2E2	CR	$1(2) \times 10^{-2}$	1.1(4)	0.951
14	Perrett (36)	Ure2p	Ure2 agitated	ThT	$3(2) \times 10^{-3}$	$4.1(7) \times 10^{-2}$	0.975
15			15Ure2 agitated	ThT	$1(1) \times 10^{-6}$	$3.3(4) \times 10^{-2}$	0.999
16			Ure2 non-agitated	ThT	$2(1) \times 10^{-4}$	$1.3(2) \times 10^{-3}$	0.975
17			Δ15-42Ure2 non-agitated	ThT	$1.4(4) \times 10^{-3}$	$4(1) \times 10^{-2}$	0.931
18	Perrett (38)	Ure2p	25 μM	ThT	$4.2(9) \times 10^{-3}$	$4.2(2) \times 10^{-2}$	0.997
19			38 μM	ThT	$5(2) \times 10^{-3}$	$3.3(3) \times 10^{-2}$	0.987
20			pH 7, 25 °C	ThT	$3.2(9) \times 10^{-3}$	$3.7(3) \times 10^{-2}$	0.996
21			pH 7, 37 °C	ThT	$3(1) \times 10^{-3}$	$6.3(7) \times 10^{-2}$	0.992
22			pH 8, 25 °C	ThT	$1.8(7) \times 10^{-3}$	$5.0(4) \times 10^{-2}$	0.995
23			pH 8, 37 °C	ThT	$6(2) \times 10^{-3}$	$5.1(5) \times 10^{-2}$	0.993
24	Liebman (12)	Rnq1p		ThT	$8.0(9) \times 10^{-5}$	$6.7(1) \times 10^{-3}$	0.999
25	Cohen (39)	Mouse prion		ThT	$1(1) \times 10^{-6}$	$2.9(2) \times 10^{-3}$	0.999
26	Baskakov (40)	Human prion	20 μM	ThT	$1(2) \times 10^{-4}$	$5(1) \times 10^{1}$	0.977
27			50 μM	ThT	$4(2) \times 10^{-3}$	$1.2(2) \times 10^{1}$	0.988

[a] The error bars are for the fit and are determined by the square root of the reduced χ^2 based on a modified Levenberg-Marquardt algorithm. [b] The concentration units on k_2 were determined by multiplying by the maximum intensity and dividing by the initial starting concentration of the respective protein being measured. [c] The coefficient of determination R^2 is a measure of the quality of the regression model; the closer the value of R^2 is to 1, the more precise the fit is to the data.

Discussion

Additional Background on the 2-Step F-W Model. The F-W model has been shown to apply to a variety of aggregation processes in nature, including (first) nanocluster formation (42,43,44,45,46,47) and more recently protein fibrillation (31). It consists, typically, of slow nucleation followed by fast autocatalytic growth, as displayed in Scheme 1, where k_1 and k_2 represent the rate constants for nucleation and growth, respectively. In the case of protein aggregation, A corresponds to a monomeric form of the protein while B corresponds to protein in oligomers that are past the nucleation stage. As species B is both a product and a catalyst for growth, the second step of the F-W model is referred to as autocatalytic (the strict definition for autocatalysis being $A+B \rightarrow 2B$).

The F-W model results from the use of the concept of pseudo-elementary kinetics, an approach pioneered in the 1970s by R. Noyes (48). With this approach, the many substeps involved in nucleation and growth, are combined into two pseudo-elementary steps of nucleation and growth, which behave kinetically as (pseudo) elementary steps.[5] The two pseudo-elementary steps of nucleation and growth in the F-W model yield an integrable rate equation (eq. 1) which can be used to fit concentrations as a function of time, as shown here and elsewhere (31,42,43,44,45,46,47). The resulting rate constants k_1 and k_2 for nucleation and growth, respectively, can thus be seen as pseudo-elementary rate constants (i.e., where k_1 and k_2 are pseudo elementary step "rate constants" that may vary with concentration, and therefore, may not be true "constants").

[5] The pseudo-elementary step concept involves summing faster reactions with one or more slower reactions, giving an overall (sum) reaction that can be treated kinetically as an elementary step.

Because of the convolution of many steps into (kinetically useful) average k_1 and k_2 steps, the F-W model *is at best a starting point in attempts to elucidate the molecular picture of the aggregation process.*[6] The primary use of the F-W 2-step kinetic model is to deconvolute (average) nucleation from (average) growth. It is obviously a minimalistic kinetic model, and its limitations (vide infra) derive from that Ockham's razor nature.

Fitting Prion Aggregation Data with the F-W Model: Effect of the N-terminal and C-terminal Regions of the Sup35p Prion Domain on Nucleation and Growth. The prion domain (PrD) of the yeast prion protein Sup35p consists of two subdomains: (i) the N-terminal portion (amino acids 1-39), which is rich in glutamine/asparagine (Gln/Asn) residues and is thought to drive prion aggregation (35), and (ii) the C-terminal region (amino acids 40-114), which contains a series of imperfect oligopeptide repeats and is believed to play a role in prion propagation. Experiments with both Sup35p and Ure2p prion proteins have demonstrated that it is the overall amino acid composition (rather than the sequence) of the prion domain that is the predominant feature driving prion formation (37,49).

Among yeast species, the overall composition of the prion domain of Sup35p is relatively similar. Little homology in sequence is found for the N-terminal portion of

---

[6] An alternative form of the model, which provides more physical insight, is shown below, where B is replaced by the "full oligomer" B_n. This descriptive form does not however, allow for a simple mathematical treatment with analytically integrable rate equations. As such, it cannot be easily used to fit experimental aggregation kinetic data.

$$n\,A \xrightarrow{k_1'} B_n$$

$$A + B_n \xrightarrow{k_2'} B_{n+1}$$

PrD, but the amino acid composition is generally conserved, whereas in the C-terminal portion of PrD, both the amino acid composition and the overall repeat structure (thus the sequence) are generally conserved among species. Figure 2 (vide supra) shows aggregation curves obtained by Weissman and co-workers for the prion domain of Sup35p from *SC* and *CA* yeast species (32). Fits by the F-W model indicate that while nucleation kinetics (i.e., values of the nucleation rate constant k_1) vary considerably between species, growth kinetics (and the growth rate constant k_2) vary much less. Since the main difference between the prion domains of these yeast species lies in the sequence of the N-terminal portion, our results suggest the hypothesis that the N-terminal portion of the PrD in Sup35p affects nucleation more than growth, an observation consistent with Osterovish et al. (35).

When Weissman and co-workers screened the prion domain of Sup35p to look for mutations causing defects in amyloid formation, they found that (i) all these mutations were contained within a short region between residues 8 and 24 (a region with a high content of Gln and Asn residues); and (ii) most mutations were the result of a Gln or Asn mutation (33). Therefore, the authors suspected "a critical role for [the] Gln and Asn residues" in the N-terminal portion of the Sup35p PrD (33). Figure 3 shows the aggregation curves measured for the variants wild type (WT) and Q15R mutant (with Q→R at position 15) (33). Fits by the F-W model show that the mutation resulted in appreciably slower nucleation (k_1) kinetics, and in somewhat slower growth (k_2) kinetics. Since the mutation only involved substituting a glutamine (Q) by arginine (R) at position 15, our results indicate that the (Gln/Asn rich) N-terminal portion of the PrD in Sup35p affects nucleation kinetics more than growth kinetics. These results are also consistent

with subsequent single-fiber growth studies, which indicate that the Q15R mutant exhibits average growth kinetics similar to those of WT (50). The results also allow an intriguing prediction that merits future testing: although the Q15R mutant does not efficiently add to WT prion fibers in vivo (33), when expressed as the sole copy of Sup35p, the Q15R mutant should have a low frequency of prion formation, but should still be able to efficiently propagate prions once they form. It should be noted that this mutant was isolated based on its inability to propagate wild-type [*PSI+*]. However, this may reflect a species barrier between WT and Q15R rather than the intrinsic ability of Q15R to propagate prions.

The significance with respect to prion aggregation of oligopeptide repeat sequences, located in the C-terminal portion of the Sup35p prion domain, came into focus when Lindquist and co-workers created variants of Sup35p that differed only by the number of repeats (34). Figure 4a (vide supra) shows the aggregation curves obtained for Sup35p WT, a deletion variant RΔ2-5 (with repeats 2-5 deleted), and an expansion variant R2E2 (with repeat 2 expanded twice). Fits by the F-W model suggest that the number of repeats affects growth kinetics, as indicated by a correlation between values of the growth rate constant k_2 and the (difference in) number of repeats, Figure 4c. On the other hand (in part due to large error bars), we could not detect a reliable correlation between the nucleation rate constant k_1 and the (difference in) number of repeats (Figure 4b). Since the variants differed by the number of oligopeptide repeat sequences, our analysis implies the hypothesis that the (repeats-containing) C-terminal portion of the PrD in Sup35p plays a role that affects growth kinetics, an observation consistent with the proposed role of the repeats in prion propagation (35).

Fitting Prion Aggregation Data with the F-W Model: Effect of a Gln/Asn Rich Region of the Ure2p Prion Domain on Nucleation and Growth. The prion domain of Ure2p (residues 1-90) shows consistently a high Gln/Asn content. As there is no evidence of separate aggregation or propagation subdomains, the Ure2p PrD is generally not subdivided. Perrett and co-workers were interested in the role of specific regions of the Ure2p PrD that are particularly rich in Gln and Asn, and created a 15Ure2 mutant with residues 1-14 (a Gln/Asn rich region) removed (36). Figure 5a, vide supra, shows aggregation curves measured for the WT and 15Ure2 variants. Fits by the F-W model indicate that the mutant displays substantially slower nucleation (k_1) kinetics, but rather similar growth (k_2) kinetics. Because the mutation results from the removal of a Gln/Asn rich region, our results suggest the hypothesis that this Gln/Asn rich region in the PrD of Ure2p plays a role that affects nucleation kinetics more than growth kinetics (37). Further tests of this and the other hypotheses provided above are, of course, needed.

A Return to the Issue of Fragmentation. We began these studies expecting that the F-W 2-step model would *not* be able to fit prion aggregation kinetic data due to the fact that it does not specifically account for fragmentation step(s), yet fragmentation is generally believed to be an integral part of prion aggregation (51,52,53,54,55). However, the generally good to excellent fits to the literature prion aggregation kinetic data (i.e., specifically to the data which shows sigmoidal behavior) by the F-W 2-step model demonstrated herein were achieved without specifically including fragmentation. These two seemingly inconsistent facts (the literature evidence for fragmentation, *op. cit.*, vs the lack of a specific step to account for fragmentation in the F-W 2-step model) require thought and subsequent comment.

Three logical hypotheses seem possible here: (i) fragmentation is not really a part of at least the kinetic pathway for prion aggregation as previously believed; we believe, however, that this idea cannot be right in light of the extant evidence for fragmentation (51,52,53,54,55). (ii) A second thought is that too little is known at present about the true products of prion aggregation and their structures versus time to make a definitive statement about the role of fragmentation in the mechanism of prion aggregation. A related issue here is that fibers are believed to grow from one or both ends only (56,57,58). If fibers grow only from their ends, and if B is the growth surface, then one would expect the growth step to be expressed as "A+B→B" instead of "A+B→2B". However, this stoichiometry per se can be ruled out as it is inconsistent with the observed autocatalytic curves [i.e., with the need for A+B→2B (the kinetic definition of autocatalysis) as the second step of the F-W kinetic model]. (iii) A third, more likely hypothesis in our opinion is that the F-W 2-step model may be "hiding" fragmentation somehow. For example, if the number of fiber ends closely parallels the total amount of aggregated protein, which was predicted in one model of Sup35p aggregation (54), then B would be approximately proportional to the growth surface, effectively hiding fragmentation in "B". A related issue here is that the way a particular physical method monitors "growth" may be hiding fragmentation within the description of "B". Again, the need for additional studies is apparent (41).

Examination of Four Prior Mechanisms for Prion Transmission and Propagation. Four prior mechanisms have been suggested for the transmission and propagation of prions, as briefly discussed in the introductory section and summarized in Table 1. Attempts have been made in the prion literature to test the validity of these mechanisms

(23,59).[7] In our view, all four mechanisms present valid and alternative descriptions of prion aggregation at the molecular level; however, none can be ruled out at the present. Attempts to rule out any of these four mechanisms by using kinetic data (a) have both not been done, and (b) will not be possible until these word/picture mechanisms are turned into specific reactions, defining individual, specific rate constants that can then be used to derive differential or integrable rate equations for actual data fitting.[7] Nevertheless, in time, the true picture of prion aggregation at the molecular level may turn out to contain features from all four mechanisms.

Caveats and Limitations of the F-W Model. The F-W model appears to fit a broad range of kinetic aggregation data (31,42,43,44,45,46,47) including amyloid-β, α-synuclein, polyglutamine, and now the prion protein aggregation systems. However, and as was previously addressed (31), the F-W model has limitations:

i) The F-W 2-step model is obviously a highly condensed, oversimplified, *phenomenological kinetic model* of the real, molecular prion aggregation process that probably consists of hundreds if not thousands of steps in most cases.

[7] Attempts have been made in the yeast prion literature to test the validity of these four (molecular) mechanisms by looking at the effect of yeast protein concentration on the observed lag times and aggregation rates of the protein in its prion form. There, for each mechanism, the authors state expected correlations between the observed lag times/aggregation rates and initial protein concentration that are essentially *qualitative* in nature. The authors use these stated expectations *quantitatively*, as a proposed way to test the likelihood of each mechanism. We will argue that these correlations were derived in a manner that is not strictly proper kinetically, since a kinetically rigorous treatment would require that: (i) chemical equations be written; (ii) appropriate rate equations be given; and where possible, that (iii) either numerical integration be used, or integrated rate equations be provided and used.

ii) The resultant k_1 and k_2 values are, therefore, averages over all of the true underlying steps. As such, important kinetic and mechanistic information must be hidden in the average $A \rightarrow B$ and $A + B \rightarrow 2B$ pseudo-elementary steps and their resultant, average k_1 and k_2 rate constants.

iii) Another issue is that during the nucleation process, [A] is approximately constant. Therefore, a higher kinetic order in [A], that is $k_1'[A]^n$, is easily hidden kinetically and can appear as an apparent $k_{1(apparent)}[A]$ dependence, where in fact $k_{1(apparent)} = k_1'[A]^{n-1}$ (60). A study of any concentration dependence of $k_{1(apparent)}$ can be used to overcome this issue, however.

iv) Finally, the fact that all sizes of the growing aggregate are hidden behind the general descriptor "B", while a great advantage for extracting an average k_2, hides important molecular details that will need to be the focus of separate studies. The possible role of fragmentation in the proper description of "B" vs. time is a related issue here, again one requiring future studies.

In light of the above limitations, we have attempted to be cautious in the interpretations herein and urge others to use caution when interpreting results from the F-W model.

Conclusions

In conclusion, the main contributions from this paper can be summarized as follows:

i) Twenty-seven prion aggregation curves from the literature, measured by seven different physical methods, were successfully fit using the F-W model, allowing

for the deconvolution of average nucleation (k_1) and growth (k_2) kinetic parameters for the first time in all cases.

ii) The F-W model was able to fit the in vitro aggregation curves of three different types of yeast prions, as well as that of mouse and human prion.

iii) The aggregation of Sup35p measured by 6 different physical methods under the same experimental conditions (23), fit by the F-W model, provided values of the nucleation and growth rate constants that were equivalent within experimental error. This is an important study (23), and the k_1 and k_2 analysis provided herein yields considerably more confidence than before that true aggregation rates (and now, k_1 and k_2 rate constants) can be faithfully measured using different methods, at least for this Sup35p system.

iv) Fits of Sup35p aggregation curves by the F-W model indicate the hypotheses requiring further study that (i) the Glu/Asn rich N-terminal region of the prion domain appears to influence the kinetics of *nucleation* more than growth and (ii) the oligopeptide repeat containing C-terminal region of the prion domain appears to affect the kinetics of *growth* more than nucleation.

v) Fits of Ure2p aggregation curves by the F-W model indicate the hypothesis that Gln/Asn rich regions in the prion domain of this protein appear to influence the kinetics of nucleation more than growth.

vi) Examination of four prior mechanisms suggested for prion transmission and propagation reveals that these prior mechanisms, while useful molecular descriptions, have not yet been turned into useful kinetic models that can be

used for fitting data and, thereby, distinguishing those postulated (molecular) mechanisms from one another.

vii) The F-W model provides, therefore, the presently available *kinetic* model/ minimal kinetic mechanism of prion aggregation and its average nucleation (k_1) and growth (k_2) rate constants.

viii) Caveats of the F-W model were discussed, particularly the facts that (a) since each pseudo-elementary step hides many (sub)steps at the molecular level, the rate constants obtained for nucleation and growth are actually *average* rate constants, and (b) the average nature of B also hides potentially important data about fibril size vs. time effects, including the possible contributions of fragmentation effects.

ix) Hence, the need for future experimental work on the issues of agitation and fragmentation was noted.

x) Also included in the needed future experiments are more direct experiments on what "A" and "B" are in prion aggregation, and B vs. time. How do those "A" and "B" relate to the molecular descriptions in the previously suggested mechanisms in Table 1? These issues, and the three hypotheses given in the Discussion section, should provide fertile ground for future studies.

Acknowledgement. It is important to specifically acknowledge and thank all the research groups and authors listed in the references and Table 1. Those authors, and not us, provided the primary experimental and other data which were analyzed herein. Without their original efforts, nothing reported herein would have been possible.

Supporting Information Available: Raw, digitized data (i.e., not normalized to an initial minimum of zero) corresponding to Figures 2-7, and 9-10 of the main text along with their fits to the F-W model and ensuing discussion on the observed effects of agitation on the aggregation curve of Sup35p is provided.

References

1. Prusiner, S. B., Scott, M. R., DeArmond, S. J., and Cohen, F. E. (1998) Prion protein biology, *Cell 93*, 337-348.

2. Prusiner, S. B. (1982) Novel proteinaceous infectious particles cause scrapie, *Science 216*, 136-144.

3. Somerville, R. A., Bendheim, P. E., and Bolton, D. C. (1991) The transmissible agent causing scrapie must contain more than protein, *Rev. Med. Virol. 1*, 131-139.

4. Weissmann, C. (1991) A 'unified theory' of prion propagation, *Nature 352*, 679-683.

5. Lieberman, M., Marks, A. D., and Smith, C. (2007) *Marks' Essentials of Medical Biochemistry. A Clinical Approach*, Lippincott Williams & Wilkins, Baltimore.

6. Wickner, R. B., Edskes, H. K., Shewmaker, F., and Nakayashiki, T. (2007) Prions of fungi: inherited structures and biological roles, *Nature Reviews Microbiology 5*, 611-618.

7. Ross, E. D., Minton, A., and Wickner, R. B. (2005) Prion domains: Sequences, structures and interactions, *Nat. Cell Biol. 7*, 1039-1044.

8. Wickner, R. B., Edskes, H. K., Roberts, B. T., Baxa, U., Pierce, M. M., Ross, E. D., and Brachmann, A. (2004) Prions: Proteins as genes and infectious entities, *Genes Dev. 18*, 470-485.

9. Wickner, R. B. (1994) [URE3] as an altered URE2 protein: evidence for a prion alnalog in *Saccharomyces cerevisiae*, *Science 264*, 566-569.

10. Chernoff, Y. O., Derkach, I. L., and Ingevechtomov, S. G. (1993) Multicopy Sup35 gene induces *de novo* appearance of *PSI*-like factors in the yeast *Saccharomyces Cerevisiae*, *Current Genetics 24*, 268-270.

11. Derkatch, I. L., Bradley, M. E., Hong, J. Y. and Liebman, S. W. (2001) Prions affect the appearance of other prions: the story of [PIN(+)], *Cell 106*, 171-182.

12. Patel, B. K., and Liebman, S. W. (2007) "Prion-proof" for [*PIN*+]: infection with *in vitro*-made amyloid aggregates of Rnq1p-(132-405) induces [*PIN*+], *J. Mol. Biol. 365*, 773-782.

13. Brachmann, A., Baxa, U., and Wickner, R. B. (2005) Prion generation *in vitro*: amyloid of Ure2p is infectious, *EMBO J. 24*, 3082-3092.

14. Tanaka, M., Chien, P., Naber, N., Cooke, R., and Weissman, J. S. (2004) Conformational variations in an infectious protein determine prion strains differences, *Nature 428*, 323-328.

15. King, C. Y., and Diaz-Avalos, R. (2004) Protein-only transmission of three yeast prion strains, *Nature 428*, 319-323.

16. Atwood, C. S., Obrenovich, M. E., Liu, T. B., Chan, H., Perry, G., Smith, M. A., and Martins, R. N. (2003) Amyloid-beta: A chameleon walking in two worlds: A review of the trophic and toxic properties of amyloid-beta, *Brain Res. Rev. 43*, 1-16.

17. Walsh, D. M., and Selkoe, D. J. (2007) Aβ oligomers—A decade of discovery, *J. Neurochem. 101*, 1172-1184.

18. Ferreira, S. T., Vieira, M. N. N., and Felice, F. G. D. (2007) Soluble protein oligomers as emerging toxins in Alzheimer's and other amyloid diseases, *IUBMB Life 59*, 332-345.

19. Fink, A. L. (2006) The aggregation and fibrillation of α-synuclein *Acc. Chem. Res.*, 628-634.

20. Walsh, D. M., Lomakin, A., Benedek, G. B., Condron, M. M., and Teplow, D. B. (1997) Amyloid β-protein fibrillogenesis: Detection of a protofibrillar intermediate, *J. Biol. Chem. 272*, 22364-22372.

21. Harper, J. D., Wong, S. S., Lieber, C. M., and P.T. Lansbury, J. (1997) Observation of metastable Aβ amyloid protofibrils by atomic force microscopy, *Chem. Biol. 4*, 119-125.

22. Gosal, W. S., Morten, I. J., Hewitt, E. W., Smith, D. A., Thomson, N. H., and Radford, S. E. (2005) Competing pathways determine fibril morphology in the self-assembly of beta(2)-microglobulin into amyloid, *J. Mol. Biol. 351*, 850-864.

23. Serio, T. R., Cashikar, A. G., Kowal, A. S., Sawicki, G. J., Moslehi, J. J., Serpell, L., Arnsdorf, M. F., and Lindquist, S. L. (2000) Nucleated conformational conversion and the

replication of conformational information by a prion determinant, *Science 289*, 1317-1321.

24. Eigen, M. (1996) Prionics or the kinetic basis of prion diseases, *Biophys. Chem. 63*, A1-A18.

25. Prusiner, S. B. (1991) Molecular-biology of prion diseases, *Science 252*, 1515-1522.

26. Cohen, F. E., Pan, K.-E., Huang, Z., Baldwin, M., Fletterick, R. J., and Prusiner, S. B. (1994) Structural clues to prion replication, *Science 264*, 530-531.

27. Come, J. H., Fraser, P. E., and Lansbury, P. T. (1993) A kinetic model for amyloid formation in the prion diseases. Importance of seeding, *Proc. Natl. Acad. Sci. USA 90*, 5959-5963.

28. Jarrett, J. T., and Lansbury, P. T. (1993) Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie?, *Cell 73*, 1055-1058.

29. Griffith, J. S. (1967) Self-replication and scrapie, *Nature 215*, 1043-1044.

30. Morris, A. M., Watzky, M. A., and Finke, R. G., Protein aggregation kinetics, mechanism, and curve-fitting: a critical review of the literature, *Biochim. Biophys. Acta,* submitted.

31. Morris, A. M., Watzky, M. A., Agar, J. N., and Finke, R. G. (2008) Fitting neurological protein aggregation kinetic data via a 2-step, minimal/"Ockham's razor" model: the Finke-Watzky mechanism of nucleation followed by autocatalytic surface growth, *Biochemistry 47*, 2413-2427.

32. Santoso, A., Chien, P., Osherovich, L. Z., and Weissman, J. S. (2000) Molecular basis of a yeast prion species barrier, *Cell 100*, 277-288.

33. DePace, A. H., Santoso, A., Hillner, P., and Weissman, J. S. (1998) A critical role for amino-terminal glutamine/asparagines repeats in the formation and propagation of a yeast prion, *Cell 93*, 1241-1252.

34. Liu, J.-J., and Lindquist, S. L. (1999) Oligopeptide-repeat expansions modulate 'protein-only' inheritance in yeast, *Nature 400*, 573-576.

35. Osherovich, L. Z., Cox, B. S., Tuite, M. F., and Weissman, J. S. (2004) Dissection and design of yeast prions, *PLoS Biol. 2*, 442-451.

36. Jiang, Y., Li, H., Zhu, L., Zhou, J.-M., and Perrett, S. (2004) Amyloid nucleation and hierarchical assembly of Ure2p fibrils, *J. Biol. Chem. 279*, 3361-3369.

37. Ross, E. D., Edskes, H. K., Terry, M. J., and Wickner, R. B. (2005) Primary sequence independence for prion formation, *Proc. Natl. Acad. Sci. USA 102*, 12825-12830.

38. Zhu, L., Zhang, X. J., Wang, L. Y., Zhou, J. M., and Perrett, S. (2003) Relationship between stability of folding intermediates and amyloid formation for the yeast prion Ure2p: A quantitative analysis of the effects of pH and buffer system, *J. Mol. Biol. 328*, 235-254.

39. Baskakov, I. V., Legname, G., Prusiner, S. B., and Cohen, F. E. (2001) Folding of prion protein to its native α-helical conformation is under kinetic control, *J. Biol. Chem. 276*, 19687-19690.

40. Baskakov, I. V. and Bocharova, O. V. (2005) In vitro conversion of mammalian prion protein into amyloid fibrils displays unusual features, *Biochemistry 44*, 2339-2348.

41. Watzky, M. A., Morris, A. M., Ross, E. D. and Finke, R. G., *studies in progress.*

42. Watzky, M. A., and Finke, R. G. (1997) Transition metal nanocluster formation kinetic and mechanistic studies. A new mechanism when hydrogen is the reductant: slow, continuous nucleation and fast autocatalytic surface growth, *J. Am. Chem. Soc. 119*, 10382-10400.

43. Aiken III, J. D., and Finke, R. G. (1998) Nanocluster formation synthetic, kinetic, and mechanistic studies. The detection of, and then methods to avoid, hydrogen mass-transfer limitations in the synthesis of polyoxoanion- and tetrabutylammonium-stabilized, near-monodisperse 40 ± 6 Å Rh(0) nanoclusters, *J. Am. Chem. Soc. 120*, 9545-9554.

44. Widegren, J. A.; Aiken III, J. D., Özkar, S., and Finke, R. G. (2001) Additional investigations of a new kinetic method to follow transition-metal nanocluster formation, including the discovery of heterolytic hydrogen activation in nanocluster nucleation reactions, *Chem. Mater. 13*, 312-324.

45. Hornstein, B. J., and Finke, R. G. (2004) Transition-metal nanocluster kinetic and mechanistic studies emphasizing nanocluster agglomeration: demonstration of a kinetic method that allows monitoring of all three phases of nanocluster formation and aging, *Chem. Mater. 16*, 139-150. See also correction Hornstein, B. J., and Finke, R. G. (2004) *Chem. Mater. 16*, 3972-3972.

46. Besson, C., Finney, E. E., and Finke, R. G. (2005) A mechanism for transition-metal nanoparticle self-assembly, *J. Am. Chem. Soc. 127*, 8179-8184.

47. Besson, C., Finney, E. E., and Finke, R. G. (2005) Nanocluster nucleation, growth, and then agglomeration kinetic and mechanistic studies: a more general, four-step mechanism involving double autocatalysis, *Chem. Mater. 17*, 4925-4938.

48. Field, R. J., and Noyes, R. M. (1977) Oscillations in chemical systems. 18. Mechanisms of chemical oscillators: Conceptual bases, *Acc. Chem. Res. 10*, 214-221.

49. Ross, E. D., Baxa, U., and Wickner, R. B. (2004) Scrambled prion domains form prions and amyloid, *Mol. Cell. Biol. 24*, 7206-7213.

50. Chien, P., DePace, A. H., Collins, S. R., and Weissman, J. S. (2003) Generation of prion transmission barriers by mutational control of amyloid conformations, *Nature 424*, 948-951.

51. Paushkin, S. V., Kushnirov, V. V., Smirnov, V. N., and Ter-Avanesyan, M. D. (1996) Propagation of the yeast prion-like [psi+] determinant is mediated by oligomerization of the SUP35-encoded polypeptide chain release factor, *EMBO J. 15*, 3127-3134.

52. Tanaka, M., Collins, S. R., Toyama, B. H., and Weissman, J. S. (2006) The physical basis of how prion conformations determine strain phenotypes, *Nature 442*, 585-589.

53. Ness, F., Ferreira, P., Cox, B. S., and Tuite, M. F. (2002) Guanidine hydrochloride inhibits the generation of prion "seeds" but not prion protein aggregation in yeast, *Mol. Cell. Biol. 22*, 5593-605.

54. Collins, S. R., Douglass, A., Vale, R. D., and Weissman, J. S. (2004) Mechanism of prion propagation: amyloid growth occurs by monomer addition, *PLoS Biol. 2*, 1582-1590.

55. Wegrzyn, R. D., Bapat, K., Newnam, G. P., Zink, A. D., and Chernoff, Y. O. (2001) Mechanism of prion loss after Hsp104 inactivation in yeast, *Mol. Cell. Biol. 21*, 4656-4669.

56. Scheibel, T., Kowal, A. S., Bloom, J. D., and Lindquist, S. L. (2001) Bidirectional amyloid fiber growth for a yeast prion determinant, *Curr. Biol. 11*, 366-369.

57. Inoue, Y., Kishimoto, A., Hirao, J., Yoshida, M., and Taguchi, H. (2001) Strong growth polarity of yeast prion fiber revealed by single fiber imaging, *J. Biol. Chem. 276*, 35227-35230.

58. DePace, A. H., and Weissman J. S. (2002) Origins and kinetic consequences of diversity in Sup35 yeast prion fibers, *Nat. Struct. Biol. 9*, 389-396.

59. Kelly, J. W. (2000) Mechanisms of amyloidogenesis, *Nat. Struct. Biol. 7*, 824-826.

60. Watzky, M. A., Ott, L. S., Finney, E. E., and Finke, R. G. *studies in progress.*

Supporting Information for:

Fitting Yeast and Mammalian Prion Aggregation Kinetic Data with the Finke-Watzky 2-Step Mechanism of Nucleation and Autocatalytic Growth

Murielle A. Watzky, Aimee M. Morris, Eric D. Ross, and Richard G. Finke

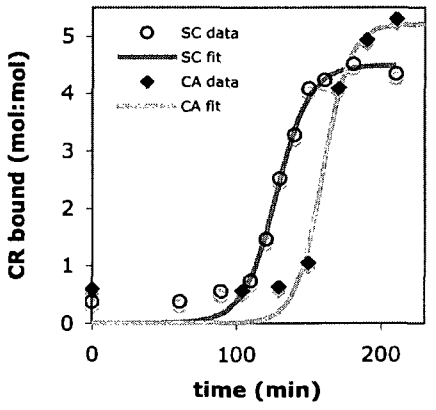


Figure S1. Weissman and co-workers' raw, digitized data (1) for the aggregation of Sup35p PrDM that was expressed in different yeast species, measured by Congo Red binding at [PrD]= 2.5 μM with continuous rotation along with the fits to the F-W model. Resultant rate constants for *Saccharomyces cerevisae* (*SC*): k_1= 1(2) × 10^{-6} min^{-1}, k_2= 3.5(7) × 10^{-2} $\mu M^{-1}.min^{-1}$, R^2= 0.983; for *Candida albicans* (*CA*): k_1= 3(21) × 10^{-9} min^{-1}, k_2= 4(2) × 10^{-2} min^{-1}, R^2= 0.973. See Figure 2 of the main text for the corrected data and fits.

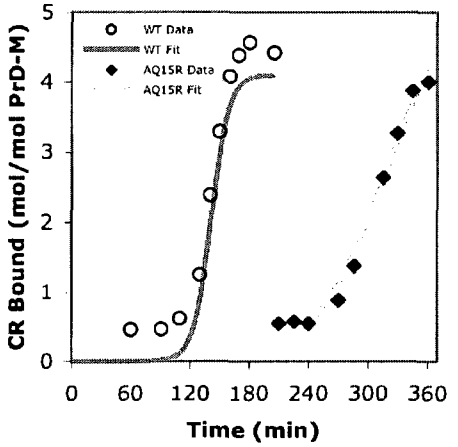


Figure S2. Weissman and coworkers' (2) Sup35p and mutant AQ15R raw, digitized aggregation data measured by Congo Red binding. The data were fit to the F-W mechanism. Resultant rate constants for WT: k_1=1(2) × 10^{-6} min^{-1},k_2=3.3(9) × 10^{-2} μM^{-1} min^{-1}, R^2=0.978; and AQ15R: k_1=4(4) × 10^{-6} min^{-1}, k_2=8(3) × 10^{-3} μM^{-1} min^{-1}, R^2=0.983. For the corrected data fit, see Figure 3 of the main text.

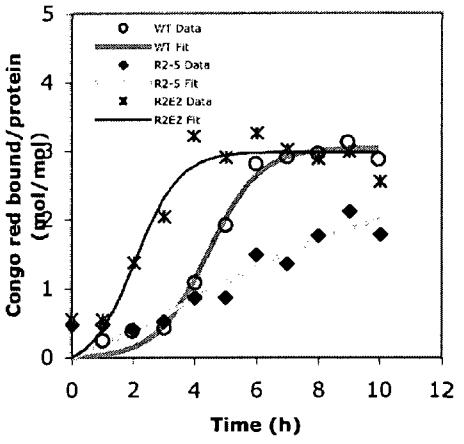


Figure S3. WT Sup35p along with mutants, RΔ2-5 and R2E2, raw aggregation data measured by Congo Red from Lui and Lindquist (3). The RΔ2-5 mutant contains four less repeat sequences than the WT, while R2E2 has two more repeat sequences (3). The data were digitized and fit to the F-W mechanism. The following rate constants and R^2 values were obtained: WT, $k_1=6(5) \times 10^{-3}$ h^{-1}, $k_2=3.9(8) \times 10^{-1}$ μM^{-1} h^{-1}, and $R^2=0.982$; RΔ2-5, $k_1=7(3) \times 10^{-2}$ h^{-1}, $k_2=1(2) \times 10^{-1}$ μM^{-1} h^{-1}, and $R^2=0.862$; R2E2, $k_1=1(2) \times 10^{-2}$ h^{-1}, $k_2=1.1(4)$ μM^{-1} h^{-1} and $R^2=0.918$. See Figure 4 of the main text for the corrected data and fits.

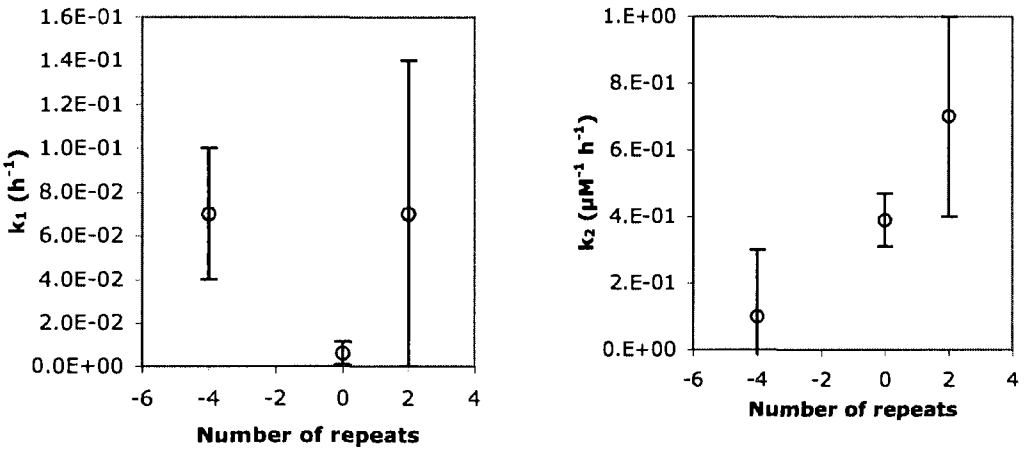


Figure S4. Correlations between the number of repeat sequences, if you assume WT has zero repeats, and the k_1, k_2 values obtained from the raw data fits in Figure S3.

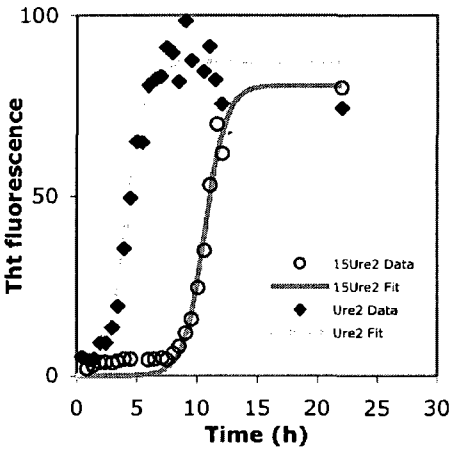


Figure S5. Perrett and coworkers' (4) Ure2 and mutant 15Ure2 aggregation data measured by fluorescence. The raw (uncorrected) data were digitized and fit to the F-W mechanism. Rate constants for Ure2 are k_1=7(4) × 10^{-2} h^{-1}, k_2=3.5(5) × 10^{-2} μM^{-1} h^{-1}, R^2=0.972. Rate constants for 15Ure2 are k_1= 6(7) × 10^{-6} h^{-1}, k_2=2.8(4) × 10^{-2} μM^{-1} h^{-1}, R^2=0.974.

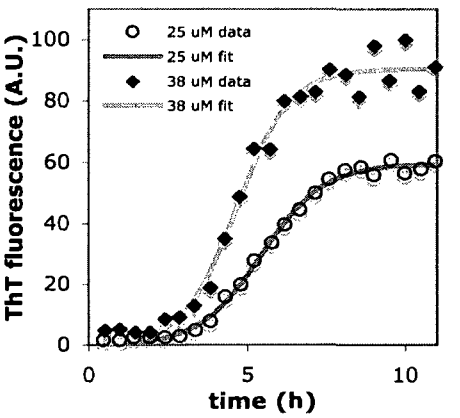


Figure S6. Perrett and co-workers' Ure2p aggregation measured by ThT fluorescence for different [Ure2p] concentrations at pH = 7.5 and T = 25°C, in 50 mM Tris buffer *with shaking* (5). The uncorrected data was digitized and fit to the F-W model. Resultant rate constants for [Ure2p]=25 μM: k_1=4.0(8) × 10^{-3} h^{-1}, k_2= 4.0(2) × 10^{-2} $\mu M^{-1}.h^{-1}$, R^2= 0.996; [Ure2p]=38 μM: k_1= 4(2) × 10^{-3} h^{-1}, k_2= 3.1(4) × 10^{-2} $\mu M^{-1}.h^{-1}$, R^2= 0.984. The corrected data and resultant fits are shown in Figure 6 of the main text.

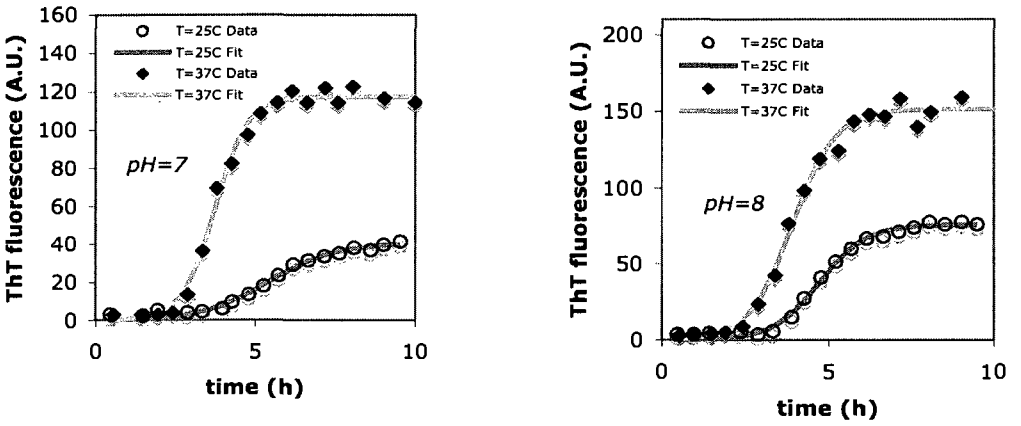


Figure S7. (a) Perrett and co-workers' Ure2p aggregation measured by ThT fluorescence for different temperatures at pH = 7 and [Ure2p]= 30μM, in 50 mM phosphate buffer *with shaking* (5). The data was digitized (not corrected) and fit to the F-W model. Resultant rate constants for T = 25 °C: k_1= 5(2) × 10^{-3} h^{-1}, k_2= 3.2(4) × 10^{-2} $\mu M^{-1}.h^{-1}$, R^2= 0.988; for T = 37 °C: k_1= 1.3(7) × 10^{-3} h^{-1}, k_2= 6.5(6) × 10^{-2} $\mu M^{-1}.h^{-1}$, R^2= 0.994. (b) Perrett and co-workers' Ure2p aggregation measured by ThT fluorescence for different temperatures at pH = 8 and [Ure2p]= 30μM, in 50 mM phosphate buffer with shaking (5). The data was digitized and fit to the F-W model. Resultant rate constants for T = 25 °C: k_1= 1.6(7) × 10^{-3} h^{-1}, k_2= 4.8(4) × 10^{-2} $\mu M^{-1}.h^{-1}$, R^2= 0.994; for T = 37 °C : k_1= 4(1) × 10^{-3} h^{-1}, k_2= 5.1(5) × 10^{-2} $\mu M^{-1}.h^{-1}$, R^2= 0.993.

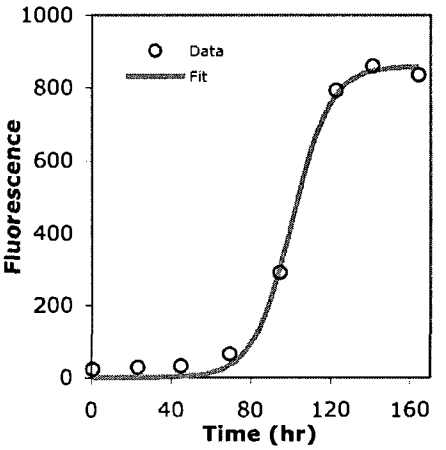


Figure S8. Uncorrected kinetic data of the assembly of 40 μM mouse prion protein from Cohen and coworkers (6) and fit to the F-W mechanism. The resultant rates constants are k_1=3(5) × 10^{-6} hr^{-1} and k_2=2.6(4) × 10^{-3} μM^{-1} hr^{-1} with R^2=0.996.

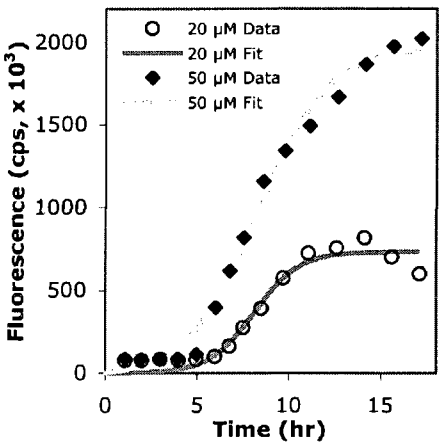


Figure S9. Baskakov et al.'s (7) raw kinetic data of human recombinant prion protein at varying concentrations fit to the F-W mechanism. The following rate constants and R^2 values are obtained for each concentration: 20 μM, $k_1=1(1) \times 10^{-3}$ hr^{-1}, $k_2=4(1) \times 10^{1}$ μM^{-1} hr^{-1}, $R^2=0.955$; 50 μM, $k_1=7(2) \times 10^{-3}$ hr^{-1}, $k_2=1.0(1) \times 10^{1}$ μM^{-1} hr^{-1}, $R^2=0.990$.

Attempting to Analyze the Effects of Agitation (and Associated Fragmentation) on Nucleation and Growth. A major issue associated with prion aggregation (but not necessarily with other amyloid aggregation) is fragmentation (defined as the breakage of existing fibrils to smaller, multiple fibrils). It has been shown in the literature that *in vivo* fragmentation of prions occurs during propagation (8,9,10,11). In order to explore the effect of fragmentation on yeast prion aggregation, we searched the literature for an *unseeded* prion aggregation data set under identical experimental conditions, with and without agitation that can induce fragmentation. We found only two data sets that could be readily digitized: one from Liebman and co-workers that unfortunately appeared to be truncated at the base (12), and one from Weissman and co-workers for both an agitated and a non-agitated sample of the prion protein Sup35p under otherwise identical *aggregation* conditions, Figure S10 (16).

The data from Weissman and co-workers was fit using the F-W model yielding the results of the nucleation rate constant (k_1) for the agitated prion system being apparently 8 orders of magnitude larger than that obtained for the non-agitated system. This large difference is also visually reflected in the large differences in the lag times observed (Figure S10), the larger lag time corresponding to much slower nucleation (13,14,15). There is also an observed difference of approximately half an order of magnitude in the apparent growth rate constant, k_2, its value being larger for the agitated sample than for the non-agitated sample. On the surface, these observations would seem to suggest that the agitation effects in this system are primarily in the k_1 nucleation step.

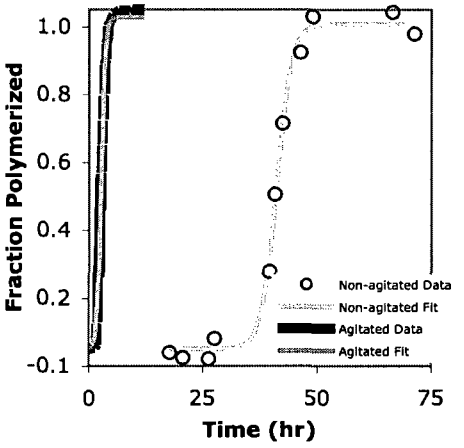


Figure S10. Weismann and co-workers' kinetics of aggregation of the NM domain of Sup35p with and without agitation fit to the F-W model (16). With agitation: k_1(apparent)=5.2(6) × 10^{-3} h^{-1}; k_2=2.26(7) × 10^{-2} μM^{-1} h^{-1}; R^2=0.986. Without agitation: k_1=1(6) × 10^{-11} h^{-1}; k_2=6(2) × 10^{-3} μM^{-1} h^{-1}; R^2=0.995.

However, the dramatic difference in k_1 values is not expected to arise simply from the effects of agitation. Overall we could not disprove that one or more of the following experimental artifacts may be affecting this data set. (i) The aggregation of the agitated

sample was measured with continuous ThT fluorescence, and that of the non-agitated sample with CR binding. It is not clear how the binding properties of these dyes might be affected by agitation. (ii) It has previously been observed that a larger number of shorter fibrils might cause an increase in ThT fluorescence intensity (12). (iii) The fibril growth in the agitated reaction may be too fast for the response time of the ThT fluorescent dye, so that the desired kinetics of aggregation are not what is actually being measured.

Obviously, more studies where fragmentation is involved, and where a F-W analysis can be performed (that will need to include fragmentation as a discrete step), are needed.

References

1. Santoso, A., Chien, P., Osherovich, L. Z., and Weissman, J. S. (2000) Molecular basis of a yeast prion species barrier, *Cell 100*, 277-288.

2. DePace, A. H., Santoso, A., Hillner, P., and Weissman, J. S. (1998) A critical role for amino-terminal glutamine/asparagines repeats in the formation and propagation of a yeast prion, *Cell 93*, 1241-1252.

3. Liu, J.-J., and Lindquist, S. L. (1999) Oligopeptide-repeat expansions modulate 'protein-only' inheritance in yeast, *Nature 400*, 573-576.

4. Jiang, Y., Li, H., Zhu, L., Zhou, J.-M., and Perrett, S. (2004) Amyloid nucleation and hierarchical assembly of Ure2p fibrils, *J. Biol. Chem. 279*, 3361-3369.

5. Zhu, L., Zhang, X. J., Wang, L. Y., Zhou, J. M., and Perrett, S. (2003) Relationship between stability of folding intermediates and amyloid formation for the yeast prion Ure2p: A quantitative analysis of the effects of pH and buffer system, *J. Mol. Biol. 328*, 235-254.

6. Baskakov, I. V., Legname, G., Prusiner, S B., and Cohen, F. E. (2001) Folding of prion protein to its native α-helical conformation is under kinetic control, *J. Biol. Chem. 276*, 19687-19690.

7. Baskakov, I. V. and Bocharova, O. V. (2005) In vitro conversion of mammalian prion protein into amyloid fibrils displays unusual features, *Biochemistry 44*, 2339-2348.

8. Paushkin, S. V., Kushnirov, V. V., Smirnov, V. N., and Ter-Avanesyan, M. D. (1996) Propagation of the yeast prion-like [psi+] determinant is mediated by oligomerization of the SUP35-encoded polypeptide chain release factor, *EMBO J. 15*, 3127-3134.

9. Wegrzyn, R. D., Bapat, K., Newnam, G. P., Zink, A. D., and Chernoff, Y. O. (2001) Mechanism of prion loss after Hsp104 inactivation in yeast, *Mol. Cell. Biol. 21*, 4656-4669.

10. Ness, F., Ferreira, P., Cox, B. S., and Tuite, M. F. (2002) Guanidine hydrochloride inhibits the generation of prion "seeds" but not prion protein aggregation in yeast, *Mol. Cell. Biol. 22*, 5593-5605.

11. Tanaka, M., Collins, S. R., Toyama, B. H., and Weissman, J. S. (2006) The physical basis of how prion conformations determine strain phenotypes, *Nature 442*, 585-589.

12. Patel, B. K., and Liebman, S. W. (2007) "Prion-proof" for [*PIN*+]: Infection with *in vitro*-made amyloid aggregates of Rnq1p-(132-405) induces [*PIN*+], *J. Mol. Biol. 365*, 773-782.

13. Watzky, M. A., and Finke, R. G. (1997) Transition metal nanocluster formation kinetic and mechanistic studies. A new mechanism when hydrogen is the reductant: slow, continuous nucleation and fast autocatalytic surface growth, *J. Am. Chem. Soc. 119*, 10382-10400.

14. Morris, A. M., Watzky, M. A., and Finke, R. G., Protein aggregation kinetics, mechanism, and curve-fitting: a critical review of the literature, *in progress*.

15. Watzky, M. A., Finney, E. E., and Finke, R. G. *studies in progress*.

16. Collins, S. R., Douglass, A., Vale, R. D., and Weissman, J. S. (2004) Mechanism of prion propagation: amyloid growth occurs by monomer addition, *PLoS Biol. 2*, 1582-1590.

CHAPTER V

α-SYNUCLEIN AGGREGATION VARIABLE TEMPERATURE AND VARIABLE pH KINETIC DATA: A RE-ANALYSIS USING THE FINKE-WATZKY 2-STEP MODEL OF NUCLEATION AND AUTOCATALYTIC GROWTH

This dissertation chapter contains a full paper published in *Biophysical Chemistry* **2009**, *140*, 9-15. Insights gained from the analysis of literature α-synuclein aggregation kinetic data using the Finke-Watzky 2-step kinetic model are presented and discussed.

α-Synuclein AggregationVariable Temperature and Variable pH Kinetic Data: A Re-analysis using the Finke-Watzky 2-Step Model of Nucleation and Autocatalytic Growth

Aimee M. Morris and Richard G. Finke

Abstract

The aggregation of proteins is believed to be intimately connected to many neurodegenerative disorders. We recently reported an "Ockham's razor"/minimalistic approach to analyze the kinetic data of protein aggregation using the Finke-Watzky (F-W) 2-step model of nucleation ($A \rightarrow B$, rate constant k_1) and autocatalytic growth ($A+B \rightarrow 2B$, rate constant k_2). With that kinetic model we have analyzed 41 representative protein aggregation data sets in two recent publications, including amyloid β, α-synuclein, polyglutamine, and prion proteins (Morris, A. M., et al. (2008) Biochemistry 47, 2413-2427; Watzky, M. A., et al. (2008) Biochemistry 47, 10790-10800). Herein we use the F-W model to reanalyze protein aggregation kinetic data

obtained under the experimental conditions of variable temperature or pH 2.0 to 8.5. We provide the average nucleation (k_1) and growth (k_2) rate constants and correlations with variable temperature or varying pH for the protein α-synuclein. From the variable temperature data, activation parameters $\Delta G^{\ddagger}$, $\Delta H^{\ddagger}$, and $\Delta S^{\ddagger}$ are provided for nucleation and growth, and those values are compared to the available parameters reported in the previous literature determined using an empirical method. Our activation parameters suggest that nucleation and growth are energetically similar for α-synuclein aggregation ($\Delta G^{\ddagger}_{nucleation}$=23(3) kcal/mol; $\Delta G^{\ddagger}_{growth}$=22(1) kcal/mol at 37 °C). From the variable pH data, the F-W analyses show a maximal k_1 value at pH ~3, as well as minimal k_1 near the isoelectric point (pI) of α-synuclein. Since solubility and net charge are minimized at the pI, either or both of these factors may be important in determining the kinetics of the nucleation step. On the other hand, the k_2 values increase with decreasing pH (i.e., do not appear to have a minimum or maximum near the pI) which, when combined with the k_1 vs. pH (and pI) data, suggest that solubility and charge are less important factors for growth, and that charge is important in the k_1, nucleation step of α-synuclein. The chemically well-defined nucleation (k_1) rate constants obtained from the F-W analysis are, as expected, *different than the 1/lag-time empirical constants previously obtained.* However, $k_2 \times [A]_0$ (where k_2 is the rate constant for autocatalytic growth and $[A]_0$ is the initial protein concentration) is related to the empirical constant, k_{app} obtained previously. Overall, the average nucleation and average growth rate constants for α-synuclein aggregation as a function of pH and variable temperature have been quantitated. Those values support the previously suggested formation of a partially folded intermediate that promotes aggregation under high temperature or acidic conditions.

Introduction

The processes of nucleation and growth are of central importance to many different neurological disorders since the aggregation of certain proteins is hypothesized in the literature to be intimately linked to the cause of numerous neurodegenerative diseases [1,2,3,4,5,6]. For instance, the aggregation of α-synuclein has been hypothesized to be the underlying cause of Parkinson's disease [5]. The ability to monitor and detect protein aggregation is, therefore, of importance and may lead to possible therapeutic treatments.

Recently, we reviewed the literature relevant to the mechanism(s) of protein aggregation and found five classes of proposed kinetic mechanisms for the aggregation of proteins [7]. Experimentally, we found that the "Ockham's razor" [8] / minimalistic, Finke-Watzky (F-W) 2-step model, Scheme 1, is able to fit a wide variety of kinetic data and provide quantitative rate constants corresponding to the nucleation and growth of protein aggregates [9,10].

Scheme 1. The minimalistic, "Ockham's razor" Finke-Watzky (F-W) 2-step kinetic model of nucleation followed by autocatalytic growth [11].

$$A \xrightarrow{k_1} B$$
$$A + B \xrightarrow{k_2} 2B$$

The F-W 2-step model was originally developed for transition-metal nanocluster nucleation and growth [11] but has recently been applied to a wider range of natural growth phenomena [12] including, protein aggregation [9,10] kinetics. In protein aggregation, A is the monomeric form of the protein, B is the polymeric form of the

protein aggregate, and k_1 and k_2 correspond to the average rate constants for nucleation and growth, respectively, rigorously defined kinetically by the F-W 2-step kinetic model in Scheme 1.

Recently we reported that the F-W model is able to fit the kinetic data in 14 different data sets of protein aggregation relevant to Alzheimer's, Parkinson's, and Huntington's diseases [9], as well as 27 data sets for yeast and mammalian prion aggregation kinetic data [10]. The F-W fits to these protein aggregation data sets yielded the first quantitative rate constants for both nucleation (k_1) and growth (k_2) in these systems [9,10]. Those individual rate constants in turn yielded insights into factors such as the N-terminal portion and Gln/Asn rich regions that affect nucleation more than growth, while other factors such as the C-terminal portion affects growth and not nucleation in model yeast prion systems relevant to prion diseases [10].

Herein we examine α-synuclein (140 amino acids; ~14,000 MW) aggregation variable temperature and variable pH kinetic data from a classic study from A. L. Fink's laboratories[1] using the F-W model. This protein system was previously analyzed using one method available at the time, namely an empirical approach [13] (see the Results and Discussion section, *vide infra*). Specifically, herein we address the following questions: (i) Is there any correlation between the k_1 and k_2 values for nucleation and growth and the empirical constants determined previously, 1/lag-time and k_{app} [13]? (ii) Do the activation

[1] We were unaware until after we completed this project that Professor A. L. Fink had passed away. We originally picked his laboratory's data set and paper because of the high quality of data and science in that paper; indeed, we had hoped to initiate a collaboration by sending him a copy of this paper and sharing co-authorship with him. We dedicate this paper to Professor Fink's highly productive career and superb science. We wish we'd had the chance to know him personally and collaborate with him and his research group. We thank Professor G. Millhauser as UC Santa Cruz for his help and advice in constructing this dedication to Professor Fink.

parameters obtained from the variable temperature analysis using the F-W model correspond to the activation parameters reported using empirical methods? (iii) What is the correlation between the pH value and the nucleation and growth rate constants obtained from the F-W analysis? (iv) How do any correlations observed and resultant insights compare to those observed from empirical treatments of the variable pH data? (v) Are our kinetic analyses consistent with the prior finding [13] that acidic pH or increased temperature causes the formation of a partially folded intermediate that promotes α-synuclein aggregation? And lastly, (vi) what physical/biochemical insights can we derive from the deconvoluted, well-defined nucleation (k_1) and autocatalytic growth (k_2) rate constants?

Experimental

Selection of Data Sets for Analysis. The data fit in the Results and Discussion section were selected from searches of the literature (using Scifinder Scholar) that displayed variable temperature or variable pH kinetic data for α-synuclein aggregation and also provided the original raw data (i.e., the temperature or pH kinetic curves). The system chosen was Fink and co-workers' scholarly, expert biophysical studies of α-synuclein aggregation [13].

Data Analysis and Curve Fitting. Data were extracted (digitized) from published kinetic curves using Engauge Digitizer 2.12 and fit by the analytical equation shown in eq. (1) corresponding to the F-W model (Scheme 1) using Origin 7.0, all as previously described [9]. The resultant rate constants for each data set were then plotted using Excel 11.3.3 and the activation parameters were calculated using the Eyring equation, eq. (2),

where k is the rate constant, T is temperature, R is the universal gas constant, k_b is Boltzmann's constant, and h is Planck's constant.

$$[B]_t = [A]_0 - \frac{\frac{k_1}{k_2} + [A]_0}{1 + \frac{k_1}{k_2[A]_0}\exp(k_1 + k_2[A]_0)t} \quad (1)$$

$$\ln\frac{k}{T} = \frac{-\Delta H^{\ddagger}}{RT} + \ln\frac{k_b}{h} + \frac{\Delta S^{\ddagger}}{R} \quad (2)$$

Results and Discussion

Previous α-Synuclein Aggregation Kinetic Data and Analysis by an Empirical Method. Previously α-synuclein aggregation was induced by increasing temperature or decreasing pH and monitoring the aggregation kinetics using ThT fluorescence [13]. The goal of that study [13] was to test the hypothesis that higher temperature or lower pH would induce the natively unfolded α-synuclein into a partially folded, more aggregation prone, intermediate state. After following the kinetics of α-synuclein aggregation and performing other biophysical studies [13] the authors noted, "such [kinetic] curves are consistent with a nucleation-dependent polymerization model" [13].

The variable temperature and pH kinetic data was previously published and analyzed via the empirical equation shown in eq. (3) where Y is the fluorescence intensity, $(y_i + m_i x)$ is the initial slope of the line during the lag phase, $(y_f + m_f x)$ is the final slope of the line after the growth has ended, and χ_0 is the time to reach 50% maximum fluorescence intensity [13]. From eq. (3) the empirical constants[2] of $k_{app}=1/\tau$

[2] The prior work refers to k_{app} as an apparent first order rate constant. However, since k_{app} is not defined in the usual way by a well-defined and balanced chemical reaction, it is

and 1/lag-time=1/(χ_0-2τ) are calculated [13]. As these equations show, the lag time and k_{app} are both dependent upon 1/τ in the empirical model. While eq. (3) is able to fit the published kinetic data [13], the authors were careful to note, "This expression is unrelated to the underlying molecular events, but provides a convenient method for comparison of the kinetics of fibrillation" [13].

$$Y = (y_i + m_i x) + \frac{(v_f + m_f x)}{1 + e^{-\left(\frac{\chi - \chi_0}{\tau}\right)}} \tag{3}$$

It is of interest, therefore, to analyze this α-synuclein aggregation data by the F-W 2-step nucleation and autocatalytic growth model to (i) see if the F-W model will fit the data, and if so (ii) to obtain discrete, chemically well-defined nucleation (k_1) and autocatalytic growth (k_2) rate constants, as well as (iii) to see if our quantitative kinetic results support or refute the original hypothesis [13] of increased temperature or decreased pH causing the formation of a partially folded intermediate.

Analysis of Variable Temperature α-Synuclein Aggregation Kinetic Data with the F-W 2-Step Model. Variable temperature data of α-synuclein aggregation [13] were analyzed using the F-W 2-step model. The F-W analysis yields the rate constants for

---

not a true rate constant in a rigorous sense. Hence, we have referred to both 1/lag-time and k_{app} as empirical constants throughout the text. This point is not trivial. In fact, a review of the protein aggregation kinetic and modeling literature [7] shows that a failure to connect words, concepts, and their associated rate constants to *balanced chemical reactions* (as is the protocol of rigorous chemical kinetics) has contributed to considerable confusion in the protein aggregation literature [7]. Obviously, Professor Fink and co-workers were aware of this point as the quotation in the text of the original paper indicates ("This expression is unrelated to the underlying molecular events, but provides a convenient method for comparison of the kinetics of fibrillation") [13].

nucleation (k_1) and growth (k_2), along with the coefficient of determination (R^2) for each fit.

Sample F-W fit and rate constants obtained. An example of the graphic fit obtained from the F-W analysis of the variable temperature α-synuclein aggregation kinetic data is given in Figure 1 (all 4 graphical fits of this data are available in Figure S1 of the Supporting Information). Shown in Table 1 are the nucleation (k_1) and growth (k_2) rate constants obtained from the F-W analysis. Also given in Table 1 are $k_2[A]_0$ values, where $[A]_0$ is the initial concentration of α-synuclein, along with the coefficient of determination (R^2) values for the fits at each temperature. The $R^2 \geq 0.989$ as well as the visually good fits (Figure 1 and Figure S1 of the Supporting Information) indicate that the F-W model is able to fit the kinetic data well at each temperature. For comparative purposes, the empirical constants[1] previously obtained via eq. (3) are also provided in Table 1.

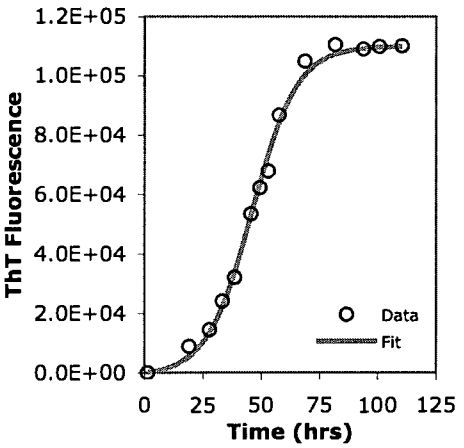


Figure 1. Literature α-synuclein aggregation at 37 °C kinetic data [13] and the fit by the F-W 2-step model. The resultant rate constants are $k_1=8(2) \times 10^{-4}$ hr^{-1} and $k_2=1.5(1) \times 10^{-3}$ μM^{-1} hr^{-1} with an $R^2=0.9957$.

Table 1. The average rate constants and R^2 values obtained from fitting the variable temperature α-synuclein aggregation data [13] with the F-W 2-step model along with comparisons to the empirical constants reported [13] using eq. (3).

		27° C	37° C	47° C	57° C
For the F-W analysis:	k_1[a] (hr^{-1})	$2.4(8) \times 10^{-4}$	$8(2) \times 10^{-4}$	$2.3(7) \times 10^{-3}$	$6(2) \times 10^{-3}$
	k_2[a] (μM^{-1} hr^{-1})	$6.4(4) \times 10^{-4}$	$1.5(1) \times 10^{-3}$	$3.6(3) \times 10^{-3}$	$9.0(7) \times 10^{-3}$
	$k_2[A]_0$ (hr^{-1})	$4.5(3) \times 10^{-2}$	$1.05(7) \times 10^{-1}$	$2.5(2) \times 10^{-1}$	$6.3(5) \times 10^{-1}$
	R^2[b]	0.9891	0.9957	0.9921	0.9914
From the previous analysis [13]:	1/lag-time[c,d] (hr^{-1})	4.1×10^{-2}	9.2×10^{-2}	2.5×10^{-1}	6.1×10^{-1}
	k_{app}[c,d] (hr^{-1})	1.4×10^{-2}	3.4×10^{-2}	1.1×10^{-1}	2.8×10^{-1}

[a] The error bars for k_1 and k_2 are determined by the square root of the reduced χ^2 based on a modified Levenberg-Marquardt algorithm [25].

[b] R^2 is the coefficient of determination and is the square of the correlation coefficient (r). The closer this value is to 1.0, the more precise the analysis is to fitting the data.

[c] The 1/lag-time and k_{app} values were found by digitizing a ln(1/lag-time) and ln(k_{app}) vs. 1/T plot from the original reference [13].

[d] While digitizing the data it became apparent that in order to correspond with the raw variable temperature kinetic data the ln(1/lag-time) and k_{app} vs. 1/T graphs had to be in units of reciprocal seconds. The 1/lag-time and k_{app} were converted to units of hr^{-1} for comparison purposes to the analyses presented herein.

As expected, the k_1 and k_2 *rate constants* obtained from the F-W model are different than the *empirical constants*[1] obtained from eq. (3), Table 1. However, examination of Table 1 reveals that k_{app} from eq. (3) appears to be proportional to $k_2[A]_0$ obtained by from the F-W model (*vide infra*). From the k_1 and k_2 rate constants, we observe that both nucleation and growth are faster at higher temperatures, as of course expected and as seen with the previous empirical analysis. The increases in k_1 and k_2 are consistent with the previous idea of an increased concentration and/or aggregation rate of a partially folded intermediate at higher temperatures [13].

Activation parameter determination. From the Eyring plots shown in Figs. 2a and 2b, the activation parameters $\Delta G^{\ddagger}$, $\Delta H^{\ddagger}$, and $\Delta S^{\ddagger}$ were obtained for both the nucleation (k_1) and growth (k_2) steps of the α-synuclein aggregation system, Table 2. The activation parameters obtained for the k_1 nucleation and k_2 growth steps are quite similar, a finding that has also been observed in actin aggregation [14]. That said, the importance of

$\Delta H^{\ddagger}_{nucleation}$, $\Delta H^{\ddagger}_{growth}$, $\Delta S^{\ddagger}_{nucleation}$, and $\Delta S^{\ddagger}_{growth}$ obtained from the F-W model, and shown in Table 2, is that they refer to the well-defined chemical steps of nucleation and growth, respectively. Application of the F-W model also gives previously unavailable $\Delta S^{\ddagger}$ values.

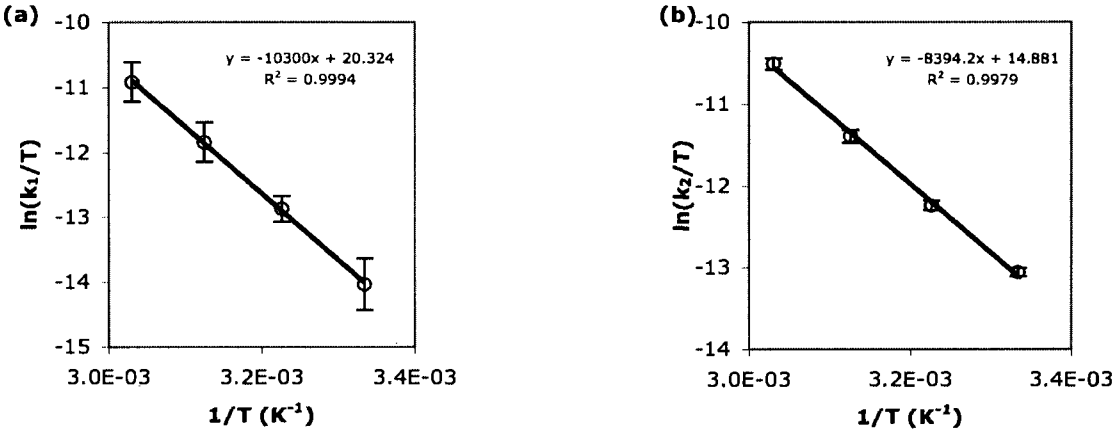


Figure 2. Eyring plots of (a) k_1 and (b) k_2 for the α-synuclein aggregation system [13]. The error bars were propagated assuming a possibly too large (i.e., upper limit) error in the temperature measurement of ± 0.5 °C and since no temperature measurement error bars were given in the original work.

The activation parameters for α-synuclein aggregation are of interest and merit comment. The values of $\Delta H^{\ddagger}$ for nucleation and growth are both positive for α-synuclein ($\Delta H^{\ddagger}_{nucleation}$=20(3) kcal/mol; $\Delta H^{\ddagger}_{growth}$=17(1) kcal/mol), while the $\Delta S^{\ddagger}$ for nucleation and growth are compensatingly more negative (i.e., compensationally less favorable) ($\Delta S^{\ddagger}_{nucleation}$=-10(10) e.u.; $\Delta S^{\ddagger}_{growth}$=-18(3) e.u.). The enthalpy vs. entropy compensation might be real or might be artifactual [15] due to the moderate temperature range (30 °C) examined; large temperature ranges are known to be required to accurately deconvolute $\Delta G^{\ddagger}$ into $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ [16].

Table 2. Activation parameters calculated for α-synuclein aggregation from the Eyring plots shown in Figure 2 using the F-W analysis and comparison to the empirical parameters previously obtained in the literature.

		$\Delta G^{\ddagger}$ (kcal/mol at 37 °C)	$\Delta H^{\ddagger}$ (kcal/mol)	$\Delta S^{\ddagger}$ (cal/mol•K)
From the F-W analysis:	Nucleation, k_1[a]	23(3)	20(3)	-10(10)
	Growth, k_2[a]	22(1)	17(1)	-18(3)
			$\Delta H_{empl}^{\ddagger}$[b] (kcal/mol)	
From the literature empirical analysis (13)[c,d]:	1/lag-time		17.3(8)[e,f]	NA[g]
	k_{app}		19.5(8)[e,f]	NA[g]

[a] The error bars are calculated by propagating the error through the Eyring equation using the following equations from Girolami and co-workers [26]:

$$(\sigma\Delta H^{\ddagger})^2 = \frac{R^2 T_{max}^{\ 2} T_{min}^{\ 2}}{(T_{max}-T_{min})^2}\left\{\left(\frac{\sigma T}{T}\right)^2\left[\left(1+T_{min}\frac{\Delta L}{(T_{max}-T_{min})}\right)^2+\left(1+T_{max}\frac{\Delta L}{(T_{max}-T_{min})}\right)^2\right]+2\left(\frac{\sigma k}{k}\right)^2\right\} \text{ and}$$

$$(\sigma\Delta S^{\ddagger})^2 = \frac{R^2}{(T_{max}-T_{min})^2}\left\{\left(\frac{\sigma T}{T}\right)^2\left[T_{max}^{\ 2}\left(1+T_{min}\frac{\Delta L}{T_{max}-T_{min}}\right)^2+T_{min}^{\ 2}\left(1+T_{max}\frac{\Delta L}{T_{max}-T_{min}}\right)^2\right]+\left(\frac{\sigma k}{k}\right)^2\left(T_{max}^{\ 2}+T_{min}^{\ 2}\right)\right\}, \text{ where}$$

$\Delta L = [\ln(k_{max}/T_{max}) - \ln(k_{min}/T_{min})]$. For σk, we used the largest error bars in the k_1 or k_2 that was observed and assumed an error of 0.5 K for σT. Alternatively, we calculated error bars based on the errors at the 95% (2σ) confidence level for the slope and intercept obtained from the linear regression analysis. However, those resultant error bars are much smaller (by ca. 2-10 fold in $\Delta H^{\ddagger}$ and 3-30 fold in $\Delta S^{\ddagger}$). Therefore, we have reported the larger error bars herein.

[b] We denote the enthalpy of activation from the previous analysis as $\Delta H_{empl}^{\ddagger}$ since it is derived from empirical constants[l] and is therefore, not a true $\Delta H^{\ddagger}$ value in the unambiguous chemical mechanism sense.

[c] In the literature, eq. **3** is used to calculate the empirical constants[l] of $k_{app}=1/\tau$ and 1/lag-time$=1/(\chi_0-2\tau)$ [13].

[d] α-Synuclein aggregation was measured from 27 °C to 57 °C; a 30 °C temperature range.

[e] These values were converted from the reported E_a by the following equation: $E_a = \Delta H^{\ddagger} + mRT_{mean}$, where m is the molecularity (taken as 1 in the present case), and T_{mean} is the mean temperature range over which measurements were taken [27].

[f] The error bars reported come from the original analysis [13] and are calculated from the error in the slope of the linear regression fits of the Arrhenius plots.

[g] NA=not available.

Examination of Table 2 shows that the activation parameter determination using the F-W model yields values that are the same within experimental error as eq. (3) for $\Delta H^{\ddagger}$. Both the previous [13] and present results are consistent with the idea that $\Delta H^{\ddagger}$ for nucleation and growth are similar.

Analysis of Variable pH α-Synuclein Aggregation Kinetic Data with the F-W 2-Step Model. The second aggregation variable examined herein is pH. Interestingly, it

has been hypothesized in the literature that a pathological decrease in pH occurs in neurons affected by neurodegenerative diseases, resulting in apoptosis [17]. It has also been suggested that lowering the pH causes a reduction in the overall net charge of proteins and thereby increases the aggregation propensity [13]. While selectively maintaining in vivo affected cells at a slightly alkaline pH has been suggested as a possible therapeutic strategy for the treatment of neurodegenerative disorders [17], it is probably not a viable treatment option. However, studying variable pH in vitro may give insights into in vivo factors that could increase the aggregation propensity and intermediate formation of α-synuclein.

Hence, we have reanalyzed the α-synuclein variable pH kinetic data that was previously analyzed using eq. (3) to give empirical constants[1] 1/lag-time and k_{app} [13]. Figure 3 shows a representative fit and Figure S2 of the Supporting Information shows the graphical fits to each of the five pH values examined using the F-W 2-step model. Table 3 shows the *rate* constants obtained from the F-W analyses along with the associated R^2 values. The rate constants obtained by the analytic equation for the 2-step model (eq. (1)) are one to two orders of magnitude different than the empirical constants obtained with eq. (3), Table 3.

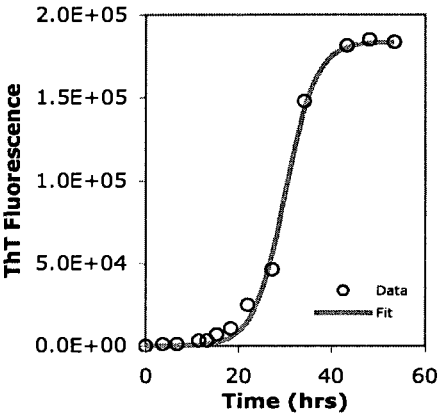


Figure 3. Literature α-synuclein aggregation kinetic data at pH 5.82 [13] and the fit using the F-W model. Resultant rate constants (and coefficient of determination): $k_1=4(2) \times 10^{-5}$ hr^{-1}; $k_2=4.2(3) \times 10^{-3}$ μM^{-1} hr^{-1} ($R^2=0.9965$).

Table 3. The resulting rate constants and R^2 values from the published [13] α-synuclein variable pH data fit by the F-W 2-step model and comparison to the empirical constants reported previously.

	pH 1.92	pH 2.79	pH 4.08	pH 5.82	pH 7.23/8.52[a]
k_1 (hr^{-1})	$3(2) \times 10^{-3}$	$1.2(4) \times 10^{-2}$	$6(2) \times 10^{-5}$	$4(2) \times 10^{-5}$	$3.5(6) \times 10^{-4}$
k_2 ($\mu M^{-1} hr^{-1}$)	$1.3(3) \times 10^{-2}$	$1.0(1) \times 10^{-2}$	$1.17(5) \times 10^{-2}$	$4.2(3) \times 10^{-3}$	$1.7(1) \times 10^{-3}$
$k_2[A]_0$ (hr^{-1})	$9(2) \times 10^{-1}$	$7.0(7) \times 10^{-1}$	$8.2(3) \times 10^{-1}$	$2.9(2) \times 10^{-1}$	$1.19(7) \times 10^{-1}$
R^2	0.9898	0.9902	0.9993	0.9965	0.9984
1/lag-time (hr^{-1})[b]	2.5×10^{-1}	4.4×10^{-1}	1.1×10^{-1}	4.2×10^{-2}	$3.5 \times 10^{-2}/3.3 \times 10^{-2}$
k_{app} (hr^{-1})[b]	8.2×10^{-1}	1.0	7.0×10^{-1}	3.4×10^{-1}	$1.1 \times 10^{-1}/8.6 \times 10^{-2}$

[a] While digitizing this data set we were unable to discern between the data points for pH 7.23 and pH 8.52
[b] These values were obtained by digitizing the pH vs. 1/lag-time and pH vs. k_{app} graphs from the original reference [13].

Correlations between pH values and nucleation and growth rate constants obtained from the F-W analysis. The correlations obtained using the F-W model for the varying pH data are shown in Figure 4. Figure 4a reveals that the k_1 rate constant is increased at pH ~2 and increased by another order of magnitude at pH ~3. From pH ~3 to ~4 the nucleation rate constant is decreased by 3 orders of magnitude. The *slowest* nucleation rate constants occur from pH 4-6 which coincide well with the isoelectric point (pI) of 4.0-4.7 given for α-synuclein [13,18,19]. The k_1 rate constant also increases

by an order of magnitude but now from pH ~6 to pH ~7. The data suggest that there is an optimum pH of ca. 3 at which the nucleation rate is greatly enhanced. The result of an optimum pH ~3 is consistent with the previous structural analyses by Fourier transform infrared spectroscopy and small angle X-ray scattering, suggesting that the partially folded intermediate forms between pH 7.5 and 3.0. This optimum pH ~3 is also consistent with the previous CD and ANS fluorescence results which suggest partially folded intermediate formation between pH 5.5 and 3.0 [13]. Assuming that the ThT fluorescence is providing an accurate measure of α-synuclein aggregation, our kinetic analysis suggests that the pH in which the partially folded intermediate is formed and nucleation is most increased is pH ~3. The minimum nucleation rate constants near the pI also suggest that solubility or net charge (or both) may be important factors in determining the nucleation rate constant for α-synuclein aggregation [20].

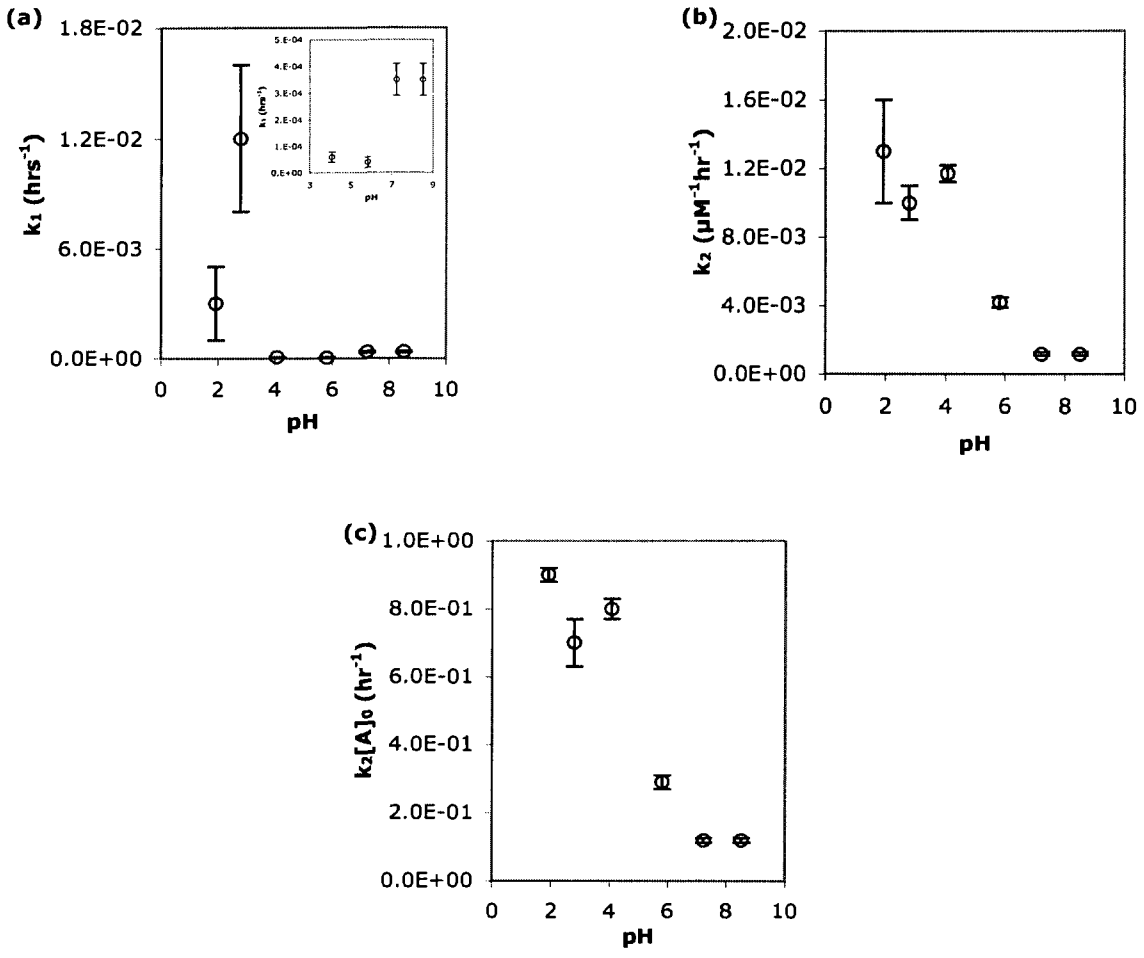


Figure 4. The (a) k_1, (b) k_2 and (c) $k_2[A]_0$ correlations determined using the F-W 2-step model to varying pH values.

The previous kinetic analysis performed via eq. (3) yielded a correlation of shorter lag-times with decreasing pH values but also revealed the shortest lag-time occurred at pH ~3 [13]. However, over the range of pH measured, the lag-time only changed by one order of magnitude. Our results herein, as well as studies in progress [21], *suggests that the lag-time in sigmoidal protein aggregation curves is not a generally reliable measure of the nucleation rate constant.*

Figure 4b shows the correlation of increasing k_2 values (i.e., faster growth) with decreasing pH. As expected since the concentration of the protein remains constant, the same correlation holds true for increasing $k_2[A]_0$ with decreasing pH, Figure 4c. This is also the same correlation previously observed for k_{app} [13]. Hence, it appears that the solubility and charge minima at the pI are not major factors in determining the *growth* rate constants of α-synuclein. More importantly, k_2 should have decreased, not increased, near the pI as did k_1 if decreased solubility of α-synuclein near the pI was the controlling variable. Hence, the combined pH/pI effects on k_2 and k_1 suggest that *charge* is the important variable in determining the α-synuclein nucleation rate constant, k_1.

In short, from the F-W reanalysis of the variable pH data, we confirm the previous findings [13] that: (i) the nucleation rate constant (k_1) is larger at lower pH values and (ii) k_{app}, $k_2[A]_0$, and k_2 increases with decreasing pH values. This supports the previous claim that the nucleation and growth rate constants are increased under acidic conditions [13]. We have, however, added the insights that: (i) the k_1 rate constant seems to be maximal at pH 3 and minimal near the pI, (ii) k_2 increases with decreasing pH and does not correlate with the pI, so that (iii) *net charge* is implicated as an important factor in determining the k_1 value. In addition, (iv) *the lag-time (i.e., 1/lag-time) is not a generally reliable measure of k_1.*

Analysis of Any Relationships of the Rate Constants Obtained from the F-W Reanalysis to the Empirical Constants from Eq. 3. Figure 5, as well as examination of Tables 1 and 3, show that the $k_2[A]_0$ values from the F-W model are proportional to the k_{app} values obtained from eq. (3), (a plot of k_2 vs. k_{app} is also available in the Supporting

Information). The $k_2[A]_0 \propto k_{app}$ reflects the fact that both k_{app} and k_2 are proportional to the slope of the growth line [11,13]. In addition, $k_2[A]_0 \propto k_{app}$ is a useful finding that helps add validity to the interpretation from the previously obtained k_{app} [13].

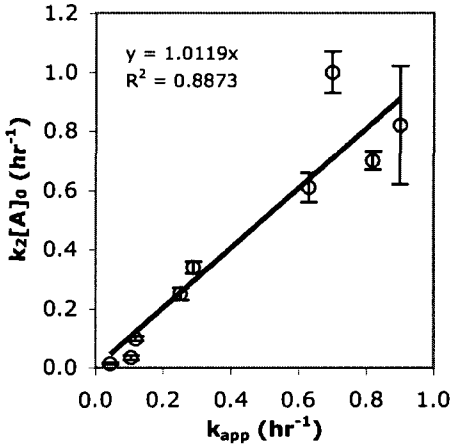


Figure 5. The k_{app} values obtained from eq. (3) for both the variable temperature and variable pH data compared to the $k_2[A]_0$ values obtained from the same data but using the F-W 2-step model. The k_{app} and $k_2[A]_0$ values can also be found in Tables 1 and 3.

We also attempted to look for any correlations between the empirical constant 1/lag-time and k_1, k_2 (or some combination of k_1 and k_2) from the F-W model. We did not find any apparent correlation (the interested reader is referred to Figs. S3-S4 in the Supporting Information to see the plots of k_1, $k_1 + k_2$, $k_1 \times k_2$, k_1/k_2, $k_1 + k_2[A]_0$, $k_1 \times k_2[A]_0$, and $k_1/k_2[A]_0$ vs. 1/lag-time). This again suggests that lag-time is not an accurate measure of the nucleation rate constant. We were also unable to find any simple mathematical equivalence correlation between eq. (1) from the F-W model and the empirical eq. (3) (Scheme S1 of the Supporting Information).

Caveats and Limitations of the F-W Kinetic Model. The weaknesses of the F-W model, which derive ultimately from its over-simplified nature, have been presented in three other publications for the interested reader [7,9,10]. Briefly, the main limitations are: (i) k_1 and k_2 are averages (although to the extent nucleation tends towards the limit of a simple misfolded protein [21], k_1 should tend towards the limit of the rate constant for that individual step); (ii) B is also an average, so that (iii) important changes in k_1, k_2, or B as a function of aggregation fibril size are hidden. In addition, (iv) other processes such as fibril fragmentation are not treated explicitly [10]. The main advantage of the F-W 2-step model is, however, its ability to deconvolute the average nucleation rate constant, k_1, from the average autocatalytic growth rate constant, k_2, in a model that has rate constants and associated concepts (words; specifically nucleation and autocatalytic growth) rigorously defined by balanced chemical reactions in the normal way of rigorous kinetic and mechanistic science.

Conclusions

The contributions from this article are the following:

- The F-W model 2-step has been used to analyze α-synuclein aggregation kinetics under variable temperature and variable pH conditions. Quantitative rate constants with well-defined physical meaning corresponding to both nucleation and growth, k_1 and k_2 respectively, were obtained.
- The k_1 rate constant found using the F-W model differs from the empirical 1/lag-time constant obtained previously. Other, work in progress confirms that 1/lag-time is not a reliable predictor of k_1 [22]. However, $k_2[A]_0$ from the F-W model is

proportional to k_{app} from the eq. (3) empirical treatment, thereby adding validity to the prior determination and interpretation of k_{app}.

- The activation parameter determination for α-synuclein performed herein is the first to our knowledge to report values for $\Delta G^{\ddagger}$, $\Delta H^{\ddagger}$, *and* $\Delta S^{\ddagger}$ for well-defined nucleation and autocatalytic growth.
- The $\Delta H^{\ddagger}$, $\Delta S^{\ddagger}$, and $\Delta G^{\ddagger}$ of nucleation and growth were found to be approximately equivalent for α-synuclein under the experimental conditions examined.
- The α-synuclein variable pH data analysis shows that the nucleation rate constant, k_1, has a maximal rate constant at pH ~3, as well as minimal rate constant at pH ~4-6. These results are consistent with the previous analysis [13] and further suggests that there is an optimal pH for the nucleation rate constant.
- The growth rate constant, k_2, (as well as $k_2[A]_0$) appears to increase with decreasing pH—that is, the average growth rate constant is faster at lower pH values in the case examined herein. This observation is again consistent with the previous analysis [13]. Furthermore, the k_2 rate constants do not correlate with the pI.
- The combined k_1 and k_2 vs. pH data near the pI of α-synuclein in turn suggest that *net charge* is an important variable in α-synuclein nucleation.
- Overall, our reanalysis quantitates the average nucleation and average growth rate constants and kinetically supports Fink and co-workers' hypothesis [13] of the formation of a partially folded intermediate species that promotes aggregation at higher temperatures or lower pH.

- Other variables that have been hypothesized to either accelerate (e.g., O_2, H_2O_2, or $O_2^{\cdot-}$ [23]), or inhibit (e.g., β- or γ-synuclein [24]), α-synuclein aggregation in vivo are of interest for deconvolution into k_1 and k_2 effects by the methods detailed herein.

Finally, it is especially important to acknowledge the original work and contributions of the Professor A. L. Fink and his co-workers [13]. It is their original efforts and α-synuclein kinetic data that have allowed the work and analysis herein to be reported. We are pleased to dedicate this work to Professor Fink's memory.[1]

Acknowledgements

Partial support from the long term NSF grant # 0611588 is gratefully acknowledged.

References

[1] M.B. Pepsys, Amyloidosis, Annu. Rev. Med. 57 (2006) 223-241.

[2] F. Chiti, and C.M. Dobson, Protein misfolding, functional amyloid, and human disease, Annu. Rev. Biochem. 75 (2006) 333-366.

[3] K. Blennow, M.J. de Leon, and H. Zetterberg, Alzheimer's disease, Lancet 368, (2006) 387-403.

[4] G.P. Bates, and C. Benn, in: Huntington's Disease, ed. G.P. Bates, P.S. Harper, and L. Jones, The polyglutamine diseases (Oxford University Press, Oxford, 2002) pp. 429-472.

[5] W. Dauer, and S. Przedborski, Parkinson's disease: mechanisms and models, Neuron 39 (2003) 889-909.

[6] S.B. Prusiner, Molecular biology of prion diseases, Science 252 (1991) 1515-1522.

[7] A.M. Morris, M.A. Watzky, and R.G. Finke, Protein aggregation kinetics, mechanism and curve-fitting: a review of the literature, Biochim. Biophys. Acta 1794 (2009) 375-397.

[8] William of Ockham, 1285-1349, as cited by E.A. Moody in: The Encyclopedia of Philosophy, vol 7 (MacMillan, New York, 1967).

[9] A.M. Morris, M.A. Watzky, J.N. Agar, and R.G. Finke, Fitting neurological protein aggregation kinetic data via a 2-step, minimal/"Ockham's razor" model: the Finke-Watzky mechanism of nucleation followed by autocatalytic surface growth, Biochemistry 47 (2008) 2413-2427.

[10] M.A. Watzky, A.M. Morris, E.D. Ross, and R.G. Finke, Fitting yeast and mammalian prion aggregation kinetic data with the Finke-Watzky two-step model of nucleation and autocatalytic growth, Biochemistry 47 (2008) 10790-10800.

[11] M.A. Watzky, and R.G. Finke, Transition metal nanocluster formation kinetic and mechanistic studies. A new mechanism when hydrogen is the reductant: slow, continuous nucleation and fast autocatalytic surface growth, J. Am. Chem. Soc. 119 (1997) 10382-10400.

[12] S.E. Smith, J.M. Sasaki, R.G. Bergman, J.E. Mondloch, and R.G. Finke, Platinum-catalyzed phenyl and methyl group transfer from tin to iridium: evidence for an autocatalytic reaction pathway with an unusual preference for methyl transfer, J. Am. Chem. Soc. 130 (2008) 1839-1841.

[13] V.N. Uversky, J. Li, and A.L. Fink, Evidence for a partially folded intermediate in α-synuclein fibril formation, J. Biol. Chem. 276 (2001) 10737-10744.

[14] F. Oosawa, and S. Asakura, Thermodynamics of the Polymerization of Protein (Academic Press, New York, 1975).

[15] S.W. Benson, Thermochemical Kinetics, 2nd Ed. (Wiley, New York, 1976) pp. 21-23.

[16] J. Halpern, Compensation effects in the activation parameters for the homolytic dissociation of transition metal-alkyl bonds, Bull. Chem. Soc. Jpn. 61 (1988) 13-15.

[17] S. Harguindey, S.J. Reshkin, G. Orive, J.L. Arranz, and E. Anitua, Growth and trophic factors, pH and the Na+/H+ exchanger in Alzheimer's disease, other neurodegenerative diseases and cancer: new therapeutic possibilities and potential dangers, Current Alzheimer Research 4 (2007) 53-65.

[18] V.N. Uversky, J.R. Gillespie, and A.L. Fink, Why are "natively unfolded" proteins unstructured under physiologic conditions?, Proteins Struct. Funct. Genet. 41 (2000) 415-427.

[19] R. Sharon, I. Bar-Joseph, M.P. Frosch, D.M. Walsh, J.A. Hamilton, and D.J. Selkoe, The formation of highly soluble oligomers of α-synuclein is regulated by fatty acids and enhanced in Parkinson's disease, Neuron 37 (2003) 583-595.

[20] D. Voet, and J.G. Voet, Biochemistry, 3rd Ed. (Wiley, New York, 2004) pp. 131-133.

[21] S. Chen, F.A. Ferrone, and R. Wetzel, Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation, Proc. Natl. Acad. Sci. USA 99 (2002) 11884-11889.

[22] A.M. Morris, E.D. Ross, and R.G. Finke, studies in progress.

[23] M.E. Götz, K. Double, M. Gerlach, M.B.H. Youdim, and P. Riederer, The relevance of iron in the pathogenesis of Parkinson's disease, Ann. N. Y. Acad. Sci. 1012 (2004) 193-208.

[24] A.L. Fink, The aggregation and fibrillation of α-synuclein, Acc. Chem. Res. 39 (2006) 628-634.

[25] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, Numerical Recipes, (Cambridge University, Cambridge, 1989).

[26] P.M. Morse, M.D. Spencer, S.R. Wilson, and G.S. Girolami, A static agostic α-CH•••M interaction observable by NMR spectroscopy: synthesis of the chromium(II) alkyl $[Cr_2(CH_2SiMe_3)_6]^{2-}$ and its conversion to the unusual "windowpane" bis(metallacycle) complex $[Cr(\kappa 2\text{-}C,C'\text{-}CH_2SiMe_2CH_2)_2]^{2-}$, Organometallics 13 (1994) 1646-1655.

[27] D.M. Golden, Standard states for thermochemical and activation parameters, J. Chem. Ed. 48 (1971) 235-237.

Supporting Information for:

α-Synuclein Aggregation Variable Temperature and Variable pH Kinetic Data: A Re-analysis using the Finke-Watzky 2-Step Model of Nucleation and Autocatalytic Growth

Aimee M. Morris and Richard G. Finke

Kinetic Data from the Literature and Fits to the Finke-Watzky (F-W) 2-Step Model of Nucleation and Autocatalytic Growth. Displayed in Figures S1-S2 are the digitized literature data sets for the aggregation of α-synuclein (1) under variable temperature or pH conditions. Each data set was fit to the Finke-Watzky (F-W) 2-step model, as detailed in the *Experimental* section of the main text. For the resultant rate constants (k_1 and k_2) and coefficients of determination (R^2) refer to Tables 1 and 3 of the main text.

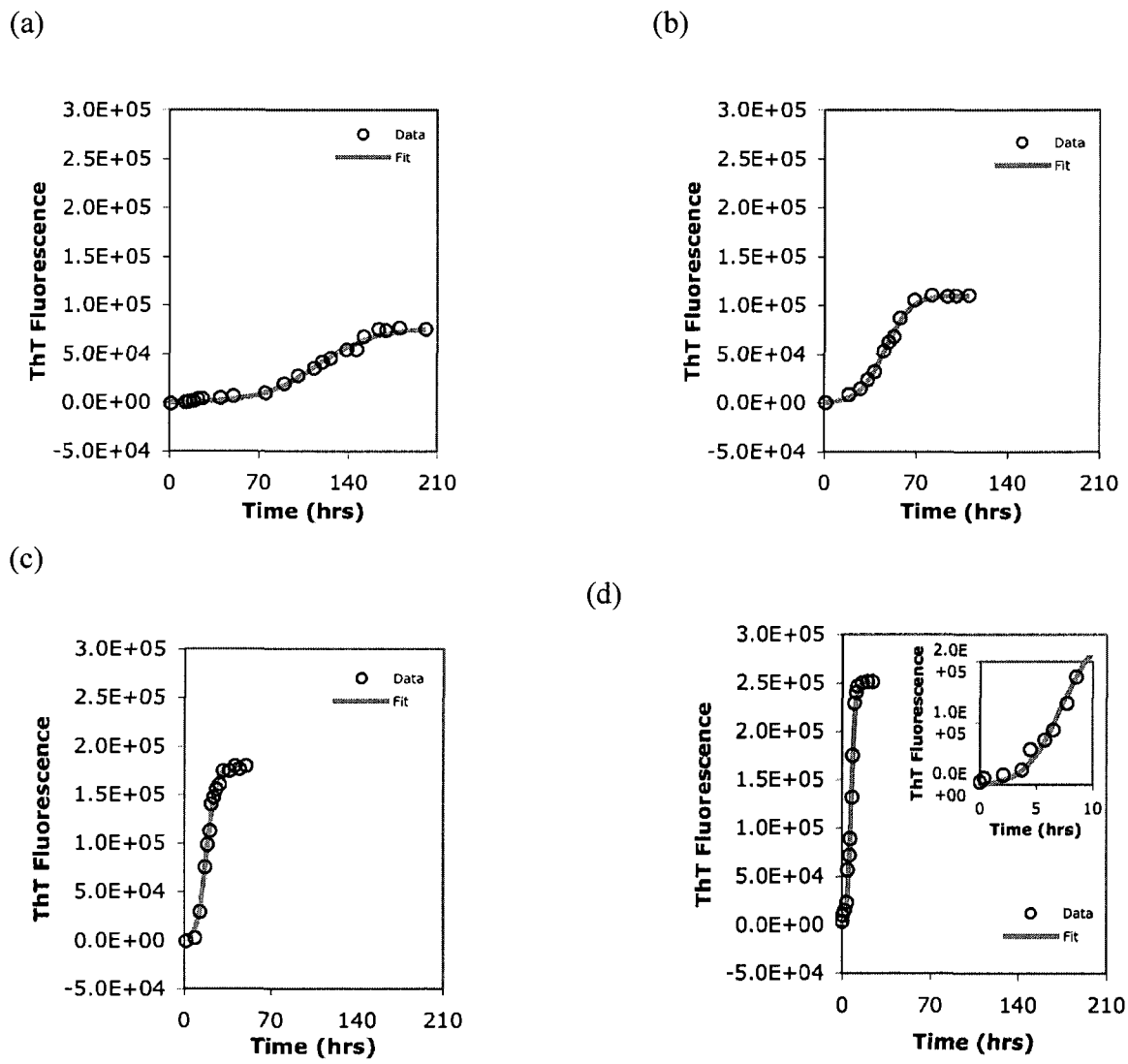


Figure S1. Literature α-synuclein aggregation variable temperature data (1) at (a) 27, (b) 37, (c) 47, and (d) 57 °C. Fits are to the F-W 2-step model cited and referenced in the main text.

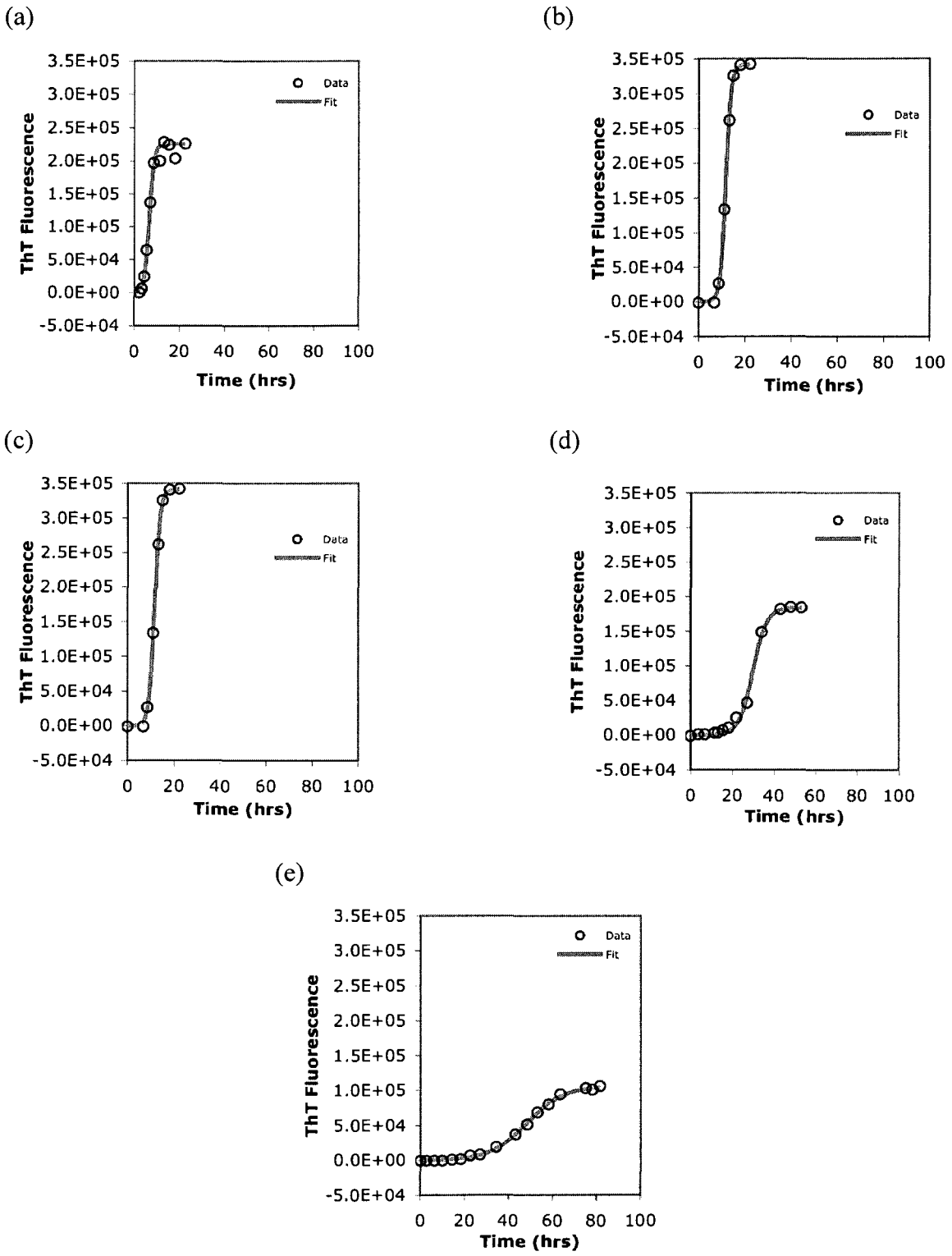


Figure S2. Literature α-synuclein aggregation variable pH data (1) at pH values of (a) 1.92, (b) 2.79, (c) 4.08, (d) 5.82, and (e) 7.23/8.52. The fits are to the F-W 2-step model cited and referenced in the main text.

Looking for correlations between the empirical constants 1/lag-time and k_{app} (1) and k_1, k_2, or some combination of k_1 and k_2, of the F-W 2-step model. In the main text it is noted, and illustrated in Fig. 4, that k_{app} from the empirical analysis elsewhere (1) is correlated to $k_2[A]_0$ from the F-W model. We then wondered if there was any correlation between k_1, k_2, or some combination of k_1 and k_2 of the F-W model with the empirical 1/lag-time. Shown below in Figure S4 are the attempts to find a correlation with 1/lag-time. No simple correlation that at least we could find is apparent.

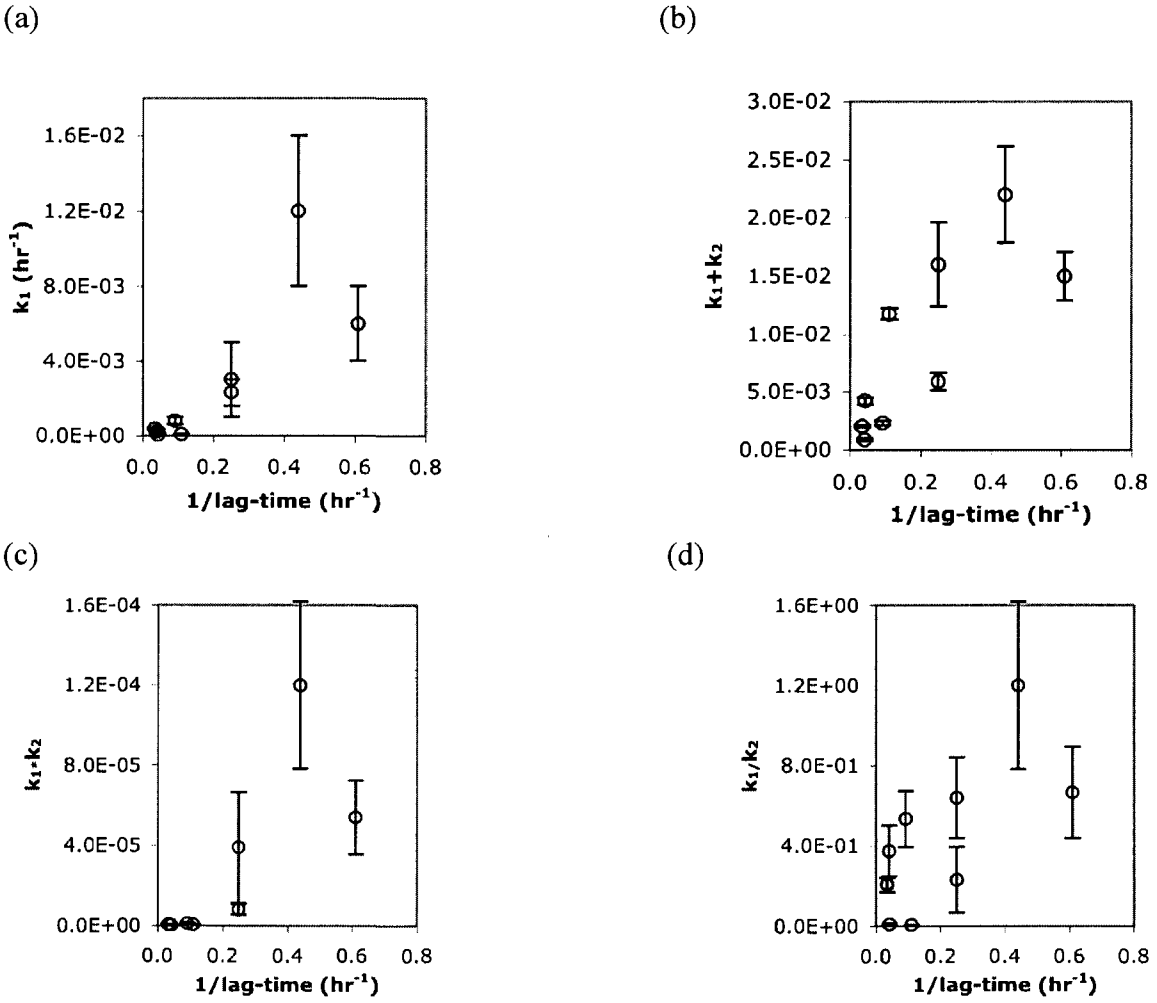


Figure S3. Attempts at finding a correlation with the literature (1) 1/lag-time and: (a) k_1, (b) k_1+k_2, (c) k_1*k_2, or (d) k_1/k_2. The data employed comes from Tables 1 and 5 of the main text.

We also investigated whether the 1/lag-time empirical constant would be correlated with some combination of k_1 and $k_2[A]_0$ as shown in Figure S4. Again no correlation is apparent between the empirical constant 1/lag-time and any of the combinations of k_1 and $k_2[A]_0$ tried.

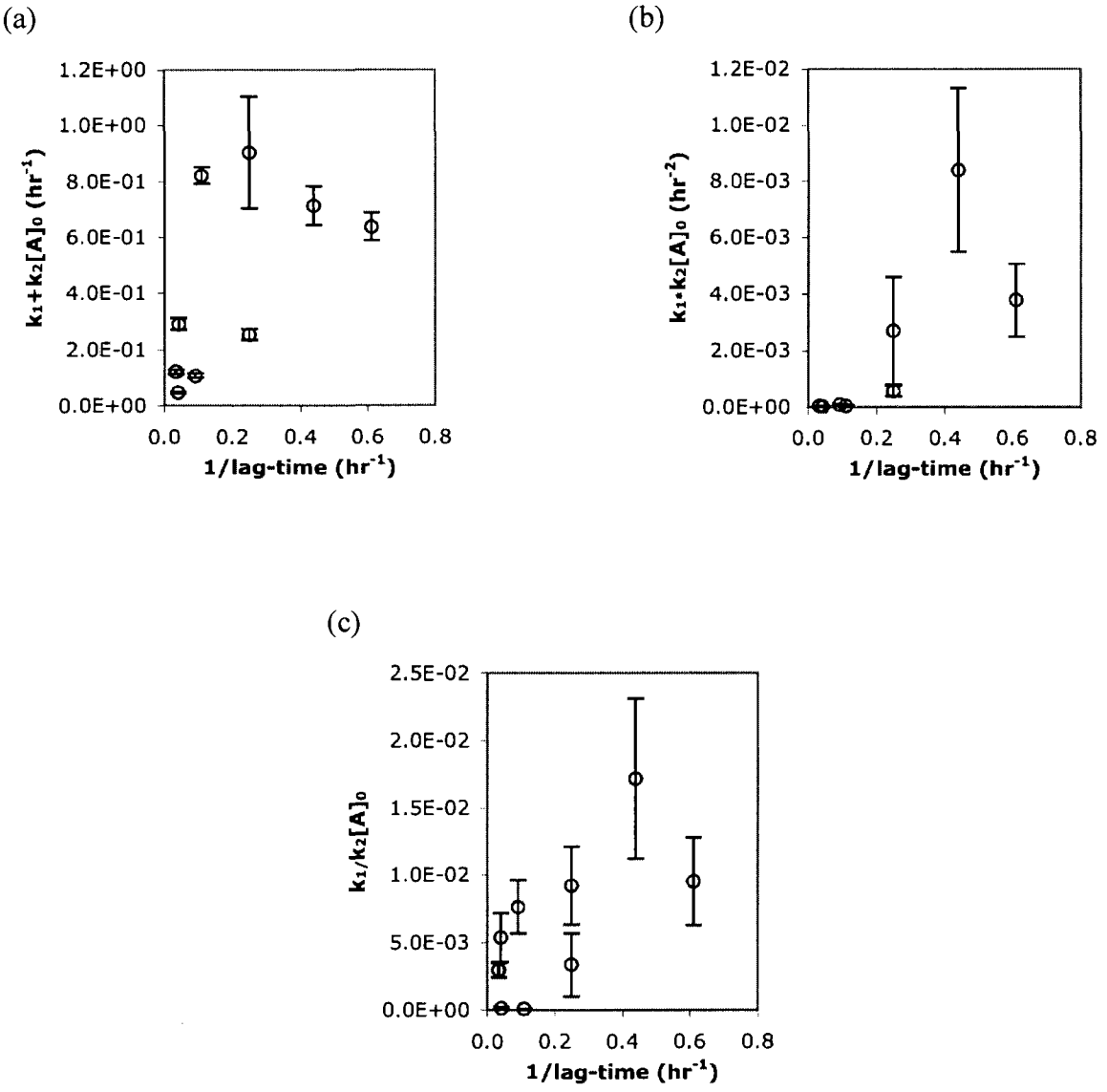


Figure S4. Attempts at finding a correlation with the empirical constant 1/lag-time and: (a) $k_1+k_2[A]_0$, (b) $k_1*k_2[A]_0$, and (c) $k_1/k_2[A]_0$. Again, the data employed in plots (a)-(c) comes from Tables 1 and 5 of the main text.

It was shown in the main text that k_{app} is correlated to $k_2[A]_0$ of the F-W model. The same linear relationship between k_{app} and k_2 can also be observed, Figure S5. However, there is an offset in the values. That is, k_{app} is equal to $100 \times k_2$. This value of 100 is the concentration of the protein (100 μM) and thus the correlation is actually between k_{app} and $k_2[A]_0$. Also multiplying k_2 by the concentration of the protein puts both k_{app} and $k_2[A]_0$ into the same units of hr^{-1}.

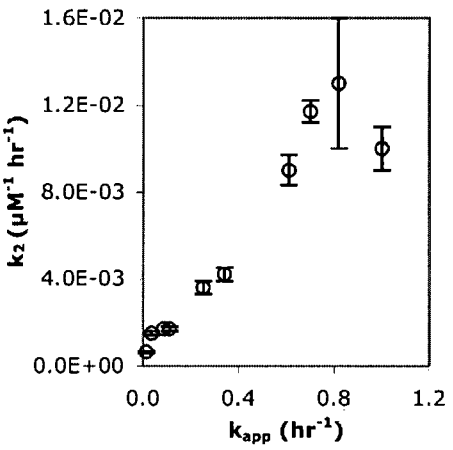


Figure S6. The correlation between the empirical constant k_{app} and k_2 of the F-W model. Note that k_{app} and k_2 are in different units and that k_{app} is approximately equal to $100 \times k_2$.

Attempts to Find a Simple Mathematical Relationship Between the Integrated Rate Equation for the F-W 2-Step Model and the Empirical Eq. 3 of the Main Text. The integrated form of the F-W 2-step model is shown below in eq. S1 (or eq. 1 of the main text) along with the previously reported (1) empirical equation, eq. S2 (or eq. 3 of the main text). We could not find any apparent, simple mathematical relationship between eqs. S1 and S2 as demonstrated below in Scheme S1.

$$[B]_t = [A]_0 - \frac{\frac{k_1}{k_2} + [A]_0}{1 + \frac{k_1}{k_2[A]_0}\exp(k_1 + k_2[A]_0)t} \tag{S1}$$

$$Y = (y_i + m_i x) + \frac{(v_f + m_f x)}{1 + e^{-\left(\frac{\chi - \chi_0}{\tau}\right)}} \tag{S2}$$

Scheme S1. Mathematical attempt at finding an equivalence between eqs. S1 and S2.

If you assume the fluoresence intensity, Y, = the concentration of aggregate at time t, $[B]_t$, then we can set eq. S1=eq. S2:

$$A0 - \frac{\frac{k1}{k2} + A0}{1 + \frac{k1}{k2 \cdot A0}\exp((k1 + k2 \cdot A0) \cdot t)} = (yi + mi \cdot x) + \frac{(vf + mf \cdot x)}{1 + \exp\left(-\left(\frac{\chi - \chi_0}{\tau}\right)\right)}$$

Since $k_{app}=1/\tau$ by definition, the right hand side can be written as:

$$A0 - \frac{\frac{k1}{k2} + A0}{1 + \frac{k1}{k2 \cdot A0}\exp((k1 + k2 \cdot A0) \cdot t)} = (yi + mi \cdot x) + \frac{(vf + mf \cdot x)}{1 + \exp\left(-\left((\chi - \chi_0) \cdot kapp\right)\right)}$$

If you also assume that $k_2 * A_0$ is proportional to k_{app} as was shown in the main text, then:

$$A0 - \frac{\frac{k1}{k2} + A0}{1 + \frac{k1}{k2 \cdot A0}\exp((k1 + k2 \cdot A0) \cdot t)} = (yi + mi \cdot x) + \frac{(vf + mf \cdot x)}{1 + \exp\left(-\left((\chi - \chi_0) \cdot k2 \cdot A0 \cdot z\right)\right)}$$

Since there are no other obvious assumptions to make, we tried to simplify the above expression using Mathematica. Unfortunately, we were unable to simplify this further to determine any correlations between eqs. S1 and S2. In fact, when trying to simplify we ended up with each variable equal to itself (i.e., $k_1=k_1$, $\chi=\chi$, etc.) and therefore, we concluded that there does not appear to be any simple equivalence between eqs. S1 and S2.

References

1. Uversky, V. N., Li, J., and Fink, A. L. (2001) *J. Biol. Chem.* **276**, 10737-10744.

CHAPTER VI

PART II INTRODUCTION: A BRIEF REVIEW OF THE IMPORTANCE OF DIOXYGENASES PLUS AN INTRODUCTION TO TWO SIGNIFICANT SYNTHETIC DIOXYGENASE SYSTEMS

This chapter is an introduction to Part II of this dissertation and introduces dioxygenase catalysis, including discussions of catecholate dioxygenases and a claimed Ru-containing polyoxometalate dioxygenase.

Part II Introduction: A Brief Review of the Importance of Dioxygenases Plus an Introduction to Two Significant Synthetic Dioxygenase Systems

I. Introduction

a. What are dioxygenases?

Dioxygenases are defined as catalysts that are able to incorporate both atoms of dioxygen into a substrate without the product of side products such as H_2O.[1] Therefore, dioxygenases are highly efficient and desirable oxidation catalysts.

b. Dioxygenases and their functions in Nature

Dioxygenases enzymes are widely distributed among plants, animals, and microorganisms in Nature.[1] Their main function in these organisms varies from amino acid metabolism to aromatic compound metabolism.[1] Dioxygenases are also involved in the metabolic disposal of a variety of drugs and foreign substances.[2] A major reaction catalyzed by dioxygenases is the cleavage of an aromatic double bond located (i) between two hydroxylated carbon atoms, (ii) adjacent to two hydroxylated carbon atoms, or (iii) in an indole ring.[2]

II. The Importance of Dioxygenases

a. O_2 activation

Activation of molecular oxygen is difficult because the ground state for dioxygen ($^3O_2$) contains two unpaired electrons in the highest occupied π^* orbitals making it spin-forbidden to react with spin-paired singlet species.[3] In addition, the higher energy of the excited singlet state of dioxygen ($^1O_2$), makes it not very accessible by dioxygenases except by radical-chain mechanisms that result in autoxidation and its low selectivity.[3] However, selective oxidation of dioxygen does occur and is initiated either by (a) a paramagnetic metal ion binding with dioxygen, or (b) an electron-rich substrate activation of a diamagnetic metal followed by an attack by dioxygen,[4,7] Figure 1.

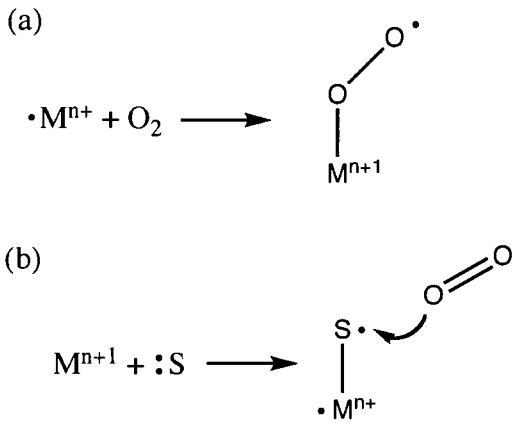


Figure 1. Oxygen activation via (a) metal activation, or (b) substrate activation.

There are so called "Holy Grail" dioxygenase reactions[5] that would have a significant impact on industry as well as on the field of catalysis, for example, those shown in Figure 2. However, these reactions currently require multi-step processes and suitable synthetic dioxygenases for these reactions have yet to be realized.

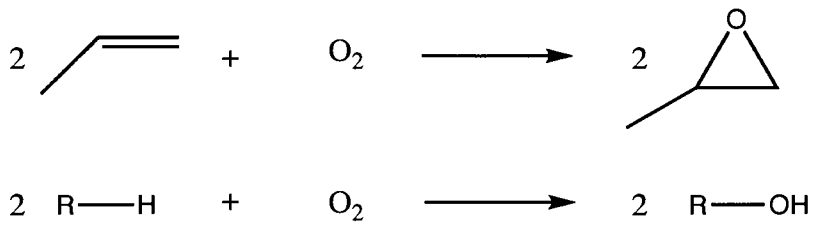


Figure 2. Desired "Holy Grail" dioxygenase reactions.

However, dioxygenase enzymes in Nature have been known since 1955.[6] Among the naturally occurring dioxygenase enzymes are the catechol dioxygenases.

b. Catechol Dioxygenases

i. Introduction. In Nature, catechol dioxygenases are Fe- or Mn-containing enzymes involved in the degradation of hydroxylated aromatics into acyclic compounds via oxygen atom insertions.[1,7] Catechol dioxygenases insert both atoms of dioxygen into catechol (1,2-dihydroxybenzene) or substituted catechol substrates while simulateously catalyzing the cleavage of the benzene ring to form acyclic, more readily degradable compounds, Figure 3. The cleavage of the benzene ring can either occur on the C–C bond between the two ortho hydroxyl groups resulting in an intradiol product (Figure 3, top), or cleavage can occur on the C–C bond adjacent to the ortho hydroxyl groups resulting in the formation of an extradiol product, Figure 3, bottom.

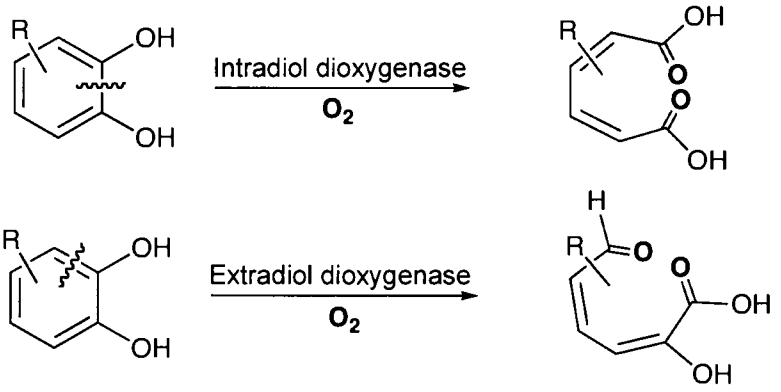


Figure 3. Intradiol and extradiol cleavage with dioxygen insertion by dioxygenases.

ii. Intra- vs. extradiol products. Interestingly, Fe(III) enzymes give intradiol products while Fe(II) enzymes have been shown to give extradiol products.[7] While both the oxidation state of the metal as well as the coordination geometry around the metal have been proposed as possible reasons for the selectivity of a dioxygenase,[8,9] how the enzyme controls the observed ring cleavage sites has yet to be fully elucidated.[10,11,12]

Synthetic dioxygenases have been shown to give only intradiol, only extradiol, or to give both intradiol and extradiol products. The ability to control and tune the catalyst for the desired dioxygenase product is a highly desirable property for a synthetic dioxygenase. Researchers have developed steric as well as electronic effects in the synthetic dioxygenases in attempts to mimic the properties of the enzymes.[10,12]

iii. The [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ catalytic-cycle resting state. One notable catechol dioxygenase precatalyst, actually catalytic-cycle resting state (*vide infra*) is [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ (where 3,5-DTBC = 3,5-di-*tert*-butylcatecholate and 3,5-DBSQ = 3,5-di-*tert*-butylsemiquinone). This complex was originally reported by Pierpont and co-workers in 1983,[13] Figure 4. In 1999,[14] a V-containing precatalyst was reported to exhibit a record catalytic lifetime of more than 100,000 total turnovers with

the substrate 3,5-di-tert-butylcatechol, H_2(3,5-DTBC). It was subsequently determined that 11 *different* V-containing catechol dioxygenase precatalysts all give, as a common component, [VO(3,5-DTBC)(3,5-DBSQ)]$_2$,[15] which was later determined to be a catalytic-cycle resting state.[15,16] Under the dioxygenase conditions, [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ produces both intradiol and extradiol oxidation products from the substrate H_2(3,5-DTBC).[16,17] The dimer catalytic-cycle resting state, [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, still remains a record synthetic dioxygenase in terms of its total turnovers for the substrate H_2(3,5-DTBC).

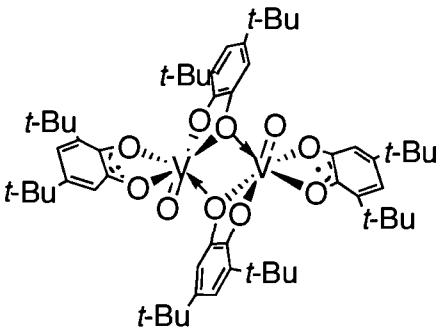


Figure 4. The catalytic-cycle resting state [VO(3,5-DTBC)(3,5-DBSQ)]$_2$.

Chapter VII of this dissertation investigates the dioxygenase catalytic activity of a related but different precatalyst, namely V(3,6-DTBC)$_2$(3,6-DBSQ). In addition Chapter VIII looks at extending the list of substrates for which [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ as well as V(3,6-DTBC)$_2$(3,6-DBSQ) and [MoO(3,5-DTBC)$_2$]$_2$ act as dioxygenase precatalysts.

c. R–H substrates

While catechols are interesting dioxygenase substrates, one of the "Holy Grail" dioxygenase reactions involve insertion of dioxygen into R–H bonds. A synthetic dioxygenase capable of performing this dioxygen insertion does not yet exist, although one claimed dioxygenase has appeared in the literature, as discussed next.

i. Ru-containing polyoxometalate claimed dioxygenase. In 1997 and 1998, the sandwich polyoxometalate $[WZn_3Ru^{III}_2(OH)(H_2O)(ZnW_9O_{34})_2]_2$ (Figure 5) was reported to be a dioxygenase catalyst for converting adamantane into 1-adamantanol and 2-adamantanone.[18,19] However, upon reinvestigation of $[WZn_3Ru^{III}_2(OH)(H_2O)(ZnW_9O_{34})_2]_2$ in 2005 by Yin and Finke,[20] the mechanism of forming 1-adamantanol and 2-adamantanone was revealed to actually be a radical-chain autoxidation, and not dioxygenase, pathway. Interestingly, the same yields of oxidation products were observed by both the prior research group and Yin and Finke, providing strong evidence that they were examining the same catalyst. Chapter IX of this dissertation investigates the actual composition of "$[WZn_3Ru^{III}_2(OH)(H_2O)(ZnW_9O_{34})_2]_2$", **1**, specifically the question of "have 2.0 equivalents of Ru been incorporated into the parent polyoxoanion $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]_2$?".

III. Summary

Dioxygenases are an important class of catalysts able to insert two oxygen atoms from O_2 into a substrate under mild conditions. One important synthetic dioxygenase in the literature is the catechol dioxygenase $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$. In prior work,

Yin and Finke showed that $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ is a catalytic-cycle resting state for $H_2(3,5\text{-DTBC})$ with a record ≥100,000 total turnovers of $H_2(3,5\text{-DTBC})$ dioxygenase catalysis using O_2. On the other hand, $[WZn_3Ru^{III}_2(OH)(H_2O)_2(ZnW_9O_{34})_2]_2$ turns out to be a nearly Ru-free autoxidation catalyst rather than the previously claimed dioxygenase catalyst.

IV. References

[1] Hayaishi, O.; Editor *Molecular Mechanisms of Oxygen Activation*; Academic Press: New York, 1974.

[2] Hayaishi, O.; Nozaki, M. *Science* **1969**, *164*, 389-396.

[3] Bugg, T. D. H. *Tetrahedron* **2003**, *59*, 7075-7101.

[4] Que, Jr., L.; Ho, R. Y. N. *Chem. Rev.* **1996**, *96*, 2607-2624.

[5] Hill, C. L.; Weinstock, I. A. *Nature* **1997**, *388*, 332-333.

[6] Hayaishi, O.; Katagiri, M.; Rothberg, S. *J. Am. Chem. Soc.* **1955**, *77*, 5450-5451.

[7] Costas, M.; Mehn, M. P.; Jensen, M. P.; Que, Jr., L. *Chem. Rev.* **2004**, *104*, 939-986.

[8] Lipscomb, J. D.; Orville, A. M.; Frazee, R. W.; Miller, M. A.; Ohlendorf, D. H. In *Keio University Symposia for Life Science and Medicine*, 1998; Vol. 1, pp 263-275.

[9] Jo, D.-H.; Que, Jr., L. *Angew. Chem., Int. Ed. Engl.* **2000**, *39*, 4284-4287.

[10] Bugg, T. D. H.; Lin, G. *Chem. Commun.* **2001**, 941-952.

[11] Vaillancourt, F. H.; Bolin, J. T.; Eltis, L. D. *Critical Rev. Biochem. Mol. Biol.* **2006**, *41*, 241-267.

[12] Bugg, T. D. H.; Ramaswamy, S. *Curr. Opin. Chem. Biol.* **2008**, *12*, 134-140.

[13] Cass, M. E.; Green, D. L.; Buchanan, R. M.; Pierpont, C. G. *J. Am. Chem. Soc.* **1983**, *105*, 2680-2686.

[14] Weiner, H.; Finke, R. G. *J. Am. Chem. Soc.* **1999**, *121*, 9831-9842.

[15] Yin, C.-X.; Finke, R. G. *J. Am. Chem. Soc.* **2005**, *127*, 9003-9013.

[16] Yin, C.-X.; Finke, R. G. *J. Am. Chem. Soc.* **2005**, *127*, 13988-13996.

[17] Yin, C.-X.; Sasaki, Y.; Finke, R. G. *Inorg. Chem.* **2005**, *44*, 8521-8530.

[18] Neumann, R.; Dahan, M. *Nature* **1997**, *388*, 353-355.

[19] Neumann, R.; Dahan, M. *J. Am. Chem. Soc.* **1998**, *120*, 11969-11976.

[20] Yin, C.-X.; Finke, R. G. *Inorg. Chem.* **2005**, *44*, 4175-4188.

CHAPTER VII

SYNTHESIS AND CHARACTERIZATION OF V^{V}(3,6-DBSQ)(3,6-DBCat)$_2$, A d^0 METAL COMPLEX WITH DIOXYGENASE CATALYTIC ACTIVITY

This dissertation chapter contains the manuscript of a communication published in *Inorganic Chemistry* **2009**, *48*, 13496-13498. This chapter presents the synthesis, characterization and initial dioxygenase catalytic activity studies of the new complex V^{V}(3,6-DBSQ)(3,6-DBCat)$_2$ (where 3,6-DBSQ = 3,6-di-tert-butylseminquinone and 3,6-DBCat = 3,6-di-tert-butylcatecholate).

The manuscript was prepared by Aimee M. Morris with near infrared, infrared, and EPR spectroscopy interpretation completed with the assistance by Prof. Cortlandt G. Pierpont. In addition, Profs. Cortlandt G. Pierpont and Richard G. Finke assisted with the writing and editing.

Synthesis and Characterization of V^V(3,6-DBSQ)(3,6-DBCat)$_2$, a d^0 Metal Complex with Dioxygenase Catalytic Activity

Aimee M. Morris, Cortlandt G. Pierpont, and Richard G. Finke

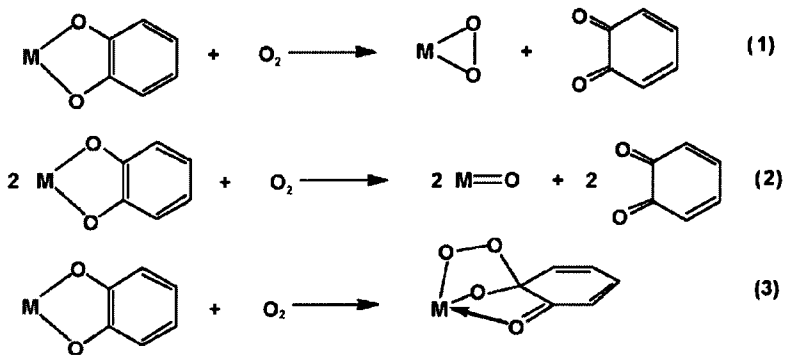
Abstract

Transition-metal complexes containing redox-active quinoid ligands are of interest because of their catalytic capabilities in multielectron, substrate-activation reactions such as dioxygenase catalysis using O_2. The new catecholate complex V^V(3,6-DBSQ)(3,6-DBCat)$_2$ (where 3,6-DBSQ = 3,6-di-*tert*-butylsemiquinone and 3,6-DBCat = 3,6-di-*tert*-butylcatecholate) was synthesized by combining VO(acac)$_2$ with 1 equiv of 3,6-DBBQ (where 3,6-DBBQ = 3,6-di-*tert*-butylbenzoquinone) and 2 equiv of H_2(3,6-DBCat) in dry methanol under an inert atmosphere. The resultant complex was characterized by single-crystal X-ray diffraction, elemental analysis, near-IR, UV/vis, and electron paramagnetic resonance (EPR) spectroscopy. The crystallography as well as the near-IR and EPR studies suggest that the radical spin is localized on the 3,6-DBSQ ligand at room temperature, making V^V(3,6-DBSQ)(3,6-DBCat)$_2$ a type 1 mixed-valence complex. Initial dioxygenase catalysis studies reveal that V^V(3,6-DBSQ)(3,6-DBCat)$_2$ is a good dioxygenase precatalyst for the substrate H_2(3,6-DBCat) with O_2 in ca. 600 total turnovers to > 93% intra- and extradiol products with only 1-2% of the undesired benzoquinone autoxidation product.

Main Text

Studies on the coordination chemistry of transition metal complexes containing redox-active quinoid ligands have recently been focused on multi-electron substrate-activation reactions that use electron density derived from the ligands without a change[1-4] in metal oxidation state.[5] By far the most common substrate is dioxygen, which is well-known to add to nucleophilic metal complexes; dioxygen also adds to complexes containing α-diamide and catecholate ligands without a change in oxidation state at the metal.[2,3] Reactions of this type are relevant to the mechanism of oxygen-activation in enzymatic and synthetic catechol oxidation reactions.[4] For example, peroxosemiquinonate complexes, proposed as intermediates in catechol oxidation reactions, have been formed reversibly for catecholate complexes of antimony(V) demonstrating that O_2 reduction is mediated by the catecholate ligand.[6]

In much earlier work with tris(catecholato) complexes of Mo(VI) and V(V), it was observed that O_2 addition leads to the formation of oxo-bis(catecholato) species with release of *o*-benzoquinone.[7,8] These reactions demonstrate that there are three potential products that may be formed by the addition of O_2 to a d^0 metal complex containing a reduced quinoid (catecholate) ligand without a change in metal oxidation state, namely reactions and products (1) – (3).



These reactions can be taken a step further with the insertion of O_2 into a catechol C-C bond to give the ring-opened products associated with the intra- and extradiol dioxygenase enzymes and related synthetic systems.[4] Recent structural characterization on an extradiol dioxygenase enzyme at various stages in the catalytic cycle has provided evidence for an iron-peroxosemiquinonate intermediate.[9]

The catechol used most frequently in reactions that model catechol dioxygenase catalytic activity is 3,5-di-tert-butylcatechol [H_2(3,5-DBCat); note once deprotonated, H_2(3,5-DBCat) becomes the catecholate 3,5-DBCat][10] due to its commercial availability and the activating effect of the tert-butyl substituents. Reactions have been studied with a wide variety of metals, but the two metals studied most extensively are iron, due to its biological relevance, and vanadium due to its catalytic efficiency.[11]

Earlier reactions between either $V(CO)_6$ and 3,5-di-tert-butyl-1,2-benzoquinone (3,5-DBBQ) or $VO(acac)_2$ and H_2(3,5-DBCat) were both observed to form "V(3,5-DBQ)$_3$" under inert conditions (where 3,5-DBQ = 3,5-di-tert-butylquinoid).[7] EPR characterization established that "V(3,5-DBQ)$_3$" had a ligand-based S= ½ spin ground state that might be associated with a V^{III}(3,5-DBSQ)$_3$ (where 3,5-DBSQ=3,5-di-tert-butylsemiquinone) species with strong V-SQ antiferromagnetic spin coupling, or with the redox isomer V^{V}(3,5-DBSQ)(3,5-DBCat)$_2$ and its mixed-charge ligands.[7] In the presence

of trace quantities of O_2, the product was observed to form dimeric [$V^{V}O$(3,5-DBSQ)(3,5-DBCat)]$_2$.[7a] In the presence of excess oxygen [VO(3,5-DBSQ)(3,5-DBCat)]$_2$ reacts further to ultimately give V_2O_5 and 3,5-DBBQ,[7a] but in the presence of excess H_2(3,5-DBCat) the dimeric complex serves as an exceptionally long-lived catalyst for catechol oxidation with $>10^5$ turnovers of H_2(3,5-DBCat) to produce well-characterized intra and extradiol dioxygenase products.[12]

However, less investigated are metal complexes containing 3,6-DBCat ligands,[13,14] or catechol oxidation catalysis[14] using H_2(3,6-DBCat). The symmetrical distribution of tert-butyl substituents should have the advantages of (i) leading to a less complicated series of organic oxidation products, relative to H_2(3,5-DBCat); and (ii) monomeric metal precursors and possibly catalysts should be favored by the disposition of tert-butyl substituents adjacent to the catecholate oxygens which, in turn, should disfavor bridging interactions to a second metal center. However, these and other advantages (or disadvantages) of 3,6-DBCat remain less investigated in comparison to 3,5-DBCat.[13,14]

Herein we report that the reaction between VO(acac)$_2$ and a 2 : 1 mixture of H_2(3,6-DBCat) and 3,6-DBBQ in dry methanol gives a new, octahedral monomeric product of formula: V^{V}(3,6-DBSQ)(3,6-DBCat)$_2$ (**1**). In addition, we find that **1** is a good H_2(3,6-DBCat) dioxygenase precatalyst,[15] leading to ~600 total catalytic turnovers (TTO) to more than 93% intra- and extra-diol dioxygenase products, with very little (1-2%) of the undesired benzoquinone autoxidation product, *vide infra*.

Complex **1** was synthesized according to equation (4). The bulk sample composition, verified by elemental analysis (Galbraith Laboratories, Inc. Knoxville, TN),

is consistent with the formula V(3,6-DBSQ)(3,6-DBCat)$_2$•CH_3OH. Calc. [found]: V, 6.85 [6.58]; C, 69.43 [69.91]; H 8.67 [8.90].

$$V^{IV}O(acac)_2 + 3,6\text{-DBBQ} + 2\ H_2(3,6\text{-DBCat}) \xrightarrow{\text{dry MeOH}} V^{V}(3,6\text{-DBSQ})(3,6\text{-DBCat})_2 + 2\ \text{acacH} + H_2O \quad (4)$$

Crystallographic characterization[16] is consistent with the bulk sample composition and reveals that **1** consists of charge-localized SQ and Cat ligands. Complex **1** is located about a crystallographic two-fold axis that bisects the semiquinonate ligand, Figure 1.

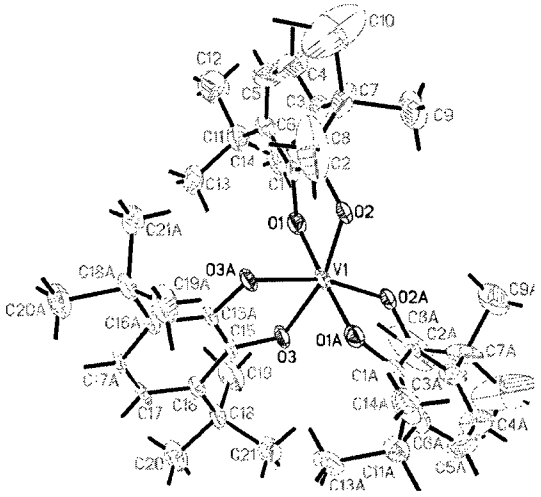


Figure 1. 50% probability thermal ellipsoid plot of V(3,6-DBSQ)(3,6-DBCat)$_2$.

Relevant to the structure of **1** is that prior structure determinations of complexes containing mixed-charge SQ and Cat ligands, C–O lengths have been diagnostic of the reduction level of the quinone ligand.[17] In this case as well, the C–O length for the ligand

located along the two-fold axis is at a SQ value, 1.298(3) Å, the C–O lengths for the independent Cat ligand are 1.330(4) and 1.336(4) Å, and the C–C length between ring C1 and C2 carbon atoms is significantly longer for the SQ ligand (1.466(6) Å) than for the Cat ligand (1.421(5) Å), Table 1. The C–C lengths at the ring 3-4 and 5-6 positions are generally found to be contracted for SQ ligands (C16–C17 of the structure determination), and this is the case for the 3,6-DBSQ ligand with a value of 1.364(4) Å. However, lengths at these positions for the Cat ligand are found to be similarly contracted (1.376(5), 1.373(6) Å, Table 1) suggesting some SQ character for this ligand that could result from either crystallographic disorder or a shift in charge distribution to, conceivably, a V^{III}(3,6-DBSQ)$_3$ redox isomer.

Table 1. Selected bond lengths (Å) for V(3,6-DBSQ)(3,6-DBCat)$_2$.

	Cat	ligands		SQ	ligands
V-O1	1.868(2)	V-O2	1.898(2)	V-O3	1.971(2)
C1-O1	1.330(4)	C2-O2	1.326(4)	C15-O3	1.298(3)
C1-C2	1.421(5)			C15-C15A	1.466(6)
C3-C4	1.376(5)	C5-C6	1.373(6)	C16-C17	1.364(4)

The related complex of manganese was found to undergo a shift in charge distribution, from Mn^{III}(3,6-DBSQ)$_3$ at temperatures just above room temperature to Mn^{IV}(3,6-DBSQ)$_2$(3,6-DBCat) at lower temperatures.[13c] This equilibrium could be monitored by the appearance of a ligand to ligand' intervalence transfer (LL'IT) transition at 2300 nm for the complex at low temperature. However, electronic spectra obtained for the vanadium complex **1** in solution (Figure 2 and Figure S4 of the SI) and in the solid state (Figure S1-S3 of the SI) show no evidence for a shift in charge between

the Cat and SQ ligands. Spectra recorded in both media show an intense absorption at 670 nm[18] (ε = 33,000-38,000 $M^{-1}cm^{-1}$, concentrations = 2.2×10^{-5}-9.4×10^{-5} M in toluene), similar to the 668 nm band ($\varepsilon \sim 27{,}000\ M^{-1}cm^{-1}$) observed for [VO(3,5-DBSQ)(3,5-DBCat)]$_2$ in toluene.[11] No other absorptions at lower energy that might be associated with a precedented[13a-c] Type 2 mixed-valence complex (where the region is ~2100 nm[13a-c]) are observed for **1**, at least at room temperature.

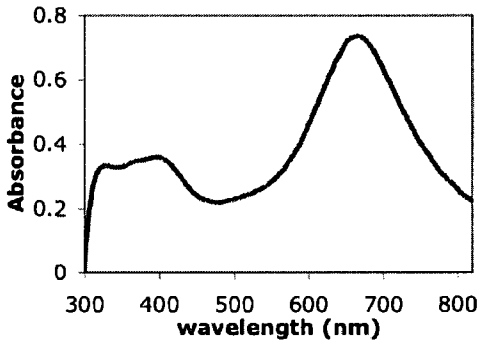


Figure 2. UV/Vis spectra of 2.2×10^{-5} M V(3,6-DBSQ)(3,6-DBCat)$_2$ in toluene and under an inert atmosphere at room temperature.

The isotropic EPR spectrum recorded for V^{V}(3,6-DBSQ)(3,6-DBCat)$_2$ in toluene confirms the charge-localized nature of the complex at room temperature. The spectrum, shown in Figure 2, is that of the 3,6-DBSQ radical, weakly coupled with the I= 7/2 ^{51}V center, with additional coupling from the two SQ ring protons at carbon atoms C4 and C5 (C17 and C17A of the structure determination). The spectrum is centered about a <g> value of 2.0058, and a reasonable simulation could be obtained with A(^{51}V) of 3.47 G and A(^{1}H) of 4.53 G. This is in contrast to the nine-line spectrum with <g> value of 2.004 along with A(^{51}V) and A(^{1}H) of 2.85 and 0.35 G, respectively, observed for

$[VO(3,5\text{-DBSQ})(3,5\text{-DBCat})]_2$.[8a] Together, the electronic and EPR spectra of **1** point to a charge localized, Type 1 mixed-valence formulation for $V^{V}(3,6\text{-DBSQ})(3,6\text{-DBCat})_2$ with no contribution from, for example, $V^{III}(3,6\text{-DBSQ})$, again at room temperature. Studies of other temperatures would be valuable and are in progress.

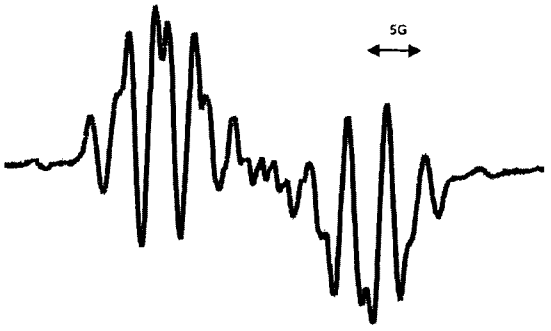
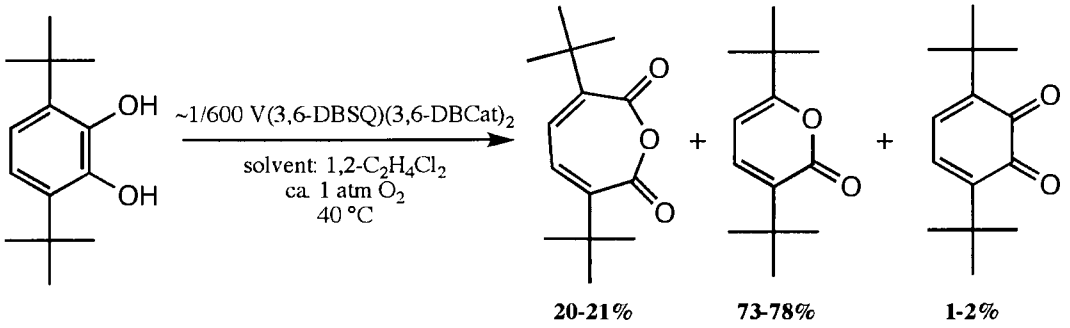


Figure 3. EPR spectrum of $V(3,6\text{-DBSQ})(3,6\text{-DBCat})_2$ in toluene and under an inert atmosphere at 23 °C.

Preliminary studies examining the catalytic activity of **1** plus the substrate $H_2(3,6\text{-DBCat})$ under O_2 show that the products and yields shown in equation (5) are obtained according to GC-MS identification and GC quantitation. Importantly, as eq (5) summarizes, >93% dioxygenase products are obtained with **1** along with only 1-2% of the undesired autoxidation product, benzoquinone (the ranges of the products shown represent the small variability in three repeat catalytic runs). In addition, even these initial catalytic results demonstrate the catalytic dioxygenase capability derived from **1** of ca. 600[19] TTOs for the substrate $H_2(3,6\text{-DBCat})$. Additional studies of the dioxygenase activity derived from **1** are currently underway.[15]



(5)

To summarize, herein we have described the synthesis of the catecholate complex V(3,6-DBSQ)(3,6-DBCat)$_2$, **1**, and its characterization by single crystal X-ray diffraction, elemental analysis, near IR, UV/Vis, and EPR. The data reveal that **1** is a type 1 mixed valence compound at room temperature with localized radical spin on the 3,6-DBSQ ligand. Catalysis of H_2(3,6-DBCat) plus O_2 plus **1** as the precatalyst gives intra and extradiol dioxygenase products along with 1-2% of the autoxidation product benzoquinone in ca. 600 TTOs of catalysis. Future studies and a full report on the dioxygenase catalytic activity of **1** as well as other V and Mo catecholate complexes will be reported in due course.[15]

Acknowledgment. We gratefully acknowledge NSF grant CHE 9531110 to RGF.

Supporting Information available: Full experimental section including materials and methods, the synthesis of H_2(3,6-DBCat), 3,6-DBBQ, and V(3,6-DBSQ)(3,6-DBCat)$_2$; X-ray crystallographic information, atomic coordinates, isotopic displacement, bond length, and bond angle tables; Near IR and IR spectra of V(3,6-DBSQ)(3,6-DBCat)$_2$; solid-state and solution UV/Vis spectra of V(3,6-DBSQ)(3,6-DBCat)$_2$.

References:

[1] (a) Blackmore, K. J.; Lal, N.; Ziller, J. W.; Heyduk, A. F. *J. Am. Chem. Soc.* **2008**, *130*, 2728-2729. (b) Ketterer, N. A.; Fan, H.; Blackmore, K. J.; Yang, X.; Ziller, J. W.; Baik, M.-H.; Heyduk, A. F. *J. Am. Chem. Soc.* **2008**, *130*, 4364-4374.

[2] Stanciu, C.; Jones, M. E.; Fanwick, P. E.; Abu-Omar, M. M. *J. Am. Chem. Soc.* **2007**, *129*, 12400-12401.

[3] Rolle III, C. J.; Hardcastle, K. I.; Soper, J. D. *Inorg. Chem.* **2008**, *47*, 1892-1894.

[4] (a) Costas, M.; Mehn, M. P.; Jensen, M. P.; Que, Jr., L. *Chem. Rev.* **2004**, *104*, 939-986. (b) Hitomi, Y.; Yoshida, M.; Higuchi, M.; Minami, H.; Tanaka, T.; Funabiki, T. *J. Inorg. Biochem.* **2005**, *99*, 755-763.

[5] For example, Heyduk has reported oxidative-addition reactions to d^0 Zr(IV) that utilize charge from reduced iminoquinone and diiminoquinone ligands.[1]

[6] (a)Abakumov, G. A.; Poddel'sky, A. I.; Grunova, E. V.; Cherkasov, V. K.; Fukin, G. K.; Kurskii, Y. A.; Abakumova, L. G. *Angew. Chem. Int. Ed.* **2005**, *44*, 2767-2771. (b) Cherkasov, V. K.; Abakumov, G. A.; Grunova, E. V.; Poddel'sky, A. I.; Fukin, G. K.; Baranov, E. V.; Kurskii, Y. V.; Abakumova, L. G. *Chem. Eur. J.* **2006**, *12*, 3916-3927.

[7] (a) Buchanan, R. M.; Pierpont, C. G. *Inorg. Chem.* **1979**, *18*, 1616-1620. (b) Cass, M. E.; Pierpont, C. G. *Inorg. Chem.* **1986**, *25*, 122-123. (c) Liu, C.-M.; Restorp, P.; Nordlander, E.; Schmeh, D.; Shoemaker, R.; Pierpont, C. G. *Inorg. Chem.* **2004**, *43*, 2114-2124.

[8] (a) Cass, M. E.; Greene, D. L.; Buchanan, R. M.; Pierpont, C. G. *J. Am. Chem. Soc.* **1983**, *105*, 2680-2686. (b) Cass, M. E.; Gordon, N. R.; Pierpont, C. G. *Inorg. Chem.* **1986**, *25*, 3962-3967.

[9] Kovaleva, E. G.; Lipscomb, J. D. *Science* **2007**, *316*, 453-457.

[10] Initial reductive activation of triplet O_2 by electron transfer from a spin singlet catecholate ligand would be spin-forbidden as an inner-sphere condensation reaction, yet basic solutions of 3,5-di-tert-butylcatechol (3,5-DBCat) are notoriously oxygen sensitive, giving ring-cleavage products upon exposure to air (Speier, G.; Tyeklar, Z. *J. Mol. Catal.* **1990**, *57*, L17-L19). It therefore seems reasonable that the initial activating Cat→O_2 electron-transfer step should take place with formation of a quasi-outer sphere intermediate that leads to one of the reduced oxygen species shown in eqs. (1)-(3).

[11] In an earlier publication, we cited 28 literature 3,5-DBCat oxidation precatalyst systems based on vanadium (Yin, C.-X.; Finke, R. G. *J. Am. Chem. Soc.* **2005**, *127*, 9003-9013). By monitoring the O_2 uptake we were able to establish that these precatalysts proceed with initial formation of H_2O_2 and 3,5-di-tert-butyl-1,2-benzoquinone (3,5-DBBQ) in the presence of excess H_2(3,5-DBCat). The catalyst resting state, [VO(3,5-DBSQ)(3,5-DBCat)]$_2$ (3,5-DBSQ=3,5-di-tert-butylsemiquinone), is

formed in an autocatalytic step and then breaks in half to a monomeric dioxygenase catalyst according to the observed kinetics.[12]

[12] (a) Yin, C.-X.; Finke, R. G. *J. Am. Chem. Soc.* **2005**, *127*, 13988-13996. (b) Yin, C.-X.; Sasaki, Y.; Finke, R. G. *Inorg. Chem.* **2005**, *44*, 8521-8530.

[13] (a) Attia, A. S.; Jung, O.-S.; Pierpont, C. G. *Inorg. Chim. Acta* **1994**, *226*, 91-98. (b) Attia, A. S. Pierpont, C. G. *Inorg. Chem.* **1995**, *34*, 1172-1179. (c) Attia, A. S.; Pierpont, C. G. *Inorg. Chem.* **1998**, *37*, 3051-3056. (d) Liu, C.-M.; Restorp, P.; Nordlander, E.; Pierpont, C. G. *Chem. Commun.* **2001**, 2686-2686. (e) Liu, C.-M.; Nordlander, E.; Schmeh, D.; Shoemaker, R.; Pierpont, C. G. *Inorg. Chem.* **2004**, *43*, 2114-2124.

[14] (a) Hitomi, Y.; Tase, Y.; Higuchi, M.; Tanaka, T.; Funabiki, T. *Chem. Lett.* **2004**, *33*, 316-317. (b) Hitomi, Y.; Yoshida, M.; Higuchi, M.; Minami, H.; Tanaka, T.; Funabiki, T. *J. Inorg. Biochem.* **2005**, *99*, 755-763.

[15] Morris, A. M.; Pierpont, C. G.; and Finke, R. G. *Inorg. Chim. Acta* submitted.

[16] Unit cell data for $C_{43}H_{64}O_7V$ (V (3,6-DBSQ)(3,6-DTBCat)$_2$•CH_3OH): orthorhombic, space group *Ccca*; Z=8. a=19.9969(17) Å, b=23.740(2) Å, c=18.1741(16) Å. $R(R_w)$=0.0540(0.0840); GOOF=1.071.

[17] (a) Bhattacharya, S.; Gupta, P.; Basuli, F.; Pierpont, C. G. *Inorg. Chem.* **2002**, *41*, 5810-5816. (b) Sun, X.; Chun, H.; Hildenbrand, K.; Bothe, E.; Weyhermuller, T.; Neese, F.; Wieghardt, K. *Inorg. Chem.* **2002**, *41*, 4295-4303.

[18] The large intensity of the band at 670 nm leads one to wonder if this band is a LL'IT band. However, this band is higher in energy than expected, and to our knowledge, unprecedented for LL'IT systems.

[19] Total catalytic turnovers (TTO) were calculated by the following formula:

$$\frac{\sum[\text{oxygenated products; (mmol)}]}{[\text{V(3,6-DBSQ)(3,6-DBCat)}_2\text{; (mmol)}]}$$

Supporting Information for:

Synthesis and Characterization of V^{V}(3,6-DBSQ)(3,6-DBCat)$_2$, a d^0 Metal Complex with Dioxygenase Catalytic Activity

Aimee M. Morris, Cortlandt G. Pierpont, and Richard G. Finke

EXPERIMENTAL

Materials

The following were obtained from the indicated sources, then used as received: $TiCl_4$ (Aldrich, 99.9%), Ag_2O (Aldrich, 99%), catechol (Aldrich, ≥99%), toluene (Aldrich, 99.8%, anhydrous), xylenes (Fisher Scientific, ACS grade), isobutylene (Aldrich, 99%), hexanes (Fisher Scientific, ACS grade), methylene chloride (Fisher, ACS grade), diethyl ether (Aldrich, HPLC grade), n-pentane (Fisher, pesticide grade), and $VO(acac)_2$ (Aldrich, 95%; stored in a Vacuum Atmospheres N_2 drybox).

Instrumentation

^1H NMR was run on a Varian Inova (JS-300) nuclear magnetic resonance (NMR) spectrometer using $CDCl_3$ as the solvent. The ^1H NMR was referenced to the residual proton impurity in the deuterated solvent. Diffraction data were collected on a Bruker APEX2 diffractometer employing Mo Kα radiation. Standard Bruker APEX2 control and integration software was employed, and Bruker SHELXTL[1] software was used for structure solution, refinement, and graphics. SADABS[2] correction was employed and the structure was solved by direct methods and refined by a full-matrix, weighted least-squares process. In addition, PLATON's SQUEEZE function was used to refine the disordered methanol solvent. The IR spectrum was collected on an Avatar 360 FT-IR spectrometer with a solid KBr sample pellet. The near IR spectrum was obtained with a solid KBr sample pellet on a Perkin-Elmer Lambda 9 UV-visible near-IR spectrophotometer. Solution UV/vis measurements were obtained on a Hewlett-Packard 8452A diode spectrometer in glass, Schlenk UV cells (i.e., cells with a Teflon valve

glass-blown onto a standard pyrex glass cuvette). Electron paramagnetic resonance (EPR) spectra were recorded on a Bruker ESP-300E spectrometer using a 5-mm o.d. quartz j-young EPR tube with 2,2-diphenyl-1-picrylhydrazyl (DPPH) as the reference compound (g=2.0037). Elemental analyses were performed by Galbraith Laboratories, Inc. (Knoxville, TN). Positive and negative ion electrospray ionization mass spectrometry analyses were performed on an Agilent model 6220 TOF mass spectrometer.

Synthesis of 3,6-di-tert-butylcatechol H_2(3,6-DBCat)

The H_2(3,6-DBCat) synthesis was carried out in a Parr pressure reactor (model 4561) made of Monel 400 alloy. The reactor is equipped with a pressure gauge and automatic temperature controller. The interior of the reactor contains an impeller, thermocouple, cooling loop, and dip tube, all of which are in contact with the reaction solution. A glass liner was dried overnight at 160 °C and used to avoid contact of the solution with the interior of the reactor. H_2(3,6-DBCat) was synthesized according to previously published literature methods,[3] but with the following changes: (i) the reaction was scaled down by a factor of five in order to accommodate our Parr bomb reactor volume (~300 mL) and not exceed a pressure of 16 atm, as described in the literature,[3] (ii) the reaction was run at 150 °C instead of 100 °C, and the reaction time was extended from 1.5 hours to 4 hours because the reaction is not complete at 1.5 hours as monitored by the pressure loss. In addition, following the vacuum distillation, the product was purified by recrystallizating twice in hot n-pentane (purified yield 4 g (12 %)). The purified product was characterized by ^{1}H NMR in $CDCl_3$ and MS. Observed [previously

reported[3]]: 6.75 [6.59]; 5.33 [5.11]; and 1.39 [1.31] ppm in a 1:1:9 ratio. GC-MS found for $C_{14}H_{21}O_2$, *m/z* 221 H(3,6-DBCat)$^-$.

Synthesis of 3,6-Di-tert-butyl-1,2-benoquinone (3,6-DBBQ)

3,6-DBBQ was synthesized according to the literature procedure.[4] Specifically, 0.899 g (4.04 mmol) of pure 3,6-DBCat was dissolved in ~10 mL of diethyl ether with stirring. Upon addition of 2.000 g (8.63 mmol) of Ag_2O (in excess), the solution immediately turns green. Stirring was continued for ~5 minutes and then the grey solid (Ag) was filtered off using a medium frit. The solution was rotovapped to dryness yielding crude 3,6-DBBQ. The crude 3,6-DBBQ was then purified by column chromatography. The column (450 × 30 mm) was packed with a slurry of silica gel (Aldrich 70-230 mesh, 100 g, suspended in ~200 mL of n-hexanes). The crude 3,6-DBBQ was dissolved in ~3 mL of CH_2Cl_2, placed on top of the column and eluted with ~500 mL of CH_2Cl_2. The green colored fractions were combined and rotovapped to dryness. The purified 3,6-DBBQ was then dried under vacuum at room temperature overnight. Yield 0.127 g (14 %). ^{1}H NMR in $CDCl_3$ observed [previously reported[4]]: 6.81 [6.17] and 1.27 [1.21] ppm in a 1:9 ratio. GC-MS found for $C_{14}H_{20}O_2Na$, *m/z* 243 (3,6-DBBQ+Na$^+$).

Synthesis of V(3,6-DBSQ)(3,6-DBCat)$_2$

V(3,6-DBSQ)(3,6-DBCat)$_2$ was synthesized under Ar by distilling at 70 °C dry methanol over the stirring mixture of $VO(acac)_2$ (68.0 mg, 0.256 mmol), H_2(3,6-DBCat) (114 mg, 0.511 mmol), and 3,6-DBBQ (58.3 mg, 0.265 mmol) over the course of ~2

hours. The resultant solution was deep blue. The entire apparatus was sealed and placed in the freezer to induce crystallization which occurred overnight. The apparatus was then transferred into a N_2 atmosphere drybox and the resultant deep-blue solid was filtered over a medium frit and dried under vacuum at room temperature overnight. Yield 121 mg (66 %). Elemental analysis (Galbraith Laboratories, Inc. Knoxville, TN) was consistent with V(3,6-DBSQ)(3,6-DBCat)$_2$•CH_3OH. Calc. [found]: V, 6.85 [6.58]; C, 69.43 [69.91]; H 8.67 [8.90].

CHARACTERIZATION

X-ray Diffraction

The single crystal was removed from the mother liquor after the freezer induced crystallization (*vide supra*), mounted on a goniometer, and data collection was started. The X-ray data, atomic coordinates, anisotropic displacement, bond length, and bond angle tables for the structure solution of V(3,6-DBSQ)(3,6-DBCat)$_2$ can be seen in Tables S1-S4 below.

Table S1. Crystal data and structure refinement for V(3,6-DBSQ)(3,6-DBCat)$_2$.

Identification code: rf110r_0m
Empirical formula: $C_{42}H_{60}O_6V$
Formula weight: 711.84
Temperature: 296(2) K
Wavelength: 0.71073
Crystal system: orthorhombic
Space group: Ccca
Unit cell dimensions:

a: = 19.9969(17) Å alpha = 90°
b: = 23.740(2) Å beta = 90°
c: = 18.1741(16) Å gamma = 90°
Volume, Z: 8627.8(13) $Å^3$, 8
Density (calculated): 1.096 mg/m^3
Absorption coefficient: 0.270 mm^{-1}
F(000): 3064
Crystal size: 0.36 x 0.10 x 0.07 mm
Eange for data collection: 1.74 to 23.53°
Limiting indices: $-22 < \theta < 22$, $-26 < \theta < 26$, $-20 < \theta < 20$
Reflections collected: 14529
Independent reflections: 3222 (R(int) = 0.0819)
Completeness to 23.53°, 99.9 %
Max. and min. transmission: 0.9826 and 0.9080
Refinement method: Full-matrix least-squares on F^2
Data / restraints / parameters: 3222 / 0 / 331
Goodness-of-fit on F^2: 1.071
Final R indices: R1 = 0.0540, wR2 = 0.1195
R indices (all data): R1 = 0.0840, wR2 = 0.1325
Largest diff. peak and hole: 0.436 and -0.356

Table 2. Atomic coordinates [x 10^4] and equivalent isotropic displacement parameters [x 10^3] for V(3,6-DBSQ)(3,6-DBCat)$_2$. U(eq) is defined as one third of the trace of the orthogonalized tensor.

	x	y	z	U(eq)
V(1)	2500	5000	385(1)	23(1)
O(1)	2995(1)	4533(1)	1008(1)	24(1)
O(2)	1924(1)	4364(1)	405(1)	25(1)
O(3)	1992(1)	5311(1)	-451(1)	26(1)
C(1)	2825(2)	3991(1)	1007(2)	30(1)
C(2)	2191(2)	3897(1)	678(2)	26(1)
C(3)	1900(2)	3359(1)	632(2)	40(1)
C(4)	2279(3)	2943(2)	956(3)	68(2)
C(5)	2906(3)	3032(2)	1278(3)	75(2)
C(6)	3207(2)	3551(2)	1307(2)	48(1)
C(7)	1240(2)	3261(2)	226(2)	44(1)
C(8)	1038(4)	2634(2)	242(4)	78(2)
C(9)	677(2)	3601(2)	590(3)	56(1)
C(10)	1314(2)	3437(2)	-581(2)	41(1)
C(11)	3921(3)	3660(2)	1598(3)	65(2)

C(12)	3880(3)	4028(2)	2285(3)	58(1)
C(13)	4257(3)	3097(2)	1815(4)	162(4)
C(14)	4344(3)	3931(4)	1004(4)	102(3)
C(15)	2204(2)	5182(1)	-1104(2)	23(1)
C(16)	1912(2)	5372(1)	-1769(2)	25(1)
C(17)	2219(2)	5185(1)	-2394(2)	29(1)
C(18)	1291(2)	5750(2)	-1757(2)	34(1)
C(19)	712(2)	5429(2)	-1392(2)	44(1)
C(20)	1078(2)	5909(2)	-2544(2)	47(1)
C(21)	1441(3)	6296(2)	-1340(2)	47(1)

Table S3. Bond lengths [Å] and angles [°] for V(3,6-DBSQ)(3,6-DBCat)$_2$.

V(1)-O(1)	1.868(2)
V(1)-O(1)#1	1.868(2)
V(1)-O(2)	1.898(2)
V(1)-O(2)#1	1.898(2)
V(1)-O(3)	1.971(2)
V(1)-O(3)#1	1.971(2)
O(1)-C(1)	1.330(4)
O(2)-C(2)	1.326(4)
O(3)-C(15)	1.298(3)
C(1)-C(6)	1.404(5)
C(1)-C(2)	1.421(5)
C(2)-C(3)	1.406(5)
C(3)-C(4)	1.376(5)
C(3)-C(7)	1.530(5)
C(4)-C(5)	1.399(7)
C(5)-C(6)	1.373(6)
C(6)-C(11)	1.543(6)
C(7)-C(10)	1.533(5)
C(7)-C(9)	1.534(6)
C(7)-C(8)	1.543(6)
C(11)-C(14)	1.514(9)
C(11)-C(12)	1.527(6)
C(11)-C(13)	1.548(6)
C(15)-C(16)	1.415(4)
C(15)-C(15)#1	1.466(6)
C(16)-C(17)	1.364(4)
C(16)-C(18)	1.533(5)
C(17)-C(17)#1	1.428(7)
C(18)-C(21)	1.532(5)

C(18)-C(19)	1.536(5)
C(18)-C(20)	1.539(5)

O(1)-V(1)-O(1)#1	105.38(12)
O(1)-V(1)-O(2)	80.66(9)
O(1)#1-V(1)-O(2)	97.99(9)
O(1)-V(1)-O(2)#1	97.99(9)
O(1)#1-V(1)-O(2)#1	80.66(9)
O(2)-V(1)-O(2)#1	177.80(13)
O(1)-V(1)-O(3)	164.25(9)
O(1)#1-V(1)-O(3)	88.39(8)
O(2)-V(1)-O(3)	90.01(9)
O(2)#1-V(1)-O(3)	91.68(9)
O(1)-V(1)-O(3)#1	88.39(8)
O(1)#1-V(1)-O(3)#1	164.25(9)
O(2)-V(1)-O(3)#1	91.68(9)
O(2)#1-V(1)-O(3)#1	90.01(9)
O(3)-V(1)-O(3)#1	79.16(12)
C(1)-O(1)-V(1)	116.01(19)
C(2)-O(2)-V(1)	115.4(2)
C(15)-O(3)-V(1)	116.67(19)
O(1)-C(1)-C(6)	125.5(3)
O(1)-C(1)-C(2)	112.3(3)
C(6)-C(1)-C(2)	122.2(3)
O(2)-C(2)-C(3)	124.9(3)
O(2)-C(2)-C(1)	112.7(3)
C(3)-C(2)-C(1)	122.4(3)
C(4)-C(3)-C(2)	113.6(4)
C(4)-C(3)-C(7)	124.9(3)
C(2)-C(3)-C(7)	121.5(3)
C(3)-C(4)-C(5)	124.3(4)
C(6)-C(5)-C(4)	123.0(4)
C(5)-C(6)-C(1)	114.4(4)
C(5)-C(6)-C(11)	124.7(4)
C(1)-C(6)-C(11)	120.7(4)
C(3)-C(7)-C(10)	109.7(3)
C(3)-C(7)-C(9)	110.2(3)
C(10)-C(7)-C(9)	109.9(4)
C(3)-C(7)-C(8)	111.3(4)
C(10)-C(7)-C(8)	107.8(4)
C(9)-C(7)-C(8)	107.9(5)
C(14)-C(11)-C(12)	111.7(5)
C(14)-C(11)-C(6)	110.1(4)
C(12)-C(11)-C(6)	109.1(4)
C(14)-C(11)-C(13)	107.8(6)
C(12)-C(11)-C(13)	107.9(4)

C(6)-C(11)-C(13)	110.2(4)
O(3)-C(15)-C(16)	124.9(3)
O(3)-C(15)-C(15)#1	113.75(17)
C(16)-C(15)-C(15)#1	121.39(18)
C(17)-C(16)-C(15)	114.9(3)
C(17)-C(16)-C(18)	124.5(3)
C(15)-C(16)-C(18)	120.5(3)
C(16)-C(17)-C(17)#1	123.66(19)
C(21)-C(18)-C(16)	110.1(3)
C(21)-C(18)-C(19)	110.8(3)
C(16)-C(18)-C(19)	109.1(3)
C(21)-C(18)-C(20)	107.8(3)
C(16)-C(18)-C(20)	110.7(3)
C(19)-C(18)-C(20)	108.4(3)

Symmetry transformations used to generate equivalent atoms#1: -x+1/2,-y+1,z

Table S4. Anisotropic displacement parameters [x 10^3] for V(3,6-DBSQ)(3,6-DBCat)$_2$.

	U11	U22	U33	U23	U13	U12
V(1)	27(1)	26(1)	16(1)	0	0	4(1)
O(1)	25(1)	28(1)	17(1)	-2(1)	0(1)	3(1)
O(2)	26(1)	27(1)	22(1)	2(1)	-4(1)	5(1)
O(3)	31(1)	28(1)	18(1)	0(1)	3(1)	8(1)
C(1)	39(2)	28(2)	21(2)	-3(2)	-7(2)	6(2)
C(2)	33(2)	28(2)	17(2)	0(1)	-3(2)	5(2)
C(3)	61(3)	27(2)	31(2)	-1(2)	-12(2)	-8(2)
C(4)	113(4)	20(2)	69(3)	3(2)	-43(3)	-9(2)
C(5)	122(5)	30(2)	73(3)	0(2)	-69(3)	12(3)
C(6)	69(3)	32(2)	42(2)	-6(2)	-29(2)	16(2)
C(7)	55(3)	40(2)	36(2)	-6(2)	-12(2)	-14(2)
C(8)	112(5)	59(4)	62(4)	-1(3)	-32(4)	-45(4)
C(9)	46(3)	80(4)	41(3)	-12(3)	-4(2)	-28(3)
C(10)	47(3)	40(3)	35(2)	-9(2)	-7(2)	-3(2)
C(11)	76(3)	50(3)	69(3)	-29(2)	-53(3)	35(2)
C(12)	64(4)	64(3)	47(3)	-13(3)	-25(3)	9(3)
C(13)	184(7)	71(4)	229(8)	-65(5)	-179(7)	80(4)
C(14)	44(4)	172(8)	90(5)	-66(5)	-24(3)	54(4)
C(15)	27(2)	20(2)	22(2)	-2(1)	2(2)	-3(1)
C(16)	25(2)	28(2)	22(2)	-1(2)	-1(2)	1(2)
C(17)	28(2)	38(2)	20(2)	5(2)	-6(2)	0(2)

C(18)	36(2)	43(2)	23(2)	0(2)	-3(2)	13(2)
C(19)	32(2)	68(3)	31(2)	2(2)	5(2)	14(2)
C(20)	47(3)	61(3)	31(2)	3(2)	-7(2)	20(3)
C(21)	61(3)	42(2)	37(3)	0(2)	-10(2)	21(2)

Near IR Spectrum

The room temperature near IR region was scanned to look for ligand-ligand charge-transfer bands; none were observed from 300 to 3000 nm, Figure S1.

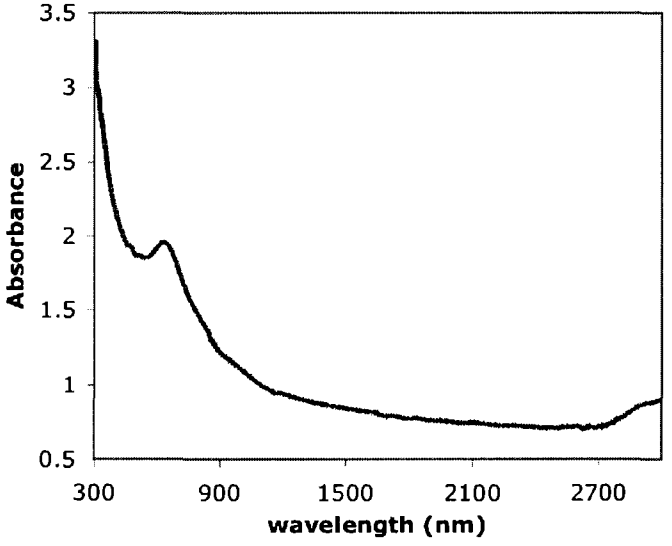


Figure S1. The near IR region of V(3,6-DBSQ)(3,6-DBCat)$_2$ at room temperature. No ligand-ligand charge transfer bands are present from 300 to 3000 nm. The small peak observed near 3000 nm is due to the residual water OH stretching from the KBr.

IR Spectrum

In order to test for a low energy ligand-ligand charge transfer band, the IR region was also scanned at room temperature. Again, no ligand-ligand charge transfer band is

observed from 400 to 4000 cm^{-1} indicating that the charge is localized on a single semiquinone ligand of V(3,6-DBSQ)(3,6-DBCat)$_2$.

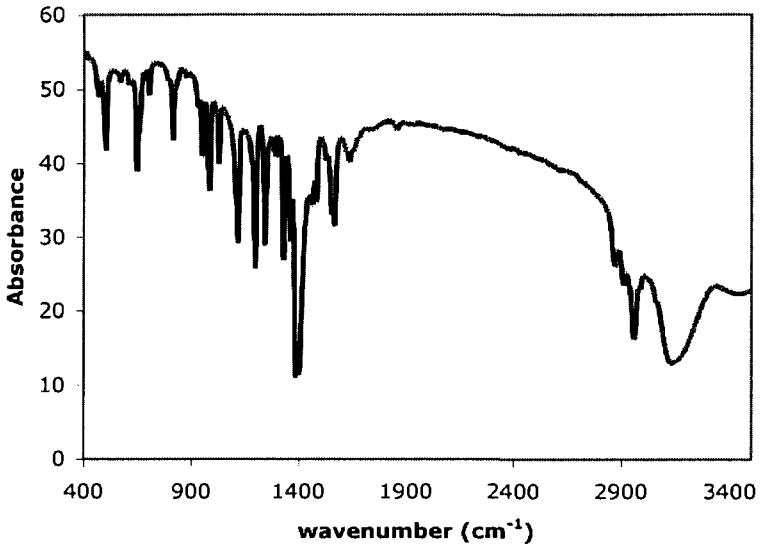


Figure S2. The IR region of V(3,6-DBSQ)(3,6-DBCat)$_2$ at room temperature. The normal C-O, C-C, and C-H stretching bands are present, but no ligand-ligand charge transfer band is observed, at least at room temperature.

Solid-State and Solution UV/Visible Spectra

The solid-state UV/Visible spectrum is shown in Figure S3. Note that the same peaks are present as in the solution spectrum (Figure 2 of the main text and Figure S4), but the peaks around ~300 and 405 nm are not as pronounced in the solid-state spectrum.

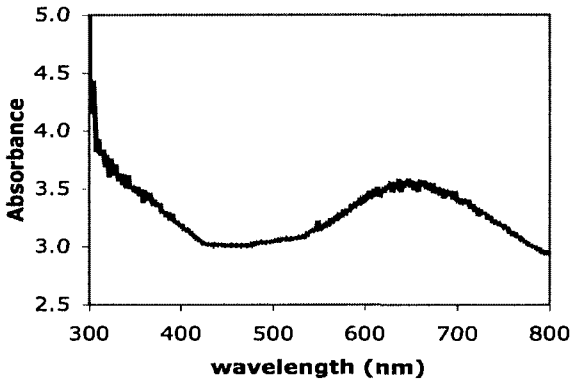


Figure S3. UV/Vis spectrum of V(3,6-DBSQ)(3,6-DBCat)$_2$ prepared as a KBr pellet and at room temperature.

The UV/Vis of V(3,6-DBSQ)(3,6-DBCat)$_2$ was taken in toluene, both under an inert N_2 atmosphere and after bubbling the solution with O_2, all at room temperature. The solution spectra can be seen in Figure S3 below. Note that upon exposure to O_2, the peak at ~670 nm diminishes, but the peaks at ~405 and 300 nm remain. This is analogous to what is observed with [VO(3,5-DBSQ)(3,5-DBCat)]$_2$.[5] That is, the peak at 668 nm diminishes upon exposure to O_2 while the peak at 294 nm remains. The products of this reaction with just O_2 (i.e., and in the absence of additional, non-bound substrate) have not been characterized, in part since they are under non-catalytic conditions, but still may be of interest for further studies.

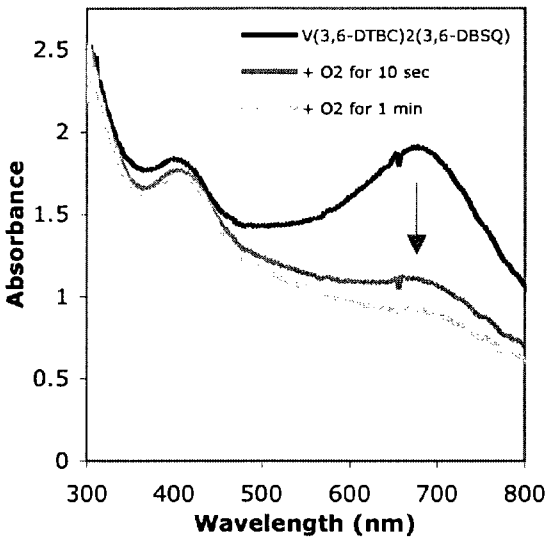


Figure S4. UV/Vis spectra of V(3,6-DBSQ)(3,6-DBCat)$_2$ in toluene under an inert atmosphere and after bubbling with O_2, all at room temperature. The spectra show the disappearance of the 670 nm band and also the retention of the 300 and 405 nm bands with exposure to O_2. The products of this reaction under non-catalytic conditions remains to be identified.

References:

[1] Sheldrick, G. M. *Acta Cryst. A* **2008**, *64*, 112-122.

[2] Sheldrick, G. M. *SADABS* (a program for Siemens Area Detection Absorption Correction), 2000.

[3] Belostotskaya, I. S.; Komissarova, N. L.; Dzhuaryan, E. V.; Ershov, V. V. *Isv. Akad. Nauk SSSR* **1972**, 1594-1596.

[4] Wheeler, D. E.; McCusker, J. K. *Inorg. Chem.* **1998**, *37*, 2296-2307.

[5] Yin, C.-X.; Finke, R. G. *J. Am. Chem. Soc.* **2005**, *127*, 9003-9013.

CHAPTER VIII

DIOXYGENASE CATALYSIS BY d^0 METAL-CATECHOLATE COMPLEXES CONTAINING VANADIUM AND MOLYBDENUM FOR H_2(3,5-DTBC) and H_2(3,6-DTBC) SUBSTRATES

This dissertation chapter contains a *Journal of Molecular Catalysis A* manuscript that is currently *in press*. The chapter presents the dioxygenase catalytic result of the three precatalysts, $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$, $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$, and $[MoO(3,5\text{-DTBC})_2]_2$ (where 3,5-DTBC = 3,5-di-*tert*-butylcatecholate and 3,5-DBSQ = 3,5-di-*tert*-butylsemiquinone) with the substrates H_2(3,5-DTBC) and the relatively little studied substrate, H_2(3,6-DTBC).

The experiments were carried out by Aimee M. Morris, who also prepared the manuscript with editing from Profs. Cortlandt G. Pierpont and Richard G. Finke.

Dioxygenase Catalysis by d^0 Metal-Catecholate Complexes Containing Vanadium and Molybdenum for H_2(3,5-DTBC) and H_2(3,6-DTBC) Substrates

Aimee M. Morris, Cortlandt G. Pierpont, and Richard G. Finke

Abstract

Facile synthetic dioxygenases, catalysts that can split O_2 and place both oxygen atoms selectively into 2 olefins to yield 2 epoxides or into 2 C—H bonds to yield 2 alcohols, remain a "Holy Grail" of oxidation catalysis. Recently, it was shown that [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ (where 3,5-DTBC and 3,5-DBSQ are 3,5-di-tert-butylcatecholate and 3,5-di-tert-butylsemiquinone, respectively) is the cycle resting state of a record catalytic lifetime catechol dioxygenase catalyst (100,000 total catalytic turnovers) for the substrate 3,5-di-tert-butylcatechol, H_2(3,5-DTBC) (Yin, C.-X.; Finke, R. G. *J. Am. Chem. Soc.* **2005**, *127*, 9003-9013). Herein we show that the precatalyst V(3,6-DTBC)$_2$(3,6-DBSQ) also gives dioxygenase products for the substrate H_2(3,5-DTBC), notably the same dioxygenase products in similar yields as seen for [VO(3,5-DTBC)(3,5-DBSQ)]$_2$. EPR studies show that the same g=2.003-2.004 species are present in solution throughout the oxidation reaction for both [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and

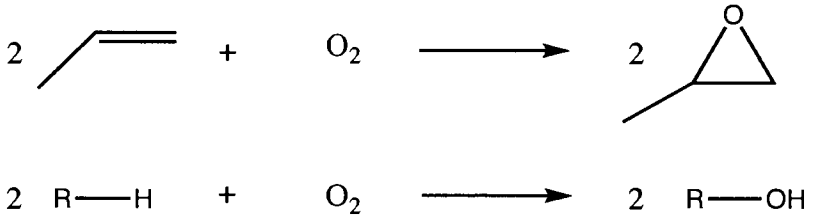
$V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$. The similar products, product yields, and EPR spectra observed suggest that both $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ and $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$ are feeding into the same catalytic cycle. In addition we have expanded the substrates to include $H_2(3,6\text{-DTBC})$, a substrate of interest for its higher symmetry, structurally simplified organic products, and steric hindrance expected to favor the formation of monomeric metal-catecholate complexes. The precatalysts $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ and $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$ were examined with $H_2(3,6\text{-DTBC})$ and found to give the same intradiol and extradiol dioxygenase products in the same yields. EPR studies of $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ and $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$ with $H_2(3,6\text{-DTBC})$ show that the same spectra are observed throughout the oxidation, but at different reaction times. These EPR observations, along with the product studies, suggest that $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ and $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$ have a common mechanistic cycle en route to the $H_2(3,6\text{-DTBC})$ dioxygenase products, albeit a mechanism different than that observed for $H_2(3,5\text{-DTBC})$ based on observed EPR spectra and differing product distributions. We also show that oxidation catalysis with $[MoO(3,5\text{-DTBC})_2]_2$ follows primarily an oxidase path leading to the undesired benzoquinone autoxidation products for both $H_2(3,5\text{-DTBC})$ and $H_2(3,6\text{-DTBC})$ and with reaction times one to two orders of magnitude longer than seen for the vanadium precatalysts. The dramatic difference in product distribution and slower rate for $[MoO(3,5\text{-DTBC})_2]_2$, along with its lack of a semiquinone ligand in at least this precatalyst, suggests the hypothesis that the d^0 vanadium(V) bonded to a semiquinone ligand in $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ and $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$ is a necessary

component of the superior dioxygenase precatalysts able to produced primarly dioxygenase products.

1. Introduction

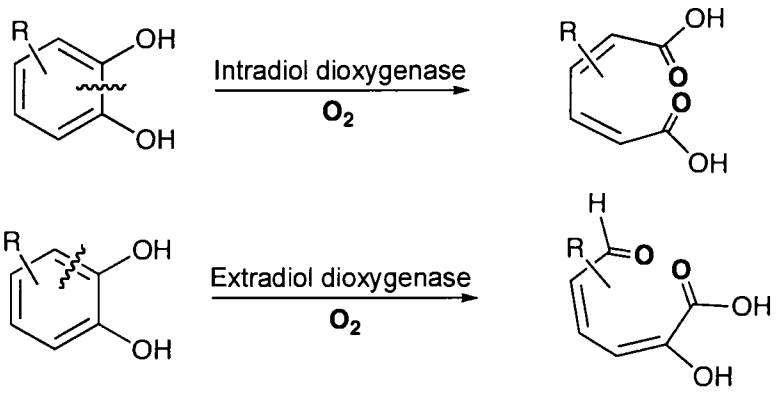
Dioxygenases are an important class of catalysts able, by definiton, to insert both atoms of O_2 into a substrate without the use of protons and electrons and hence, without the production of side products. Dioxygenase enzymes exist in nature but selective and facile synthetic dioxygenases that can do difficult reactions such as those shown in Scheme 1 remain a grand challenge or so-called "Holy Grail" in oxidation catalysis [1].

Scheme 1. Desired "Holy Grail" dioxygenase reactions.



Catechol dioxygenases are a subset of dioxygenases that catalyze the degradation of aromatic compounds via intradiol or extradiol cleavage, Scheme 2. Synthetic catechol dioxygenases have been developed containing Fe(II/III) [2], V(IV/V) [3,4,5,6,7], Ru(II) [8], Rh(III) [9], and other metals [10].

Scheme 2. Intradiol vs. extradiol cleavage by catechol dioxygenases.



In 1999 a V-containing precatalyst was reported to exhibit a record catalytic lifetime of more than 100,000 total turnovers with the substrate 3,5-di-tert-butylcatechol, H_2(3,5-DTBC) [4].[1] It was subsequently determined that 11 different V-containing catechol dioxygenase precatalysts all give a common component or catalytic-cycle resting state of $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$, where 3,5-DTBC = 3,5-di-tert-butylcatecholate (which is the deprotonated form of the catechol H_2(3,5-DTBC)) and 3,5-DBSQ = 3,5-di-tert-butylsemiquinone, Figure 1 [5]. The 11 different V-containing precatalysts were shown to form $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$ under the dioxygenase conditions [5,6] to produce both intradiol and extradiol oxidation products from H_2(3,5-DTBC) [6,7]. These observations have led to interest in whether other related d^0 metal-catecholate complexes might also be able to produce H_2(3,5-DTBC) dioxygenase

[1] Abbreviations: H_2(3,5-DTBC) = 3,5-di-*tert*-butylcatechol; 3,5-DTBC = 3,5-di-*tert*-butylcatecholate, the deprotonated form of H_2(3,5-DTBC); 3,5-DBSQ = 3,5-di-*tert*-butylsemiquinone; H_2(3,6-DTBC) = 3,6-di-*tert*-butylcatechol; 3,6-DTBC = 3,6-di-*tert*-butylcatecholate; 3,6-DBSQ = 3,6-di-*tert*-butylsemiquinone; SQ = semiquinone; spiro product = spiro[1,4-benzodioxin-2(*3H*),2'-[*2H*]-pyran]-3-one,4',6,6',8-tetrakis(1,1-dimethylethyl); rds = rate determining step.

products, what the minimum requirements might be for such catechol dioxygenases, and whether the presence of a semiquinone ligand such as 3,5-DBSQ might be important.

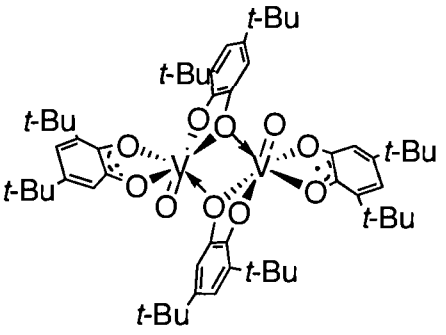


Figure 1. The catalytic-cycle resting state $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$.

We recently reported the synthesis and characterization of a related metal-catecholate complex, $V(3,6\text{-}DTBC)_2(3,6\text{-}DBSQ)$ [11]. The complex $V(3,6\text{-}DTBC)_2(3,6\text{-}DBSQ)$ is similar to $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$ in that it contains catecholate and semiquinone (SQ) ligands in addition to d^0 V(V). As noted in structural studies of complexes containing 3,6-DTBC and 3,6-DBSQ, the presence of *tert*-butyl substituents at ring positions adjacent to donor oxygen atoms blocks bridging to adjacent metals of the type found for $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$ [12]. Such blocking should, in turn, permit studies on the potential importance of the dimeric structure of $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$ on dioxygenase activity. The symmetrical $H_2(3,6\text{-}DTBC)$ should also lead to a less complicated series of organic oxidation products relative to $H_2(3,5\text{-}DTBC)$, but possibly at the expense of the catalytic activity. In fact, in an earlier study using an iron catalyst, Funabiki and co-workers found that an Fe complex containing 3,6-DTBC had

significantly lower oxidation activity than did the same Fe complex with a 3,5-DTBC ligand [11].

An additional d^0 metal-catecholate complex $[MoO(3,5\text{-}DTBC)_2]_2$ was reported in 1979 as the product of O_2 addition to $Mo(3,5\text{-}DTBC)_3$ [13]. $[MoO(3,5\text{-}DTBC)_2]_2$ is of interest since it has a dimeric structure that is identical to $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$. The metal ion (Mo(VI)) is also d^0, but the ligands are now all catecholates, a feature which permits a test of the importance of a semiquinone ligand in at least the precatalyst en route to obtaining dioxygenase products.

Herein we report the catechol dioxygenase activity for the d^0 precatalysts $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$, $V(3,6\text{-}DTBC)_2(3,6\text{-}DBSQ)$, and $[MoO(3,5\text{-}DTBC)_2]_2$ with both $H_2(3,5\text{-}DTBC)$ and $H_2(3,6\text{-}DTBC)$ as substrates. In addition to examining whether $V(3,6\text{-}DTBC)_2(3,6\text{-}DBSQ)$ and $[MoO(3,5\text{-}DTBC)_2]_2$ are able to produce dioxygenase products for both the $H_2(3,5\text{-}DTBC)$ and $H_2(3,6\text{-}DTBC)$ substrates, it is of interest to determine if they function by the dioxygenase mechanism proposed [6,7] for $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$ in which half of this dimeric complex is involved in the actual catalytic cycle.

2. Results and Discussion

2.1. O_2 Uptake by Solutions of $H_2(3,5\text{-}DTBC)$ Plus Precatalysts $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$, $V(3,6\text{-}DTBC)_2(3,6\text{-}DBSQ)$, or $[MoO(3,5\text{-}DTBC)_2]_2$. Dioxygen uptake plots for solutions containing $H_2(3,5\text{-}DTBC)$ and the precatalysts $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$, $V(3,6\text{-}DTBC)_2(3,6\text{-}DBSQ)$, and $[MoO(3,5\text{-}DTBC)_2]_2$ are shown in Figure 2. Experiments were carried out with 1.8 mmol of $H_2(3,5\text{-}DTBC)$, 1-2 μmol of

precatalyst, 8.2 mL of 1,2-dichloroethane, 40 °C, and 0.8 atm of O_2. Figure 2a shows representative O_2 uptake curves for the precatalysts [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and V(3,6-DTBC)$_2$(3,6-DBSQ) and reveals that these two V-precatalysts behave kinetically identical. In Figure 2b, comparison of the O_2 uptake time by the two V-precatalysts to that of [MoO(3,5-DTBC)$_2$]$_2$ reveals that the [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ or V(3,6-DTBC)$_2$(3,6-DBSQ) precatalysts are roughly two orders of magnitude faster than the [MoO(3,5-DTBC)$_2$]$_2$ precatalyst.

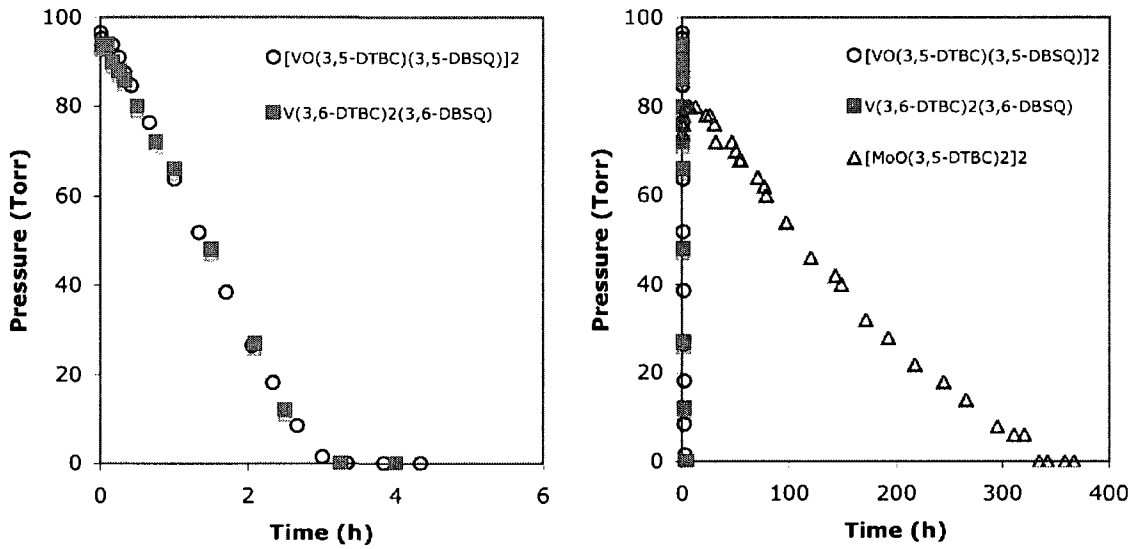
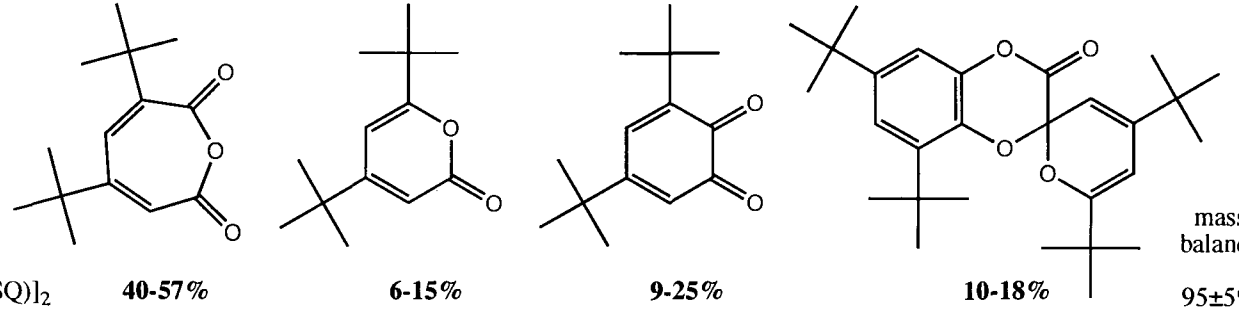


Figure 2. (a) The time region between 0 and 6 hours for the precatalysts [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and V(3,6-DTBC)$_2$(3,6-DBSQ) plus H$_2$(3,5-DTBC). (b) Three separate representative O_2-uptake curves of the substrate H$_2$(3,5-DTBC) plus the precatalyst [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, V(3,6-DTBC)$_2$(3,6-DBSQ), and [MoO(3,5-DTBC)$_2$]$_2$ on the ca. 6.6-fold expanded time scale vs Figure 2a. The greatly, ca. 102-fold faster rate of the [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and V(3,6-DTBC)$_2$(3,6-DBSQ) systems is apparent. The conditions are as follows: 1.8 mmol of substrate, 1-2 μmol of precatalyst, 8.2 mL of 1,2-dichloroethane, 40 °C, and 0.8 atm of O_2. Note that the final pressure in each case has been subtracted from each data point, so that the net pressure loss to a zero final pressure is shown here and in analogous figures hereafter.

2.2. Observed H_2(3,5-DTBC) Dioxygenase Plus Oxidase Products and Their Yields from the O_2-Uptake Reactions. At the end of each O_2-uptake reaction with H_2(3,5-DTBC), the reaction solution was analyzed by gas chromatography (GC) in comparison to authentic samples as described in the Experimental Section. Scheme 3 shows the respective H_2(3,5-DTBC) intradiol and extradiol dioxygenase products plus any benzoquinone oxidase product, along with their yields, observed for the precatalysts [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, V(3,6-DTBC)$_2$(3,6-DBSQ), and [MoO(3,5-DTBC)$_2$]$_2$.

Scheme 3. H_2(3,5-DTBC) dioxygenase products and yields observed for the precatalysts [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, V(3,6-DTBC)$_2$(3,6-DBSQ), or [MoO(3,5-DTBC)$_2$]$_2$ in comparison to calibration by authentic samples. A range of observed yields (minimum and maximum observed yields) from repeat experiments is given. In the case of the precatalyst [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, the range of yields given includes those observed from prior as well as the present work [7] (i.e., and therefore includes the full range of yields anyone repeating this work is likely to encounter, and so long as identical conditions are employed).



					mass balance	reaction time
[VO(3,5-DTBC)(3,5-DBSQ)]$_2$	**40-57%**	**6-15%**	**9-25%**	**10-18%**	95±5%	5±2 hrs
V(3,6-DTBC)$_2$(3,6-DBSQ)	**45-55%**	**10-11%**	**29-43%**	**2-4%**	95±5%	3.5±1.5 hrs
[MoO(3,5-DTBC)$_2$]$_2$	**16%**	**1%**	**56%**	**3%**	76%[a]	~330 hrs

[a] The 76% (i.e., <100%) mass balance is likely a result of product follow-up reactions and/or decomposition during the required long reaction times.

The data show that $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$ and $V(3,6\text{-}DTBC)_2(3,6\text{-}DBSQ)$ give the same dioxygenase products in similar yields for the first two products in Scheme 3, but that $V(3,6\text{-}DTBC)_2(3,6\text{-}DBSQ)$ with its sterically more crowded 3,6-DTBC ligand gives more of the benzoquinone autoxidation product and less of the (somewhat sterically crowded) spiro product, spiro[1,4-benzodioxin-2(*3H*),2'-[*2H*]-pyran]-3-one,4',6,6',8-tetrakis(1,1-dimethylethyl). On the other hand, the molybdenum complex $[MoO(3,5\text{-}DTBC)_2]_2$ exhibits quite different reactivity: in addition to taking 2 orders of magnitude longer to complete the O_2-uptake (Figure 2b), it gives a majority (56%) of the benzoquinone autoxidation product.

The results displayed in Scheme 3 show that $V(3,6\text{-}DTBC)_2(3,6\text{-}DBSQ)$ is also an effective dioxygenase precatalyst for the $H_2(3,5\text{-}DTBC)$ substrate, albeit one that gives higher yields of the undesired benzoquinone autoxidation product and lower yields of the spiro complex. More interesting is that the slow O_2-uptake and primarily non-dioxygenase products seen for $[MoO(3,5\text{-}DTBC)_2]_2$ suggest, in comparison to the results for the V-based precatalysts, that either (i) the presence of the SQ ligand in both $[VO(3,5\text{-}DTBC)(3,5\text{-}DBSQ)]_2$ and $V(3,6\text{-}DTBC)_2(3,6\text{-}DBSQ)$ is an important and necessary component of these precatalysts for their faster catalysis to the observed $H_2(3,5\text{-}DTBC)$ dioxygenase products, or that (ii) the 1st row V rather than 2nd row Mo as the d^0 metal is important for the observed catalysis since reaction rates of catechol oxidations for 2nd row metals are often orders of magnitude slower than for 1st row metals [14]. Regardless of the precise details and explanation here, the observed results demonstrate that the presence of a d^0 vanadium bonded to a semiquinone ligand is a

preferred precatalyst combination (vs d^0 Mo bonded to catecholates) for the formation of high yields of the H_2(3,5-DTBC) dioxygenase products.

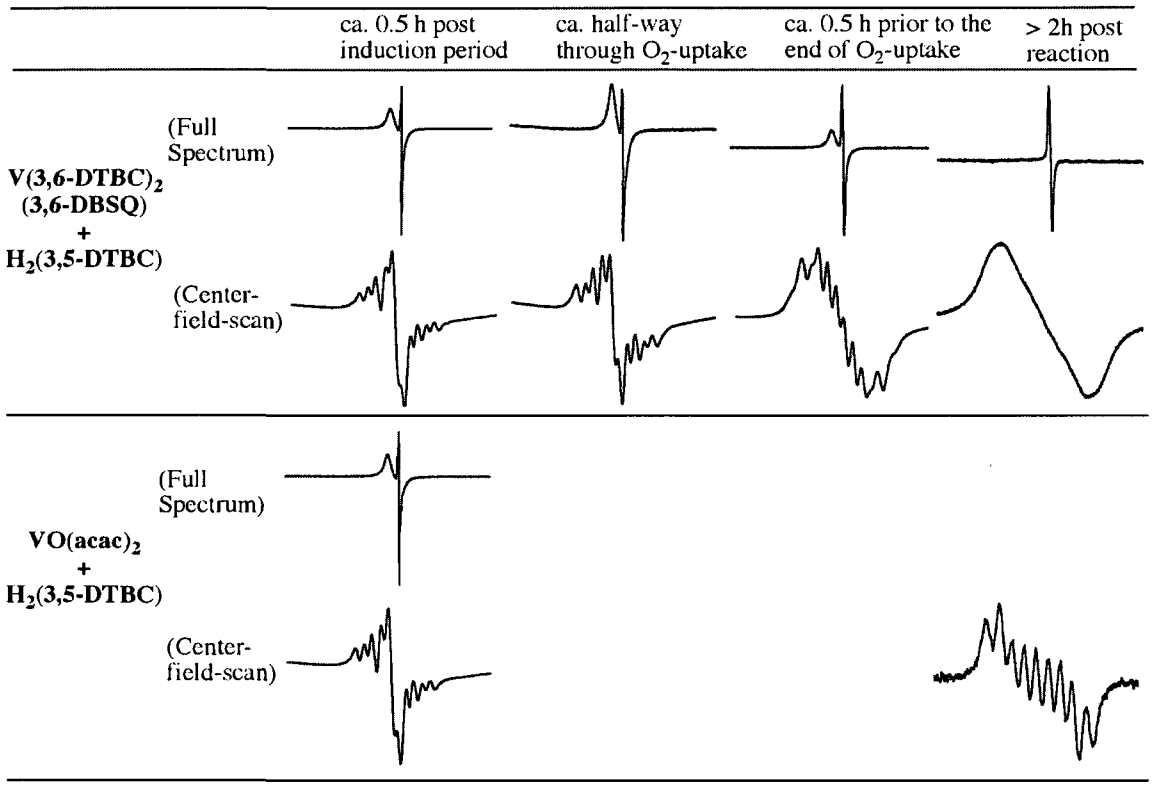
2.3. Time-dependent EPR studies of H_2(3,5-DTBC) plus V(3,6-DTBC)$_2$(3,6-DBSQ). Electron paramagnetic resonance (EPR) spectroscopy was used to follow the paramagnetic species present in solution during the O_2-uptake of H_2(3,5-DTBC) with the V(3,6-DTBC)$_2$(3,6-DBSQ) precatalyst. Table 1 displays changes in the observed EPR spectra throughout the O_2 uptake. Initially, a scan of the full 1000 G spectrum shows a broad signal at g=2.030-2.031 along with a signal at g=2.003-2.004 at its center-field, Table 1. A center-field scan at 50 G reveals a 10-line spectrum (g=2.006, A(^{51}V)=2.1 G) that matches the 10-line spectrum previously observed for the [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ precursor VO(acac)$_2$ plus H_2(3,5-DTBC) [5]. The 10-line spectrum was previously assigned to "V(3,5-DBSQ)$_3$" formed by the addition of H_2(3,5-DTBC) to either $V(CO)_6$ or VO(acac)$_2$ [13]. However, "V(3,5-DBSQ)$_3$" with a parent ion peak at m/z = 712 [13] may have actually been V(3,5-DTBC)$_2$(3,5-DBSQ) based on the recently solved structure of V(3,6-DTBC)$_2$(3,6-DBSQ) [11]. As the O_2-uptake progresses, the broad signal at g=2.030-2.031 becomes more intense and then disappears by the end of the reaction suggesting that this broad signal is associated with an active dioxygenase catalyst—that is, this signal does behave as expected for a kinetically competent intermediate. The line-width of the broad signal suggests that the species is a V(V)-bound radical with unresolved hyperfine coupling.

Furthermore, the center-field signal changes from a 10-line spectrum to a 9-line spectrum. This 9-line spectrum with g=2.004 and A(^{51}V)=3.0 G, characteristic of a single semiquinone ligand coupled to a ^{51}V center, is consistent with the EPR spectrum

due to [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ that is also observed at the end of the O_2-uptake reaction of VO(acac)$_2$ plus H_2(3,5-DTBC) [5], Table 1. This suggests that during the O_2-uptake of V(3,6-DTBC)$_2$(3,6-DBSQ) plus H_2(3,5-DTBC), the previously proposed [6] catalytic-cycle resting state of [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ is formed. This in turn suggests that the catalytic cycle previously seen, in which [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ breaks in half and then reacts with O_2 in a rate-determining step [6], *is also accessed in the V(3,6-DTBC)$_2$(3,6-DBSQ) plus H_2(3,5-DTBC) reaction*, perhaps not unexpectedly due to the presence of a large excess of the H_2(3,5-DTBC) substrate over the V(3,6-DTBC)$_2$(3,6-DBSQ) precatalyst.

Finally, after the O_2-uptake reaction is complete, a one-line spectrum of g=2.003 is observed, Table 1. This one-line spectrum may be a result of the vanadium species may be a result of a monomeric VO(DTBC)(DBSQ) species that has intramolecular electron transfer occurring on the EPR timescale between symmetrically equivalent catecholate ligands, thereby causing the loss of hyperfine coupling as has been observed in Co catecholate complexes [15].

Table 1. Time-dependent EPR spectra of $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$ plus $H_2(3,5\text{-DTBC})$ under oxidation conditions in toluene compared to the EPR spectra previously reported [5] for $VO(acac)_2$ plus $H_2(3,5\text{-DTBC})$ under otherwise identical experimental conditions.[a]

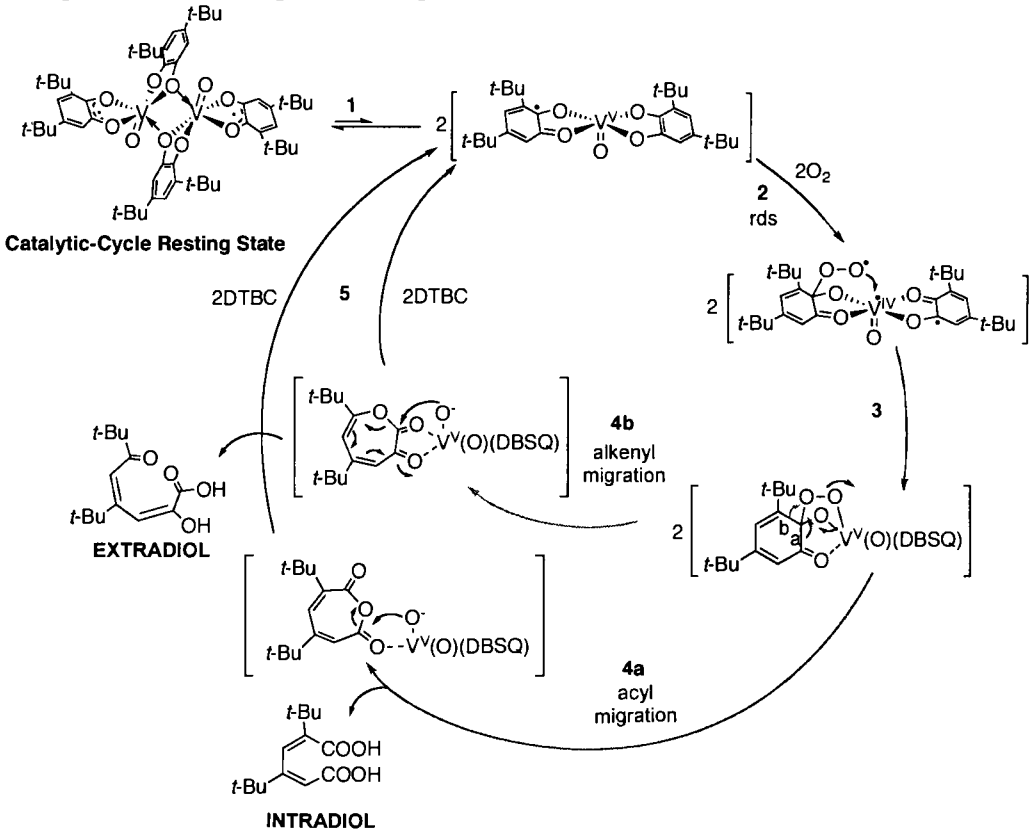


[a] The full spectra were scanned at a magnetic field width of 1000 G while the center-field scans were over 50 G.

The fact that the same 10-line and 9-line spectra are observed during the O_2-uptake of $H_2(3,5\text{-DTBC})$ with either the $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ precursor $VO(acac)_2$ or $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$, along with the same products being observed in similar yields (Scheme 2), strongly suggests that these two precatalysts lead to the same catalyst(s). The time-dependent EPR spectra observed with the $H_2(3,5\text{-DTBC})$ substrate plus $VO(acac)_2$ or $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$ as well as the previous evidence for $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ as the catalyst-cycle resting state [6], are all generally consistent with the previously proposed mechanism [6] and catalytic-cycle resting state

of [VO(3,5-DTBC)(3,5-DBSQ)]$_2$. That said, it is unknown at this time precisely how and where the V(3,6-DTBC)$_2$(3,6-DBSQ) precatalyst feeds into the previously established catalytic cycle, shown in Scheme 4 for [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ plus H$_2$(3,5-DTBC). Note that based on the product yields given as part of Scheme 3, there is at most a modest observed preference ($\leq$ 1.2-4.6) for intradiol over extradiol products, the precise origins of which are not understood [2e].

Scheme 4. The previously supported catalytic cycle for the formation of H$_2$(3,5-DTBC) dioxygenase products starting from [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ [6]. The previously published kinetic and other data revealed that the catalytic-cycle resting state dimer breaks in half and then reacts with O_2 in the rate-determining step (rds) as shown [6]. Steps after the rds (step 2) are, therefore, kinetically hidden and, therefore, had to be written by analogy with prior, seminal work of others [16,17,18] specifically the literature precedent for steps 3 and 4 [19].



2.4. O_2-Uptake Studies of H_2(3,6-DTBC) with [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, V(3,6-DTBC)$_2$(3,6-DBSQ), or [MoO(3,5-DTBC)$_2$]$_2$. We then wondered if the above d^0 metal-catecholate complexes are also dioxygenase catalysts for the more hindered substrate, H_2(3,6-DTBC). Figure 3 shows the O_2-uptake curves obtained for the precatalyst [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, V(3,6-DTBC)$_2$(3,6-DBSQ), or [MoO(3,5-DTBC)$_2$]$_2$ with the H_2(3,6-DTBC) substrate.

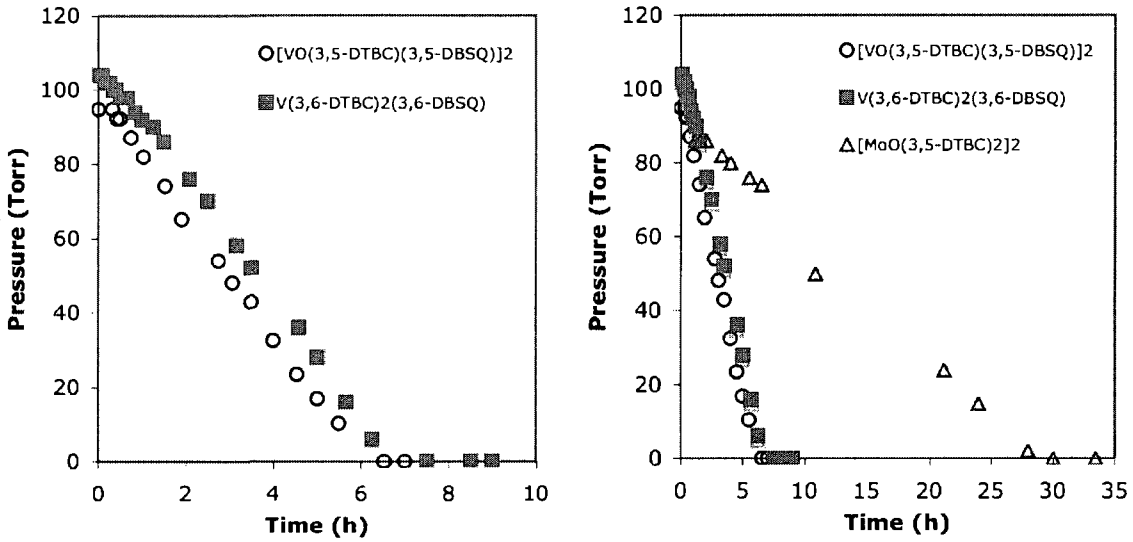


Figure 3. (a) The precatalyst [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ or V(3,6-DTBC)$_2$(3,6-DBSQ) plus H_2(3,6-DTBC) O_2-uptake curves from 0-10 hours. (b) Representative O_2-uptake curves of the substrate H_2(3,6-DTBC) plus the precatalyst [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, V(3,6-DTBC)$_2$(3,6-DBSQ), or [MoO(3,5-DTBC)$_2$]$_2$. The conditions are as follows: 1.8 mmol of substrate, 1-2 μmol of precatalyst, 8.2 mL of 1,2-dichloroethane, 40 °C, and 0.8 atm of O_2. Again, the data reveal that the two V-precatalysts are similar in their kinetics while the Mo-precatalyst is again slower, but now only ca. 5-fold.

Figure 3a illustrates that the precatalysts [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and V(3,6-DTBC)$_2$(3,6-DBSQ) show similar (slightly sigmoidal) O_2-uptake curves with similar net reaction times for the oxidation of H_2(3,6-DTBC). Interesting here is the fact

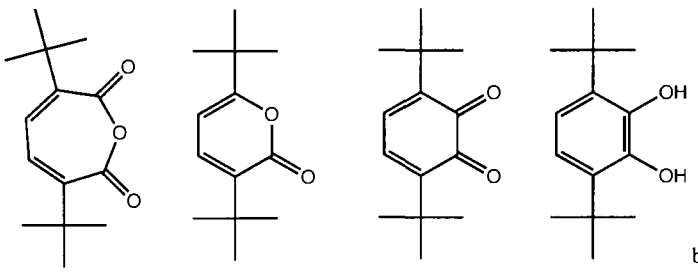
that the oxidation rates for the H_2(3,5-DTBC) and H_2(3,6-DTBC) substrates are similar contrasts with what has been previously observed for an Fe dioxygenase model system containing either a 3,5-DTBC or 3,6-DTBC ligand. The Fe 3,5-DTBC system is very active ($k_{oxidation}$=18 M^{-1} s^{-1}) while the 3,6-DTBC complex displays a ca. 820-fold lower reactivity ($k_{oxidation}$= 0.022 M^{-1} s^{-1}) [20]; hence, the high, but roughly equivalent, reactivity [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and V(3,6-DTBC)$_2$(3,6-DBSQ) with H_2(3,6-DTBC) is an interesting finding. The simplest interpretation of these results is that fairly rapid evolution (e.g., via ligand exchange or other, necessary reactions) of each precatalyst to the same catalyst is occurring. Figure 3b shows that [MoO(3,5-DTBC)$_2$]$_2$ is slower requiring ca. 1/2 an order of magnitude longer than [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ or V(3,6-DTBC)$_2$(3,6-DBSQ) for the oxidation of H_2(3,6-DTBC).

In addition to the H_2(3,6-DTBC) substrate, we briefly looked at simple catechol, but found that with the precatalyst [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ no O_2-uptake under standard conditions was observed even after > 2 days (Figure S1 of the Supporting Information). This result reaffirms the established point [21] that the electron donating *tert*-butyl groups are an important component of the substrate for facile dioxygenase catalysis to be observed, at least with precatalysts such as [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, a significant limitation of these catalyst systems in terms of the other catechols that one would like to oxygenate via dioxygenase reactions.

2.5. Observed H_2(3,6-DTBC) Dioxygenase Products and Yields from O_2-Uptake Reactions with the Precatalysts [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, V(3,6-DTBC)$_2$(3,6-DBSQ), or [MoO(3,5-DTBC)$_2$]$_2$. Similar to the dioxygenase products that were

previously observed with H_2(3,5-DTBC), both intradiol and extradiol products are also observed at the end of the O_2-uptake reaction with H_2(3,6-DTBC)—although the products are simpler as anticipated. The H_2(3,6-DTBC) dioxygenase products observed and quantified by GC at the end of the O_2-uptake reactions, in comparison to authentic standards, are shown in Scheme 5. These products were also separated from one another by column chromatography and characterized by ^1H and ^{13}C NMR, GC-MS, and X-ray diffraction (See Experimental section as well as the Supporting Information). Those characterized, authentic products were then used to calibrate the GC en route to the reported yields.

Scheme 5. H_2(3,6-DTBC) dioxygenase products and product yields. The range of yields shown represents the maximum and minimum yields that were obtained for ≥ 3 separate O_2-uptake reactions.



					mass balance	reaction time
[VO(3,5-DTBC)(3,5-DBSQ)]$_2$	**16-20%**	**69-72%**	**1-2%**	**0%**	92±8%	5±2 hrs
V(3,6-DTBC)$_2$(3,6-DBSQ)	**20-21%**	**73-78%**	**1-2%**	**0%**	95±5%	6.5±1 hrs
[MoO(3,5-DTBC)$_2$]$_2$	**2%**	**37%**	**51%**	**12%**	102%	~30 hrs

Scheme 5 shows that the same products within experimental error are observed for both the [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and V(3,6-DTBC)$_2$(3,6-DBSQ) precatalysts. Interestingly, the H_2(3,6-DTBC) substrate shows that a majority (> 69%) of the extradiol product is obtained (Scheme 5) as opposed to a majority (> 40%) of the intradiol product observed when H_2(3,5-DTBC) is the substrate, Scheme 3, *vide supra*. Furthermore, less

of the benzoquinone autoxidation product is observed with the H_2(3,6-DTBC) substrate (1-2%) as opposed to the H_2(3,5-DTBC) substrate (> 9%), Schemes 3 and 5. Another difference between the H_2(3,5-DTBC) and H_2(3,6-DTBC) substrates is that no spiro product is observed with H_2(3,6-DTBC), as expected if steric hinderance from the para t-butyl groups is inhibiting the formation of an analogous spiro product.

The results in Scheme 5 extend the substrate list for which [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and V(3,6-DTBC)$_2$(3,6-DBSQ) act as dioxygenase precatalysts. The results in Scheme 5 also demonstrate that V(3,6-DTBC)$_2$(3,6-DBSQ) is an equally good dioxygenase for H_2(3,6-DTBC) that produces the same dioxygenase products in the same yields as does [VO(3,5-DTBC)(3,5-DBSQ)]. Scheme 5 further demonstrates that with [MoO(3,5-DTBC)$_2$]$_2$, and similar to the results for H_2(3,5-DTBC), the H_2(3,6-DTBC) substrate gives primarily (ca. 51%) the undesired benzoquinone autoxidation product. Furthermore the [MoO(3,5-DTBC)$_2$]$_2$ precatalyst is slow enough and eventually deactivates so that 12% of the starting H_2(3,6-DTBC) substrate still remains even $\geq$ 2 hours after no further O_2-uptake is observed. Even to reach that point, the O_2-uptake takes half an order of magnitude longer than either of the vanadium catecholate precatalysts.

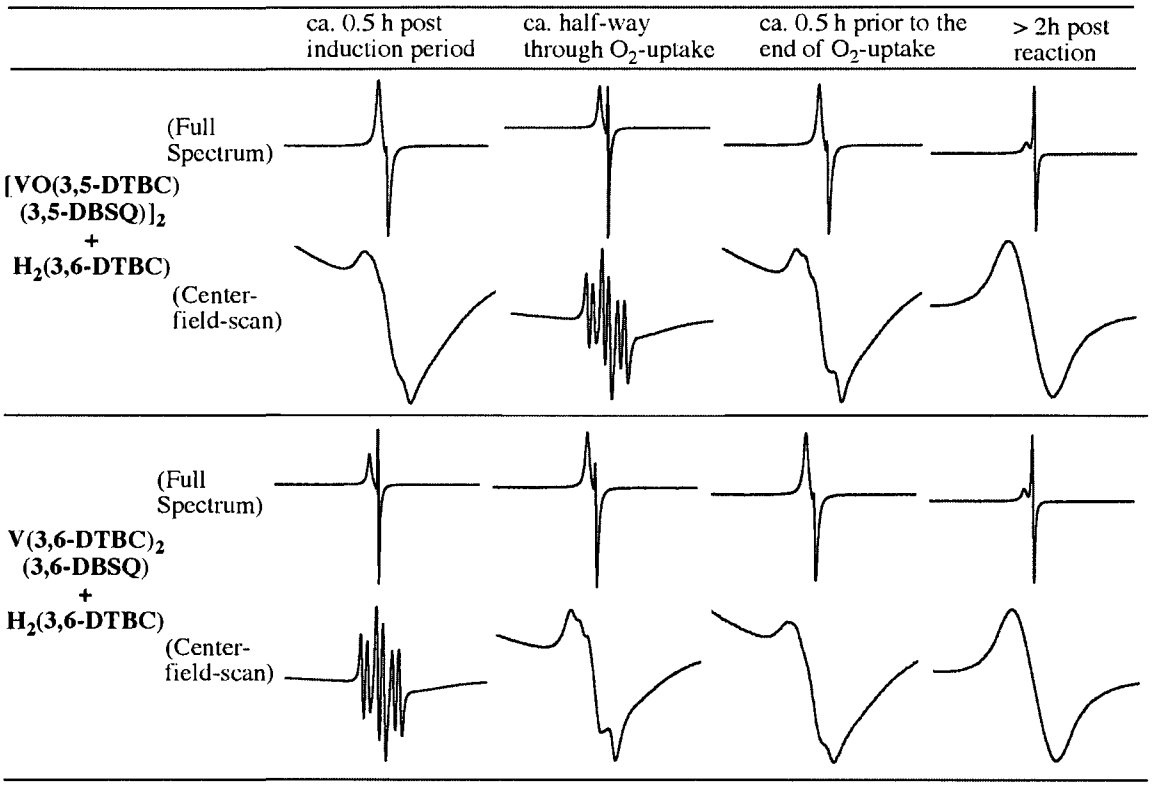
The implication of these V- vs Mo-results is, again, that either (i) a semiquinone ligand is important for the dioxygenase catalysis, or that (ii) d^0 vanadium is superior to d^0 molybdenum; or that (iii) both the d^0 vanadium and a semiquinone ligand are crucial for the observed, faster and more efficacious dioxygenase catalysis.

2.6. Time-dependent EPR Studies of H_2(3,6-DTBC) with [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, or V(3,6-DTBC)$_2$(3,6-DBSQ). EPR spectroscopy was once again used to

follow the species present in solution during the O_2-uptake of H_2(3,6-DTBC) with either the precatalyst [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ (top panel, Table 2) or V(3,6-DTBC)$_2$(3,6-DBSQ) (bottom panel, Table 2). With the [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ precatalyst, an initial scan of the full 1000 G spectrum shows a prominent broad signal at g=2.026-2.027 along with a signal at g=2.003-2.004 at its center-field, Table 2. A center-field scan at 50 G shows a broad signal with a lack of resolvable hyperfine splitting. As the reaction progresses, the broad signal at g=2.026-2.027 begins to diminish in intensity and the center-field signal at g=2.003-2.004 becomes more apparent. A scan of the center-field shows a six-line spectrum that is likely the result of a triplet split by an additional spin=1/2 nucleus. The full-spectrum after further O_2-uptake then displays another more intense signal at g=2.026-2.037 and the center-field scan begins to lose the hyperfine resolution. After the O_2-uptake is complete, the broad signal at g=2.026-2.027 is nearly absent and the prominent signal at g=2.003-2.004 shows no hyperfine splitting in a center-field scan. Interestingly, the same EPR spectra are observed throughout the O_2-uptake reaction with the V(3,6-DTBC)$_2$(3,6-DBSQ) precatalyst but, again, at different times during the reaction. That is, looking at the full spectrum scans, the broad signal at g=2.026-2.027 is also observed but it appears to increase in intensity during the reaction and then diminish after the reaction is complete as opposed to the intensity starting off large, decreasing, increasing again, and finally diminishing as seen for [VO(3,5-DTBC)(3,5-DBSQ)]$_2$. In addition, examination of the center-field scans with V(3,6-DTBC)$_2$(3,6-DBSQ) initially shows the same 6-line spectrum that is observed near the middle of the O_2-uptake reaction with [VO(3,5-DTBC)(3,5-DBSQ)]$_2$. This 6-line spectrum then gradually loses the hyperfine coupling throughout the course of the

reaction and finally the same simple one-line spectrum with g=2.004 as was observed with [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, is observed at the end of the O_2-uptake with V(3,6-DTBC)$_2$(3,6-DBSQ).

Table 2. Time-dependent EPR spectra of the precatalysts [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and V(3,6-DTBC)$_2$(3,6-DBSQ) with the H_2(3,6-DTBC) substrate under standard oxygenation conditions in toluene.[a]



[a] The full spectra were scanned at a magnetic field width of 1000 G while the center-field scans were obtained at 50 G.

The fact that the same EPR spectra are observed, but at different reaction times, for H_2(3,6-DTBC) with both [VO(3,5-DTBC(3,5-DBSQ)]$_2$ and V(3,6-DTBC)$_2$(3,6-DBSQ) suggests that these two precatalysts are feeding into the same catalytic cycle, at least in the simplest, "Ockham's razor" interpretation of the data that we observe. We hypothesize that the 6-line spectrum that is observed with the H_2(3,6-DTBC) substrate is

a result of the formation of H(3,6-DBSQ) which arises from protonation of one SQ by protons derived from H_2(3,6-DTBC). The triplet arises from coupling to the two equivalent protons at the 4 and 5 ring positions (A= 3.9 G). Weaker coupling to the proton bound to the oxygen (A= 1.6 G) gives the 6 lines. The g-value and coupling constants observed for the 6-line spectrum match those previously observed for free H(3,6-DBSQ) [22].

The other center-field spectra that are observed throughout the O_2-uptake of H_2(3,6-DTBC) look to be a result of the loss of hyperfine coupling from the 6-line spectrum. The asymmetry of the signal suggests overlap of more than one signal. Again as we suggested for the H_2(3,5-DTBC) substrate, the broad one-line signal observed after the O_2-uptake is complete for H_2(3,6-DTBC) plus [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ or V(3,6-DTBC)$_2$(3,6-DBSQ) may be a result of a monomeric VO(3,6-DTBC)(3,6-DBSQ) species that has intramolecular electron transfer between the symmetrically equivalent 3,6-DTBC and 3,6-DBSQ ligands causing the hyperfine signals to be lost.

Overall, the EPR spectra shown in Table 2 along with the observed products and yields (Scheme 5) suggest that both the precatalysts [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and V(3,6-DTBC)$_2$(3,6-DBSQ) are feeding into the same dioxygenase catalytic cycle. It also appears by the difference in EPR spectra, as well as the difference in intradiol vs. extradiol products observed, that the H_2(3,6-DTBC) substrate may be operating under a different mechanism in comparison to the H_2(3,5-DTBC) substrate (Scheme 4) for the formation of dioxygenase products. However, further kinetic, spectroscopic and other studies probing the mechanism of H_2(3,6-DTBC) with [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and

$V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$ precatalysts will be needed to gain deeper insights into the mechanism underlying the $H_2(3,6\text{-DTBC})$ oxidation pathway(s).

3. Conclusions

The main findings from this study can be summarized as follows:

- Product studies reveal that, as with $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ before [5,6,7], $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$ is an effective dioxygenase precatalyst for the substrate $H_2(3,5\text{-DTBC})$. The precatalyst $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$ generally gives two primary dioxygenase products in similar yields, but also gives a higher yield of the undesired benzoquinone autoxidation product and a lower yield of the spiro complex in comparison to beginning with $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$.
- Time-dependent EPR studies show that similar species are observed in solution throughout the O_2-uptake using either $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ or $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$ plus $H_2(3,5\text{-DTBC})$. Specifically, a 10-line spectrum corresponding to "$V(3,5\text{-DBSQ})_3$" [13], as well as a 9-line spectrum corresponding to $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$, are both observed regardless of which precatalyst is used. The EPR spectra are observed at different times during the reaction, however, results that require different kinetics en route to the observed species.
- The EPR studies along with the observed similar products and yields suggest that $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ and $V(3,6\text{-DTBC})_2(3,6\text{-DBSQ})$ plus $H_2(3,5\text{-}$

DTBC) are operating predominantly by the same, mostly likely previously established [6] mechanism.

- We have also expanded the list of substrates to include H_2(3,6-DTBC) as a substrate for which both [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and V(3,6-DTBC)$_2$(3,6-DBSQ) function as effective precatalysts en route to intradiol and extradiol dioxygenase products. The use of the H_2(3,6-DTBC) substrate gives much less (1-2% vs. >9%) of the benzoquinone autoxidation product in comparison to H_2(3,5-DTBC), suggesting that the autoxidation pathway is significantly slowed relative to the dioxygenase pathway by using this sterically more hindered substrate.
- The choice of H_2(3,5-DTBC) vs. H_2(3,6-DTBC) substrate dictates whether the major observed dioxygenase product is a result of intradiol or extradiol cleavage: the H_2(3,5-DTBC) substrate gives > 40% of an intradiol dioxygenase product while the H_2(3,6-DTBC) substrate results in > 69% of an extradiol dioxygenase product with either [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ or V(3,6-DTBC)$_2$(3,6-DBSQ). In addition, no spiro product is observed with the H_2(3,6-DTBC) substrate presumably due to steric hinderance of the para-oriented (vs. meta-oriented with the H_2(3,5-DTBC) substrate) *tert*-butyl groups.
- EPR studies of H_2(3,6-DTBC) plus [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ or V(3,6-DTBC)$_2$(3,6-DBSQ) show that the same EPR spectra are observed throughout the reaction, but they occur at different reaction times depending on which precatalyst is used. The six-line hyperfine spectrum observed with both [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and V(3,6-DTBC)$_2$(3,6-DBSQ) is hypothesized to result

from the formation of free semiquinone radical. The different EPR spectra and different major products in comparison to those observed with the H_2(3,5-DTBC) substrate, suggests that H_2(3,6-DTBC) is oxygenated by a different mechanism than seen for H_2(3,5-DTBC).

- The ca. one to two orders of magnitude slower dioxygenase activity of [MoO(3,5-DTBC)$_2$]$_2$, and much lower yields of the desired dioxygenase products plus high yields of undesired benzoquinone autoxidation product, reveal that a d^0 vanadium complex containing a SQ ligand is important for the best dioxygenase catalysis that at least we have observed from the precatalysts tested.

4. Experimental

4.1. Reagents. H_2(3,5-DTBC) (Aldrich, 99%) was recrystallized three times from n-pentane under argon and stored in a Vacuum Atmosphere drybox ($O_2 \leq 5$ppm). H_2(3,6-DTBC) was synthesized and recrystallized according to the modified literature procedure[11] and stored in the drybox. The solvent 1,2-dichloroethane (Aldrich, HPLC grade) was dried over preactivated 4 Å molecular sieves and stored in the drybox. VO(acac)$_2$ (Aldrich, 95%) and Mo(CO)$_6$ (Aldrich, 99.9+%) were used as received and stored in the drybox. The following were purchased where indicated and used as received: catechol (Aldrich, $\geq$ 99%), chloroform (Fisher, ACS grade), hexanes (Fisher, ACS grade), methanol (Fisher, HPLC grade), and n-pentane (Fisher, pesticide grade). Argon (99.985%) and oxygen (99.5%) gases were purchased from General Air and used as received.

4.2. Instrumentation. ^{1}H NMR were recorded in 5-mm o.d. tubes on a Varian Inova (JS-300) NMR spectrometer and referenced to the residual proton impurity in the deuterated solvent. EPR spectra were recorded on a Bruker EMX 200U EPR spectrometer using quartz 4-mm o.d. tubes and referenced to 2,2-diphenyl-1-picrylhydrazyl (DPPH, g=2.0037 g). GC analyses were performed on a Hewlett-Packard (HP) 5890 Series II gas chromatograph equipped with an FID detector and an SPB-1 capillary column (30 m, 0.25 mm i.d.) with the following temperature program: initial temperature, 200 °C (initial time, 2 min); heating rate, 2 °C/min; final temperature, 240 °C (final time, 3 min); FID detector temperature, 250 °C; injector temperature, 250 °C. An injection volume of 1 μL was used. Electrospray ionization mass spectrometry analyses were performed on either a Thermo Finnigan LCQ Advantage Duo mass spectrometer or an Agilent model 6220 TOF mass spectrometer. Elemental analyses were performed by Galbraith Laboratories, Inc (Knoxville, TN).

4.3. Preparation of [VO(3,5-DTBC)(3,5-DBSQ)]$_2$. The precatalyst [VO(DTBC)(DBSQ)]$_2$ was synthesized according to the published method [5,19] and characterized by EPR and UV/Vis. The synthesized product gave the characteristic 9-line EPR spectrum with a g-value of 2.006 in accord with the literature [19]. The UV/Vis under an inert atmosphere in toluene shows two absorbances at 298 and 660 nm in comparison to 294 and 668 nm reported in the literature [5]. Elemental analysis: calculated for [VO(DTBC)(DBSQ)]$_2$•CH_3OH [found]: C, 65.38 [65.26]; H, 8.09 [7.83]; V 9.73 [9.92].

4.4. Preparation of V(3,6-DTBC)$_2$(3,6-DBSQ). V(3,6-DTBC)$_2$(3,6-DBSQ) was synthesized according to a recently published method [11] and characterized by UV/Vis,

EPR, single crystal X-ray diffraction, and elemental analysis. UV/Vis under an inert atmosphere in toluene shows an intense signal at 670 nm (ε=33,000-38,000 M^{-1} cm^{-1}). An EPR spectrum also in toluene and under an inert atmosphere shows a 9-line spectrum with g=2.0058 and A(^{51}V)= 3.47 G and A(^{1}H)= 4.53 G. Elemental analysis yields: calculated for V(3,6-DTBC)$_2$(3,6-DBSQ)•CH_3OH [found]: C, 69.43 [69.91]; H, 8.67 [8.90]; V, 6.85 [6.58].

4.5. Preparation of [MoO(3,5-DTBC)$_2$]$_2$. This precatalyst was synthesized according to the literature procedure [13]. In addition, the crude product was purified by recrystallization from hot n-pentane (recrystallization yield 40%) to yield a >95% pure product (by ^{1}H NMR). Melting point: 301 to 302 °C. ^{1}H NMR ($CDCl_3$) observed: 0.95, 1.30, 1.32, 1.55, 6.40, 6.77, 7.04, and 7.20 ppm; previously reported [13]: 0.95, 1.27, 1.28, 1.57, 6.81, 6.83, 7.08, and 7.11 ppm. Elemental analysis calculated for [MoO(3,5-DTBC)$_2$]$_2$•1.5($C_6H_5CH_3$) [found]: C, 64.24 [64.03]; H, 7.46 [7.84]; Mo 15.4 [15.3].

4.6. Standard Conditions O_2-Uptake Experiments. These experiments were carried out as detailed elsewhere on a volume-calibrated oxygen-uptake line [4]. In a Vacuum Atmosphere drybox, ca. 1.8 mmol of H_2(3,5-DTBC) or H_2(3,6-DTBC) substrate were added to a 50 mL side-arm, round-bottomed, schlenk flask equipped with a septum and an egg-shaped 3/8-in × 3/16-in Teflon-coated magnetic stir bar. Using a 10 mL gas-tight syringe, 8 mL of predried 1,2-dichloroethane was added to the flask. A vacuum adapter was attached and sealed with a Teflon stopcock, and the flask was brought out of the drybox. The flask was then connected to the oxygen uptake line via an O-ring joint, and the reaction solution was frozen in a dry-ice/ethanol bath (-76 °C) for 15 minutes. Two freeze-pump-thaw-fill cycles were performed using O_2 as the filling gas. The dry-

ice/ethanol bath was then replaced with a temperature-controlled oil bath. The flask was then brought up to 40.0 ± 0.7 °C and allowed to equilibrate under stirring for 30 minutes. (Note: the catechol substrates, while initially insoluble, dissolve upon heating and stirring.) In the drybox, ca. 1-2 μmol of the selected precatalyst was weighed into a 5 mL glass vial and dissolved in 0.2 mL of 1,2-dichloroethane. The catalyst solution was then mixed by repeatedly pulling the solution into and expelling from a 1 mL gas-tight syringe and then drawn into the syringe. The syringe was then inserted into a septum-capped 5 mL glass vial to protect the solution from air and brought out of the drybox. The catalyst solution was then injected into reaction flask through the sidearm and $t = 0$ was set. Pressure readings from the manometer were used to follow the reactions. Reactions were stopped when no further pressure loss was observed for ≥ 1 hour.

4.7. Scaled-Up O_2-Uptake Experiments. The same procedure as detailed in the "Standard Conditions O_2-Uptake Experiments" was employed except that ca. 1 g (4.5 mmol) of substrate and 5-6 μmol of precatalyst is used in order to scale-up the amount of oxygenated product obtained.

4.8. H_2(3,5-DTBC) Product Identification. Product identification using GC was determined by co-injection of authentic materials [4]. In addition, mass spectrometry (MS) was employed to verify the formation of the products shown in Scheme 3.

4.9. H_2(3,6-DTBC) Product Separation and Identification. Following the "Scale-Up O_2-Uptake Experiments", the main three H_2(3,6-DTBC) oxygenated products were separated in an analogous manner to the previously detailed separation of H_2(3,5-DTBC) products [4]. For details of this separation see the Supporting Information. The H_2(3,6-DTBC) oxygenated products were then characterized by ^1H and ^{13}C NMR, MS, and in

the case of 3,6-di-*tert*-butyl-1-oxacyclohepta-3,5-diene-2,7-dione by single crystal X-ray diffraction (results are given in the Supporting Information). The three main products making up ~90% of the mass balance are presented below along with their characterization.

4.9.1. 3,6-Di-*tert*-butyl-1-oxacyclohepta-3,5-diene-2,7-dione. GC-MS found for $C_{14}H_{20}O_3$, *m/z* 236 (M^+); ^1H NMR ($CDCl_3$): δ 1.25 (s, 18H, t-butyl) 6.40 (s, 2H, ring); ^{13}C NMR ($CDCl_3$): δ 29.4 (s, primary t-butyl), 36.3 (s, quaternary t-butyl) 126.0 (s, C4 and C5), 146.7 (s, C3 and C6), 161.6 (s, C2 and C7). Single crystals suitable for X-ray diffraction were obtained from hot pentane. The crystallographic data can be found in the Supporting Information.

4.9.2. 3,6-Di-*tert*-butyl-*2H*-pyran-2-one. GC-MS found for $C_{13}H_{20}O_2$, *m/z* 208 (M^+); ^1H NMR ($CDCl_3$): δ 1.25 (d, 18H, t-butyl) 5.94 (d, 1H, hydroxyl) 7.11 (d, 1H, ring); ^{13}C NMR ($CDCl_3$): δ 28.4 (s, primary t-butyl), 28.6 (s, primary t-butyl), 34.5 (s, quaternary t-butyl), 35.8 (s, quaternary t-butyl), 99.3 (s, C5), 133.5 (s, C4) 137.2 (s, C3), 161.0 (s, C2), 170.8 (s, C6).

4.9.3. 3,6-Di-*tert*-butyl-1,2-benzoquinone. GC-MS found for $C_{14}H_{20}O_2$, *m/z* 221 ($M+H^+$); ^1H NMR ($CDCl_3$): δ 1.25 (s, 18H, t-butyl) 6.78 (s, 2H, ring); ^{13}C NMR ($CDCl_3$): δ 29.4 (s, primary t-butyl), 32.5 (s, quaternary t-butyl), 134.2 (s, C4 and C5), 149.8 (s, C3 and C6), 181.4 (s, C1 and C2).

4.10. EPR during and following O_2-Uptake Experiments. The same procedure as detailed in the "Standard Conditions O_2-Uptake" section was employed, except that: (i) ca. 5-8 μmol of the V(3,6-DTBC)$_2$(3,6-DBSQ) or [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ precatalyst was used, (ii) toluene was used as the solvent instead of 1,2-dichloroethane,

with the toluene being frozen with liquid nitrogen, and (iii) the precatalyst was added to the H_2(DTBC) solution inside the drybox instead of by injection to the pre-stirred solution. The O_2 pressue loss was monitored by a mercury manometer [4]. Aliquots of reaction solution (0.3 mL) were drawn through the septum covered side-arm of the reaction flask at ca. 0.5 hour after the pressure began to decrease and then ca. at every 10 torr of pressure loss. The solution was also sampled again while still under the O_2 atmosphere through the reaction flask side-arm after no pressure loss was observed for > 2 hours.

Acknowledgement

These studies were supported by NSF grant CHE 9531110 to R. G. F.

Supporting Information Available

O_2-uptake reaction and discussion of [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ plus catechol, including a control reaction involving the subsequent addition of H_2(3,5-DTBC); Column separation scheme for the isolation of H_2(3,6-DTBC) dioxygenase products; Crystallization conditions, X-ray diffraction experimental section, thermal ellipsoid plot, and crystallographic tables of 3,6-di-*tert*-butyl-1-oxacyclohepta-3,5-diene-2,7-dione; Reaction Schemes for the formation of intradiol and extradiol H_2(3,6-DTBC) dioxygenase products; ^{1}H and ^{13}C NMR assignments for the 3 main observed H_2(3,6-DTBC) products.

References:

[1] C.L. Hill, I.A. Weinstock, Nature 388 (1997) 332-333.

[2] (a) L. Que Jr., R.Y.N. Ho, Chem. Rev. 96 (1996) 2607-2624. (b) T. Funabiki, in: T. Funabiki (Ed.), Catalysis by Metal Complexes, Kluwer Academic Publishers Dordrecht, The Netherlands, 1997, Vol. 19, pp 105-155. (c) H.-J. Krüger, in: B. Meunier (Ed.), Biomimetic Oxidations Catalyzed by Transition Metal Complexes, Imperial College Press London, 2000, pp 363-413. (d) R. Yamahara, S. Ogo, H. Masuda, Y. Watanabe, J. Inorg. Biochem. 88 (2002) 284-294. (e) M. Costas, M.P. Mehn, M.P. Jensen, L. Que Jr., Chem. Rev. 104 (2004) 939-986. (f) H.G. Jang, D.D. Cox, L. Que Jr., J. Am. Chem. Soc. 113 (1991) 9200-9204. (g) A. Dei, D. Gatteschi, L. Pardi, Inorg. Chem. 32 (1993) 1389-1395. (h) M. Ito, L. Que Jr., Angew. Chem. Int. Ed. 36 (1997) 1342-1344. (i) D.-H. Jo, L. Que Jr., Angew. Chem. Int. Ed. 39 (2000) 4284-4287. (j) M. Pascaly, M. Duda, F. Scheweppe, K. Zurlinden, F.K. Müller, B. Krebs, J. Chem. Soc. Dalton Trans. (2001) 828-837. (k) C.-H. Wang, J.-W. Lun, H.-H. Wei, M. Takeda, Inorg. Chim. Acta 360 (2007) 2944-2952. (l) R. Mayilmurugan, H. Stoeckli-Evans, M. Palaniandacar, Inorg. Chem. 47 (2008) 6645-6658.

[3] (a) Y. Tatsuno, M. Tatsuda, S. Otsuka, J. Chem. Soc., Chem. Commun. (1982) 1100-1101. (b) Y. Tatsuno, C. Nakamura, T. Saito, J. Mol. Catal. 42 (1987) 57-66. (c) Y. Tatsuno, M. Tatsuda, S. Otsuka, K. Tani, Chem. Lett. (1984) 1209-1212. (d) U. Casellato, S. Tamburini, P.A. Vigato, M. Vidali, D.E. Fenton, Inorg. Chim. Acta 84 (1984) 101-104. (e) E. Roman, F. Tapia, M. Barrera, M.T. Garland, J.Y. Le Marouille, C. Giannotti, J. Organomet. Chem. 297 (1985) C8-C12. (f) B. Galeffi, M. Postel, A. Grand, P. Rey, Inorg. Chim. Acta 129 (1987) 1-5. (g) B. Galeffi, M. Postel, A. Grand, P. Rey, Inorg. Chim. Acta 160 (1989) 87-91. (h) Y. Nishida, H.Z. Kikuchi, Naturforsch., B: Chem. Sci. 44 (1989) 245-247. (i) U. Russo, B. Zarli, P. Zanonato, M. Vidali, Polyhedron 10 (1991) 1353-1361. (j) B.Y. Zhang, Y. Zhang, B.W. Chen, K. Wang, Chin. Chem. Lett. 8 (1997) 547-550.

[4] H. Weiner, R.G. Finke, J. Am. Chem. Soc. 121 (1999) 9831-9842.

[5] C.-X. Yin, R.G. Finke, J. Am. Chem. Soc. 127 (2005) 9003-9013.

[6] C.-X. Yin, R.G. Finke, J. Am. Chem. Soc. 127 (2005) 13988-13996.

[7] C.-X. Yin, Y. Sasaki, R.G. Finke, Inorg. Chem. 44 (2005) 8521-8530.

[8] M. Matsumoto, K. Kuroda, J. Am. Chem. Soc. 104 (1982) 1433-1434.

[9] (a) C. Bianchini, P. Frediani, F. Laschi, A. Meli, F. Vizza, P. Zanello, Inorg. Chem. 29 (1990) 3402-3409. (b) A. Vlcek Jr., Chemtracts: Inorg. Chem. 3 (1991) 275-280.

[10] A. Nishinaga, in: Catalysis by Metal Complexes, Kluwer Academic Publishers Dordrecht, The Netherlands, 1997, Vol. 19, pp 157-194.

[11] A.M. Morris, C.G. Pierpont, R.G. Finke, Inorg. Chem. 48 (2009) 3496-3498.

[12] (a) C.W. Lange, B.J. Conklin, C.G. Pierpont, Inorg. Chem. 33 (1994) 1276-1283. (b) C.-M. Liu, P. Restorp, E. Nordlander, C.G. Pierpont, Chem. Commun. (2001) 2686-2687. (c) C.-M. Liu, E. Nordlander, D. Schmeh, R. Shoemaker, C.G. Pierpont, Inorg. Chem. 43 (2004) 2114-2124.

[13] R.M. Buchanan, C.G. Pierpont, Inorg. Chem. 18 (1979) 1616-1620.

[14] C.G. Pierpont, C.W. Lange, Prog. Inorg. Chem. 41 (1994) 331-442.

[15] R.M. Buchanan, C.G. Pierpont, J. Am. Chem. Soc. 102 (1980) 4951-4957.

[16] H.G. Jang, D.D. Cox, L. Que Jr., J. Am. Chem. Soc. 113 (1991) 9200-9204.

[17] (a) C.J. Winfield, Z. Al-Mahrizy, M. Gravestock, T.D.H. Bugg, Perkin 1 (2000) 3277-3289. (b) T.D.H. Bugg, G. Lin, Chem. Commun. (2001) 941-952.

[18] P. Barbaro, C. Bianchini, K. Linn, C. Mealli, A. Meli, F. Vizza, Inorg. Chim. Acta 198-200 (1992) 31-56.

[19] M.E. Cass, D.L. Green, R.M. Buchanan, C.G. Pierpont, J. Am. Chem. Soc. 105 (1983) 2680-2686.

[20] Y. Hitomi, M. Yoshida, M. Higuchi, H. Minami, T. Tanaka, T. Funabiki, J. Inorg. Biochem. 99 (2005) 755-763.

[21] M. Costas, M.P. Mehn, M.P. Jensen, L. Que Jr., Chem. Rev. 104 (2004) 939-986.

[22] V.B. Vol'eva, A.I. Prokof'ev, A.Y. Karmilov, N.L. Komissarova, I.S. Belostotskaya, T.I. Prokof'eva, V. V. Ershov, Russ. Chem. Bull. 47 (1998) 1920-1923.

Supporting Information for:

Dioxygenase Catalysis by d^0 Metal-Catecholate Complexes Containing Vanadium and Molybdenum for H_2(3,5-DTBC) and H_2(3,6-DTBC) Substrates

Aimee M. Morris, Cortlandt G. Pierpont, and Richard G. Finke

Attempted O_2-Uptake Reaction of [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ Plus Catechol. An oxygenation experiment was attempted using simple catechol as the substrate, under the standard conditions outlined in the main text, with the precatalyst [VO(3,5-DTBC)(3,5-DBSQ)]$_2$. No uptake of O_2 was observed even after 2 days. As a control, at the end of those 2 days, we then injected a solution containing the H_2(3,5-DTBC) substrate to make sure that O_2-uptake could in fact be observed and that the precatalyst was not defective. Figure S1 shows this reaction beginning with just the catechol substrate where no O_2-uptake is observed, but that the injection of H_2(3,5-DTBC) into the reaction solution is quickly followed by active uptake of O_2. Product verification after O_2-uptake was complete (ca. 150 hrs) by GC showed only H_2(3,5-DTBC) dioxygenase products plus the starting, unreacted catechol.

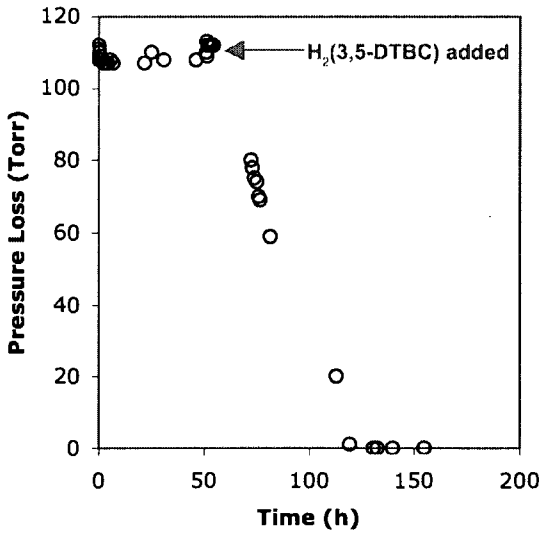


Figure S1. O_2-uptake graph for [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ plus catechol. After no O_2-uptake was observed even after 2 days, 403.2 mg (1.81 mmol) of H_2(3,5-DTBC) dissolved in 1.3 mL of 1,2-dichloroethane was injected into the reaction flask; O_2-uptake was then observed as shown above. Reaction conditions: 203.7 mg (1.85 mmol) of catechol, 5.9 mg (6 μmol) of [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, 8.2 mL of 1,2-dichloroethane, 40 °C, and ca. 1 atmosphere of O_2.

Separation of H_2(3,6-DTBC) Dioxygenase Products from a Resultant O_2-Uptake Reaction with [VO(3,5-DTBC)(3,5-DBSQ)]$_2$. Following the "Scaled-up O_2-Uptake" procedure outlined in the main text, the products were separated by column chromatography as follows: column size, 450 x 30 mm; filled with~100 g of 60 Å silica gel (Aldrich)/hexanes slurry; eluent, ~1000 mL of $CHCl_3$; elution rate, ~1.5 mL/min. Each fraction was then run through the GC column and temperature program outlined in the main text to determine which fractions contained the desired products. A schematic of the column separation, one analogous to that employed for separating H_2(3,5-DTBC) oxygenated products [1], is shown below in Figure S2.

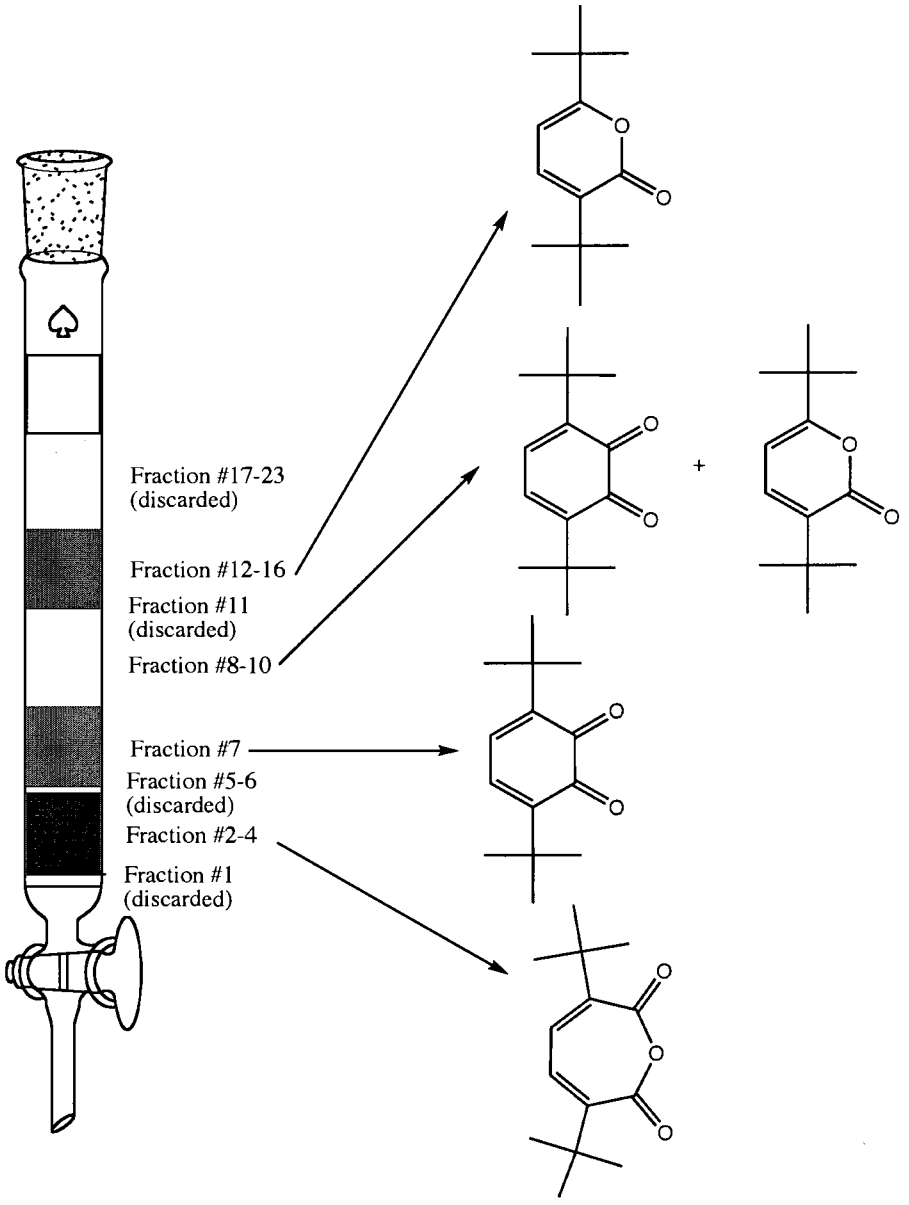


Figure S2. Column chromatography product separation of the 3 main H_2(3,6-DTBC) dioxygenase products. Column: 450 x 30 mm; silica gel (Aldrich, 60Å); suspended in n-hexane. Eluent: ~1000 mL $CHCl_3$; ~1.5 mL/min. Product identification: GC, GC-MS, ^1H and ^{13}C NMR, X-ray crystallography as detailed in the main text (or, for the X-ray crystallography, in this Supporting Information, *vide infra*).

X-ray Crystal Structure of 3,6-di-*tert*-butyl-1-oxacyclohepta-3,5-diene-2,7-dione (1). Following the column separation of the H_2(3,6-DTBC) dioxygenase products, 3,6-di-*tert*-butyl-1-oxacyclohepta-3,5-diene-2,7-dione (**1**) was isolated by

rotoevaporation of the solvent and redissolving the residue in a minimum (ca. 2 ml) of hot n-pentane. The solution was then capped and placed in the freezer. After a couple of days clear, colorless single crystals suitable for X-ray diffraction were obtained.

Diffraction data were collected on a Bruker APEX2 diffractometer employing Mo Kα radiation. Standard Bruker APEX2 control and integration software was employed. Bruker SHELXTL [2] software used for structure solution, refinement, and graphics. SADABS [3] correction was employed and the structure was solved by direct methods and refined by a full-matrix, weighted least-squares process.

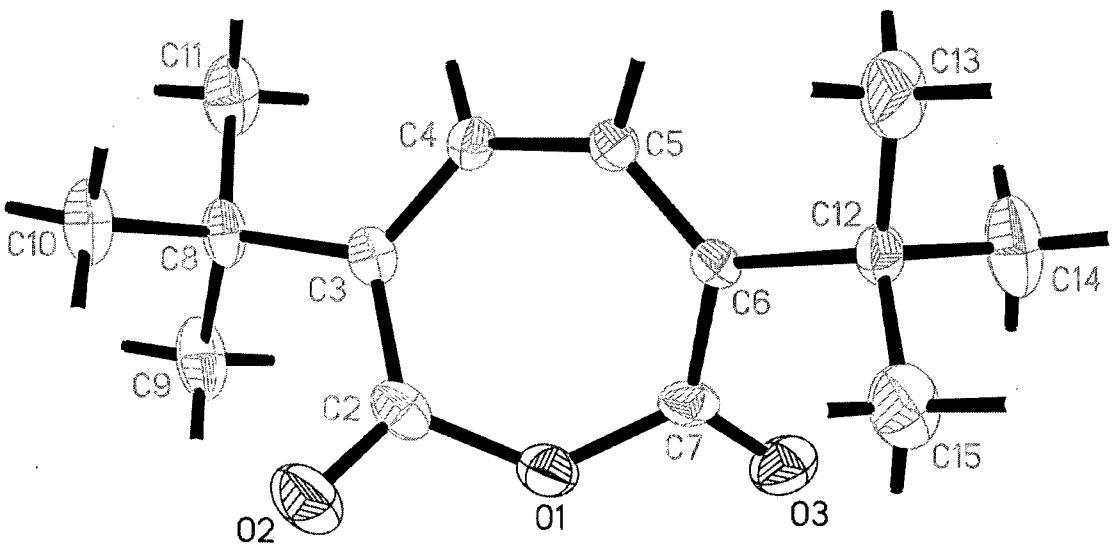


Figure S3. 50% Thermal Ellipsoid plot of 3,6-di-*tert*-butyl-1-oxacyclohepta-3,5-diene-2,7-dione (**1**).

Table S1. Crystal data and structure refinement for **1**.

Empirical formula: $C_{14}H_{20}O_3$
Formula weight: 236.30
Temperature: 100(2) K
Wavelength: 0.71073
Crystal system: Monoclinic
Space group: $P2_1$
Unit cell dimensions:

a = 6.4032(2) Å alpha = 90 °
b = 10.3625(3) Å beta = 100.8250(10) °
c = 10.4866(3) Å gamma =90 °
Volume: 683.44(4) $Å^3$ Z: 2
Density (calculated): 1.148 Mg/m^3
Absorption coefficient: 0.079 mm^{-1}
F(000): 256
Range for data collection: 1.98 to 31.49 °
Limiting indices: $-9 < h < 9$, $-10 < k < 15$, $-15 < l < 15$
Reflections collected: 7778
Independent reflections: 3786 (Rint = 0.0195)
Completeness = 31.49 °, 98.5 %
Refinement method: Full-matrix least-squares on F^2
Data / restraints / parameters: 3786 / 1 / 160
Goodness-of-fit on F^2: 1.059
Final R indices: R1 = 0.0490, wR2 = 0.1253
R indices (all data): R1 = 0.0576, wR2 = 0.1326
Absolute structure parameter: 1.5(11)
Largest diff. peak and hole: 0.382 and -0.156

Table S2. Atomic coordinates [$\times 10^4$] and equivalent isotropic displacement parameters [$Å^2 \times 10^3$] for **1**. U(eq) is defined as one third of the trace of the orthogonalized tensor.

	x	y	z	U(eq)
C(2)	1681(3)	6736(2)	1841(1)	26(1)
C(3)	-464(3)	7254(2)	1284(1)	22(1)
C(4)	-1374(2)	8103(2)	1984(2)	22(1)
C(5)	-535(2)	8593(2)	3281(1)	23(1)
C(6)	1490(2)	8832(2)	3840(1)	19(1)
C(7)	3182(2)	8723(2)	3059(1)	24(1)
C(8)	-1624(3)	6654(2)	1(1)	25(1)
C(9)	-83(4)	6552(2)	-963(2)	38(1)
C(10)	-2421(4)	5304(2)	280(2)	35(1)
C(11)	-3527(4)	7487(2)	-608(2)	40(1)
C(12)	2172(3)	9378(2)	5217(1)	23(1)
C(13)	318(4)	9329(3)	5956(2)	52(1)
C(14)	2856(5)	10795(2)	5197(2)	49(1)
C(15)	4005(4)	8572(2)	5946(2)	44(1)
O(1)	3275(2)	7632(1)	2330(1)	31(1)
O(2)	2273(2)	5640(1)	1811(1)	40(1)

O(3)	4587(2)	9500(2)	3047(1)	34(1)

Table S3. Bond lengths [Å] and angles [°] for **1**.

C(2)-O(1)	1.403(2)
C(2)-O(2)	1.199(2)
C(2)-C(3)	1.488(2)
C(3)-C(4)	1.347(2)
C(3)-C(8)	1.540(2)
C(4)-C(5)	1.457(2)
C(5)-C(6)	1.342(2)
C(6)-C(7)	1.480(2)
C(6)-C(12)	1.535(2)
C(7)-O(3)	1.209(2)
C(7)-O(1)	1.373(2)
C(8)-C(11)	1.532(3)
C(8)-C(10)	1.536(3)
C(8)-C(9)	1.543(2)
C(12)-C(15)	1.524(3)
C(12)-C(14)	1.533(3)
C(12)-C(13)	1.536(3)

O(2)-C(2)-O(1)	115.08(16)
O(2)-C(2)-C(3)	127.14(17)
O(1)-C(2)-C(3)	117.36(15)
C(4)-C(3)-C(2)	118.91(14)
C(4)-C(3)-C(8)	123.53(15)
C(2)-C(3)-C(8)	117.14(14)
C(3)-C(4)-C(5)	127.97(15)
C(6)-C(5)-C(4)	128.97(14)
C(5)-C(6)-C(7)	119.45(13)
C(5)-C(6)-C(12)	124.05(13)
C(7)-C(6)-C(12)	116.08(13)
O(3)-C(7)-O(1)	115.59(14)
O(3)-C(7)-C(6)	125.37(16)
O(1)-C(7)-C(6)	118.96(15)
C(11)-C(8)-C(10)	109.01(16)
C(11)-C(8)-C(3)	110.65(14)
C(10)-C(8)-C(3)	108.90(13)
C(11)-C(8)-C(9)	108.79(15)
C(10)-C(8)-C(9)	109.79(16)
C(3)-C(8)-C(9)	109.69(14)
C(15)-C(12)-C(14)	109.63(17)

C(15)-C(12)-C(6)	109.05(14)
C(14)-C(12)-C(6)	111.84(15)
C(15)-C(12)-C(13)	108.89(17)
C(14)-C(12)-C(13)	106.76(19)
C(6)-C(12)-C(13)	110.61(13)
C(7)-O(1)-C(2)	130.04(13)

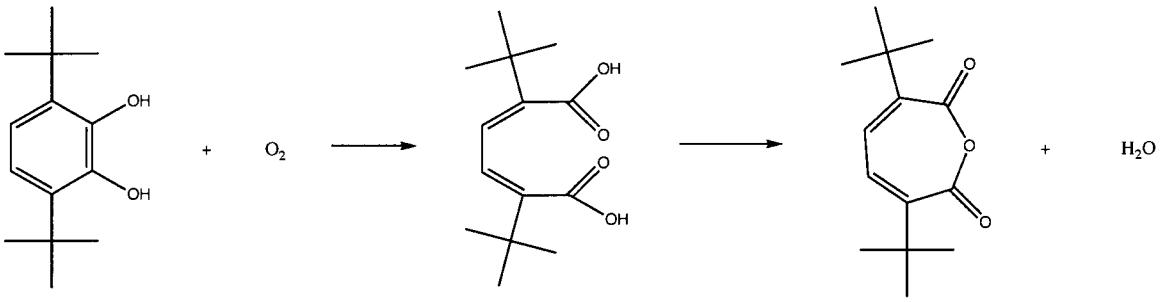
Table S4. Anisotropic displacement parameters [$Å^2 \times 10^3$] for **1**.

	U11	U22	U33	U23	U13	U12
C(2)	36(1)	25(1)	18(1)	1(1)	8(1)	7(1)
C(3)	30(1)	19(1)	16(1)	2(1)	4(1)	-1(1)
C(4)	21(1)	24(1)	21(1)	-4(1)	2(1)	-2(1)
C(5)	21(1)	29(1)	20(1)	-7(1)	6(1)	-1(1)
C(6)	21(1)	19(1)	18(1)	-1(1)	5(1)	2(1)
C(7)	21(1)	31(1)	18(1)	2(1)	1(1)	10(1)
C(8)	41(1)	20(1)	14(1)	-2(1)	1(1)	-5(1)
C(9)	65(1)	32(1)	20(1)	-7(1)	13(1)	-12(1)
C(10)	55(1)	23(1)	24(1)	-1(1)	2(1)	-10(1)
C(11)	54(1)	33(1)	24(1)	1(1)	-12(1)	0(1)
C(12)	26(1)	24(1)	19(1)	-4(1)	2(1)	-2(1)
C(13)	51(1)	79(2)	27(1)	-26(1)	13(1)	-15(1)
C(14)	80(2)	26(1)	37(1)	-6(1)	1(1)	-7(1)
C(15)	56(1)	45(1)	23(1)	-3(1)	-11(1)	14(1)
O(1)	27(1)	37(1)	30(1)	-7(1)	4(1)	7(1)
O(2)	48(1)	29(1)	42(1)	-1(1)	4(1)	13(1)
O(3)	26(1)	42(1)	36(1)	1(1)	8(1)	-3(1)

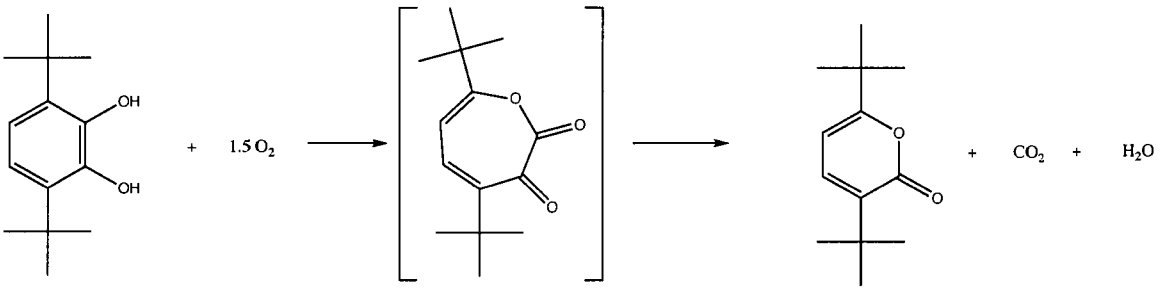
Reactions for Observed H_2(3,6-DTBC) Dioxygenase Product Formation.

Shown below in Schemes S1 and S2 are the proposed reactions for the formation of the intra and extradiol dioxygenase products formed from the H_2(3,6-DTBC) substrate.

Scheme S1. Reactions for the formation of 3,6-Di-*tert*-butyl-1-oxacyclohepta-3,5-diene-2,7-dione.



Scheme S2. Reactions for the formation of 3,6-di-*tert*-butyl-*2H*-pyran-2-one.



[1] Weiner, H.; Finke, R. G. *J. Am. Chem. Soc.* **1999**, *121*, 9831-9842.

[2] Sheldrick, G. M. *Acta Cryst. A* **2008**, *64*, 112-122.

[3] Sheldrick, G. M. *SADABS* (a program for Siemens Area Detection Absorption Correction), 2000.

CHAPTER IX

RE-INVESTIGATION OF A Ru_2-INCORPORATED POLYOXOMETALATE DIOXYGENASE PRECATALYST, "$[WZnRu_2^{III}(H_2O)(OH)(ZnW_9O_{34})_2]^{11-}$": EVIDENCE FOR MARGINAL, ≤0.2 EQUIVALENTS OF Ru INCORPORATION PLUS FASTER CATALYSIS BY PHYSICAL MIXTURES OF $[Ru^{II}(DMSO)_4Cl_2]$ AND THE PARENT POLYOXOMETALATE $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$

This chapter contains the manuscript of a full article currently *in press* in *Inorganic Chemistry*. This manuscript re-analyzes the composition of "$[WZnRu_2^{III}(H_2O)(OH)(ZnW_9O_{34})_2]^{11-}$" (**1**), a previously claimed all-inorganic dioxygenase. The findings are that in our hands **1** is composed of primarily the parent polyoxometalate $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ with small amounts of Ru either incorporated into the polyoxometalate structure or more likely as Ru^{n+} impurity. Furthermore, mixtures of the parent polyoxometalate $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ plus $[Ru(DMSO)_4Cl_2]$ gives faster catalysis to the same autoxidation products than does "**1**".

This manuscript was prepared, and the experiments contained herein were performed by Aimee M. Morris, with assistance on the structure determination and critical analysis of previous structures from Prof. Oren P. Anderson and editing by Prof. Richard G. Finke.

Re-Investigation of a Ru_2-Incorporated Polyoxometalate Dioxygenase Precatalyst, "$[WZnRu_2^{III}(H_2O)(OH)(ZnW_9O_{34})_2]^{11-}$": Evidence For Marginal, ≤0.2 Equivalents of Ru Incorporation Plus Faster Catalysis by Physical Mixtures of $[Ru^{II}(DMSO)_4Cl_2]$ and the Parent Polyoxometalate $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$

Aimee M. Morris, Oren P. Anderson, and Richard G. Finke

Abstract

A 1997 *Nature* paper (*Nature* **1997**, *388*, 353-355) and subsequent 1998 *J. Am. Chem. Soc.* paper (*J. Am. Chem. Soc.* **1998**, *120*, 11969-11976) reported that a putative Ru_2-substituted polyoxoanion, "$[WZnRu_2^{III}(H_2O)(OH)(ZnW_9O_{34})_2]^{11-}$", (**1**), is an all-inorganic dioxygenase able to incorporate 1 O_2 into 2 adamantane C–H bonds to yield 2 equivs. of 1-adamantanol as the primary product. In a subsequent 2005 *Inorg. Chem.* publication (*Inorg. Chem.* **2005**, *44*, 4175-4188), strong evidence was provided that the putative dioxygenase chemistry is, instead, the result of classic autoxidation catalysis. That research raised the question of whether the reported Ru_2 precatalyst, **1**, was pure or even if it contained two Ru atoms, since Ru is known to be difficult to substitute into polyoxoanion structures (Nomiya, K.; Torii, H.; Nomura, K.; Sato, Y. *J. Chem. Soc. Dalton Trans.* **2001**, 1506-1521). After our research group had contact with three other

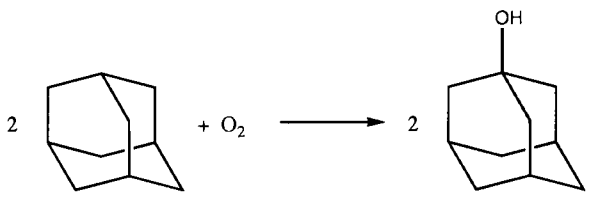
groups who also had difficulties reproducing the reported synthesis and composition of **1**, we decided to re-examine **1** in some detail. Herein we provide evidence that the claimed **1** actually appears to be the parent polyoxoanion $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ with small amounts of Ru (≤ 0.2 atoms) either substituted into the parent complex or present as a small amount of a Ru^{n+} impurity, at least in our and two other group's hands. The evidence obtained, on three independent samples prepared from two research groups including ours, includes elemental analysis on the bulk samples, single-crystal X-ray diffraction, elemental analysis on single crystals from the same batch used for X-ray diffraction, ^{183}W NMR, and adamantane oxidation oxygen uptake and product determination studies. Also re-examined herein are the two previously reported crystal structures of **1** that appear to be very similar to the structure of the parent polyoxoanion, $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$. Furthermore, we report that trace Ru alone, in the form of $[Ru(DMSO)_4Cl_2]$, or the parent polyoxoanion $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ alone, are capable of producing the same products. More significantly, a simple physical mixture of $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ plus the average 0.13 equiv. of Ru found by analysis added as the $[Ru(DMSO)_4Cl_2]$ starting material is a ca. 2 fold kinetically more competent catalyst than "$[WZnRu_2^{III}(H_2O)(OH)(ZnW_9O_{34})_2]^{11-}$", (**1**). In short, the evidence is strong that the putative "$[WZnRu_2^{III}(H_2O)(OH)(ZnW_9O_{34})_2]^{11-}$", (**1**), which underlies the previously reported all-inorganic dioxygenase catalysis claim, is probably not correct. That does not mean that **1** cannot or even does not exist, but just that (a) no reliable synthesis of it exists if it has actually been made before, and (b) that a simple mixture of the $[Ru(DMSO)_4Cl_2]$ plus $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ precursors gives ca. 2-fold faster catalysis of adamantane hydroxylation that occurs by, the evidence suggests, a radical-

chain autoxidation mechanism rather than via the previously claimed, novel all-inorganic-based dioxygenase catalysis.

Introduction

A 1997 *Nature* paper[1] and subsequent 1998 *J. Am. Chem. Soc.* paper[2] by others claimed to have synthesized the first "all inorganic dioxygenase", $Q_{11}[ZnWRu_2^{III}(OH)(H_2O)(ZnW_9O_{34})_2]$ (referred to hereafter as, Q_{11}-**1**), where Q = tricaprylmethylammonium ion; $\{CH_3N[(CH_2)_7CH_3]_3\}^+$. This claim involving Q_{11}-**1** is significant and of interest on several accounts: (i) first, only a few catalytically active 2nd row transition-metal di-M_2-substituted polyoxoanions derived from the parent compound $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ are known in the literature;[3] and second and more importantly, no other all-inorganic dioxygenase that will do unactivated C–H bond activation has been reported in the literature, to our knowledge. Such a dioxygenase is a current "Holy Grail" of oxidation catalysis due to its ability, by definition, to insert both atoms of O_2 into 2 C–H bonds to yield 2 C–OH, for example, all without the use of protons or electrons and, hence, without the formation of even trace H_2O.[4,5,6] Therefore, knowledge of the true composition and purity of **1**, as well as the critical question of the true nature of the underlying catalyst, is an important question.

The main evidence reported by others[2] for Q_{11}-**1** being a dioxygenase was hydroxylated alkane product studies, leading to the claimed stoichiometry shown in equation (1), plus the finding that the addition of the alkyl radical scavenger, 4-*tert*-butylcatechol, caused "only a very slight decrease in catalytic activity."[7]



(1)

However, in 2005 Yin et al. reinvestigated the claimed dioxygenase, Q_{11}-**1**.[8] Their results provided compelling evidence that the claimed Q_{11}-**1** is not a dioxygenase but, rather, is a classic autoxidation catalyst for which a detailed autoxidation mechanism could be written consistent with the re-evaluated kinetics.[8] The reported evidence for an autoxidation catalyst was extensive and included: (i) detection of trace amounts of peroxide, ROOH, the initial product of any autoxidation (peroxide would of course not be present if Q_{11}-**1** were in fact a dioxygenase as in eq. 1 above); (ii) kinetic studies demonstrating a fractional rate law quantitatively consistent with, and highly characteristic of, a radical-chain mechanism but inconsistent with the previously suggested dioxygenase pathway[1]; (iii) a 1:1 products:O_2 stoichiometry (in distinct contrast to the previously claimed 2:1 dioxygenase stoichiometry, eq. (1)), evidence that by itself rules out a pure dioxygenase reaction; (iv) detection of $H_2{}^{18}O$ as a product, a product expected for autoxidation (but that could not be present if Q_{11}-**1** was a true dioxygenase); (v) initiation of the reaction by known radical initiators, AIBN and *t*-BuOOH, again indicative of a radical (and not a dioxygenase) mechanism; and (vi) complete inhibition of the progress of the reaction by four radical scavengers including 4-*tert*-butylcatechol that was previously[7] claimed to not be a strong inhibitor, evidence again consistent with and highly supportive of a radical-chain mechanism.[8] The results were so different from those reported in the 1997 Nature paper[1] that Yin, et al. stated in

their paper "at times during our studies it seemed as if we were studying a different system [than the prior workers had studied]."[8] Hence, a closer look at the precatalyst Q_{11}-**1** also became important from that perspective and since others began using "Q_{11}-**1**" for additional chemistry.[9]

Contact with three other research groups indicated that they, too, had concerns regarding the reported Q_{11}-**1**, either from trying to repeat its preparation or from trying to confirm its reported Ru_2 composition. Specifically, Nomiya and co-workers reported[10,11] that they were not able to reproduce the reported preparation of the Q_{11}-**1** precursor, K_{11}-**1**.[8] C. L. Hill's group shared that they, too, had issues with the original crystal structure of K_{11}-**1** and concerns over the reported composition of Q_{11}-**1**.[12] And, in 2004, Shannon and co-workers[9] reported the structure of Na_{11}-**1** to be $Na_{14}[Ru^{III}{}_2Zn_2(H_2O)_2(ZnW_9O_{34})_2]$, although their crystallographic data indicated only partial incorporation of Ru into the parent $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ polyoxoanion. Their work reported that Ru replaces one zinc atom and a partial W atom (as opposed to replacing two Zn atoms as originally suggested[7]), resulting in a composition of $Na_{14}[W_{0.76}Ru_{1.24}Zn_2(H_2O)_2(ZnW_9O_{34})_2]$ (*vide supra*). In short, the issues surrounding K_{11}-**1** and Q_{11}-**1** raised independently by three research groups,[9,11,12] along with the importance of claims of a dioxygenase catalyst and the key question of what is the true catalyst or at least precatalyst, led us to re-scrutinize **1** in some detail.

Reported herein is our reinvestigation of **1**, $[ZnWRu_2^{III}(OH)(H_2O)(ZnW_9O_{34})_2]^{11-}$, and its claimed Ru_2 composition.[1,2] We find by elemental analysis of the bulk samples, single crystal X-ray structure determination, elemental analysis on single crystals from the same batch used for the X-ray diffraction structure, ^{183}W NMR, oxygen uptake, and

product identification studies that there is little to no evidence of di-substitution of Ru(III) into the parent compound $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ by the reported synthesis[7,21] using $[Ru(DMSO)_4Cl_2]$ regardless if the source is Yin's,[8] Nomiya's,[11] Hill's,[12] Shannon's[9] or our present preparation of **1**. Instead, we provide evidence for the composition of "**1**" being primarily $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ in at least three independent samples examined herein with only a small amount (≤ 0.2 atoms) of Ru being observed in any of the three independent samples—that low Ru level due presumably to either partial incorporation of Ru into the parent polyoxoanion *or* a small level of a Ru^{n+} impurity. Significantly, we also provide O_2 uptake and product identification studies herein that show the same products are obtained in similar yields and ratios for: (i) the original[1,2] "Q_{11}-**1**" (i.e., based on the published yields[1,2]); (ii) three independent samples of "Q_{11}-**1**" containing ≤ 0.2 atoms (which is ≤ 0.2 equiv.) of Ru, (iii) the parent compound $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ containing no Ru, (iv) 0.13 equiv. of Ru added as the precursor $[Ru(DMSO)_4Cl_2]$, and (v) a mixture of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ plus 0.13 equiv. of Ru added as $[Ru(DMSO)_4Cl_2]$. The significance of these studies is at least four-fold: (i) in providing evidence on whether Ru_2-substituted "K_{11}-**1**" and "Q_{11}-**1**" actually exist; (ii) in providing further evidence bearing on the prior claim of an all-inorganic dioxygenase[8]; (iii) in clarifying what the true catalyst probably is in the prior report of adamantane hydroxylation by Ru and polyoxoanion-containing precatalysts—the true catalyst being a central, but often perplexing part of modern catalysis,[13,14,15,16,17] and (iv) in bearing on the generally difficult problem of synthesizing Ru[7,10,18] (and other later transition-metal[19]) -containing polyoxoanions cleanly. The present results also (v) tie, in a general way, into other areas

of synthetic inorganic chemistry such as the difficulties historically in putting Ru into other ligands (e.g. porphyrins).[20]

Results and Discussion

Synthesis and Characterization of Independent Samples of "Q_{11}-1." Three independent samples of both "K_{11}-**1**" and "Q_{11}-**1**" were prepared according to the literature procedure,[7,21] shown in Scheme 1,[22] in two different laboratories by three individual researchers (A. M. Morris at Colorado State University (CSU) (*vide infra*), C.-X. Yin also at CSU,[8] and T. Anderson in C. L. Hill's laboratories at Emory University[12]), and then used in this study. Each of the three independent "**1**" samples was analyzed by: elemental analysis, IR, comparison of O_2 uptake curves to the literature,[2] and adamantane oxidation product studies. The "**1**" sample prepared herein was also analyzed by single crystal X-ray diffraction and ^{183}W NMR.

Scheme 1. The literature[7,21] synthesis of "**1**."

$Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2] + 2[Ru(DMSO)_4]Cl_2$

↓ Reflux in deionized water at 90 °C for 18 hrs
Cool, expose to air, and add 2 g KCl
Filter crude product
Recrystallize twice from hot water
Dry overnight under vacuum at room temp

"$K_{11}[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]$", K_{11}-**1**

↓ Dissolve in deionized water
Add 11QCl
Extract three times with 1,2-dichloroethane
Remove 1,2-dichloroethane *in vacuo*

"$Q_{11}[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]$", Q_{11}-**1**

Elemental Analysis Results. Elemental analysis on each of the three samples reported herein of "K_{11}-**1**," the "Q_{11}-**1**" precursor, and in comparison to the "K_{11}-**1**" originally reported,[2] are shown in Table 1. Three independent samples of "K_{11}-**1**," each prepared by the same procedure[7,21] in so far as possible, *show different compositions.* While no two analyses are the same (indicating that an impure sample is obtained as previously suggested[8]), each of the three independent samples *are low in Ru and two of the three are high in Zn content*, Table 1.

Table 1. Elemental Analyses for Three Independent Samples of "K_{11}-**1**" vs. the Literature's Reported Analysis.[7,21]

	Calculated for K_{11}-**1**•15H_2O	Neumann's[7,21] "K_{11}-**1**•15H_2O"	"K_{11}-**1**" prepared herein	Yin's8 "K_{11}-**1**"	Anderson's12 "K_{11}-**1**"
K	7.57	7.06	4.69	4.29	NM[a]
Na	0.00	NM[a]	0.59	1.39	NM[a]
Ru	*3.56*	*3.56*	*0.75*	*0.24*	*0.29*
Zn	*3.45*	*3.24*	*7.13*	*6.65*	*3.08*[b]
W	61.5	61.3	NM[a]	61.2	NM[a]
H_2O	4.76	4.69	NA[c]	NA[c]	NA[c]

[a] NM = not measured. [b] This Zn analysis appears to be low/in error in light of the low amount of Ru that is observed in this same sample. [c] NA = not applicable, since the number of hydrated waters in each sample found by TGA studies can vary depending upon the preparation and especially drying conditions.

Crystal Structure of "Na_{11}-1." The X-ray structure of "Na_{11}-**1**," shown in Figure 1, confirms the previous findings[7,9,21] that "**1**" has a sandwich type polyoxoanion structure. In the previous study, it was reported that a ring containing WRuZnRu is sandwiched between the two (ZnW_9O_{34}) polyoxoanion units.[21] However, an issue with that prior structure is the unusually large thermal coefficients reported for the two Ru atoms. The authors rationalized this by assuming that the two positions at which the Ru atoms are located are relatively labile.[21] However, upon examination of the two sites which Ru is thought to occupy, these sites contain 5 internal oxygen bonds as well as one bond to a H_2O or OH^- ligand. Given that the estimated bond dissociation energy of a Ru–O bond is ca. 110 kcal/mol,[23] the probability of the atom in this site being labile is low. A simpler explanation for the unusually large thermal coefficient observed is that *the atom assignment is incorrect*, a hypothesis not just consistent with, but required by, three of the four elemental analysis results in Table 1 revealing relatively little incorporation of Ru into the polyoxoanion.

In a second, literature study of the crystal structure of "Na_{11}-**1**", the crystallographic data indicated a $Ru^{III}_{1.24}W_{0.76}Zn_2$ ring.[9] However, the reported composition in the main text of the previous publication[9] suggests a $Ru^{III}_2Zn_2$ ring is present in "Na_{11}-**1**" giving the composition $Na_{14}[Ru^{III}_2Zn_2(H_2O)_2(ZnW_9O_{34})_2]$. Moreover, the structure suggests that the Ru atoms substitute into the Zn and W positions of the ring[9] and not the two Zn atom positions containing terminal H_2O ligands as previously suggested.[7,21] In addition, in that study,[9] the composition of "Na_{11}-**1**" is reported *to vary* from six independent crystals studied by both X-ray diffraction and energy dispersive X-ray (EDX) spectroscopy. The average EDX composition was consistent with Ru_2Zn_2 but not with the $Ru_{1.24}W_{0.76}Zn$ ring of the crystal structure. This second study[9] again illustrates the variation and inconsistency of "**1**."

Table 2. Crystallographic data for the structure solved in this publication compared to structures previously solved by Tourné,[10] Neumann,[21] and Shannon.[9]

	Tourné's[10] $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]\cdot 46H_2O$	Neumann's[21] $Na_{11}[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]\cdot 42H_2O$	Shannon's[9] $Na_{14}[W_{0.76}Ru_{1.24}Zn_2(H_2O)_2(ZnW_9O_{34})_2]\cdot 46H_2O$	$Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]\cdot 39H_2O$ solved herein
Space Group	$P2_1/c$	$P2_1/c$	$P2_1/n$	$P2_1/n$
M	6047.9	6023.9	6110.4	5995.0
a/Å	13.027(4)	13.069(3)	13.0374(6)	13.0017(6)
b/Å	17.788(5)	17.827(4)	17.7859(8)	17.8485(7)
c/Å	24.124(4)	24.182(5)	21.1030(10)	21.0260(9)
β/°	118.94(2)	118.97(2)	93.3800(10)	93.223(3)
U/Å^3	4892(2)	4929(1)	4884.9(4)	4871.6(4)
Crystal size /mm	0.21 × 0.16 × 0.34	NR[a]	0.088 × 0.102 × 0.326	0.067 × 0.166 × 0.263
R	0.039	0.051	0.0380	0.0509
GOF	NR[a]	NR[a]	1.061	1.120

[a] NR = not reported

In order to further test the hypothesis that there is little to no Ru incorporation into the parent polyoxoanion, we grew crystals of "Na_{11}-**1**" according to the literature procedure[21] and performed an independent single crystal X-ray structure determination. Our structural results agree with our bulk elemental analysis results above (Table 1), *as well as our elemental analysis results on the single crystals from the same batch used for X-ray diffraction* (Table 3, *vide infra*), in that there appears to be *no Ru* present in at least this sample of "Na_{11}-**1**", Figure 1. Attempts to introduce even partially occupied Ru sites into the refinement were unsuccessful. Instead our reported structure suggests that "Na_{11}-**1**" is primarily, if not completely, the parent precursor polyoxoanion, $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$, Table 2, last column. In order to rule out the possibility of having some crystals which contain Ru, three separate crystals from two different crystal batches were analyzed and each of the three data sets refined to the same structure, Ru-free $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$.

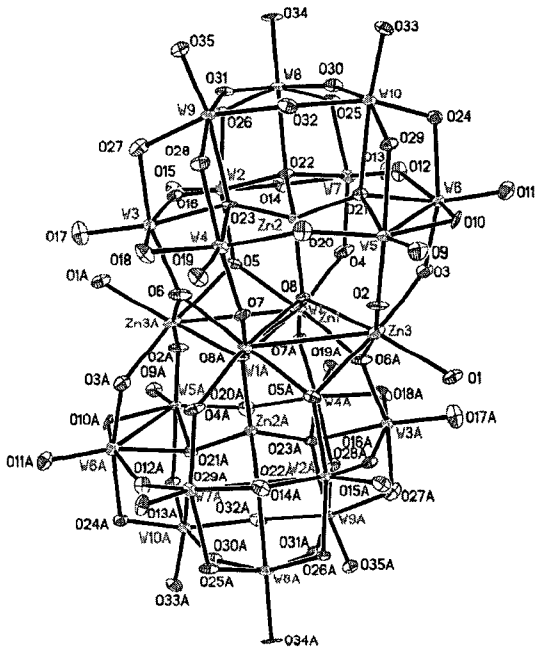


Figure 1. Thermal ellipsoid plot of $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ (50% ellipsoids). Solvent (H_2O) and counter cations (Na^+) have been removed for clarity.

While reinvestigating the crystal structure of "Na_{11}-**1**," similarities between the two previous structure determinations[9,21] of "Na-**1**" and Tourné's[10] determination of $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ became apparent (Table 2). The structure determination reported herein appears at first glance to be different from the reported structures of "Na_{11}-**1**"[21] and $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$[10] as the space group is different (we chose to solve the structure in the International Union of Crystallographers' preferred space group of $P2_1/n$). However, Table S1 of the Supporting Information shows that the alternative $P2_1/c$ space group for the structure reported herein gives a unit cell that exhibits similar cell constants to those shown for the structures of "Na_{11}-**1**" and $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ in Table 2. This provides further support consistent with the hypothesis that **1** made in at least our hands is the parent polyoxoanion,

$[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$; our crystallography result also raises the possibility that the atom assignments in the previous structure may have been misinterpreted.

As noted above, elemental analysis of the crystals from the batch used for X-ray diffraction gave the composition $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]\cdot 13H_2O$. This composition is identical to the polyoxoanion precursor used (Table 3), and supports the results from the crystal structure determination. Like the bulk sample, the crystal batch sample also indicates no detectable Ru substitution into the parent polyoxoanion. That there is little to no Ru substitution is not surprising as Ru has been shown to be difficult to substitute into polyoxoanions.[10] Furthermore, the synthesis published[7,21] uses a neutral pH, whereas acidic (or sometimes basic) pH is often required for polyoxoanion substitutions.[24] An alternative hypothesis here is that it is simply our failure (i.e., in our hands) to be able to incorporate Ru into the desired structure. However, the same low to no Ru results for 2 independent researchers in our labs, in C. L. Hill's labs, and Nomiya's inability to repeat the reported synthesis of Ru_2-**1**,[11] all argue strongly *that there is a problem with the reported synthesis*, Scheme 1, and the claimed composition of **1**. The catalytic results provide additional, highly suggestive evidence that "Ru_2-**1**" may have never been made, *vide infra*.

Table 3. Elemental analysis data for the single crystals of "Na_{11}-**1**" from the batch used for the X-ray diffraction structural analysis.

	Calculated for Na_{11}-**1** $\cdot 13H_2O$	Calculated for $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})] \cdot 13H_2O$	Observed for "Na_{11}-**1**" sample prepared herein
Na	4.60%	5.06%	4.62%
Zn	3.57%	5.99%	5.99%
Ru	3.67%	0%	<0.1%
W	63.5%	64.1%	62.8%[a]
O	24.1%	24.4%	24.2%

[a] Low W analyses are unfortunately prevalent in the polyoxometalate area (for example see the repeat W analyses required in the experimental section of ref 25); W analyses can vary by ca. ±2% even when performed (as was done here) by inductively coupled plasma (ICP) emission spectroscopy.[26]

^{183}W NMR Studies. Tourné previously reported the ^{183}W NMR spectrum for the Li^+ salt of $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ taken at 22 °C. In order to collect further evidence on the composition of the bulk sample, ^{183}W NMR on the Li^+ exchanged "K_{11}-**1**" prepared herein (i.e., by A. M. Morris) was taken at 52 °C in order to increase the solubility of "**1**", Figure 2. The spectrum shown in Figure 2, although with a relatively poor signal/noise, is the same within experimental error as the spectrum of Li^+ exchanged $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ once the difference in temperature of the two spectra are taken into account.[3] Both spectra contain the expected ten W signals, nine of which correspond to the W atoms in the two (ZnW_9O_{34}) units. One (the tenth) signal at ca. +20 ppm has half the intensity of the other signals and can be assigned to the single W atom contained in the WZn_3 ring of the polyoxoanion and labeled W1/W1A (each having 50% occupancy) in Figure 1. The ^{183}W NMR spectrum of Li^+-"**1**" is consistent with the elemental analysis studies as well as the crystal structure in showing that "**1**" is at least mostly composed of the parent polyoxoanion, $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$. The caveat here is, however, that ^{183}W NMR is notoriously insensitive to smaller, trace components, especially in non-ideal signal/noise spectra such as we were able to obtain (Figure 2).

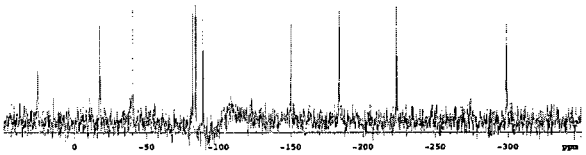


Figure 2. ^{183}W NMR spectrum of Li^+ exchanged "K_{11}-**1**" in D_2O at 52 °C.

O_2-Uptake and Product Identification and Quantitation Studies: General Procedure. Following the previously published literature procedure for O_2 uptake studies,[8] 6.0 mmol of adamantane, 6.0 μmol of the precatalyst, and 12.0 mL of 1,2-dichloroethane were combined and equilibrated at 80 °C under ~1.0 atm of O_2. The oxygen uptake curves for each of the experiments discussed below can be seen in Figures 3 and S3 to S8 of the Supporting Information. Control experiments of the appropriate added Zn and $[Ru(DMSO)_4Cl_2]$ plus QCl, along with discussion, are also provided in the Supporting Information for the interested reader. Following the O_2 uptake experiments, the products of each reaction were identified and quantified by comparison to authentic samples with GC.

O_2 Uptake and Product Identification and Quantitation Studies with "Q_{11}-1." For each of the three independently prepared samples of "Q_{11}-**1**," there was approximately 2.5-3 psig of O_2 taken up during the course of 24 hours over which the reaction was monitored (following the previous literature procedure[2]). A sample O_2 uptake is shown in Figure 3. The resulting products from the O_2-uptake curves for each of the three independent samples of "Q_{11}-**1**" are the same within experimental error, Table 4, entries 1, 2, and 6, despite the varying amounts of Ru present in each of the samples. In addition, the products obtained from each of the three independent samples

(entries 1, 2, and 6 of Table 4) of "Q_{11}-**1**" gave the same yields and selectivities as the previously reported "Q_{11}-**1**" reported to contain > 1.8 more Ru atoms, entry 7, Table 4. This implies that "Q_{11}-**1**" reported previously[1,2,7,21] (i.e., and with its claimed two incorporated Ru atoms) either (i) coincidently gives the same oxidation products in the same yields as we see with "**1**" that has ≤ 0.2 equiv of Ru, *or* more likely we believe (ii) that the previously reported[1,2,7,21] "Q_{11}-**1**" is similar to our "Q_{11}-**1**" samples—samples prepared by the same method[7,21] and reported herein to contain ≤ 0.2 Ru. The latter is certainly the simplest, "Ockham's razor" interpretation of the extant data.

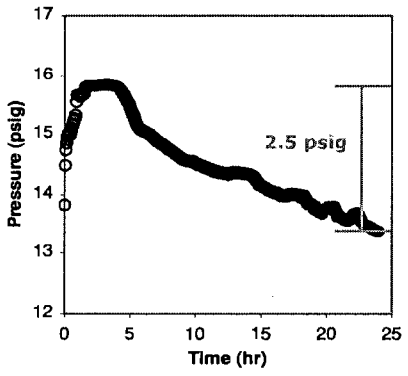


Figure 3. Pressure vs. time curves recorded by a O_2 pressure transducer. The initial pressure rise is due to the 1,2-$C_2H_4Cl_2$ solvent pressure equilibration following flushing the Fischer-Porter bottle with O_2 and then pressurizing it with ca. 14 psig of O_2. (The noise in the pressure curve arises from temperature variations since the reaction temperature is at 80 °C yet the rest of the apparatus is at ambient temperature with ±3 °C variation.) However, the net pressure loss shown matters, so that this experiment more than suffices for its intended purpose of monitoring the net O_2 uptake over the time (24 hours) that the adamantane oxidation products were also monitored. Reaction conditions: 6.0 mmol of adamantane, 6.0 μmol of "Q_{11}-**1**," 12 mL of 1,2-$C_2H_4Cl_2$, 80 °C, and ca. 14 psig of O_2 pressure.

Table 4. Products Quantitation and Other Relevant Data for Adamantane Hydroxylation Under 1 atmosphere of O_2, at 80 °C, and after 24 hours.

Entry	Sample	equiv Ru	1-adamantanol yield	2-adamantanone yield	Mass Balance[a]	Selectivity[b]	Induction period?
1	"Q_{11}-**1**" sample prepared herein	0.21	13 ± 1%	2.2 ± 0.4%	98 ± 4%	5.9	Yes
2	C. Hill's "Q_{11}-**1**" sample	<0.1	13 ± 1%	2.3 ± 0.4%	96 ± 4%	5.7	Yes
3	$[Ru(DMSO)_4Cl_2]$	0.13	5.4%	1.3%	98%	4.2	Yes
4	$Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$[c]	0	0%	0%	98 ± 2%	0	Yes
	At 120 ± 20 hours[d]	0	30 ± 5%	6 ± 1%	96 ± 4%	5.0	Yes
5	$Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2$ + $[Ru(DMSO)_4Cl_2]$	0.13	23%	5.6%	95%	4.1	Yes
6	Yin et al's Literature[8]	0.05	12 ± 1%	2.2 ± 0.4%	95 ± 5%	5.5	Yes
7	Neumann et al's Literature[2]	2.0	12.3%	NR[e]	NR[e]	NR[e]	Yes

[a] The mass balance is calculated from the product mass observed by GC in comparison to the initial amount of adamantane substrate used. [b] The selectivity calculated refers to the preference of 1-adamantanol over 2-adamantanone. [c] $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ alone exhibits variable and long induction periods and hence, yields no products at 24 hours. However, after longer reaction times of ~ 120 hours, the products are observed. [d] The yields reported are a reflection of two averaged data sets where the product analysis was performed at ~ 120 hours. [e] NR=not reported.

O_2-Uptake and Product Identification and Quantitation Studies with Just $[Ru(DMSO)_4Cl_2]$ as the Precatalyst. By using just 0.13 equiv of Ru (by adding the 0.77 μmol of Ru, which is the same amount of Ru if one could add 1.0 equiv of $Q_{12}[WRu_{0.13}Zn_{2.87}(H_2O)_2(ZnW_9O_{34})_2]$) in the form of $[Ru(DMSO)_4Cl_2]$ (an amount that is the same as the average Ru seen in the elemental analysis of "Q_{11}-**1**" from our laboratories, entries 1 and 6 of Table 4), an O_2 uptake curve the same within experimental error of Figure 3 was produced, see Figure S3 of the Supporting Information. Furthermore, 1-adamantanol and 2-adamantanone products are again obtained as reported for[2] "Q_{11}-**1**" in a similar selectivity (ca. the same within experimental error), albeit in a bit less than half as much yield (entry 3, Table 4, in comparison to entries 1 and 2)—half the yield that may simply reflect that there is ca. half the Ru in this control experiment as in entries 1 and 2. Overall, this control experiment shows that starting with $[Ru(DMSO)_4Cl_2]$ gives similar adamantane hydroxylation products, but more slowly in ca. half the yield.

O_2-Uptake and Product Identification and Quantitation Studies with $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$•nQCl. The next obvious control experiment was to use the Ru-free Q^+ salt of the polyoxoanion precursor, $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$, instead of "$Q_{11}$-**1**". Again, the identical two main 1-adamantanol and 2-adamantanone products were observed for $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ (entry 4, Table 4 in comparison to entries 1 and 2). However, no products were observed after 24 hours for $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$•nQCl since the reaction time is variable (as expected for a radical-chain autoxidation reaction[8]) and requires 95 to 140 hrs for completion (Figures S4a and b of the Supporting Information). In combination with the prior control of just

$[Ru(DMSO)_4Cl_2]$ alone, these results would seem to indicate that (i) both $[Ru(DMSO)_4Cl_2]$ and Ru-free $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$•nQCl are able to exhibit the observed (radical-chain[8]) catalysis, but (ii) that their combination is more effective. This is in fact the case as the third control reaction described next demonstrates.

O_2-Uptake and Product Identification and Quantitation Studies with Pre-mixed $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$•nQCl plus $[Ru(DMSO)_4Cl_2]$. The third control of pre-mixing 0.13 equiv $[Ru(DMSO)_4Cl_2]$ with $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ produced the anticipated accelerated hydroxylation reaction. The induction period was shortened as expected, the same two main products were produced as with "Q_{11}-**1**", but the yields were higher beyond experimental error (entry 5, Table 4). *This is an important result, one of the most telling results in this report.* It indicates higher kinetic competence of low levels of Ru not necessarily *in* the polyoxoanion structure as would be the case for authentic **1**; it argues for enhanced catalysis by a perhaps polyoxoanion-supported[27] (and not polyoxoanion-framework incorporated,[28] such as **1**) form of Ru^{n+} (e.g., $Ru^{n+}/[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ is one possibility). It also makes an important connection between our results and the work in question[2]: we get *the same catalysis in terms of yields (rates) and selectivities with just low levels of Ru and the parent polyoxoanion*. This strongly implies that *if* the prior complex **1** really contains the 2.0 equiv of Ru as listed in Table 4, then that Ru is not essential to catalysis! The alternative hypothesis is that the previously reported[1,2,7] "**1**" does not contain 2 Ru but, instead, is really the same as our low-Ru material.

O_2-Uptake and Product Identification and Quantitation Studies with Pre-mixed $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$•nQCl plus $[Ru(DMSO)_4Cl_2]$ plus 4-*tert*-

butylcatechol. The above experiments, along with our prior demonstration[8] of a radical-chain mechanism for the formation of 1-adamantanol and 2- adamantanone beginning with "Q_{11}-**1**" as the precatalyst, strongly imply that the results in Table 4 are all due to radical-chain chemistry. However, to provide evidence for this in the now important case of pre-mixed $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ plus $[Ru(DMSO)_4Cl_2]$, these two precursors were combined under the standard reaction conditions employed herein. After ca. 8 hours when O_2 uptake was observed to be underway, 4-*tert*-butylcatechol (a radical inhibitor used in both our and the literature previous studies[2,8]) was added. As shown in Figure S8 of the Supporting Information, the addition of 4-*tert*-butylcatechol *completely halted* the otherwise on-going O_2 uptake; no further hydroxylated adamantane products were observed by GC. While only a single experiment, in light of the now extensive evidence for a radical-chain mechanism when using "Q_{11}-**1**" as the precatalyst[8] and given that 4-*tert*-butylcatechol is an established inhibitor of that chain,[8] even this single result provides prima facie evidence of a similar mechanism, and a close connection, between "Q_{11}-**1**" and a physical mixture of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ plus $[Ru(DMSO)_4Cl_2]$.

Conclusions

The main findings from the present study can be summarized as follows:

- "Q_{11}-**1**" prepared in at least our laboratories by two independent researchers (reported herein and ref 8) and in Hill's[12] laboratories (as well as in Nomiya's laboratories[11]) is *not* consistent with the previously claimed[1,2] $Q_{11}[Ru^{III}_2ZnW(H_2O)(OH)(ZnW_9O_{34})_2]$ on the basis of elemental analysis, single crystal X-ray diffraction, ^{183}W NMR, and oxygen uptake and product

quantification studies plus appropriate control experiments. In particular, to date no one has reported being able to independently prepare a sample of **1** that contains the claimed 2 Ru atoms.

- Instead "Q_{11}-**1**", at least as we or others can synthesize, it appears to be primarily the parent polyoxoanion precursor, $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ with small amounts of Ru (≤ 0.2 atoms) either substituted into the complex, or perhaps more likely variable amounts of Ru^{n+} present as an impurity.
- The two previous crystal structure determinations of "Na_{11}-**1**" as $Na_{11}[Ru^{III}{}_2ZnW(H_2O)(OH)(ZnW_9O_{34})_2]$ or $Na_{14}[W_{0.76}Ru_{1.24}Zn_2(H_2O)_2(ZnW_9O_{34})_2]$ show strong similarities to the parent polyoxoanion, $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$. While this is to be expected based on the similarities of their structures, this does re-illustrate[19e] the need for other physical methods able to show the amount and location of low-valent metals in polyoxoanion structures—a difficult topic.[19e]
- Control O_2-uptake and product identification studies using $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ alone, $[Ru(DMSO)_4Cl_2]$ alone, or a combination of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ and $[Ru(DMSO)_4Cl_2]$ provide evidence that pre-mixed $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ and 0.13 equiv of $[Ru(DMSO)_4Cl_2]$ is a ca. *2-fold faster, kinetically more competent catalyst to the same products* that were reported for the original "Q_{11}-**1**" with its claimed two Ru atoms. This is an important new piece of evidence, one as important as any reported herein, evidence that either (i) raises the hypothesis that the claimed 2-Ru containing "Q_{11}-**1**" may actually be a mixture of low levels of Ru plus the parent

polyoxoanion (and since their catalysis is basically the same), or (ii) alternatively suggests that if **1** exists, then the 2 framework-incorporated Ru atoms are not relevant to catalysis.

- Another conclusion is that the *catalysis* is an important additional handle, one that allows an independent, albeit indirect, probe of the amount of Ru^{n+} and its location. Such an independent handle would be valuable in other areas where metal content and location in late-transition-metal-substituted polyoxometalates remains controversial.[19e]
- The radical-chain mechanism previously elucidated for "Q_{11}-**1**" in our prior work[8] still appears to be operative. That mechanism can now be updated slightly by replacing the "Ru_2^{III}" and "$Ru_2^{III/IV}$" species in Scheme 3 elsewhere[8] with a more general Ru^{n+} and Ru^{n+1} species in the chain mechanism (see Figure S1 of the present Supporting Information for the slightly revised radical-chain mechanistic scheme).

Overall, we find no independent evidence for Ru_2-containing "**1**." We emphasize that the evidence herein does not mean that **1** cannot or even does not exist, but just that (a) *no reliable synthesis of it exists*. Another important implication of our results is that if authentic Ru_2-containing **1** were made with its framework-incorporated[28] Ru_2, it might still be an inferior oxidation catalyst vs simply beginning with a physical mixture of low levels of $[Ru(DMSO)_4Cl_2]$ plus 1.0 equiv $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$.

An overall message of the present and our prior[8] work is that claims of Ru and other late-metal substituted polyoxometalates will need to consider carefully the

problems and issues present in getting pure samples of fully substituted, non-disordered, non-late-metal contaminated products. Then, catalysis with those precatalyst complexes[13,14,15,16,17] will need to rely on the appropriate kinetic and other, ideally operando,[17] studies needed to rule out the conceivable catalyst possibilities en route to providing evidence for the true catalyst and true reaction type. In oxidation chemistry, ruling out "omni-present autoxidation"[29] continues to be job #2, while obeying Platt's scientific method of the disproof of all alternative hypotheses[30] continues to be job #1 here and in science in general.

Experimental

Materials. All reaction solutions were prepared under oxygen and moisture-free conditions in a Vacuum Atmosphere drybox (O_2 level $\leq$ 5 ppm, as continuously monitored by an oxygen sensor). 1,2-Dichloroethane (Aldrich, HPLC grade) was dried with preactivated 4 Å molecular sieves and stored in the drybox. Adamantane (Aldrich, 99+%) was used as received and stored in the drybox. Na2WO4•2H2O (Aldrich, 99%), Zn(NO3)2•6H2O (Fisher Chemicals), RuCl3•xH2O (Aldrich, 99.98%), NaCl (Fisher Chemicals, ACS grade) and Aliquat 336 (Aldrich) were used as received. 4-*tert*-butylcatechol (Aldrich, 97%) was stored in the freezer. Deionized water was used for solution preparations.

Instrumentation. Infrared spectra were obtained on a Nicolet 5DX spectrometer using neat samples in a press-fit KBr cell. The 1H nuclear magnetic resonance (NMR) spectra in D2O or CDCl3 were obtained on a Varian Inova (JS-300) NMR spectrometer and referenced to the residual proton impurity in the deuterated solvent. The 183W NMR

spectrum was obtained on a Varian Inova 500 MHz NMR and was referenced to an external Na_2WO_4 sample. The 500 MHz NMR was kept at 52.0 ± 0.2 °C by heating the carrier gas. A high precision (±0.015 psig) oxygen pressure transducer (Omegadyne Inc., model PX02C1-100G10T-OX) connected to an analog-to-digital converter (Omega D1131) and integrated with LabView 6.1 software was used for O2-uptake data collection, all as previously described.31 GC analyses were performed on a Hewlett-Packard 5890 series II gas chromatograph equipped with a FID detector and a SPB-1 capillary column (30 m, 0.25 mm i.d.) with the following temperature program, as before8 for adamantane hydroxylation products: initial temperature, 140 °C (initial time, 4 min.); heating rate, 5 °C/min; final temperature, 180 °C (final time, 3 min.); FID detector temperature, 250 °C; injector temperature, 250 °C. An injection volume of 1 μL was used. Product peaks were identified by comparison to authentic sample peaks. Elemental analyses were performed by Mikroanalytisches Labor Pascher (Remagen-Bandorf, Germany) or, where indicated, by Galbraith Laboratories, Inc (Knoxville, TN). Diffraction data were collected on a Bruker APEX2 diffractometer employing Mo Kα radiation for three independent crystals. Standard Bruker APEX2 control and integration software was employed, and Bruker SHELXTL32 software was used for structure solution, refinement, and graphics. A face-indexed absorption correction yielded results that were not better than obtained using SADABS33 alone. The structure was solved by direct methods and refined by a full-matrix, weighted least-squares process. Residual electron density representing solvent molecules (H_2O) and the countercations (Na^+) were included in the model.

Preparation of Putative "$\{[WZnRu_2(OH)(H_2O)](ZnW_9O_{34})_2\}^{11-}$", 1, Precatalyst. The precursor $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]\cdot xH_2O$ (precrystallization yield 37 g, 27%) was synthesized and recrystallized according to the literature.[3,8] ^{183}W NMR (D_2O) observed at 52 °C [reported[3], estimated from the spectrum provided at 32 °C]: δ +24 [+21], -14.2 [-20], -37 [-55], -77 [-95], -83 [-100], -89 [-105], -147 [-155], -178 [-190], -218 [-230], -313 [-310]. Elemental analysis on a sample dried at room temperature under vacuum yields: Calculated [found]: Na 5.02 [4.99]; W 63.6 [63.0]; O 24.8 [24.6]; Zn 5.95 [5.99].

Similarly, $[Ru(DMSO)_4Cl_2]$ was synthesized and recrystallized according to the literature[10,34] (recrystallized yield 70-75%). The ^1H NMR of recrystallized $[Ru(DMSO)_4Cl_2]$ was identical within experimental error to a recent published preparation.[10] ^1H NMR ($CDCl_3$): δ (major peaks) 2.67, 2.78, 3.37, 3.49, 3.55, and 3.58; literature[10] ($CDCl_3$): δ (major peaks) 2.60, 2.72, 3.32, 3.43, 3.48, and 3.50. Elemental analysis from Galbraith Laboratories, Inc. on a recrystallized sample dried overnight at room temperature under vacuum yields: Calculated [found]: Ru 20.9 [20.7]; C 19.8 [20.0]; H 4.99 [5.23]; S 26.5 [27.0]; Cl 14.6 [14.3].

Next, "$K_{11}[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]\cdot 15H_2O$" was prepared and recrystallized twice from hot water according to the literature[21]: specifically and in brief, 22.5 g (3.7 mmol) of $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]\cdot$46-48 H_2O was dissolved in ~10 mL of deionized water under Ar. Under a flow of Ar, 4.0 g (8.3 mmol) of $[Ru(DMSO)_4Cl_2]$ was added and the solution was refluxed under Ar at 90 °C for 18 hours. Following the reflux, the solution was exposed to air and 2 g (27 mmol) of KCl was added causing an immediate formation of yellow precipitate. The product was

filtered and recrystallized twice from hot water and finally dried under vacuum overnight at room temperature. Yield 4.1 g (18%); literature: 3.7 g (16%)[8] and 24%.[21] Neat IR: 720 (s), 770 (s), 881 (m) 924 (m). Elemental analysis, calculated [found]: K 7.53 [4.69]; Na 0 [0.59]; ***Zn 3.43 [7.13]; Ru 3.54 [0.75]***.

Next, the organic counterpart of the polyoxometalate salt "$Q_{11}[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]$•nQCl" was synthesized via the published method,[7,21] specifically and briefly as follows: equimolar amounts of "$K_{11}[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]$•$15H_2O$" (0.18 mmol) and QCl (2.0 mmol) were combined in deionized water and the product was extracted three times using ~5 mL of 1,2-dichloroethane. The 1,2-dichloroethane was then removed *in vacuo* to yield of the gum-like, red-orange compound. Yield 51%; literature[8] 50-65%. IR data (as a neat sample): 714 (s), 797 (s), 874 (m), 924 (m) cm^{-1}; literature[8]: 723 (s), 766 (s), 871 (m), 921 (m) cm^{-1}. The IR spectrum can be seen in Figure S1 of the Supporting Information. Elemental analysis of the above prepared sample which had been dried under vacuum overnight at room temperature gave the following analysis, calculated [found]: C, 39.91 [39.75]; H, 7.28 [7.10]; N, 1.86 [1.62]; ***Ru, 0.21 [0.21]; Zn, 3.21 [3.50],*** which in turn yields the composition $Q_{12}[WRu_{\sim 0.2}Zn_{\sim 2.8}(H_2O)_2(ZnW_9O_{34})_2]$•QCl. An independent sample of "Q_{11}-**1**" and "K_{11}-**1**" was also generously donated by C. Hill's lab; elemental analysis of the Hill K^+ salt yields: ***Ru, <0.1; Zn, 2.92***; Elemental analysis results of a third independent sample prepared by C.-X. Yin from the Finke group yields the composition, $Q_{12}[WRu_{\sim 0.05}Zn_{\sim 3}(H_2O)_2(ZnW_9O_{34})_2]$•2.5QCl. Calculated [found]: C, 41.93 [41.44]; H, 7.64 [7.92]; N, 1.96 [1.83]; O, 10.8 [10.3]; W, 33.6 [32.7]; ***Ru, 0.05 [0.05]; Zn, 3.15 [3.49]***. Unless otherwise indicated, the "K_{11}-**1**" and "Q_{11}-**1**" samples

prepared at Colorado State University by A. M. Morris were used in subsequent experiments.

Preparations of crystals of "$Na_{11}[WZnRu^{III}_2(OH)(H_2O)(ZnW_9O_{34})_2]\cdot 42H_2O$."

X-ray quality crystals of "$Na_{11}[WZnRu^{III}_2(OH)(H_2O)(ZnW_9O_{34})_2]\cdot 42H_2O$" were grown according to the literature procedure.[2] Briefly, "$K[WZnRu^{III}_2(OH)(H_2O)(ZnW_9O_{34})_2]\cdot 15H_2O$" was dissolved in a 0.5 M solution of NaCl and allowed to slowly evaporate in air at room temperature. After approximately one week, spear-shaped, clear-orange crystals were obtained. Elemental analysis of this crystalline sample gives: calculated [found]: Na, 5.06 [4.62]; Zn, 5.99 [5.99]; W, 64.0 [62.8]; O, 24.4 [24.2], consistent with. the composition, $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]\cdot 13H_2O$.

Preparation of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]\cdot nQCl$.

$Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]\cdot nQCl$ was prepared in the analogous manner to the literature preparation of "$Q_{11}[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]\cdot nQCl$."[7,21] Specifically, 410-520 mg (70-90 μmol) of $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]\cdot 46\text{-}48H_2O$ was dissolved in ca. 10 mL of deionized water with heating and stirring. This solution was cooled to room temperature and 320-420 mg (0.8-1.0 mmol) of Aliquat 336 (Aldrich) was added. The solution was then extracted three times with ca. 5 mL of 1,2-dichloroethane, each time collecting the lower, organic layer. The 1,2-dichloroethane solvent was then removed *in vacuo* resulting in a colorless to orange gel. Yield 260-380 mg (37-38%). IR data (as a neat sample): 722 (s), 765 (s), 873 (m), and 923 (m) cm^{-1}. The full IR spectrum can be seen in Figure S2 of the Supporting Information.

Preparation of the ^{183}W NMR Sample. 1.0 g (0.18 mmol) of "K_{11}-**1**" was dissolved with heating and stirring in ca. 3 mL of D_2O. To this solution, 0.17 g (1.6 mmol) of $Li(ClO_4)_2$ was added in order to exchange the potassium cation and improve solublility. The solution was filtered to remove the solid $KClO_4$ and transferred into a 10 mm NMR tube. The solution was kept at or above 50 °C at all times in order to prevent the product from recrystallizing. ^{183}W NMR (D_2O, 52 °C): δ +25.2, -17.6, -40.0, -81.5, -83.7, -88.8, -149.7, -183.1, -222.9, -298.5; literature (D_2O, 22 °C):[3] δ (estimated from spectrum) +20, -20, -55, -95, -100, -105, -155, -190, -230, -310.

General Procedures for Oxygen-Uptake Experiments. Adamantane hydroxylation was monitored by oxygen pressure loss on a computer-interfaced, high precision O_2-pressure transducer apparatus (*vide supra*). The reaction flask was a pressurized Fischer-Porter bottle attached via Swagelock quick-connects and flexible stainless steel tubing to both an oxygen tank and to the pressure transducer. (A schematic of this general apparatus, but where instead a H_2-compatible pressure transducer is instead used, is provided elsewhere.[31]) In the drybox, adamantane (ca. 815 ± 8 mg, 6.0 mmol) was weighed into a 5 dram vial and transferred using the 1,2-dichloroethane solvent into a new 22 mm x 175 mm Pyrex culture tube along with a new 5/8 in. x 5/16 in. Teflon stir bar. The precatalyst, "Q_{11}-**1**" (ca. 60 ± 5 mg, 6.0 μmol) was dissolved in ca. 5 mL of 1,2-dichloroethane, and then quantitatively transferred into the culture tube via a disposable pipet. The remaining solvent (total = 12 mL) was added to the culture tube and the culture tube was then placed inside the Fischer-Porter bottle, sealed, and brought out of the drybox. The bottle was connected via the quick-connects, placed in a temperature controlled oil bath (80 °C), and stirring was initiated. The solution was

equilibrated with stirring at 80 °C for 40 minutes under the inert N_2 atmosphere of the Fischer-Porter bottle. The Fischer-Porter bottle was then purged 15 times with ~14 psig of O_2; 15 s/purge, equilibrate 1 min 15 s; 5 min total time elapsed before the pressure recordings were initiated. The reaction vessel pressure was set at 14 ± 2 psig and $t = 0$ was set. The reactions shown here were stopped after 24 hours following literature precedent unless otherwise noted.[1,2] At the end of the reaction, GC analysis was used for the determination of the final products following the hydroxylation. The Fischer-Porter bottle was vented, opened in air, and the contents were poured into a 5 dram vial. A 0.2 mL sample from the vial was dissolved in 8.2 mL of 1,2-dichloroethane (41× dilution) and used for the GC analyses.

O_2-Uptake Experiments with $[Ru(DMSO)_4Cl_2]$. The same procedure as above ("General Procedures for Oxygen-Uptake Experiments") was used except, instead of dissolving 60 mg of the precatalyst ,"Q_{11}-**1**," in ca. 5 mL of 1,2-dichloroethane, 0.1 mL of a solution containing 3.6 mg (7.7 μmol) of $[Ru(DMSO)_4Cl_2]$ in 1.0 mL of 1,2-dichloroethane was added. This $[Ru(DMSO)_4Cl_2]$ solution permitted 0.77 μmol of Ru to be added to the reaction which is the average amount (0.13 equiv) of Ru present in the "Q_{11}-**1**" sample by elemental analysis from the two samples prepared in our laboratories (entries 1 and 6, Table 4).

O_2-Uptake Experiments with $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$•nQCl. The same procedure as above ("General Procedures for Oxygen-Uptake Experiments") was used except instead of 60 mg of "Q_{11}-**1**," 60 ± 4 mg (~6 μmol) of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$•nQCl was dissolved in ca. 5 mL of 1,2-dichloroethane and the reaction

time was lengthened to ~120 hours since no O_2 uptake was observed after the normal 24 hours.

O_2-Uptake Experiments with Pre-mixed $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$•nQCl and $[Ru(DMSO)_4Cl_2]$. The same procedure as above ("General Procedures for Oxygen-Uptake Experiments") was used except instead of 60 mg of "Q_{11}-**1**," 60 ± 3 mg of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$•nQCl and 0.1 mL of the $[Ru(DMSO)_4Cl_2]$ solution described in "O_2-Uptake Experiments with $[Ru(DMSO)_4]Cl_2$" were combined and dissolved in ca. 5 mL of 1,2-dichloroethane. In a separate experiment, the pre-mixed $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$•nQCl and $[Ru(DMSO)_4Cl_2]$ solution was monitored by its O_2 uptake and at t = 8.3 hours, 35 mg (0.21 mmol) of the previously employed[2,8] radical trap, 4-*tert*-butylcatechol, was added. The results are shown in Figure S8 of the Supporting Information.

Acknowledgment. The authors would like to thank Ms. Susie Miller for assistance with the crystal structure data collection and determination, as well as Dr. Travis M. Anderson and Prof. Craig L. Hill for supplying their samples of "K_{11}-**1**" and "Q_{11}-**1**". Prof. Kenji Nomiya and his co-workers[11] are also gratefully acknowledged for allowing us to cite in the Supporting Information the results and details of their six attempted preparations of **1**. This work was supported by NSF grant CHE 9531110.

Supporting Information Available: IR spectrum of "Q_{11}-**1**" prepared herein; IR spectrum of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$; Crystallographic data table for the structure solved herein in the alternative $P2_1/c$ space group and comparison to Neumann and

Tourné's structures; Adamantane O_2-uptake curve using $[Ru(DMSO)_4Cl_2]$; Adamantane O_2-uptake curves using $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$; Adamantane O_2-uptake curve and discussion of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ plus $[Ru(DMSO)_4Cl_2]$; Experimental, O_2-uptake, and discussion of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ plus 1 equivalent of Zn; Experimental, O_2 uptake, and discussion of QCl plus $[Ru(DMSO)_4Cl_2]$; Adamantane O_2-uptake and discussion of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ plus $[Ru(DMSO)_4Cl_2]$ with added 4-tert-butylcatechol; Slightly updated mechanism for the radical chain initiated adamantane hydroperoxylation plus concurrent Ru-catalyzed ROOH-based adamantane reaction; Details of six attempts to prepare **1** by Nomiya and co-workers; Crystallographic tables of atomic coordinates, isotropic displacement, bond lengths, and bond angles for $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ solved herein.

References

[1] Neumann, R.; Dahan, M. *Nature* **1997**, *388*, 353-355.

[2] Neumann, R.; Dahan, M. *J. Am. Chem. Soc.* **1998**, *120*, 11969-11976.

[3] Tourné, C. M.; Tourné, G. F.; Zonnevijlle, F. *J. Chem. Soc. Dalton Trans.* **1991**, 143-155.

[4] Hayaishi, O.; Katagiri, M.; Rothberg, S. J. *J. Am. Chem. Soc.* **1955**, *77*, 5450-5451.

[5] Nozaki, M. *Top. Curr. Chem.* **1979**, *78*, 145-186.

[6] Funabiki, T., Ed. Oxygenases and model systems. In *Catalysis of Metal Complexes*; Vol. 19; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1997.

[7] Neumann, R.; Khenkin, A. M.; Dahan, M. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1587-1589.

[8] Yin, C.-X.; Finke, R. G. *Inorg. Chem.* **2005**, *44*, 4175-4188.

[9] Howells, A. R.; Sankarraj, A.; Shannon, C. *J. Am. Chem. Soc.* **2004**, *126*, 12258-12259.

[10] Nomiya, K.; Torii, H.; Nomura, K.; Sato, Y. *J. Chem. Soc. Dalton Trans.* **2001**, 1506-1521.

[11] Nomiya, K.; Torii, H.; Nomura, K.; Sato, Y. These worker's results, cited with permission in the Supporting Information herein for the first time and, detail six attempted preparations of **1** by the reported procedure[7,21] followed by characterization of "yellow-brown materials, ...obtained sometimes as a mixture with crystals, which were characterized by IR, ^{1}H NMR, TG/DTA and UV-visible absorption spectra, and magnetic susceptibility plus CV measurements." In no case could a pure sample of the putative **1** be obtained, however, and "the IR spectra were almost unchanged after the reaction, that is, they only showed a pattern due to the original Keggin fragments". For the full details, see the Supporting Information (see also the *Note Added in Proof* in our earlier publication[8]).

[12] Hill, C. L.; Anderson, T. M.; Hardcastle, K.; Fang, X., cited with permission.

[13] (a) Yin, C.-X.; Sasaki, Y.; Finke, R. G. *Inorg. Chem.* **2005**, *44*, 8521-8530. (b) Yin, C.-X.; Finke, R. G. *J. Am. Chem. Soc.* **2005**, *127*, 9003-9013. (c) Yin, C.-X.; Finke, R. G. *J. Am. Chem. Soc.* **2005**, *127*, 13988-13996.

[14] (a) Widegren, J. A.; Finke, R. G. *J. Mol. Catal. A* **2003**, *198*, 317-341. (b) Phan, N. T. S.; Van Der Sluys, M.; Jones, C. W. *Adv. Synth. Catal.* **2006**, *348*, 609-679.

[15] (a) Lin, Y.; Finke, R. G. *Inorg. Chem.* **1994**, *33*, 4891-4910. (b) Widegren, J. A.; Bennett, M. A.; Finke, R. G. *J. Am. Chem. Soc.* **2003**, *125*, 10301-10310. (c) Hagen, C. M.; Widegren, J. A.; Maitlis, P. M.; Finke, R. G. *J. Am. Chem. Soc.* **2005**, *127*, 4423-4432. (d) Finney, E. E.; Finke, R. G. *Inorg. Chim. Acta* **2006**, *359*, 2879-2887.

[16] (a) Halpern, J.; Okamoto, T.; Zakhariev, A. *J. Mol. Catal.* **1976**, *2*, 65-68. (b) Halpern, J.; Riley, D. P.; Chan, A. S. C.; Pluth, J. J. *J. Am. Chem. Soc.* **1977**, *99*, 8055-8057.

[17] For this reason, so-called operando (from the Latin for "working" or "operating") methods, along with (by definition already operando) kinetic studies, are leading the way to the identification of true active catalysts: (a) Weckhuysen, B. M. *Chem. Commun.* **2002**, 97-110. (b) Guerrero-Perez, M. O.; Banares, M. A. *Chem. Commun.* **2002**, 1292-1293. (c) Meunier, F.; Daturi, M. *Catal. Today* **2006**, *113*, 1-2.

[18] (a) Rong, C. Y.; Pope, M. T. *J. Am. Chem. Soc.* **1992**, *114*, 2932-2938. (b) Randall, W. J.; Weakley, T. J. R.; Finke, R. G. *Inorg. Chem.* **1993**, *32*, 1068-1071. (c) Filipek, K. *Inorg. Chim. Acta* **1995**, *231*, 237-239. (d) Chen, S.-W.; Villaneau, R.; Li, Y.; Chamoreau, L.-M.; Boubekeur, K.; Thouvenot, R.; Gouzerh, P.; Proust, A. *Eur. J. Inorg. Chem.* **2008**, 2137-2142.

[19] For example, Rh, Pd, and Pt containing polyoxoanions: (a) Kuznetsova, N. I.; Detusheva, L. G.; Kuznetsoca, L. I.; Fefotov, M. A.; Likholobov, V. A. *J. Mol. Catal. A* **1996**, *114*, 131-139. (b) Wei, X.; Bachman, R. E.; Pope, M. T. *J. Am. Chem. Soc.* **1998**, *120*, 10248-10253. (c) Bi, L.-H.; Reicke, M.; Kortz, U.; Keita, B.; Nadjo, L.; Clark, R. J. *Inorg. Chem.* **2004**, *43*, 3915-3920. (d) Kortz, U.; Bi, L. U.S. Pat. Appl. Publ. **2007**, US 2006-445073 20060531. (e) See also the following two citations which independently reveal the non-trivial nature of making, and then adequately characterizing, late-transition-metal-substituted polyoxometalates: (i) Cao, R.; Anderson, T. M.; Hillesheim, D. A.; Kögerler, P.; Hardcastle, K. I.; Hill, C. L. *Angew. Chem. Int. Ed.* **2008**, *47*, 9380-9382. (ii) Kortz, U.; Lee, U.; Joo, H.-C.; Park, K.-M.; Mal, S. S.; Dickman, M. H.; Jameson, G. B. *Angew. Chem. Int. Ed.* **2008**, *47*, 9383-9384.

[20] (a) Camenzind, M. J.; James, B. R.; Dolphin, D. *J. Chem. Soc., Chem. Commun.* **1986**, 1137-1139. (b) Ke, M.; Rettig, S. J.; James, B. R.; Dophin, D. *J. Chem. Soc., Chem. Commun.* **1987**, 1110-1112. (c) Collman, J. P.; Rose, E.; Venburg, G. D. *J. Chem. Soc., Chem. Commun.* **1994**, 11-12. (d) Cheng, S. Y. S.; Rajapakse, N.; Rettig, S. J.; James B. R. *J. Chem. Soc., Chem. Commun.* **1994**, 2669-2670. (e) Reboucas, J. S.; Cheu, E. L. S.; Ware, C. J.; James, B. R.; Skov, K. A. *Inorg. Chem.* **2008**, *47*, 7894-7907.

[21] Neumann, R.; Khenkin, A. M. *Inorg. Chem.* **1995**, *34*, 5753-5760.

[22] Interestingly, the transition metal substituted derivatives of $[WZnM_2(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ with M= Mn(II), Co(II), Ni(II), and Cu(II), have previously been synthesized using "temperatures near the boiling point, a large [8-12 fold molar] excess of M^{II}, and, typically, a long heating period.[3]" Other derivatives such as M= Mn(III), Fe(II) or Fe(III), Pd(II), Pt(II), and V(IV) di-substituted derivatives have also been formed using molar to slightly excess molar M precursors. In the cases of M(III) substitution, the previous syntheses either begin with a M(III) precursor or the M(II) substituted polyoxoanion is first prepared and then oxidized the M(II) to M(III).[3] However, in the synthesis of "**1**",[7,21] a molar ratio of $[Ru(DMSO)_4]Cl_2$ is used even though Ru has historically been difficult to substitute into polyoxoanions[10] or other ligands,[20] and a Ru(II) precursor is used rather than a Ru(III) or Ru(IV) precursor. Overall, the proposed synthesis of "**1**" appears to be inconsistent with the synthetic precedent[3] that might allow **1** to be formed with 2 Ru(III) atoms incorporated into the structure in high yield.

[23] Meier, U. C.; Scopelliti, R.; Solari, E.; Merbach, A. E. *Inorg. Chem.* **2000**, *39*, 3816-3822.

[24] Pope, M. T. Preparation, structural principles, properties and applications. In *Heteropoly and Isopoly Oxometalates*; Springer-Verlag: New York, 1983, pp. 15-32.

[25] Lyon, D. K.; Miller, W. K.; Novet, T.; Domaille, P. J.; Evitt, E.; Johnson, D. C.; Finke, R. G. *J. Am. Chem. Soc.* **1991**, *113*, 7209-7221.

[26] Brenner, I. B.; Erlich, S. *Appl. Spectrosc.* **1984**, *38*, 887-890.

[27] (a) Finke, R. G.; Droege, M. W. *J. Am. Chem. Soc.* **1984**, *106*, 7274-7277. (b) Rapko, B. M.; Pohl, M.; Finke, R. G. *Inorg. Chem.* **1994**, *33*, 3625-3634. (c) Pohl, M.; Lin, Y.; Weakley, T. J. R.; Nomiya, K.; Kaneko, M.; Weiner, H.; Finke, R. G. *Inorg. Chem.* **1995**, *34*, 767-777. (d) Nagata, T.; Pohl, M.; Weiner, H.; Finke, R. G. *Inorg. Chem.* **1997**, *36*, 1366-1377. (e) Finke, R. G., Polyoxoanions in Catalysis: From Record Catalytic Lifetime Nanocluster Catalysis to Record Catalytic Lifetime Catechol Dioxygenase Catalysis. In *Polyoxometalate Chemistry*; Pope, M. T.; Müller, A., Eds.; Kluwer: Netherlands, 2001; pp 363-390.

[28] For the definitions of, and differences between, POM-supported vs POM-incorporated transition metal catalysts, see: Weiner, H.; Hayashi, Y.; Finke, R. G. *Inorg. Chem.* **1999**, *38*, 2579-2591.

[29] Limburg, C. *Angew. Chem., Int. Eng. Ed.* **2003**, *42*, 5932-5954.

[30] Platt, J. R. *Science*, **1964**, *146*, 347. As Platt notes, "for exploring the unknown, there is no faster way".

[31] Lin, Y.; Finke, R. G. *Inorg. Chem.* **1994**, *33*, 4891-4910.

[32] Sheldrick, G. M. *Acta Cryst. A* **2008**, *64*, 112-122.

[33] Sheldrick, G. M. *SADABS* (a program for Siemens Area Detection Absorption Correction), 2000.

[34] Evans, I. P.; Spencer, A.; Wilkinson, G. *J. Chem. Soc. Dalton Trans.* **1973**, 204-209.

Supporting Information for:

Re-Investigation of a Ru_2-Incorporated Polyoxometalate Dioxygenase Precatalyst, "$[WZnRu_2^{III}(H_2O)(OH)(ZnW_9O_{34})_2]^{11-}$": Evidence For Marginal, ≤0.2 Equivalents of Ru Incorporation Plus Faster Catalysis by Physical Mixtures of $[Ru^{II}(DMSO)_4Cl_2]$ and the Parent Polyoxometalate $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$

Aimee M. Morris, Oren P. Anderson, and Richard G. Finke

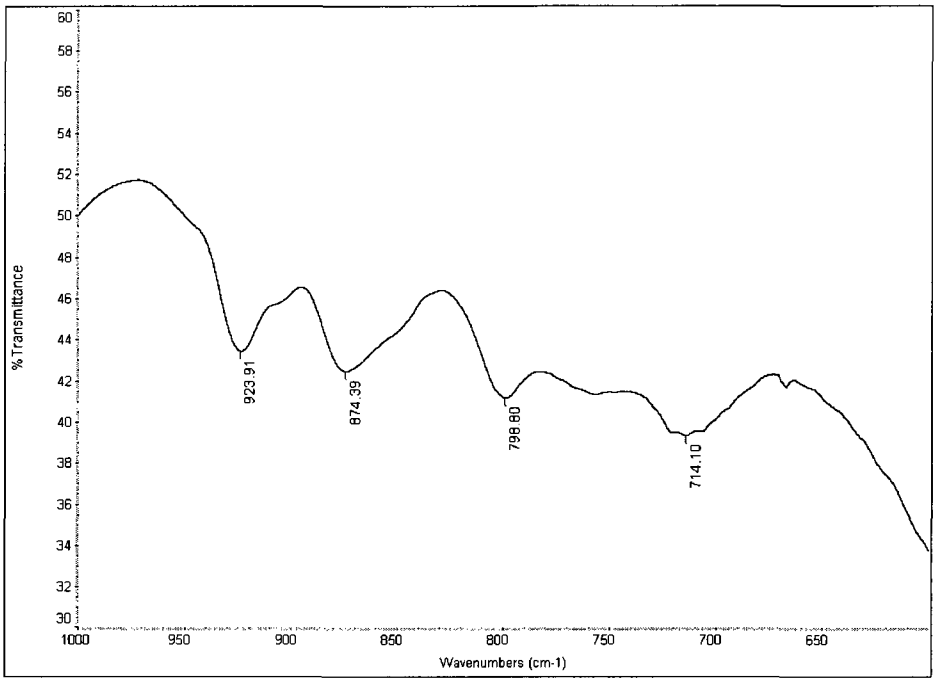


Figure S1. IR spectrum of neat "Q_{11}-**1**" prepared herein.

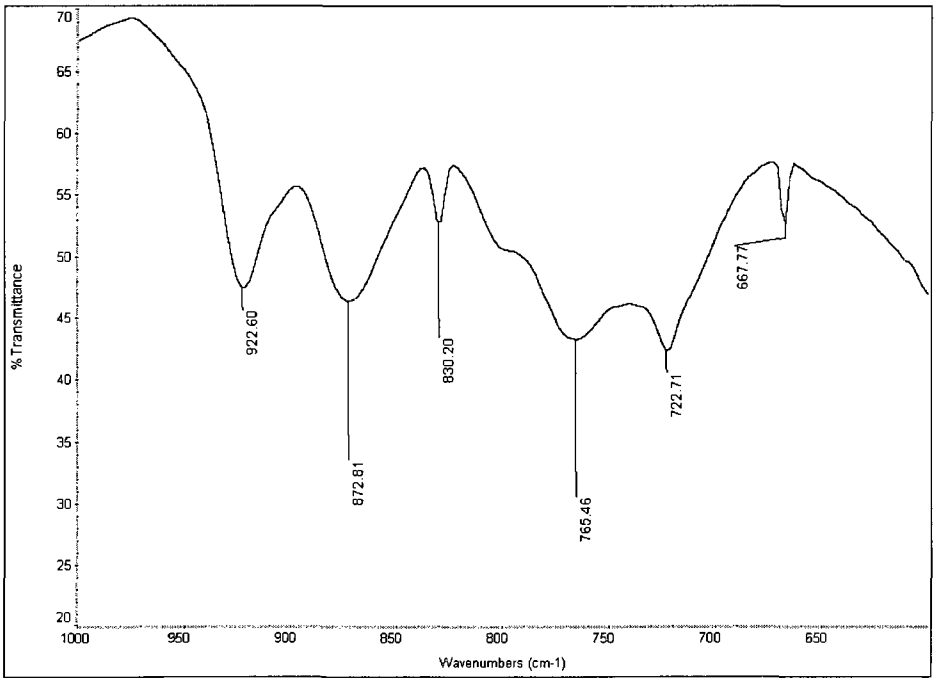


Figure S2. IR spectrum of neat $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$

Figures S1 and S2 show that similar peaks are observed for the "Q_{11}-**1**" and $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ prepared herein. Specifically, they both contain peaks at 923-924 cm^{-1} and 873-874 cm^{-1}. However, in the ~800 to ~700 cm^{-1} region, differences in the broad peaks between Figures S1 and S2 are seen.

Table S1. Crystallographic data for the structure solved herein in the alternative $P2_1/c$ space group compared to structures previously solved by Neumann[2] and Tourné.[1]

	$Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ $\cdot 34H_2O$	Neumann's[2] $Na_{11}[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]$ $\cdot 42H_2O$	Tourné's[1] $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ $\cdot 46H_2O$
Space Group	$P2_1/c$	$P2_1/c$	$P2_1/c$
M	5814.4	6023.9	6047.9
a/Å	13.0017(6)	13.069(3)	13.027(4)
b/Å	17.8485(7)	17.827(4)	17.788(5)
c/Å	24.0914(11)	24.182(5)	24.124(4)
B/°	119.381(3)	118.97(2)	118.94(2)
U/Å^3	4871.6(4)	4929(1)	4892(2)
Crystal size /mm	0.067 x 0.166 x 0.263	NR[b]	0.21 x 0.16 x 0.34
R	NA[a]	0.051	0.039

[a] NA means not applicable as the structure was not solved in the above $P2_1/c$ space group. [b] NR stands for not reported in the previous publications.

The results in Table S1 show the similarities exhibited for the unit cell constants for the structures of $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ solved herein and by Tourné, as well as Neumann's $Na_{11}[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]$ as discussed further in the main text.

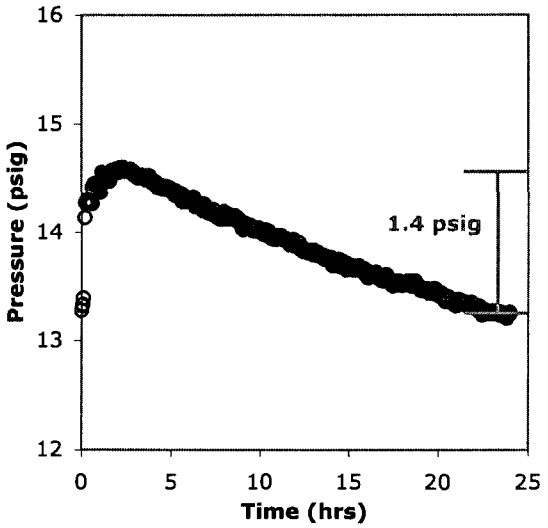


Figure S3. Pressure vs. time curve recorded by an O_2 pressure transducer. The initial pressure rise is due to the solvent pressure equilibration following the addition of O_2 to the reaction flask. Reaction conditions: 6.0 mmol of adamantane, 0.74 μmol of $[Ru(DMSO)_4Cl_2]$, 12 mL of 1,2-$C_2H_4Cl_2$, 80 °C, and ca. 14 psig of O_2 pressure. The net pressure difference of 1.4 ± 0.1 psig corresponds to 0.39 ± 0.02 mmol of O_2 that was taken up over 24 hours. The product yields from this O_2-uptake are given and discussed in the main text.

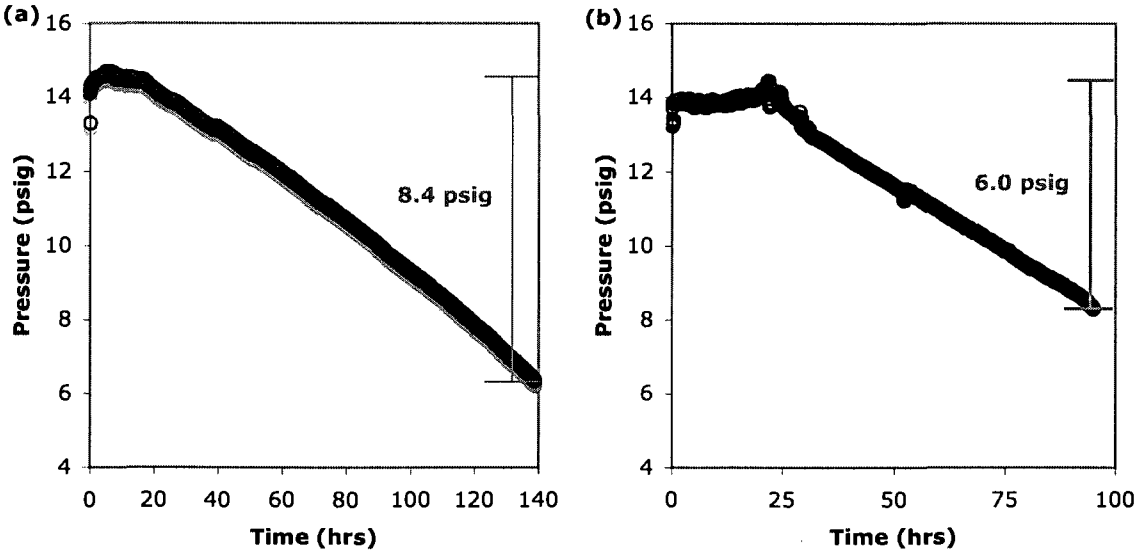


Figure S4. Two different pressure vs. time curves recorded by an O_2 pressure transducer for the reaction conditions: 6.0 mmol of adamantane, 6.0 μmol of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$, 12 mL of 1,2-$C_2H_4Cl_2$, 80 °C, and ca. 14 psig of O_2 pressure. The two figures illustrate the length and variation in induction period times and amount of O_2 that was consumed for the Ru-free polyoxoanion. The initial pressure rise in (a) and (b) are due to the solvent pressure equilibration following the addition of O_2 to

the reaction flask. The small fluctuations in pressure arise from temperature fluctuations during the long reaction times. The product yields were determined by GC and averaging the two curves at the end of the reaction, ~120 hours; these results are given and discussed in the main text.

Experimental Conditions for the O_2-Uptake Experiments with Pre-mixed $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$•nQCl and Zn. The same procedure as described in "O_2-Uptake Experiments with $Q_{12}[Zn_3W(H_2O)_2(ZnW_9O_{34})_2]$•nQCl" of the main text was used, only 6 μmol (2 equiv) of Zn metal was also added to the ca. 5 mL of 1,2-dichloroethane inside the drybox and the reaction time was lengthened to 95 hours.

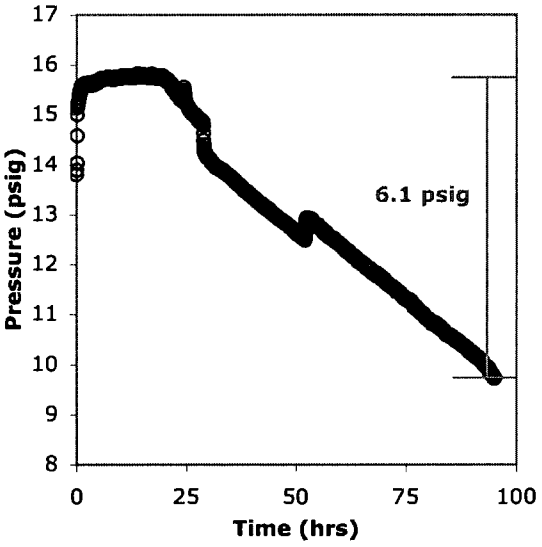


Figure S5. Pressure vs. time curves recorded by an O_2 pressure transducer. The initial pressure rise is due to the solvent pressure equilibration following the addition of O_2 to the reaction flask. The features seen at ~25 and ~50 hours are most likely due to in temperature fluctuations and temperature gradients present during the long reaction times. Reaction conditions: 6.0 mmol of adamantane, 6.0 μmol of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$, 6 μmol of Zn, 12 mL of 1,2-$C_2H_4Cl_2$, 80 °C, and ca. 14 psig of O_2 pressure. The important part of the Figure is the net pressure drop of 6.1 psig as shown.

O_2-Uptake Studies with Pre-mixed $Q_{12}[Zn_3W(H_2O)_2(ZnW_9O_{34})_2]$•nQCl + Zn.

In the literature report,[2] it is claimed that the addition of $Zn(0)_n$ to the reaction containing "Q_{11}-**1**," adamantane, and solvent eliminates the induction period observed when no Zn is added. Contrarily, in Yin et al's previous report,[3] they observed the exact opposite effect, namely that the addition of Zn causes a *lengthening* in the induction period of the reaction. When Zn was added to the reaction containing $Q_{12}[Zn_3W(H_2O)_2(ZnW_9O_{34})_2]$ as part of the present work, the same lengthening of the induction period (from ~10 hours to ~30 hours) was also observed, Figures S4a and S4b). This finding is consistent with the previous hypothesis[3] that Zn may be acting as a weak radical inhibitor, thereby causing a longer induction period to be observed.

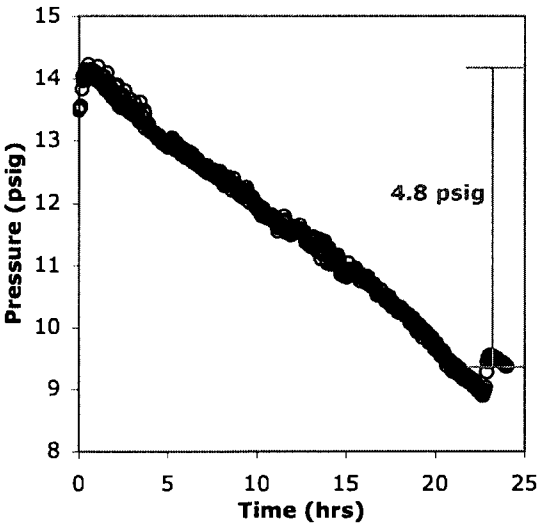


Figure S6. Pressure vs. time curves recorded by an O_2 pressure transducer. The initial pressure rise is due to the solvent pressure equilibration following the addition of O_2 to the reaction flask. The small fluctuations in pressure arise from temperature fluctuations. Reaction conditions: 6.0 mmol of adamantane, 6.0 μmol of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$, 0.74 μmol of $[Ru(DMSO)_4Cl_2]$, 12 mL of 1,2-$C_2H_4Cl_2$, 80 °C, and ca. 14 psig of O_2 pressure. The premixed $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ and $[Ru(DMSO)_4]Cl_2$ conditions allow for more O_2 uptake to occur in the 24 hour period (4.8 psig) in comparison to $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ alone (0-2.2 psig), $[Ru(DMSO)_4Cl_2]$ alone (1.4 psig), or "Q_{11}-**1**" (2.0 ± 0.2 psig).

Experimental Conditions for Pre-mixed QCl + $[Ru(DMSO)_4Cl_2]$. The same procedure as described in "O_2-Uptake Experiments with $[Ru(DMSO)_4Cl_2]$" of the publication was used except 67 μmol of QCl (equivalent to the amount of Q^+ present in 6 μmol of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$) was added to the ca. 5 mL of 1,2-dichloroethane inside the drybox.

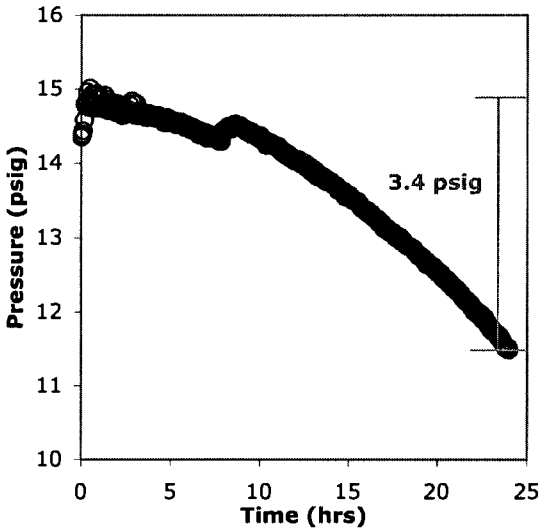


Figure S7. Pressure vs. time curve recorded by an O_2 pressure transducer. The initial pressure rise is due to the solvent pressure equilibration following the addition of O_2 to the reaction flask. The small fluctuations in pressure arise from temperature fluctuations. Reaction conditions: 6.0 mmol of adamantane, 67 μmol of QCl, 0.74 μmol of $[Ru(DMSO)_4Cl_2]$, 12 mL of 1,2-$C_2H_4Cl_2$, 80 °C, and ca. 14 psig of O_2 pressure.

O_2-Uptake Studies with Pre-mixed QCl + $[Ru(DMSO)_4Cl_2]$. In order to test the possibility that the increased yields and shortening of the induction period is not caused by *QCl* + $[Ru(DMSO)_4Cl_2]$ rather than by the pre-mixed polyoxoanion, $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$, and $[Ru(DMSO)_4Cl_2]$, the control experiment in Figure S7 was run using 0.13 equivs $[Ru(DMSO)_4Cl_2]$ plus 1.0 equiv QCl. As can be seen Figure

S7, there is an increase in the O_2 uptake in the 24 hour period to 3.4 psig (vs. 1.4 psig from $[Ru(DMSO)_4Cl_2]$ alone, Figure S3), but a slight decrease from the pre-mixed $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ and $[Ru(DMSO)_4Cl_2]$ (4.8 psig) shown in Figure S6. In short, added QCl without the parent polyoxoanion does have some affect on the product yields.

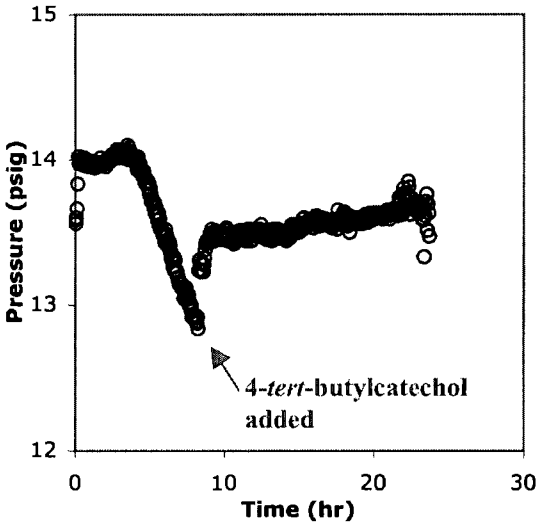


Figure S8. Pressure vs. time curves recorded by a high precision pressure transducer. The initial pressure rise and re-rise following the addition of 4-tert-butylcatechol are due to the solvent pressure equilibration following the addition of O_2 to the reaction flask. The small fluctuations in pressure arise from temperature fluctuations. Reaction conditions: 6.0 mmol of adamantane, 6.0 μmol of $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$, 0.74 μmol $[Ru(DMSO)_4Cl_2]$, 12 mL of 1,2-$C_2H_4Cl_2$, 80 °C, and ca. 14 psig of O_2 pressure. At t=8.3 hours, 0.21 mmol of 4-*tert*-butylcatechol was added under a positive flow of O_2. The important observation here is that adding 4-*tert*-butylcatechol completely stops the previously active adamantane oxidation catalysis.

O_2-Uptake Studies with Premixed $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ + $[Ru(DMSO)_4Cl_2]$ Along with the Addition of the Radical Inhibitor 4-tert-butylcatechol. The radical inhibitor, 4-tert-butylcatechol is added to the O_2-uptake of

premixed $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ + $[Ru(DMSO)_4Cl_2]$ at t=8.3 hours which is after the uptake of O_2 has begun. Figure S8 provides evidence that premixed $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ plus $[Ru(DMSO)_4Cl_2]$ are operating via a radical mechanism. This is an important control, as "$Q_{11}[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]$" was shown previously[3] to operate via a radical mechanism. Taken with the results in the main text, the implication is that $Q_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$•nQCl plus trace amounts of Ru^{n+} are a highly plausible explanation for the previously observed[3] autoxidation catalysis.

Scheme S1. Slightly revised mechanism for the radical-chain autoxidation of adamantane plus concurrent Ru-catalyzed ROOH-based adamantane reaction; adapted from reference 3. The mechanism is identical to that previously given,[3] except that the Ru atoms are not necessarily incorporated into the polyoxoanion but can be trace Ru^{n+} species the current evidence suggests.

Initiation

(1) $Adm-H + Ru^{n+} + O_2 \xrightarrow{k_{1a}} Ru^{n+/n+1}OOH + \underset{\text{(mostly 1-adm•; but some 2-adm•)}}{Adm•}$

Propagation

(2) $Adm• + O_2 \xrightarrow[\text{fast}]{k_2} AdmOO•$

(3) $AdmOO• + Adm-H \xrightarrow{k_3} \underset{\text{(primary autoxidation product)}}{AdmOOH} + Adm•$

Concurrent Ru - Catalyzed AdmOOH - Based Reaction

(4) $AdmOOH + Ru^{n+1} \xrightarrow{k_4} AdmOO• + H^+ + Ru^{n+}$

(5) $AdmOOH + Ru^{n+} \xrightarrow{k_5} AdmO• + OH^- + Ru^{n+1}$

(6) $OH^- + H^+ \xrightarrow[\text{fast}]{k_6} H_2O$

(7) $AdmO• + AdmOOH \xrightarrow{k_7} AdmOH + AdmOO•$

(8) $AdmO• + Adm-H \xrightarrow{k_8} AdmOH + Adm•$

(9) $2AdmOO• \xrightarrow{k_9} 2AdmO• + O_2$

Termination

(10a) $2AdmOO• \xrightarrow{k_{10}} AdmOOAdm + O_2$

(10b) $2AdmOO• \xrightarrow{k'_{10}} 1\text{-}AdmOH + 2\text{-}Adm=O + O_2$
(possible when one AdmOO• is 2-AdmOO•)

As detailed elsewhere,[3] this mechanism accounts quantitatively for the radical-chain kinetics, and steady-state rate law (equation S1, where "Ru source" is for example equal to "$Q_{11}\{WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2\}$") observed.[3]

$$\frac{d[1-\text{adamantanol}]}{dt} = k_{obs}[\text{adamantane}]^{3/2}[\text{"Ru source"}]^{1/2}[O_2]^{\approx 1/2} \qquad \textbf{(S1)}$$

The Results of Six Attempted Preparations of 1 by Nomiya and Co-Workers.[4] We thank Prof. Nomiya and his colleagues for sharing their following results and permitting us to cite them herein for the first time. The write-up which follows is from Prof. Nomiya.

In the preparation of the Neumann compound, $K_{11}[WZnRu^{III}_2(OH)(H_2O)(ZnW_9O_{34})_2]$,[5] the two precursor complexes, that is, *cis*-$[RuCl_2(DMSO)_4]$ [6a] and $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$,[6b] were prepared without any problems. The former was characterized with FT-IR, CH analysis, TG/DTA, ^1H and ^{13}C NMR, and the latter with FT-IR, ^{183}W NMR, TG/DTA, and cyclic voltammetry (CV) measurements. The resultant characterization data were satisfactory and in accord with the literature.[6a,b]

Using these well-characterized precursors, we have repeated (6 times) the preparation of the Neumann compound according to the reported procedure; "the solution was stirred for 18 h at 90 °C under argon and cooled and opened to air", that is, we have performed the preparation by refluxing the aqueous solution, with the only difference that we performed our studies under *N_2 gas*. Yellow-brown materials were obtained, sometimes as a mixture with crystals, with that product then being characterized by IR, ^1H NMR, TG/DTA and UV-visible absorption spectra, and magnetic susceptibility plus CV measurements. However, the IR spectra were almost unchanged after the reaction, that is, *they only showed a pattern due to the original Keggin fragments.* The ^1H NMR

spectra in D_2O showed that the materials contained several S-bonding DMSO molecules showing signals at 3.44, 3.47, 3.51, 3.68, 3.69, 3.71 and 4.00 ppm, and O-bonding DMSO molecule showing a signal at 2.78 ppm. (The original paper[5] does not report the C, H, or S analyses nor ^1H NMR measurements.) ^{183}W NMR measurements in D_2O were unsuccessful, because of the insufficient concentration of the saturated solution of the K_{11} salt. TG/DTA showed the presence of ca. 30 water. The UV-visible absorption spectra did not show the reported peak at 430 nm due to the reported[5] O → Ru charge transfer. Furthermore, apparent magnetic moments of our six materials were observed in the range of 0.32 - 1.15 BM (specifically, 0.32, 0.53, 0.65, 0.81, 0.93 and 1.15 BM), with estimated diamagnetic corrections based on the assumed formula of "$K_{11}[WZnRu^{III}_2(OH)(OH_2)(ZnW_9O_{34})_2]\cdot 30H_2O$ (MW 5985)". These results suggest that the products we obtain are almost diamagnetic. Also, in the CV measurements of these compounds, we did not find reversible redox peaks due to $Ru^{III/II}$ and/or $Ru^{IV/III}$ at positive potentials, as reported.[5]

We have examined the parallel control experiments based on the reactions of *cis*-$[RuCl_2(DMSO)_4]$ in aqueous solution with the closely related, two sandwich-polyoxotungstates constructed with two Keggin or two Dawson fragments, both with P heteroatoms, specifically $K_{10}[Zn_4(H_2O)_2(PW_9O_{34})_2]$ and $Na_{16}[Zn_4(H_2O)_2(P_2W_{15}O_{56})_2]$, respectively.[7] These two Zn_4 derivatives have been prepared according to the literature and characterized with FT-IR, TG/DTA, ^{31}P and ^{183}W NMR, and CV measurements. Using Neumann's procedure (as well as with other modifications) we obtained many materials. However, neither di-Ru^{III}-substituted paramagnetic compounds nor mono-Ru^{III} compounds were never obtained in pure form. In every case, including the attempted

preparation of Neumann's compound, a color change from colorless to yellow-brown was observed. Probably, a partial or incomplete substitution occurs, but the magnetic susceptibility measurements showed that the products were always close to diamagnetic. These results were also consistent with ^{31}P NMR spectra characteristic of diamagnetic species, that is, the two phosphorus resonances were almost unchanged after the reactions.

In short, all compounds we have obtained according to the literature method[5] are a mixture of materials with incomplete substitution. The only difference in our synthesis is the fact that our experiments have been performed under *N_2 gas*, while the previous work under *Ar gas*. Our results indicate that two-site substitution of the WZn_3 moiety with *cis*-$[RuCl_2(DMSO)_4]$, or even one-site substitution, will be very difficult under the reported conditions. While our work does not deny the existence of the di-Ru^{III}-substituted sandwich polyoxotungstate $K_{11}[WZnRu^{III}_2(OH)(H_2O)(ZnW_9O_{34})_2]$, it certainly echos the findings of Finke and co-workers as well as Hill and co-workers reported in the main text that the reported synthesis[5] does not produce pure, di-Ru substituted product if *cis*-$[RuCl_2(DMSO)_4]$ is used as the ruthenium source.

Crystallographic Data For $Na_{12}[WZn_3(H_2O)_2(ZnW_9O_{34})_2]$ or Na_{12}-1. Shown below in Tables S1-S3 are the crystallographic data tables from the structure solved and discussed in the main text of Na_{12}-**1**.

Table S1. Crystal data and structure refinement for Na_{12}-**1**.

Empirical formula	$H_{82}\ Na_{12}\ O_{109}\ W_{19}\ Zn_5$	
Formula weight	5922.51	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	P2(1)/n	
Space group	P2(1)/n	
Unit cell dimensions	a = 13.0017(6) Å	a= 90°.
	b = 17.8485(7) Å	b= 93.223(3)°.
	c = 21.0260(9) Å	g = 90°.
Volume	4871.6(4) $Å^3$	
Z	2	
Density (calculated)	4.012 Mg/m^3	
Absorption coefficient	23.716 mm^{-1}	
F(000)	5210	
Crystal size	0.263 x 0.166 x 0.067 mm^3	
Theta range for data collection	1.50 to 34.34°.	
Index ranges	-20<=h<=20, -28<=k<=28, -33<=l<=33	
Reflections collected	84169	
Independent reflections	20295 [R(int) = 0.0630]	
Completeness to theta = 36.36°	98.8 %	
Absorption correction	Multi	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	20295 / 0 / 690	
Goodness-of-fit on F^2	1.120	
Final R indices [I>2sigma(I)]	R1 = 0.0509, wR2 = 0.1079	
R indices (all data)	R1 = 0.0717, wR2 = 0.1153	
Largest diff. peak and hole	6.596 and -4.092 $Å^{-3}$	

Table S2. Atomic coordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$x 10^3) for Na_{12}-**1**. U(eq) is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	U(eq)
Na(1)	10289(4)	895(2)	8191(2)	19(1)
Na(2)	8596(4)	2268(2)	9045(2)	19(1)
Na(3)	10599(4)	6184(3)	8229(2)	25(1)
Na(4)	10448(4)	7468(2)	7152(2)	19(1)
Na(5)	10455(4)	9393(2)	6972(2)	20(1)
Na(6)	12058(13)	609(10)	5020(8)	67(5)
Na(7)	11434(14)	4154(15)	9039(9)	104(9)
O(1)	6735(6)	6065(4)	5324(3)	16(1)
O(2)	7978(6)	5202(4)	6356(3)	16(1)
O(3)	7015(6)	4488(4)	5192(3)	15(1)
O(4)	8497(6)	3811(4)	4281(3)	15(1)
O(5)	10660(5)	3842(4)	4455(3)	13(1)
O(6)	11602(6)	4523(4)	5609(3)	17(1)
O(7)	10398(5)	5163(3)	6509(3)	13(1)
O(8)	9343(5)	4680(3)	5333(3)	8(1)
O(9)	7214(6)	4877(4)	7524(3)	15(1)
O(10)	6423(5)	4210(4)	6405(3)	12(1)
O(11)	5361(6)	3556(4)	5352(3)	18(1)
O(12)	7291(6)	2998(4)	5018(3)	15(1)
O(13)	7495(5)	2531(4)	3768(3)	15(1)
O(14)	9608(6)	2624(4)	4024(3)	13(1)
O(15)	11756(6)	2524(4)	4124(3)	16(1)
O(16)	11574(5)	3050(3)	5387(3)	11(1)
O(17)	13304(6)	3646(5)	6022(3)	20(2)
O(18)	11981(6)	4280(4)	6864(3)	15(1)
O(19)	10947(6)	4958(4)	7816(3)	15(1)
O(20)	9149(6)	4346(4)	7193(3)	16(1)
O(21)	8081(5)	3620(3)	6174(3)	10(1)
O(22)	9446(5)	2882(3)	5187(3)	9(1)
O(23)	10454(5)	3612(3)	6390(3)	9(1)
O(24)	6584(5)	2719(4)	6198(3)	13(1)
O(25)	8451(6)	1753(3)	4772(3)	14(1)
O(26)	10558(5)	1782(3)	4962(3)	13(1)
O(27)	11965(6)	2788(4)	6679(3)	18(1)
O(28)	10771(6)	3446(4)	7587(3)	14(1)
O(29)	7499(5)	3381(3)	7278(3)	11(1)
O(30)	8365(6)	2021(4)	6040(3)	14(1)
O(31)	10363(6)	1991(4)	6232(3)	14(1)
O(32)	9201(6)	2634(4)	7039(3)	16(1)
O(33)	7367(6)	1833(4)	7186(3)	16(1)
O(34)	9394(6)	763(3)	5606(4)	19(2)
O(35)	10953(6)	1890(4)	7530(3)	15(1)
O(36)	10419(6)	2005(4)	8950(4)	18(1)
O(37)	8496(6)	1323(4)	8279(3)	19(1)
O(38)	9537(6)	542(4)	7095(3)	16(1)
O(39)	12098(6)	688(4)	8294(4)	20(2)
O(40)	10391(6)	-533(4)	8066(3)	17(1)

O(41)	10053(7)	352(5)	9291(4)	24(2)
O(42)	11867(6)	-503(4)	5642(3)	18(1)
O(43)	9089(8)	2733(6)	10067(4)	34(2)
O(44)	8865(7)	3279(5)	8378(4)	26(2)
O(45)	11808(6)	6648(4)	7517(5)	31(2)
O(46)	9443(6)	6385(4)	7254(4)	18(1)
O(47)	12078(12)	6548(6)	9027(6)	57(4)
O(48)	10311(6)	7514(4)	8296(4)	20(2)
O(49)	9166(6)	8387(5)	6950(4)	21(2)
O(50)	11748(6)	8445(4)	7121(3)	18(1)
O(51)	10151(9)	9381(5)	5875(4)	37(2)
O(52)	13100(20)	420(11)	3893(16)	90(13)
O(53)	11891(9)	10191(6)	6705(6)	50(3)
O(54)	11043(10)	3432(6)	10152(5)	51(3)
O(55)	8910(30)	5849(13)	8655(13)	34(7)
O(56)	9680(20)	5582(16)	8976(11)	30(7)
O(57)	10690(50)	5420(18)	9270(20)	100(20)
W(1)	9522(6)	4611(4)	4375(3)	8(1)
W(2)	10770(1)	2785(1)	4604(1)	9(1)
W(3)	11983(1)	3666(1)	6089(1)	8(1)
W(4)	10601(1)	4418(1)	7158(1)	10(1)
W(5)	7742(1)	4442(1)	6876(1)	10(1)
W(6)	6647(1)	3677(1)	5619(1)	11(1)
W(7)	8274(1)	2769(1)	4430(1)	10(1)
W(8)	9414(1)	1731(1)	5532(1)	11(1)
W(9)	10622(1)	2594(1)	6993(1)	10(1)
W(10)	7749(1)	2582(1)	6728(1)	10(1)
Zn(1)	9469(17)	4600(12)	4383(10)	8(1)
Zn(2)	9339(1)	3722(1)	5755(1)	7(1)
Zn(3)	8056(1)	5342(1)	5386(1)	16(1)

Table S3. Bond lengths [Å] and angles [°] for Na_{12}-**1**.

Na(1)-O(39)	2.378(9)
Na(1)-O(35)	2.442(8)
Na(1)-O(37)	2.471(9)
Na(1)-O(38)	2.533(9)
Na(1)-O(41)	2.543(9)
Na(1)-O(36)	2.545(9)
Na(1)-O(40)	2.565(8)
Na(1)-Na(5)#1	3.723(6)
Na(1)-Na(2)	3.812(6)
Na(2)-O(44)	2.324(9)
Na(2)-O(37)	2.331(9)
Na(2)-O(43)	2.359(10)
Na(2)-O(15)#2	2.435(9)
Na(2)-O(36)	2.436(9)
Na(2)-O(1)#3	2.573(8)
Na(3)-O(56)	2.29(2)
Na(3)-O(45)	2.380(10)
Na(3)-O(19)	2.406(8)
Na(3)-O(48)	2.409(9)
Na(3)-O(55)	2.49(3)
Na(3)-O(46)	2.498(10)
Na(3)-O(47)	2.563(16)
Na(3)-O(57)	2.58(4)
Na(3)-Na(4)	3.220(6)
Na(3)-Na(7)	4.12(2)
Na(4)-O(46)	2.351(9)
Na(4)-O(49)	2.359(10)
Na(4)-O(45)	2.389(10)
Na(4)-O(48)	2.424(8)
Na(4)-O(50)	2.432(9)
Na(4)-O(14)#4	2.476(7)
Na(4)-Na(5)	3.456(6)
Na(5)-O(40)#5	2.310(8)
Na(5)-O(51)	2.318(10)
Na(5)-O(50)	2.393(9)
Na(5)-O(38)#5	2.394(8)
Na(5)-O(53)	2.439(12)
Na(5)-O(49)	2.454(9)
Na(5)-Na(1)#5	3.723(6)
Na(6)-O(42)	2.396(17)
Na(6)-O(47)#6	2.80(2)
Na(6)-O(52)	2.82(3)
Na(6)-O(26)	2.86(2)
Na(6)-Na(7)#6	3.75(3)
Na(7)-O(42)#7	2.35(2)
Na(7)-O(57)	2.52(5)
Na(7)-O(54)	2.745(18)
Na(7)-O(19)	2.981(15)
Na(7)-Na(6)#7	3.75(3)
O(1)-Zn(3)	2.147(7)
O(1)-Na(2)#8	2.573(8)
O(2)-W(5)	1.780(6)

O(2)-Zn(3)	2.061(7)
O(3)-W(6)	1.783(7)
O(3)-Zn(3)	2.064(8)
O(4)-Zn(1)	1.90(2)
O(4)-W(7)	1.912(6)
O(4)-W(1)	1.956(10)
O(5)-W(2)	1.918(6)
O(5)-W(1)	2.017(10)
O(5)-Zn(1)	2.05(2)
O(5)-Zn(3)#4	2.227(7)
O(6)-W(3)	1.883(7)
O(6)-Zn(1)#4	2.10(2)
O(6)-W(1)#4	2.129(11)
O(6)-Zn(3)#4	2.177(7)
O(7)-W(1)#4	1.910(10)
O(7)-W(4)	1.912(6)
O(7)-Zn(1)#4	1.94(2)
O(8)-Zn(2)	1.926(6)
O(8)-W(1)#4	2.013(9)
O(8)-Zn(1)	2.02(2)
O(8)-W(1)	2.045(9)
O(8)-Zn(3)	2.057(6)
O(8)-Zn(1)#4	2.07(2)
O(9)-W(5)	1.741(7)
O(10)-W(6)	1.942(7)
O(10)-W(5)	1.976(7)
O(11)-W(6)	1.747(8)
O(12)-W(7)	1.873(7)
O(12)-W(6)	1.971(7)
O(13)-W(7)	1.727(7)
O(14)-W(2)	1.908(7)
O(14)-W(7)	1.992(7)
O(14)-Na(4)#4	2.476(7)
O(15)-W(2)	1.739(7)
O(15)-Na(2)#9	2.435(9)
O(16)-W(3)	1.892(6)
O(16)-W(2)	1.959(7)
O(17)-W(3)	1.731(8)
O(18)-W(4)	1.946(7)
O(18)-W(3)	1.964(6)
O(19)-W(4)	1.726(7)
O(20)-W(4)	1.897(8)
O(20)-W(5)	1.920(8)
O(21)-Zn(2)	1.911(6)
O(21)-W(5)	2.144(6)
O(21)-W(6)	2.146(7)
O(21)-W(10)	2.243(6)
O(22)-Zn(2)	1.926(6)
O(22)-W(7)	2.151(7)
O(22)-W(2)	2.176(6)
O(22)-W(8)	2.181(6)
O(23)-Zn(2)	1.925(7)
O(23)-W(3)	2.122(7)
O(23)-W(4)	2.163(6)

O(23)-W(9)	2.219(6)
O(24)-W(10)	1.847(7)
O(24)-W(6)	2.103(6)
O(25)-W(7)	1.959(6)
O(25)-W(8)	1.973(7)
O(26)-W(8)	1.964(7)
O(26)-W(2)	1.967(6)
O(27)-W(9)	1.932(8)
O(27)-W(3)	2.000(7)
O(28)-W(4)	1.962(6)
O(28)-W(9)	1.970(7)
O(29)-W(10)	1.876(6)
O(29)-W(5)	2.104(6)
O(30)-W(8)	1.853(7)
O(30)-W(10)	1.966(7)
O(31)-W(8)	1.924(7)
O(31)-W(9)	1.942(7)
O(32)-W(9)	1.857(8)
O(32)-W(10)	1.965(8)
O(33)-W(10)	1.736(6)
O(34)-W(8)	1.733(6)
O(35)-W(9)	1.728(7)
O(38)-Na(5)#1	2.394(8)
O(40)-Na(5)#1	2.310(8)
O(42)-Na(7)#6	2.35(2)
O(47)-Na(6)#7	2.80(2)
O(55)-O(56)	1.27(4)
O(56)-O(57)	1.45(7)
W(1)-O(7)#4	1.910(10)
W(1)-O(8)#4	2.013(9)
W(1)-O(6)#4	2.129(11)
W(1)-Zn(3)#4	3.161(7)
W(1)-W(1)#4	3.165(15)
W(1)-Zn(1)#4	3.184(15)
W(5)-W(6)	3.2339(5)
Zn(1)-O(7)#4	1.94(2)
Zn(1)-O(8)#4	2.07(2)
Zn(1)-O(6)#4	2.10(2)
Zn(1)-W(1)#4	3.184(15)
Zn(3)-O(6)#4	2.177(7)
Zn(3)-O(5)#4	2.227(7)
Zn(3)-W(1)#4	3.161(7)

O(39)-Na(1)-O(35)	77.6(3)
O(39)-Na(1)-O(37)	166.8(3)
O(35)-Na(1)-O(37)	100.5(3)
O(39)-Na(1)-O(38)	111.9(3)
O(35)-Na(1)-O(38)	78.1(3)
O(37)-Na(1)-O(38)	80.1(3)
O(39)-Na(1)-O(41)	91.6(3)
O(35)-Na(1)-O(41)	149.2(3)
O(37)-Na(1)-O(41)	83.5(3)
O(38)-Na(1)-O(41)	132.3(3)
O(39)-Na(1)-O(36)	91.9(3)

O(35)-Na(1)-O(36)	77.1(3)
O(37)-Na(1)-O(36)	75.0(3)
O(38)-Na(1)-O(36)	140.6(3)
O(41)-Na(1)-O(36)	74.5(3)
O(39)-Na(1)-O(40)	78.4(3)
O(35)-Na(1)-O(40)	130.0(3)
O(37)-Na(1)-O(40)	111.7(3)
O(38)-Na(1)-O(40)	71.3(3)
O(41)-Na(1)-O(40)	74.0(3)
O(36)-Na(1)-O(40)	146.7(3)
O(39)-Na(1)-Na(5)#1	81.7(2)
O(35)-Na(1)-Na(5)#1	95.5(2)
O(37)-Na(1)-Na(5)#1	111.5(2)
O(38)-Na(1)-Na(5)#1	39.54(19)
O(41)-Na(1)-Na(5)#1	111.5(2)
O(36)-Na(1)-Na(5)#1	171.2(3)
O(40)-Na(1)-Na(5)#1	37.73(18)
O(39)-Na(1)-Na(2)	130.6(2)
O(35)-Na(1)-Na(2)	91.8(2)
O(37)-Na(1)-Na(2)	36.2(2)
O(38)-Na(1)-Na(2)	112.6(2)
O(41)-Na(1)-Na(2)	73.4(2)
O(36)-Na(1)-Na(2)	39.01(19)
O(40)-Na(1)-Na(2)	136.2(2)
Na(5)#1-Na(1)-Na(2)	147.65(17)
O(44)-Na(2)-O(37)	98.6(3)
O(44)-Na(2)-O(43)	103.6(4)
O(37)-Na(2)-O(43)	152.3(4)
O(44)-Na(2)-O(15)#2	96.0(3)
O(37)-Na(2)-O(15)#2	98.1(3)
O(43)-Na(2)-O(15)#2	95.9(3)
O(44)-Na(2)-O(36)	85.4(3)
O(37)-Na(2)-O(36)	79.6(3)
O(43)-Na(2)-O(36)	85.8(3)
O(15)#2-Na(2)-O(36)	177.5(3)
O(44)-Na(2)-O(1)#3	173.9(3)
O(37)-Na(2)-O(1)#3	75.4(3)
O(43)-Na(2)-O(1)#3	82.4(3)
O(15)#2-Na(2)-O(1)#3	84.2(3)
O(36)-Na(2)-O(1)#3	94.2(3)
O(44)-Na(2)-Na(1)	96.0(2)
O(37)-Na(2)-Na(1)	38.7(2)
O(43)-Na(2)-Na(1)	121.4(3)
O(15)#2-Na(2)-Na(1)	136.5(2)
O(36)-Na(2)-Na(1)	41.1(2)
O(1)#3-Na(2)-Na(1)	79.8(2)
O(56)-Na(3)-O(45)	169.1(10)
O(56)-Na(3)-O(19)	86.4(7)
O(45)-Na(3)-O(19)	86.9(3)
O(56)-Na(3)-O(48)	109.4(8)
O(45)-Na(3)-O(48)	78.7(3)
O(19)-Na(3)-O(48)	161.9(3)
O(56)-Na(3)-O(55)	30.4(10)
O(45)-Na(3)-O(55)	159.6(8)

O(19)-Na(3)-O(55)	95.7(6)
O(48)-Na(3)-O(55)	94.1(6)
O(56)-Na(3)-O(46)	108.4(9)
O(45)-Na(3)-O(46)	79.8(3)
O(19)-Na(3)-O(46)	87.2(3)
O(48)-Na(3)-O(46)	79.5(3)
O(55)-Na(3)-O(46)	80.1(8)
O(56)-Na(3)-O(47)	94.0(8)
O(45)-Na(3)-O(47)	80.0(4)
O(19)-Na(3)-O(47)	108.5(4)
O(48)-Na(3)-O(47)	79.9(4)
O(55)-Na(3)-O(47)	118.0(8)
O(46)-Na(3)-O(47)	153.5(4)
O(56)-Na(3)-O(57)	33.9(15)
O(45)-Na(3)-O(57)	136.0(15)
O(19)-Na(3)-O(57)	80.0(7)
O(48)-Na(3)-O(57)	118.0(7)
O(55)-Na(3)-O(57)	64.2(17)
O(46)-Na(3)-O(57)	140.3(15)
O(47)-Na(3)-O(57)	65.2(14)
O(56)-Na(3)-Na(4)	143.2(9)
O(45)-Na(3)-Na(4)	47.6(3)
O(19)-Na(3)-Na(4)	113.5(2)
O(48)-Na(3)-Na(4)	48.4(2)
O(55)-Na(3)-Na(4)	113.8(6)
O(46)-Na(3)-Na(4)	46.5(2)
O(47)-Na(3)-Na(4)	107.1(3)
O(57)-Na(3)-Na(4)	166.4(7)
O(56)-Na(3)-Na(7)	56.0(9)
O(45)-Na(3)-Na(7)	113.5(4)
O(19)-Na(3)-Na(7)	45.5(4)
O(48)-Na(3)-Na(7)	151.6(4)
O(55)-Na(3)-Na(7)	81.8(6)
O(46)-Na(3)-Na(7)	126.7(4)
O(47)-Na(3)-Na(7)	77.4(4)
O(57)-Na(3)-Na(7)	35.5(8)
Na(4)-Na(3)-Na(7)	157.0(4)
O(46)-Na(4)-O(49)	101.4(3)
O(46)-Na(4)-O(45)	82.7(3)
O(49)-Na(4)-O(45)	170.5(4)
O(46)-Na(4)-O(48)	82.2(3)
O(49)-Na(4)-O(48)	93.7(3)
O(45)-Na(4)-O(48)	78.3(3)
O(46)-Na(4)-O(50)	169.5(3)
O(49)-Na(4)-O(50)	89.0(3)
O(45)-Na(4)-O(50)	86.8(3)
O(48)-Na(4)-O(50)	95.3(3)
O(46)-Na(4)-O(14)#4	92.9(3)
O(49)-Na(4)-O(14)#4	83.4(3)
O(45)-Na(4)-O(14)#4	105.0(3)
O(48)-Na(4)-O(14)#4	173.8(3)
O(50)-Na(4)-O(14)#4	90.1(3)
O(46)-Na(4)-Na(3)	50.4(2)
O(49)-Na(4)-Na(3)	129.5(3)

O(45)-Na(4)-Na(3)	47.4(2)
O(48)-Na(4)-Na(3)	48.0(2)
O(50)-Na(4)-Na(3)	120.9(3)
O(14)#4-Na(4)-Na(3)	130.7(2)
O(46)-Na(4)-Na(5)	146.4(3)
O(49)-Na(4)-Na(5)	45.2(2)
O(45)-Na(4)-Na(5)	129.6(3)
O(48)-Na(4)-Na(5)	94.3(2)
O(50)-Na(4)-Na(5)	43.8(2)
O(14)#4-Na(4)-Na(5)	87.5(2)
Na(3)-Na(4)-Na(5)	141.64(17)
O(40)#5-Na(5)-O(51)	167.8(4)
O(40)#5-Na(5)-O(50)	88.6(3)
O(51)-Na(5)-O(50)	101.7(3)
O(40)#5-Na(5)-O(38)#5	78.3(3)
O(51)-Na(5)-O(38)#5	93.2(3)
O(50)-Na(5)-O(38)#5	160.5(3)
O(40)#5-Na(5)-O(53)	105.4(4)
O(51)-Na(5)-O(53)	82.3(4)
O(50)-Na(5)-O(53)	84.4(4)
O(38)#5-Na(5)-O(53)	85.2(4)
O(40)#5-Na(5)-O(49)	89.9(3)
O(51)-Na(5)-O(49)	84.1(4)
O(50)-Na(5)-O(49)	87.7(3)
O(38)#5-Na(5)-O(49)	106.5(3)
O(53)-Na(5)-O(49)	162.5(4)
O(40)#5-Na(5)-Na(4)	87.0(2)
O(51)-Na(5)-Na(4)	95.7(3)
O(50)-Na(5)-Na(4)	44.7(2)
O(38)#5-Na(5)-Na(4)	146.6(3)
O(53)-Na(5)-Na(4)	127.8(3)
O(49)-Na(5)-Na(4)	43.0(2)
O(40)#5-Na(5)-Na(1)#5	42.8(2)
O(51)-Na(5)-Na(1)#5	132.7(3)
O(50)-Na(5)-Na(1)#5	119.1(2)
O(38)#5-Na(5)-Na(1)#5	42.3(2)
O(53)-Na(5)-Na(1)#5	79.2(3)
O(49)-Na(5)-Na(1)#5	118.3(2)
Na(4)-Na(5)-Na(1)#5	129.78(15)
O(42)-Na(6)-O(47)#6	98.9(6)
O(42)-Na(6)-O(52)	115.5(8)
O(47)#6-Na(6)-O(52)	118.8(10)
O(42)-Na(6)-O(26)	122.6(7)
O(47)#6-Na(6)-O(26)	80.6(6)
O(52)-Na(6)-O(26)	114.2(8)
O(42)-Na(6)-Na(7)#6	37.5(5)
O(47)#6-Na(6)-Na(7)#6	81.9(5)
O(52)-Na(6)-Na(7)#6	95.7(8)
O(26)-Na(6)-Na(7)#6	149.8(7)
O(42)#7-Na(7)-O(57)	94.4(19)
O(42)#7-Na(7)-O(54)	95.4(6)
O(57)-Na(7)-O(54)	99.3(11)
O(42)#7-Na(7)-O(19)	105.8(6)
O(57)-Na(7)-O(19)	70.8(11)

O(54)-Na(7)-O(19)	157.1(8)
O(42)#7-Na(7)-Na(6)#7	38.3(5)
O(57)-Na(7)-Na(6)#7	58.1(17)
O(54)-Na(7)-Na(6)#7	89.5(7)
O(19)-Na(7)-Na(6)#7	101.6(7)
O(42)#7-Na(7)-Na(3)	96.3(7)
O(57)-Na(7)-Na(3)	36.6(9)
O(54)-Na(7)-Na(3)	135.1(9)
O(19)-Na(7)-Na(3)	35.2(3)
Na(6)#7-Na(7)-Na(3)	74.2(6)
Zn(3)-O(1)-Na(2)#8	128.8(4)
W(5)-O(2)-Zn(3)	136.1(4)
W(6)-O(3)-Zn(3)	133.7(4)
Zn(1)-O(4)-W(7)	144.1(8)
Zn(1)-O(4)-W(1)	1.4(9)
W(7)-O(4)-W(1)	143.4(5)
W(2)-O(5)-W(1)	137.0(4)
W(2)-O(5)-Zn(1)	135.2(7)
W(1)-O(5)-Zn(1)	1.8(8)
W(2)-O(5)-Zn(3)#4	124.8(3)
W(1)-O(5)-Zn(3)#4	96.2(3)
Zn(1)-O(5)-Zn(3)#4	97.8(6)
W(3)-O(6)-Zn(1)#4	139.2(7)
W(3)-O(6)-W(1)#4	137.8(4)
Zn(1)#4-O(6)-W(1)#4	1.8(7)
W(3)-O(6)-Zn(3)#4	122.9(3)
Zn(1)#4-O(6)-Zn(3)#4	95.5(6)
W(1)#4-O(6)-Zn(3)#4	96.4(3)
W(1)#4-O(7)-W(4)	145.9(4)
W(1)#4-O(7)-Zn(1)#4	2.0(9)
W(4)-O(7)-Zn(1)#4	145.6(7)
Zn(2)-O(8)-W(1)#4	116.0(4)
Zn(2)-O(8)-Zn(1)	113.3(7)
W(1)#4-O(8)-Zn(1)	104.3(7)
Zn(2)-O(8)-W(1)	113.8(3)
W(1)#4-O(8)-W(1)	102.5(4)
Zn(1)-O(8)-W(1)	2.0(8)
Zn(2)-O(8)-Zn(3)	117.5(3)
W(1)#4-O(8)-Zn(3)	101.9(3)
Zn(1)-O(8)-Zn(3)	101.9(7)
W(1)-O(8)-Zn(3)	103.1(3)
Zn(2)-O(8)-Zn(1)#4	116.0(7)
W(1)#4-O(8)-Zn(1)#4	1.4(8)
Zn(1)-O(8)-Zn(1)#4	103.2(8)
W(1)-O(8)-Zn(1)#4	101.3(7)
Zn(3)-O(8)-Zn(1)#4	102.9(7)
W(6)-O(10)-W(5)	111.3(3)
W(7)-O(12)-W(6)	152.6(4)
W(2)-O(14)-W(7)	112.6(3)
W(2)-O(14)-Na(4)#4	128.7(3)
W(7)-O(14)-Na(4)#4	117.3(3)
W(2)-O(15)-Na(2)#9	139.3(4)
W(3)-O(16)-W(2)	155.4(4)
W(4)-O(18)-W(3)	112.2(3)

W(4)-O(19)-Na(3)	138.7(4)
W(4)-O(19)-Na(7)	117.1(7)
Na(3)-O(19)-Na(7)	99.3(6)
W(4)-O(20)-W(5)	155.7(4)
Zn(2)-O(21)-W(5)	117.8(3)
Zn(2)-O(21)-W(6)	119.0(3)
W(5)-O(21)-W(6)	97.9(3)
Zn(2)-O(21)-W(10)	120.9(3)
W(5)-O(21)-W(10)	98.9(2)
W(6)-O(21)-W(10)	97.8(3)
Zn(2)-O(22)-W(7)	117.3(3)
Zn(2)-O(22)-W(2)	119.9(3)
W(7)-O(22)-W(2)	97.2(2)
Zn(2)-O(22)-W(8)	121.6(3)
W(7)-O(22)-W(8)	97.6(2)
W(2)-O(22)-W(8)	98.2(2)
Zn(2)-O(23)-W(3)	118.0(3)
Zn(2)-O(23)-W(4)	118.8(3)
W(3)-O(23)-W(4)	98.5(3)
Zn(2)-O(23)-W(9)	121.8(3)
W(3)-O(23)-W(9)	98.2(3)
W(4)-O(23)-W(9)	96.7(2)
W(10)-O(24)-W(6)	113.6(3)
W(7)-O(25)-W(8)	111.9(3)
W(8)-O(26)-W(2)	113.8(3)
W(8)-O(26)-Na(6)	118.5(4)
W(2)-O(26)-Na(6)	124.9(4)
W(9)-O(27)-W(3)	113.2(4)
W(4)-O(28)-W(9)	112.7(3)
W(10)-O(29)-W(5)	113.7(3)
W(8)-O(30)-W(10)	155.5(4)
W(8)-O(31)-W(9)	147.0(4)
W(9)-O(32)-W(10)	157.1(4)
W(9)-O(35)-Na(1)	144.9(4)
Na(2)-O(36)-Na(1)	99.9(3)
Na(2)-O(37)-Na(1)	105.1(3)
Na(5)#1-O(38)-Na(1)	98.1(3)
Na(5)#1-O(40)-Na(1)	99.5(3)
Na(7)#6-O(42)-Na(6)	104.3(9)
Na(3)-O(45)-Na(4)	84.9(3)
Na(4)-O(46)-Na(3)	83.2(3)
Na(3)-O(47)-Na(6)#7	125.0(6)
Na(3)-O(48)-Na(4)	83.5(3)
Na(4)-O(49)-Na(5)	91.7(3)
Na(5)-O(50)-Na(4)	91.5(3)
O(56)-O(55)-Na(3)	66.2(16)
O(55)-O(56)-O(57)	166(3)
O(55)-O(56)-Na(3)	83.4(17)
O(57)-O(56)-Na(3)	84.0(17)
O(56)-O(57)-Na(7)	116(3)
O(56)-O(57)-Na(3)	62.1(18)
Na(7)-O(57)-Na(3)	107.9(12)
O(7)#4-W(1)-O(4)	97.5(4)
O(7)#4-W(1)-O(8)#4	94.9(4)

O(4)-W(1)-O(8)#4	166.7(5)
O(7)#4-W(1)-O(5)	98.4(4)
O(4)-W(1)-O(5)	90.2(4)
O(8)#4-W(1)-O(5)	83.2(4)
O(7)#4-W(1)-O(8)	164.0(5)
O(4)-W(1)-O(8)	91.6(4)
O(8)#4-W(1)-O(8)	77.5(4)
O(5)-W(1)-O(8)	94.7(4)
O(7)#4-W(1)-O(6)#4	86.4(4)
O(4)-W(1)-O(6)#4	93.8(4)
O(8)#4-W(1)-O(6)#4	91.8(4)
O(5)-W(1)-O(6)#4	173.3(5)
O(8)-W(1)-O(6)#4	79.9(3)
O(7)#4-W(1)-Zn(3)#4	92.4(3)
O(4)-W(1)-Zn(3)#4	134.6(4)
O(8)#4-W(1)-Zn(3)#4	39.5(2)
O(5)-W(1)-Zn(3)#4	44.4(2)
O(8)-W(1)-Zn(3)#4	90.5(3)
O(6)#4-W(1)-Zn(3)#4	131.1(4)
O(7)#4-W(1)-W(1)#4	132.5(5)
O(4)-W(1)-W(1)#4	129.6(4)
O(8)#4-W(1)-W(1)#4	39.1(2)
O(5)-W(1)-W(1)#4	88.7(4)
O(8)-W(1)-W(1)#4	38.4(2)
O(6)#4-W(1)-W(1)#4	84.6(4)
Zn(3)#4-W(1)-W(1)#4	61.0(2)
O(7)#4-W(1)-Zn(1)#4	131.5(3)
O(4)-W(1)-Zn(1)#4	130.7(3)
O(8)#4-W(1)-Zn(1)#4	37.9(4)
O(5)-W(1)-Zn(1)#4	88.1(4)
O(8)-W(1)-Zn(1)#4	39.6(4)
O(6)#4-W(1)-Zn(1)#4	85.3(4)
Zn(3)#4-W(1)-Zn(1)#4	59.8(3)
W(1)#4-W(1)-Zn(1)#4	1.3(5)
O(15)-W(2)-O(14)	100.0(3)
O(15)-W(2)-O(5)	102.7(3)
O(14)-W(2)-O(5)	89.7(3)
O(15)-W(2)-O(16)	100.1(3)
O(14)-W(2)-O(16)	159.9(3)
O(5)-W(2)-O(16)	86.0(3)
O(15)-W(2)-O(26)	95.9(3)
O(14)-W(2)-O(26)	89.1(3)
O(5)-W(2)-O(26)	161.3(3)
O(16)-W(2)-O(26)	88.7(3)
O(15)-W(2)-O(22)	168.8(3)
O(14)-W(2)-O(22)	75.5(3)
O(5)-W(2)-O(22)	87.7(3)
O(16)-W(2)-O(22)	84.7(3)
O(26)-W(2)-O(22)	74.0(2)
O(17)-W(3)-O(6)	101.8(4)
O(17)-W(3)-O(16)	99.3(3)
O(6)-W(3)-O(16)	90.0(3)
O(17)-W(3)-O(18)	97.2(3)
O(6)-W(3)-O(18)	88.8(3)

O(16)-W(3)-O(18)	163.3(3)
O(17)-W(3)-O(27)	94.7(3)
O(6)-W(3)-O(27)	163.2(3)
O(16)-W(3)-O(27)	90.9(3)
O(18)-W(3)-O(27)	85.5(3)
O(17)-W(3)-O(23)	166.8(3)
O(6)-W(3)-O(23)	88.5(3)
O(16)-W(3)-O(23)	88.8(3)
O(18)-W(3)-O(23)	74.5(3)
O(27)-W(3)-O(23)	74.7(3)
O(19)-W(4)-O(20)	102.8(3)
O(19)-W(4)-O(7)	101.7(3)
O(20)-W(4)-O(7)	88.7(3)
O(19)-W(4)-O(18)	97.0(3)
O(20)-W(4)-O(18)	160.2(3)
O(7)-W(4)-O(18)	87.4(3)
O(19)-W(4)-O(28)	96.2(3)
O(20)-W(4)-O(28)	90.5(3)
O(7)-W(4)-O(28)	161.9(3)
O(18)-W(4)-O(28)	87.3(3)
O(19)-W(4)-O(23)	168.1(3)
O(20)-W(4)-O(23)	86.5(3)
O(7)-W(4)-O(23)	85.8(3)
O(18)-W(4)-O(23)	74.0(3)
O(28)-W(4)-O(23)	76.0(2)
O(9)-W(5)-O(2)	103.3(3)
O(9)-W(5)-O(20)	100.3(3)
O(2)-W(5)-O(20)	95.0(3)
O(9)-W(5)-O(10)	96.7(3)
O(2)-W(5)-O(10)	91.4(3)
O(20)-W(5)-O(10)	159.9(3)
O(9)-W(5)-O(29)	90.8(3)
O(2)-W(5)-O(29)	165.3(3)
O(20)-W(5)-O(29)	86.7(3)
O(10)-W(5)-O(29)	82.4(3)
O(9)-W(5)-O(21)	161.5(3)
O(2)-W(5)-O(21)	92.8(3)
O(20)-W(5)-O(21)	87.1(3)
O(10)-W(5)-O(21)	73.5(3)
O(29)-W(5)-O(21)	72.7(2)
O(9)-W(5)-W(6)	130.7(3)
O(2)-W(5)-W(6)	84.4(2)
O(20)-W(5)-W(6)	127.85(19)
O(10)-W(5)-W(6)	34.03(19)
O(29)-W(5)-W(6)	83.07(18)
O(21)-W(5)-W(6)	41.09(18)
O(11)-W(6)-O(3)	102.5(3)
O(11)-W(6)-O(10)	98.5(3)
O(3)-W(6)-O(10)	95.0(3)
O(11)-W(6)-O(12)	98.7(3)
O(3)-W(6)-O(12)	92.3(3)
O(10)-W(6)-O(12)	159.3(3)
O(11)-W(6)-O(24)	91.0(3)
O(3)-W(6)-O(24)	166.4(3)

O(10)-W(6)-O(24)	83.9(3)
O(12)-W(6)-O(24)	84.4(3)
O(11)-W(6)-O(21)	162.9(3)
O(3)-W(6)-O(21)	93.6(3)
O(10)-W(6)-O(21)	74.2(2)
O(12)-W(6)-O(21)	86.1(3)
O(24)-W(6)-O(21)	73.0(3)
O(11)-W(6)-W(5)	133.2(2)
O(3)-W(6)-W(5)	87.2(2)
O(10)-W(6)-W(5)	34.7(2)
O(12)-W(6)-W(5)	126.9(2)
O(24)-W(6)-W(5)	84.25(19)
O(21)-W(6)-W(5)	41.06(16)
O(13)-W(7)-O(12)	101.3(3)
O(13)-W(7)-O(4)	101.2(3)
O(12)-W(7)-O(4)	90.6(3)
O(13)-W(7)-O(25)	96.9(3)
O(12)-W(7)-O(25)	91.7(3)
O(4)-W(7)-O(25)	160.9(3)
O(13)-W(7)-O(14)	96.2(3)
O(12)-W(7)-O(14)	162.5(3)
O(4)-W(7)-O(14)	85.0(3)
O(25)-W(7)-O(14)	87.1(3)
O(13)-W(7)-O(22)	168.1(3)
O(12)-W(7)-O(22)	88.4(3)
O(4)-W(7)-O(22)	85.5(3)
O(25)-W(7)-O(22)	75.6(3)
O(14)-W(7)-O(22)	74.4(3)
O(34)-W(8)-O(30)	102.2(3)
O(34)-W(8)-O(31)	100.6(3)
O(30)-W(8)-O(31)	87.3(3)
O(34)-W(8)-O(26)	96.8(3)
O(30)-W(8)-O(26)	161.0(3)
O(31)-W(8)-O(26)	88.9(3)
O(34)-W(8)-O(25)	94.7(3)
O(30)-W(8)-O(25)	90.4(3)
O(31)-W(8)-O(25)	164.7(3)
O(26)-W(8)-O(25)	88.4(3)
O(34)-W(8)-O(22)	165.7(3)
O(30)-W(8)-O(22)	87.5(3)
O(31)-W(8)-O(22)	90.2(3)
O(26)-W(8)-O(22)	73.9(2)
O(25)-W(8)-O(22)	74.6(3)
O(35)-W(9)-O(32)	101.8(3)
O(35)-W(9)-O(27)	99.0(3)
O(32)-W(9)-O(27)	158.7(3)
O(35)-W(9)-O(31)	99.2(3)
O(32)-W(9)-O(31)	86.4(3)
O(27)-W(9)-O(31)	86.3(3)
O(35)-W(9)-O(28)	97.7(3)
O(32)-W(9)-O(28)	90.0(3)
O(27)-W(9)-O(28)	91.2(3)
O(31)-W(9)-O(28)	163.0(3)
O(35)-W(9)-O(23)	169.1(3)

O(32)-W(9)-O(23)	86.1(3)
O(27)-W(9)-O(23)	73.8(3)
O(31)-W(9)-O(23)	88.6(2)
O(28)-W(9)-O(23)	74.6(2)
O(33)-W(10)-O(24)	100.7(3)
O(33)-W(10)-O(29)	100.3(3)
O(24)-W(10)-O(29)	96.1(3)
O(33)-W(10)-O(32)	98.7(3)
O(24)-W(10)-O(32)	159.2(3)
O(29)-W(10)-O(32)	87.5(3)
O(33)-W(10)-O(30)	99.1(3)
O(24)-W(10)-O(30)	88.7(3)
O(29)-W(10)-O(30)	158.8(3)
O(32)-W(10)-O(30)	81.1(3)
O(33)-W(10)-O(21)	173.3(3)
O(24)-W(10)-O(21)	75.6(3)
O(29)-W(10)-O(21)	74.7(2)
O(32)-W(10)-O(21)	85.7(3)
O(30)-W(10)-O(21)	86.6(2)
O(4)-Zn(1)-O(7)#4	98.5(11)
O(4)-Zn(1)-O(8)	94.2(9)
O(7)#4-Zn(1)-O(8)	163.4(13)
O(4)-Zn(1)-O(5)	90.8(10)
O(7)#4-Zn(1)-O(5)	96.1(9)
O(8)-Zn(1)-O(5)	94.4(10)
O(4)-Zn(1)-O(8)#4	167.1(13)
O(7)#4-Zn(1)-O(8)#4	92.2(9)
O(8)-Zn(1)-O(8)#4	76.8(8)
O(5)-Zn(1)-O(8)#4	80.8(9)
O(4)-Zn(1)-O(6)#4	96.7(10)
O(7)#4-Zn(1)-O(6)#4	86.6(9)
O(8)-Zn(1)-O(6)#4	81.3(8)
O(5)-Zn(1)-O(6)#4	171.6(13)
O(8)#4-Zn(1)-O(6)#4	91.1(9)
O(4)-Zn(1)-W(1)#4	131.5(8)
O(7)#4-Zn(1)-W(1)#4	130.0(8)
O(8)-Zn(1)-W(1)#4	37.8(5)
O(5)-Zn(1)-W(1)#4	87.5(7)
O(8)#4-Zn(1)-W(1)#4	39.0(5)
O(6)#4-Zn(1)-W(1)#4	84.7(7)
O(21)-Zn(2)-O(23)	107.5(3)
O(21)-Zn(2)-O(8)	108.8(3)
O(23)-Zn(2)-O(8)	112.9(3)
O(21)-Zn(2)-O(22)	107.8(3)
O(23)-Zn(2)-O(22)	105.8(3)
O(8)-Zn(2)-O(22)	113.8(3)
O(8)-Zn(3)-O(2)	94.0(3)
O(8)-Zn(3)-O(3)	95.3(3)
O(2)-Zn(3)-O(3)	92.2(3)
O(8)-Zn(3)-O(1)	173.1(3)
O(2)-Zn(3)-O(1)	92.8(3)
O(3)-Zn(3)-O(1)	85.3(3)
O(8)-Zn(3)-O(6)#4	78.5(3)
O(2)-Zn(3)-O(6)#4	171.1(3)

O(3)-Zn(3)-O(6)#4	93.4(3)
O(1)-Zn(3)-O(6)#4	94.6(3)
O(8)-Zn(3)-O(5)#4	77.2(3)
O(2)-Zn(3)-O(5)#4	90.5(3)
O(3)-Zn(3)-O(5)#4	172.2(3)
O(1)-Zn(3)-O(5)#4	101.9(3)
O(6)#4-Zn(3)-O(5)#4	83.1(3)
O(8)-Zn(3)-W(1)#4	38.5(2)
O(2)-Zn(3)-W(1)#4	87.1(3)
O(3)-Zn(3)-W(1)#4	133.4(2)
O(1)-Zn(3)-W(1)#4	141.2(2)
O(6)#4-Zn(3)-W(1)#4	84.0(3)
O(5)#4-Zn(3)-W(1)#4	39.4(2)

References

[1] Tourné, C. M.; Tourné, G. F.; Zonnevijlle, F. *J. Chem. Soc. Dalton Trans.* **1991**, 143-155.

[2] Neumann, R.; Dahan, M. *J. Am. Chem. Soc.* **1998**, *120*, 11969-11976.

[3] Yin, C.-X.; Finke, R. G. *Inorg. Chem.* **2005**, *44*, 4175-4188.

[4] (a) The cited work was performed by Kenji Nomiya,[4b] Hideki Torii, [4b] Keiichi Nomura,[4b] and Yuichi Sato[4c]. Again, we thank them for permission to cite their work herein. (b) Department of Materials Science, Faculty of Science, Kanagawa University, Hiratsuka, Kanagawa 259-1293, Japan. E-mail: nomiya@chem.kanagawa-u.ac.jp. (c) Department of Applied Chemistry, Faculty of Engineering, Kanagawa University, Rokkakubashi, Yokohama, Kanagawa 221-8686, Japan

[5] Neumann, R.; Khenkin, A.M. *Inorg. Chem.* **1995**, *34*, 5753.

[6] (a) Evans, I. P.; Spencer, A.; Wilkinson, G. *J. Chem. Soc., Dalton Trans.* **1973**, 204. (b) Tourne, C. M.; Tourne, G. F.; Zonnevijlle, F. *J. Chem. Soc., Dalton Trans.* **1991**, 143.

[7] Randall, W. J.; Droege, M. W.; Mizuno, N.; Nomiya, K.; Weakley T. J. R.; Finke, R. G. *Inorg. Synth.* **1997**, *31*, 167.

CHAPTER X

SUMMARY

Part I of this dissertation has focused on fitting protein aggregation kinetic data relevant to neurodegenerative disease using an "Ockham's razor" model and mechanistic approach, namely the Finke-Watzky (F-W) 2-step nucleation and autocatalytic growth model. First, a review was provided of the various models that have been used to fit protein aggregation kinetic data with a focus on curve-fitting the existing data. Then, 41 various representative kinetic aggregation curves from the literature of amyloid-β, α-synuclein, polyglutamine, and prion proteins were fit using the F-W model to give, for the first time, quantitative nucleation and growth rate constants for these 41 aggregation data sets related to the above noted neurological disorders. From these curve fits we were able to gather information on different factors, such as glutamine/asparagines rich regions in prion proteins, that lead to testable hypotheses for strategies towards slowing or possible stopping the protein aggregation underlying the neurodegenerative disorders. In addition, we have compiled evidence that the lag-time observed in the kinetic curves does not correspond to the nucleation rate, as previously widely believed in the literature.

Part I of this dissertation opens many potential avenues of future research. For

example, useful possible future studies include: (i) testing to see if other aggregation prone proteins are also well fit by the F-W model; and importantly, (ii) careful and controlled experimental measurement of protein aggregation kinetic data to test hypotheses of what factors influence nucleation over growth, or vice versa, in hopes of elucidating therapeutic strategies against neurodegenerative disorders. In fact (ii) is currently being pursued for yeast prion systems by a collaboration between the Finke and Ross labs at Colorado State University and is a direct outgrowth of the research detailed in this dissertation.

Part II focused on the investigation of dioxygenase catalysts beginning with a brief overview of the definition and importance of dioxygenase systems plus an introduction to two important synthetic dioxygenases. Then, the synthesis and characterization of a new dioxygenase precatalyst V(3,6-DBSQ)(3,6-DBCat)$_2$ was given. Next, a full report on the dioxygenase activity was provided for three d^0 metal precatalysts: [VO(3,5-DTBC)(3,5-DBSQ)]$_2$, V(3,6-DBSQ)(3,6-DBCat)$_2$, and [MoO(3,5-DTBC)$_2$]$_2$. It was determined that both [VO(3,5-DTBC)(3,5-DBSQ)]$_2$ and V(3,6-DBSQ)(3,6-DBCat)$_2$ give the same dioxygenase products in similar yields for the H$_2$(3,5-DTBC) and H$_2$(3,6-DTBC) substrates tested, while [MoO(3,5-DTBC)$_2$]$_2$ gives a majority of the less desirable autoxidation product, benzoquinone. Overall this suggests that the presence of a d^0 vanadium bound to a semiquinone ligand is important for obtaining intradiol and extradiol dixoygenase products. Finally, a reinvestigation of the composition of a previously claimed polyoxometalate dioxygenase, $[WZnRu_2(OH)(H_2O)(ZnW_9O_{34})_2]^{11-}$ (**1**), was examined. Three independent samples from two different laboratories of **1** are consistent with "**1**" actually being the parent

polyoxometalate $[WZn_3(H_2O)_2(ZnW_9O_{34})_2]^{12-}$ (**2**) with ≤ 0.2 equivalents of Ru present either as partially incorporated or as a Ru^{n+} impurity. In addition, a simple mixture of **2** plus 0.13 equiv of Ru added as $[Ru(DMSO)_4Cl_2]$ is a ca. 2-fold faster catalyst for the conversion of adamantane to 1-adamantanol and 2-adamantanone as compared to "**1**", casting considerable doubt on the *Nature* paper from others claiming that "**1**" is the first, all-inorganic, Ru-polyoxometalate-based dioxygenase catalyst.

Part II also leads to possible further research using the $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ and $V(3,6\text{-DBSQ})(3,6\text{-DBCat})_2$ dioxygenase precatalysts . Interesting future studies include: (i) extending the list of dioxygenase substrates that $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ and $V(3,6\text{-DBSQ})(3,6\text{-DBCat})_2$ may serve as precatalysts for by testing substrates such as: 3,5-di-methylcatechol, 3,6-dimethylcatechol, and various chloro-substituted catechols; and (ii) determining the mechanism by which $[VO(3,5\text{-DTBC})(3,5\text{-DBSQ})]_2$ and $V(3,6\text{-DBSQ})(3,6\text{-DBCat})_2$ convert $H_2(3,6\text{-DTBC})$ into the observed intradiol and extradiol dioxygenase products.

APPENDIX A

GENERAL STATEMENT ON "JOURNALS-FORMAT" THESES

(Written by Professor Richard G. Finke)

The Graduate School at Colorado State University allows, and the Finke Group in particular encourages, so-called journals-format theses. Journals-format theses, such as the present one, consist of a student written and lightly edited literature background section, chapters corresponding (in the limiting, ideal case) to final-form papers either accepted or at least submitted for publication, a summary or conclusions chapter, and short bridge or transition sections between the chapters as needed to make the thesis cohesive and understandable to the reader. The "bridge" sections and summary are crucial so that the thesis fulfills the requirement that the thesis be an entity (an official requirement of most Graduate Schools). All chapters (manuscripts) in a journals-format thesis must of course be written initially by the student, with subsequent (ideally light) editing by the Professor, the student's committee, and even the student's colleagues where appropriate and productive.

The advantages for doing a journals-format thesis are several-fold and compelling. Specifically, some of the major advantages are: the level of science (i.e., of refereed, accepted publications) is at the highest level; the student and Professor must

interact closely and vigorously (i.e., to bring both the science and the writing to their highest level), hence the student is getting the best education possible and is being at least exposed to (if not held to) the highest standards; the needed clean-up or control experiments that invariably come up have all been identified and completed before the student leaves; there are no further time demands once the student has left the University (since all publication are at least submitted; it is terribly inefficient to try to complete either writing or often specialized experiments once the student has left); and the American tax payers, who ultimately pay the bill for the research, are getting their money's worth since all the research is published and thus widely disseminated in the highest form, as refereed science. Professorial experience teaches that a student who has achieved a journals-format thesis has indeed received a better education and has learned critical thinking and clear writing skills that will serve them well for a lifetime.

Experience also teaches, however, that much more than light editing is often needed in at least some student theses; it follows, then, that considerable professorial writing and editing might be needed for at least the initial chapters of most journals-format thesis. Indeed, a journals-format thesis is not recommended (and may not even be possible) for less strong students. Hence, the issue arises of exactly how much of the science and the writing, in the final (or submittable) chapters, is due to the student vs. the Professor and whether or not this level of contribution constitutes that acceptable of a new Ph.D. and independent investigator.

To deal with this issue, several recommendations are made; the recommendations below have been discussed with the committee signing Aimee M. Morris's dissertation. (Ms. Morris's dissertation is the fifteenth such thesis form the Finke group following Dr.

C. Garr's, Dr. Y. Lin's, Dr. M. Pohl's, Dr. J. Sirovatka's, Dr. J. Aiken's, Dr. R. Suto's, Dr. J. Widegren's, Dr. K. Doll's, Dr. C.-X. Yin's, and Dr. L. Ott's dissertations, and Ms. K. Weddle's, Mr. W. White's, Mr. C. Hagen's, and Mr. C. Graham's Masters theses.)

Recommendations

The recommendations are:

(i) That the present pages be enclosed in the thesis until such a time as it is no longer needed (i.e., when the policies and procedures for journals-format theses become routine);

(ii) That for each chapter it is detailed, and to the satisfaction of the committee and the advisor, who made what contributions, both of intellectual substance and writing. [Substantial contributions of other students or Professors should of course be acknowledged. In the case of disagreements, the various drafts (i.e., as their electronic files) can be examined by the committee (in light of a knowledge of who wrote which draft) to easily determine who contributed what. In possible borderline or controversial cases it may even be advisable to keep all (electronic) drafts of the papers as a record];

(iii) That it be specifically stated whether or not all the experimental work is the Ph.D. candidate's [as is usually the case, although the increasing (desirable) collaboration among scientists worldwide makes this a non-trivial point].

(iv) Furthermore, it is recommended that allowances be made for the expectation that a greater degree of involvement of the professorial advisor is likely in a journals-format thesis than in a traditional thesis. [That this is reasonable follows from the fact that some Professors write 100% of all their papers; this, unfortunately, robs the student

of the valuable experience of participating in the science and the end product as practiced at the highest levels. (It also creates an unmanageable writing burden for Professors involved in all but the narrowest of research areas or for Professors involved in more than one competitive research area)];

(v) Notwithstanding (iv), there needs to be ideally no more than ca. 40% Professorial writing contribution in a given *early* chapter in the thesis, and there should be a clear evolution in the thesis of a decreasing professorial involvement to, say, a 10-20% direct contribution in the last chapter or two.

(vi) As a further aid towards separating out the candidate's and the professorial (and other) contributions, it is recommended that the Introductory (usually literature background) chapter(s) and at least the final chapter be lightly edited only, so that authentic examples of the student's contributions are documented in an unambiguous form.

APPENDIX B

RESEARCH PROPOSAL

Synthesis, Characterization, and Mechanical and Biocompatibility Testing of Biologically Relevant Metal-Doped Nanocrystalline Hydroxyapatite for Improved Bone Regeneration Materials

Abstract and Specific Aims

The need for better bone regeneration materials is escalating with continuing advances in healthcare along with a growing elderly population. Current implants are focused on bone replacement, but have limited lifespans. Recent research has been directed toward the development of bone regeneration materials that assist in promoting new bone growth in needed areas. Hydroxyapatite ($Ca_{10}(PO_4)_6(OH)_2$, HA) has been the focus of many bone regeneration materials due to its high composition in human bone (~65%) and excellent biocompatibility properties. However, pure HA displays higher than desired dissolution rates, as well as poor mechanical properties that limit its use to

non-loadbearing applications. It has been suggested that doping HA with different metal ions could increase both the mechanical and biocompatibility properties of HA. Previous studies have focused on the use of monovalent, divalent, some trivalent, and tetravalent Ti and Zr doped HA. Higher valency metal dopants have been suggested to improve the mechanical properties of HA, but in some cases the biocompatibility is reduced. It was also previously suggested that HA doped with trivalent metal dopants of size $\leq Ca^{2+}$ would produce superior bone regeneration materials.

Herein the following hypotheses are proposed to be tested: (i) that HA doped with 2.5 wt% of each of the following metals: V^{3+}, V^{4+}, V^{5+}, Mn^{3+}, Mn^{4+}, Mo^{3+}, Mo^{4+}, Mo^{5+}, and Mo^{6+}, can be synthesized with metal incorporation into the HA lattice; (ii) that the metal doped HA will be nontoxic to cells and will stimulate the growth of new osteoblasts, making it a safe and desirable bone regeneration material; and (iii) that trivalent or higher valency metal-doped and sintered HA will provide better bone regeneration materials in terms of mechanical properties, as evidenced by Vicker's hardness and uniaxial compression tests, as well as in terms of their biocompatibility.

Background and Significance

With the continuing advances in healthcare and the average life expectancy of humans continually lengthening, there becomes a high demand for bone replacement materials with lifetimes greater than that of the patient.[1] Presently, autografts and allografts, which are bones obtained from another anatomic site of the same patient and those obtained from another subject, respectively, are the gold standards for bone repair and substitution.[2,3] However, autografts can cause trauma at the donation site and are in

limited supply.[2,3,4,5] Allografts are also problematic due to the high cost, limited supply, possible immune rejection, and possible pathogen transmission.[2,3,4,5] Therefore, there is a need to develop synthetic alternatives for bone replacement and repair to improve the quality of human life.

The use of both bioinert (i.e., materials that do not stimulate bone formation or bond directly to bone such as inert metal implants) and bioactive (i.e., materials that stimulate bone formation and directly bond to bone such as calcium carbonates) synthetic implants has increased dramatically in the past 25 years.[2,6] The most common bone replacements occur in the hip[7] and knee joints and are often a result of Osteoporosis, where bone resorption occurs faster than new bone can be produced causing a reduction in bone density and eventually fracture.[1] Most currently used prostheses are composed of bioinert materials that often have a limited lifespan of ≤ 25 years and also tend to degrade healthy bone to which they are attached due to frictional forces.[1] In addition, all currently used orthopedic implants lack three critical properties: (i) the ability to self-repair, (ii) the ability to maintain a blood supply, and (iii) the ability to modify their structure in response to environmental factors.[1,8] Due to these missing properties, current research has focused on exploring bone *regeneration* instead of bone *replacement* materials.[1,2] It has been suggested that the best regeneration materials will attempt to mimic both the physical structure and composition of bone.[2,6,8]

In the body, bone has four main functions. It provides structural support, protects and stores healing cells, maintains ion homeostasis, and repairs damage caused by locomotion or trauma.[9,10] The multifunctional macroporous and microporous networks that comprise bone have proven to be hard to mimic synthetically. According to Ben-

Nissan, "many have tried to mimic the unique qualities of bone, however nature has proved to be nearly impossible to adequately copy or emulate."[11] The actual structure of bone is complex and estimated to be (Ca, Mg, $Na)_{10}(PO_4, CO_3, HPO_4)_6(OH, Cl, F)_2$ with 65 wt% of the composition being hydroxyapatite (HA, $Ca_{10}(PO_4)_6(OH)_2$).[2,11] HA is osteoconductive, meaning that its surface can serve as a scaffold or template to guide the formation of new bone[2,11,12] (a highly desirable property for a bone regeneration material). HA is also a bioresorbable material that degrades over time as it is replaced by newly formed bone tissue.[6] In addition, HA is light-weight, resistant to microbial attack and pH changes, and importantly, is able to directly bond with existing bone.[6] Therefore, HA has drawn considerable attention as a possible bone regeneration compound.[1,2,6,11,13,14]

However, the often too stiff and brittle mechanical properties of HA limit the load bearing capabilities unless combined with other supports such as metal plates.[2,6] In addition, HA is highly soluble in the body. This causes problems with durability, tissue integration, and longevity due to faster than desired dissolution.[11] It is well known that the complex composition of bone and the presence of trace elements play an important role in the overall performance.[6,11,15] Hence, researchers have been interested in doping HA with metal ions other than Ca^{2+} in hopes of improving both the mechanical and biocompatibility properties.[15,16,17,18,19,20,21,22,23,24]

Researchers have mainly focused on the use of main group monovalent dopants such as Na^+ and K^+,[17,18] as well as many divalent dopants including Zn^{2+}, Mg^{2+}, Cu^{2+} and Fe^{2+}.[15,16,19,20] In an interesting study by Webster and co-workers, the trivalent metals Y^{3+}, La^{3+}, In^{3+}, and Bi^{3+} were doped into the HA lattice.[21] The main finding in this study was

that both the Y^{3+}- and In^{3+}- doped HA increased osteoblast adhesion by 50 and 56%, respectively, while the other metal dopants' osteoblast adhesion was not statistically different than that of undoped HA.[21] Overall, this study suggested that trivalent metal dopants of smaller size than Ca^{2+} encourage osteoblast adhesion and hence, may provide better biocompatible materials for bone regeneration. In addition, another study suggested that Cr^{3+}-doped HA was a better dopant than divalent Ni^{2+}, Co^{2+}, and Cu^{2+} due to its ability to retain metal–OH groups on the important reactive surface of the doped HA.[19] Interestingly, the only metal dopants tested with oxidation states higher than trivalent are Ti^{4+} and Zr^{4+}, which have both been shown to improve the mechanical properties but decrease biocompatibility in comparison to undoped HA.[22,23]

The amount of dopant used in HA has varied greatly from 1.0 wt%[20] to 40 wt%.[23] However, in studies using both 1.0 and 2.5 wt% dopant, it has been shown that the 2.5 wt% dopant further improved the mechanical properties of HA, while not disrupting the biocompatibility.[20] Use of 40 wt% Zr^{4+} increased the mechanical properties of the HA, but, unfortunately, also greatly decreased the biocompatibility properties.[23] It has also been demonstrated that 5 and 10 wt% dopant increased the mechanical properties while maintaining the biocompatibility,[23,25] but the majority of studies involving metal-doped HA have found desirable results in terms of both the mechanical and biocompatibility properties using approximately 2.5 wt% dopant.[15,16,17,20,24]

In addition to variations in the amount of dopant used, the relative size of the doped HA in these materials is also varied. Some studies focus on the formation of micron sized metal-doped HA, while others have made nanocrystalline doped HA. The advantage to making nanosized HA is that it mimics the structural and dimensional

characteristics of bone.[2,6,20] In addition, the large surface area of nano-HA can lead to better densification and hence, improvements in the mechanical properties over larger micron-HA.[6] Therefore, the trend in using HA as a bone regeneration material has shifted toward the use of nanometer-sized HA.[3,6,20]

Herein, the aims are to synthesize and characterize HA doped with 2.5 wt% of each of the following: V^{3+}, V^{4+}, V^{5+}, Mn^{3+}, Mn^{4+}, Mo^{3+}, Mo^{4+}, Mo^{5+}, and Mo^{6+}. The goal of the synthesis is to incorporate the metal dopants into the HA lattice. Dopant incorporation and homogeneity will be tested by powder X-ray diffraction (XRD), elemental analysis, and high-angle annular dark field scanning transmission electron microscopy (HAADF-STEM). In addition, the morphology and size of the doped HA will be probed using transmission electron microscopy (TEM). The oxidation state of the incorporated metal will then be analyzed using X-ray photoelectron spectroscopy (XPS) in order to determine if the original oxidation state is retained. Synthesis and characterization of these materials are necessary in order to test the hypothesis presented herein that HA doped with metals that are (i) trivalent or higher valency, (ii) naturally occurring in the body, and (iii) of size $\leq Ca^{2+}$, will be superior bone regeneration materials in terms of their mechanical and biocompatibility properties. Biocompatibility properties will be tested using osteoblast adhesion and growth studies in order to determine the cytotoxicity and osteoconductivity of the proposed metal doped HA. Finally, the mechanical properties of the proposed materials will be tested using Vicker's hardness and uniaxial compression measurements.

Research Design and Methods

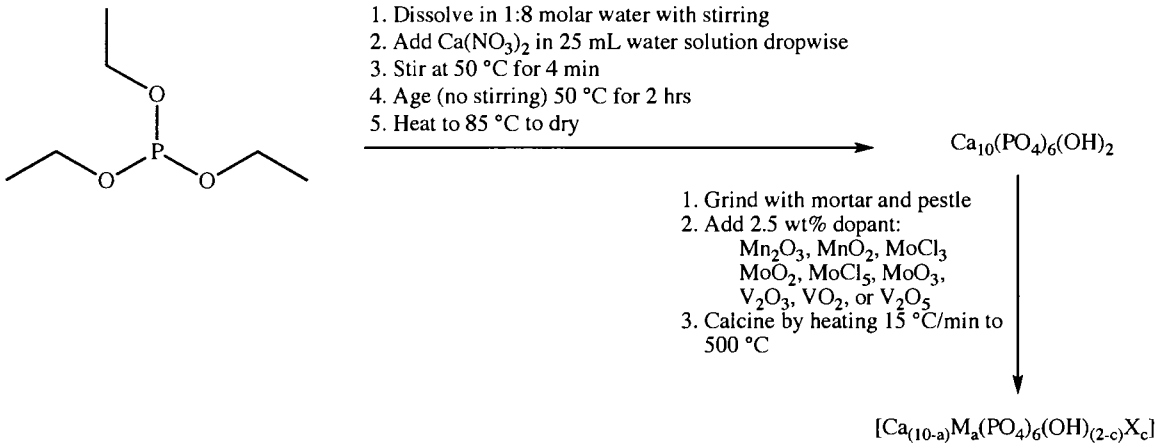
Synthesis of metal-doped hydroxyapatite

Finding the correct experimental conditions could prove challenging, but previous studies have provided a solid background to begin the synthesis of metal-doped HA. I propose using the sol-gel synthesis method over the commonly employed precipitation or hydrothermal methods.[3] Synthesis via sol-gel allows for relatively short reaction times on the order of a couple of hours at relatively low reaction temperatures of 80-85 °C in water,[20,26] as compared to 24-hour reactions in aprotic solvents for common precipitation syntheses,[27] or the required high temperature (≥ 500K) and pressure (≥ 100 MPa) required for hydrothermal techniques.[11] In addition, the sol-gel method allows for improved homogeneity[20,26] along with the formation of nanomaterials that resemble the dimension of HA observed in bone (30-50 nm long, 20-25 nm wide, and 1.5-4 nm thick).[2,3]

The water-based sol-gel synthesis proposed herein is modified from the work of Kalita and Bhatt for Mg^{2+}- and Zn^{2+}-doped nanocrystalline HA,[20] Scheme 1. That is, the literature synthesis will be followed with the exception of adding the manganese, molybdenum, and vanadium ion dopants as the commercially available oxide and chloride salts (Mn_2O_3, MnO_2, $MoCl_3$, MoO_2, $MoCl_5$, MoO_3, V_2O_3, VO_2, or V_2O_5) in the amount of 2.5 wt% prior to sintering, as was done previously.[20] It has also been noted that maintaining the pH at neutral or basic conditions is necessary to avoid the formation of other calcium phosphate phases.[28] Thus, the pH will be checked prior to the HA aging step and if necessary adjusted to physiological pH 7.4 using NH_4OH or HNO_3. The metal-doped HA will then be calcined under nitrogen to prevent the formation of metal

oxide side products, other calcium carbonate phases, or redox reactions of the metals[24] at the previously found optimal temperature of 500 °C to form nanocrystalline doped-HA.[20,26]

Scheme 1. Proposed synthesis of metal-doped HA.



Listed below in Table 1 are the ionic radii of dopants that have been shown to be incorporated into the lattice of HA along with the radii of the transition metals proposed herein as dopants for HA. Based on the precedent that 0.55 Å Fe^{3+} is incorporated into the crystal lattice of HA,[16,24] it seems plausible that Mn^{3+}, Mn^{4+}, V^{4+} and V^{5+} should also be incorporated into the HA lattice based on sizes[29] of 0.58, 0.53, 0.58, and 0.54 Å, respectively. In addition, based on the previous evidence for Co^{2+} incorporation,[19] size[29] similarities of 0.65 Å Mo^{4+} and 0.64 Å V^{3+} suggest that these metals can be incorporated in the HA lattice. Examination of the size[29] of Fe^{2+} and evidence for Fe^{2+} incorporation[16] suggests that both Mo^{5+} and Mo^{6+} will be incorporated, Table 1. Finally evidence for the incorporation[19] of Ni^{2+} at 0.69 Å[29] suggests, at least based on size arguments, that Mo^{3+}

also at 0.69 Å, can be incorporated into the HA lattice. Therefore, based on size similarity and precedence, it appears that all of the metal dopants suggested represent viable metals for investigation. Furthermore, the fact that higher valence Ti^{4+} and Zr^{4+} have been used in conjunction with HA to increase the mechanical properties[22,23] provides further support for the proposed tetra-, penta-, and hexavalent metal dopants suggested.

Table 1. Ionic radii[29] of previously examined[15,16,17,18,19,20,21,22,23] HA dopants along with radii[29] of the metals proposed herein for doping HA.

Previously Examined						Proposed Herein			
Main Group Dopant	Ionic radius (Å) [a]	Transition Metal Dopant	Ionic radius (Å) [a]	Transition Metal Dopant	Ionic radius (Å) [a]	Metal Dopants	Ionic radius (Å) [a]	Metal Dopants	Ionic radius (Å) [a]
Ca^{2+}	1.00	Ag^{+}	1.15	Zr^{4+}	0.72	Mn^{3+}	0.58	V^{3+}	0.64
K^{+}	1.38	Co^{2+}	0.65	Bi^{3+}	1.03	Mn^{4+}	0.53	V^{4+}	0.58
Mg^{2+}	0.72	Cu^{2+}	0.73	Cr^{3+}	0.62	Mo^{3+}	0.69	V^{5+}	0.54
Na^{+}	1.02	Fe^{2+}	0.61	Fe^{3+}	0.55	Mo^{4+}	0.65		
Si^{2+}	unk [b]	Ni^{2+}	0.69	In^{3+}	0.80	Mo^{5+}	0.61		
Ti^{4+}	0.61	Sr^{2+}	1.18	La^{3+}	1.03	Mo^{6+}	0.59		
		Zn^{2+}	0.74	Y^{3+}	0.90				

[a] The ionic radii values are based on experimental crystal structure determinations, supplemented by empirical relations, and theoretical calculations. In addition, in each case the ionic radius reported is for a coordination number of 6.

[b] The ionic radius of Si^{2+} is unknown (unk). However, the reported value for Si^{4+} is 0.40 Å and suggests that the ionic radius of Si^{2+} is larger than 0.40 Å.

Characterization of the Metal-Doped HA Materials

In order to determine if the above proposed metals do indeed become incorporated into the HA crystal lattice, elemental analysis via inductively coupled plasma spectrometry (ICP) and powder XRD will be employed. The ICP results will give the bulk composition of the material made after sintering[30] and allow for the amount of dopant incorporated to be calculated vs. the amount added, as was done previously for Ni^{2+}, Cu^{2+}, Co^{2+}, and Cr^{3+} doped HA.[19] ICP is being proposed instead of the previously used energy dispersive X-ray spectroscopy (EDS)[20] as better detection limits can

achieved with ICP. Interestingly, many of the previous studies do not analyze the bulk sample content, but assume that all of the dopant added is incorporated.[15,17,18,21,22,23]

In addition to ICP, powder XRD will allow for the determination of the relative crystallinity of the metal-doped HA, as well as confirm whether or not the metal is incorporated by an observable change in the unit cell constants.[19,21,23] Shown below in Figure 1 is the previously reported XRD pattern of pure HA, as well as 1.0 wt% Mg^{2+} and Zn^{2+} doped HA.[20] These XRD patterns illustrate what is expected for the XRD patterns of the proposed metal-doped HAs. That is, little to no change will be observed in the XRD pattern from the undoped to doped HA, but the unit cell constants should be different by a statistically significant amount.

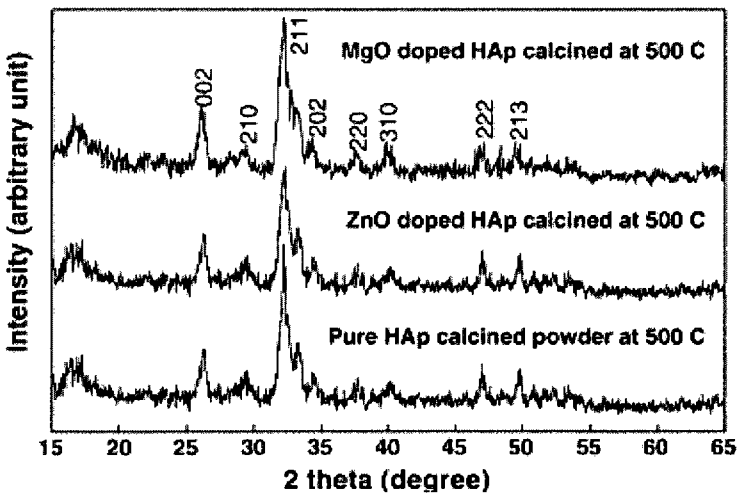


Figure 1. XRD patterns of pure HA, Zn^{2+} doped HA, and Mg^{2+} HA. Note that the XRD pattern does not change with the dopants added in 1.0 wt%. Figure modified from reference 20.

Unfortunately the unit cell parameters were not reported for the Mg^{2+} and Zn^{2+} doped samples shown in Figure 1.[20] However, the unit cell constants of Fe^{3+} doped HA

and pure HA have been reported[24] and are shown in Table 2. Table 2 illustrates that statistically different unit cell parameters are obtained with pure vs. doped HA samples.

Table 2. The *a* and *c* unit cell constants and volume measurements for pure HA and 2 atomic% Fe^{3+} doped HA.[24]

	Pure HA	Fe^{3+} doped HA
a (Å)	9.3948(4)	9.4138(5)
c (Å)	6.8906(4)	6.8896(5)
Volume ($Å^3$)	526.8	528.8

In addition to ICP and powder XRD, high-angle annular dark field scanning transmission electron microscopy (HAADF-STEM, also commonly known as Z-contrast imaging) will be used to probe the homogeneity of the metal-doped HA materials. Probing the homogeneity is important since the synthesis proposed herein adds the dopants after the sol-gel process, but before annealing, and could lead to localized doping especially on the surface. This localized doping may not be observed with XRD measurements, but HAADF-STEM is sensitive to composition and allows for small, heavy particles to be readily visible within a matrix or support.[31] In fact, iodine and gadolinium added to a hydroxyapatite mixture were recently imaged and quantitated using HAADF-STEM.[32] Hence, HAADF-STEM images should allow for the homogeneity of the metal dopants to be examined as well as for microcompositions to be determined.

The morphology of the metal-doped HA materials will also be characterized by transmission electron microscopy (TEM) in order to determine how the size of the nanocrystalline HA is affected by the addition of dopants. As shown in Figure 2 below,

different metal dopants can cause changes to the size of the nanocrystalline HA.[20] The undoped HA (Figure 2a) has an average particle size of 10-12 nm, while the Mg^{2+} doped HA (Figure 2b) is 2-5 nm and the Zn^{2+} doped HA particles are 10-20 nm.[20] Interestingly, both Mg^{2+} and Zn^{2+} are divalent and have very similar ionic radii of 0.72 and 0.74 Å, respectively,[29] yet the main group Mg^{2+} gives smaller particles while the transition metal Zn^{2+} gives larger particles than pure HA. Therefore, it cannot be predicted at this time how the metal dopants proposed will affect the size of the particles in nanocrystalline HA.

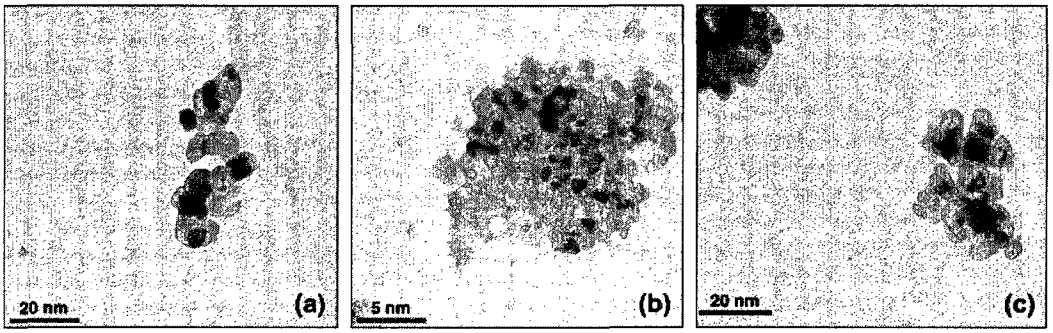


Figure 2. TEM images of (a) HA, (b) Mg^{2+} doped HA, and (c) Zn^{2+} doped HA reproduced from reference 20.

Since different oxidation states of the same metal are being suggested herein as dopants for HA, it will be important to determine the oxidation state of the metal dopants after incorporation into the HA lattice. In other words, it needs to be shown that the metal precursors retain the original oxidation state after incorporation. In order to test this, X-ray photoelectron spectroscopy (XPS) will be used to probe the oxidation states of the metals and to make sure that no redox reactions of the metal dopants have occurred during the synthesis process. XPS has also been used in similar studies to study the surface composition of metal-doped HA and it was shown that Cr^{3+} was a better dopant

than divalent Ni^{2+}, Cu^{2+}, and Co^{2+} due to its ability to maintain the surface composition of pure HA.[19] Maintaining the surface of HA is important because the surface controls further growth and incorporation at implantation sites.[6,28,33] Therefore, XPS will be used to both probe the oxidation state and to determine the composition of the surface of HA.

Biocompatibility Testing

Once the metal dopant incorporation into HA has been established by the above characterization techniques, it is important to screen the cytotoxicity of each of the metal-doped HA materials. Biocompatibility testing is proposed before mechanical properties testing in order to rule out materials that may not be good candidates as bone regeneration materials in terms of their ability to incorporate into the tissue area without causing harm. In order to test the biocompatibility, commonly used cell culture and adhesion experiments[15,16,17,18,21,23] will be run. In addition, to provide information about possible toxic effects, these experiments will also show whether further bone growth is induced in the metal-doped HA over pure HA. Specifically, I propose using an immortalized modified human osteoblast cell line (OPC-1)[34] to conduct *in vitro* studies, as has been done in other metal-doped HA studies.[15,16,17,18]

Use of an immortalized cell line is advantageous because the cells are consistent and reproducible and thus, can be used for cross comparison studies.[34] However, immortalized cell lines are used in *in vitro* studies and may not fully represent the effects *in vivo*,[35] but should provide a good preliminary toxicity screening. Cytotoxicity studies would be conducted by using staining to count the number of living healthy cells both before and after introduction of the synthesized metal-doped HA against control groups

that are not exposed to metal-doped HA. In addition, using scanning electron microscopy (SEM) will allow for analysis of the morphology along with how well the cells anchor, attach, and proliferate onto the surface of synthesized metal-doped HA.[15,16,17,18] In previous studies, biologically relevant metal-doped HA materials showed no toxicity to the cells and were able to induce the same cell growth within experimental error as pure HA.[15,16,17,18] Therefore, it is probable that the metal-doped HA materials proposed herein will have similar biocompatibility properties. In addition, it is expected that the metal-doped HAs proposed herein will have better biocompatibility properties than some of the previously proposed dopants such as Ti^{4+}, Zr^{4+}, Bi^{3+}, Y^{3+}, La^{3+}, and In^{3+}, since the metals proposed are naturally occurring in the human body, albeit in some cases in different oxidation states.

Mechanical Properties Testing

The commonly employed[15,16,17,18,20,22] Vicker's hardness and uniaxial compression tests will be used to determine the hardness and compressibility, and therefore, the mechanical properties of the non-toxic metal-doped hydroxyapatite materials proposed. Previously tested 1 wt% metal dopants such as Mg^{2+} and Zn^{2+} have improved both the hardness (Figure 3) and the compression of HA.[20] In addition, the hardness of the materials is maintained for at least five weeks under *in vitro* conditions, Figure 3.

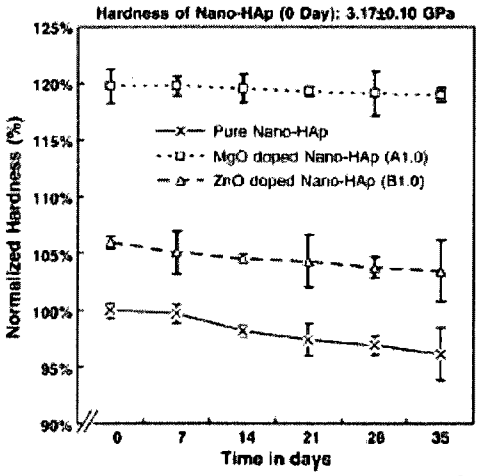


Figure 3. The normalized hardness of pure HA, 1.0 wt% Mg^{2+} and 1.0 wt% Zn^{2+} doped HA under simulated *in vivo* conditions over the course of 35 days. Reproduced from reference 20.

In fact in all cases where the mechanical properties were tested for metal-doped HA, the hardness and compression strength were greater than or equal to pure HA.[15,16,17,18,20,22] Hence, it is expected that the metal-doped HAs proposed herein will be able to improve, or at least maintain, the mechanical properties of HA.

Based on both the mechanical properties and biocompatibility tests proposed, it will be interesting to look at which materials proposed will be the best candidates for bone regeneration materials. In particular it will be of interest to look at which oxidation state of the metals of V, Mo, and Mn will be preferred. In addition, this study will allow us to test the hypothesis previously set forth[21] that trivalent metal dopants of size $\leq Ca^{2+}$ will make preferred bone regeneration materials. Also of interest is to determine if the ionic size of the metal dopants is important for incorporation into the lattice, although size precedence suggests that all the metals proposed will be incorporated. This study will hopefully provide viable material(s) for bone regeneration that have both the biocompatibility and the mechanical properties desired. However, if an outstanding

material is not found, the trends in size and oxidation state, or possible lack of trends observed from this study, may lead to a new list of possible dopants to be tested in the future.

Conclusions

This proposal will test the hypothesis that biologically relevant trivalent and higher valency 2.5 wt% metal-doped nanocrystalline HA will provide better bone regeneration materials in terms of their biocompatibility and mechanical properties. The synthesis of these materials is suggested via a precedented sol-gel method. Once prepared the metal incorporation and homogeneity into HA will be probed using elemental analysis, XRD, and HAADF-STEM. In addition, the oxidation state of the metals incorporated will be verified by XPS. Following verification of the metal incorporation into HA, the biocompatibility properties of the newly synthesized materials will be tested using *in vitro* cell adhesion and cell growth studies to determine whether or not these proposed materials are cytotoxic and if they promote cell growth on the surface. Finally, the hopefully improved mechanical properties of the proposed materials will be tested using Vicker's hardness and compressibility tests. Overall, the materials proposed herein are hopeful alternatives for bone regeneration materials that utilize the positive biocompatible properties, while improving the poor mechanical properties exhibited by pure HA.

Future Studies

Indeed, if any of the proposed metal-doped HA materials are demonstrated to be superior over the known doped and undoped HAs, further testing will be warranted. This testing will include examination of the *in vivo* biocompatibility of the material(s), along with in-depth engineering analyses. One of the significant challenges for bone regeneration materials is the ability to scale-up the material and be able to mold it into the irregular shapes of bones.[6] This is no small or easy task and significant research and effort will be needed in order to overcome the engineering challenges.

References:

[1] Jones, J. R.; Hench, L. L. *Current Opinion in Solid State and Materials Science* **2003**, *7*, 301-307.

[2] LeGeros, R. Z. *Chem. Rev.* **2008**, *108*, 4742-4753.

[3] Palmer, L. C.; Newcomb, C. J.; Kaltz, S. R.; Spoerke, E. D.; Stupp, S. I. *Chem. Rev.* **2008**, *108*, 4754-4783.

[4] Laurie, S. W.; Kaban, L. B.; Mullikan, J. B. *Plast. Reconstr. Surg.* **1984**, *73*, 933-938.

[5] Woesz, A.; Rumpler, M.; Stampfl, J.; Varga, F.; Fratzl-Zelman, N.; Roschger, P.; Klaushofer, K.; Fratzl, P.; *Mater. Sci. Eng., C* **2005**, *25*, 181-186.

[6] Hertz, A.; Bruce, I. J. *Nanomedicine* **2007**, *2*, 899-918.

[7] Dekker, R. J.; de Bruijn, J. D.; Stigter, M.; Barrere, F.; Layrolle, P.; van Blitterswijk, C. A. *Biomaterials* **2005**, *26*, 5231-5239.

[8] Wang, M. *Mat. Res. Soc. Symp. Proc.* **2002**, *724*, 83-93.

[9] Oliveira, A. L.; Mano, J. F.; Reis, R. L. *Current Opinion in Solid State and Materials Science* **2003**, *7*, 309-318.

[10] Omelon, S. J.; Grynpas, M. D. *Chem. Rev.* **2008**, *108*, 4694-4715.

[11] Ben-Nissan, B. *Current Opinion in Solid State and Materials Science* **2003**, *7*, 283-288.

[12] Davies, J. E.; Hosseini, M. M. In *Bone Engineering*; Davies, J. E., Ed.; EM Squared, Inc.: Toronto, 1999.

[13] Koutsopoulos, S. *J. Biomed. Mater. Res.* **2002**, *62*, 600-612.

[14] Kim, H.-M. *Current Opinion in Solid State and Materials Science* **2003**, *7*, 289-299.

[15] Kalita, S. J.; Rokusek, D.; Bose, S.; Hosick, H. L.; Banyopadhyay, A. *J. Biomed. Mater. Res. A* **2004**, *71*, 35-44.

[16] Bose, S.; Badyopadhyay, A. *PCT Int. Appl.* **2007**, WO 2007-US67444 20070425.

[17] Kalita, S. J.; Bose, S.; Hosick, H. L.; Bandyopadhyay, A. *Biomaterials* **2005**, *25*, 2331-2339.

[18] Seeley, Z.; Bandyopadhyay, A.; Bose, S. *Mater. Sci. Eng., C* **2008**, *28*, 11-17.

[19] Wakamura, M.; Kandori, K.; Ishikawa, T. *Colloids Surf., A* **1998**, *142*, 107-116.

[20] Kalita, S. J.; Bhatt, H. A. *Mater. Sci. Eng., C* **2007**, *27*, 837-848.

[21] Massa, E. A.; Slamovich, E. B.; Webster, T. J. *Mat. Res. Soc. Symp. Proc.* **2002**, *711*, 257-260.

[22] Gunduz, O.; Oktar, F. N.; Oz, B.; Altundal, H.; Agathopoulos, S.; Salman, S.; Ovecoglu, L. *Key Eng. Mater.* **2006**, *309-311*, 1137-1140.

[23] Evis, Z.; Sato, M.; Webster, T. J. *J. Biomed. Mater., A* **2006**, *78*, 500-507.

[24] Low, H. R.; Phonthammachai, N.; Maignan, A.; Stewart, G. A.; Bastow, T. J.; Ma, L. L.; White, T. J. *Inorg. Chem.* **2008**, ASAP web published 11/14/08, doi: 10.1021/ic801491t.

[25] Sampaio, B. V.; Goller, G.; Oktar, F. N.; Valeiro, P.; Goes, A. M.; Leite, M. F. *Key Eng. Mater.* **2005**, *284-286*, 639-642.

[26] Song, Y.; Wen, S.; Li, M. *Mat. Res. Soc. Symp. Proc.* **2002**, *724*, 135-140.

[27] Choi, D.; Marra, K. G.; Kumta, P. N. *Mater. Res. Bull.* **2004**, *39*, 417-432.

[28] Wang, L.; Nancollas, G. H. *Chem. Rev.* **2008**, *108*, 4628-4669.

[29] Lide, D. R., Ed.; In *CRC Handbook of Chemistry and Physics*, 86th Ed; Taylor and Francis: New York, 2005.

[30] Koutsopoulos, S. *J. Biomed. Mater. Res.* **2002**, *62*, 600-612.

[31] Midgley, P. A.; Ward, E. P. W.; Hungria, A. B.; Thomas, J. M. *Chem. Soc. Rev.* **2007**, *36*, 1477-1494.

[32] Schlomka, J. P.; Roessl, E.; Dorscheid, R.; Dill, S.; Martens, G.; Istel, T.; Bumer, C.; Herrmann, C.; Steadman, R.; Zeitler, G.; Livne, A.; Proksa, R. *Physics in Medicine and Biology* **2008**, *53*, 4031-4037.

[33] Ter Brugge, P. J.; Jansen, J. A. *Biomaterials* **2002**, *23*, 3269-3277.

[34] Winn, S. R.; Randolph, G.; Uludag, H.; Wong, S. C.; Hair, G. A.; Hollinger, G. O. *J. Bone Miner. Res.* **1999**, *14*, 1-13.

[35] Boskey, A. L.; Roy, R. *Chem. Rev.* **2008**, *108*, 4716-4733.