# Protein Misfolding, Functional Amyloid, and Human Disease

Fabrizio Chiti[1] and Christopher M. Dobson[2]

[1]Dipartimento di Scienze Biochimiche, Università degli Studi di Firenze, I-50134 Firenze, Italy; email: fabrizio.chiti@unifi.it

[2]Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW, United Kingdom; email: cmd44@cam.ac.uk

Annu. Rev. Biochem.
2006. 75:333–66

The *Annual Review of Biochemistry* is online at
biochem.annualreviews.org

doi: 10.1146/
annurev.biochem.75.101304.123901

Copyright © 2006 by
Annual Reviews. All rights
reserved

0066-4154/06/0707-
0333$20.00

## Key Words

aggregation mechanism, Alzheimer, Parkinson, prion, protein aggregation

## Abstract

Peptides or proteins convert under some conditions from their soluble forms into highly ordered fibrillar aggregates. Such transitions can give rise to pathological conditions ranging from neurodegenerative disorders to systemic amyloidoses. In this review, we identify the diseases known to be associated with formation of fibrillar aggregates and the specific peptides and proteins involved in each case. We describe, in addition, that living organisms can take advantage of the inherent ability of proteins to form such structures to generate novel and diverse biological functions. We review recent advances toward the elucidation of the structures of amyloid fibrils and the mechanisms of their formation at a molecular level. Finally, we discuss the relative importance of the common main-chain and side-chain interactions in determining the propensities of proteins to aggregate and describe some of the evidence that the oligomeric fibril precursors are the primary origins of pathological behavior.



Petitioner Merck, Ex. 1139, p. 333
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

## Contents

INTRODUCTION ..... 334
THE ROLE OF AMYLOID-LIKE STRUCTURES IN DISEASE AND IN NORMAL BIOLOGY .. 335
Many Human Diseases Are Associated with Protein Aggregation ..... 336
Formation of Amyloid Fibrils Is Sometimes Exploited by Living Systems ..... 339
Amyloid Structures Can Serve as Nonchromosomal Genetic Elements ..... 339
THE STRUCTURES OF AMYLOID FIBRILS ..... 341
High-Resolution Structural Studies Using Solid-State NMR ..... 341
High-Resolution Structural Studies Using X-ray Crystallography ... 343
Other Approaches to Defining the Structural Properties of Amyloid Fibrils ..... 343
Similarities and Differences in Fibrillar Structures from Various Systems ..... 344
The Polymorphism of Amyloid Fibrils ..... 345
MECHANISMS OF AMYLOID FIBRIL FORMATION ..... 346
Amyloid Formation Occurs via a Nucleated Growth Mechanism .. 347
Oligomers Preceding Amyloid Fibril Formation: Structured Protofibrils ..... 347
Oligomers Preceding Fibril and Protofibril Formation: Unstructured Aggregates ..... 348
Aggregation of Globular Proteins Can Occur via Partial Unfolding ..... 349
Aggregation of Globular Proteins Can Occur via Formation of Native-Like Oligomers ..... 349
A Multitude of Conformational States Is Accessible to Polypeptide Chains ..... 350
THE INFLUENCE OF SEQUENCE ON AMYLOID FORMATION ..... 352
Hydrophobicity, Charge, and Secondary Structure Propensities Strongly Influence Amyloid Formation ..... 353
The Amino Acid Sequence Affects Fibril Structure and Aggregation Rate ..... 353
Unfolded Regions Play Critical Roles in Promoting the Aggregation of Partially Folded States ..... 355
Variations in Fibrillar Structure Can Be Reconciled by Common Determinants of the Aggregation Process ..... 356
THE PATHOGENESIS OF PROTEIN DEPOSITION DISEASES ..... 357
The Search Is on for the Causative Agents of Protein Aggregation Diseases ..... 357
The Toxicity of Prefibrillar Aggregates Results from their Misfolded Nature ..... 358
PERSPECTIVES ..... 359

## INTRODUCTION

Writing a review focused on protein misfolding and the diseases with which it is related is both an exciting and a challenging activity. This is in part because recent interest in this topic has led to an explosion in the number of papers published across a broad spectrum of disciplines, and in part because many of the pathological features of the different diseases, and the characteristics of the proteins



Petitioner Merck, Ex. 1139, p. 334
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

with which they are associated, appear at first sight to be quite diverse. Despite this diversity, it is increasingly evident from the experimental data emerging from a wide range of studies that there are some, perhaps many, common features in the underlying physicochemical and biochemical origins of the various disorders and, indeed, of the cases in which similar processes contribute positively to biological function. It has been one of our primary objectives during the writing of this article to explore the extent to which such common features can provide the foundation on which to develop a deeper understanding of the various phenomena associated with protein misfolding and its consequences. Fortunately, within the past year or two, a variety of excellent reviews and books has appeared on the more specific features of many aspects of this complex subject, such as the two-volume book entitled *Protein Misfolding, Aggregation and Conformational Diseases* (1).

To provide a framework on which to build this article, we first describe the variety of human diseases that are now thought to arise from the misfolding of proteins, particularly those, perhaps the majority, in which misfolding results in the formation of highly organized and generally intractable thread-like aggregates termed amyloid fibrils. We point out, however, that in addition living organisms can take advantage of the inherent ability of proteins to form such structures to generate novel and diverse biological functions. Second, we describe the dramatic advances that have recently been made toward the elucidation of the structures of amyloid fibrils at a molecular level and emphasize that our knowledge of these structures is no longer limited to the notion of a fibrillar morphology and an ordered "cross-β" arrangement of the polypeptide chains of which they are composed. We then describe the progress that is being made toward understanding the mechanism of aggregation and toward identifying the nature of key intermediates in the aggregation process. Finally, we discuss some of the important ideas that are emerging about the pathogenesis of the various protein deposition diseases and show that, in at least some cases, the prefibrillar aggregates, rather than the mature and stable fibrils into which they convert, are the likely origins of pathological behavior.

From the evidence that emerges from such considerations, we have tried to pull together the various threads of this complex subject in an attempt to identify both the common features of the various disorders and the differences that lead to their individual identities. We also try to show that, in delving into the general phenomenon of protein misfolding, considerable light can be shed on the origins of some of the most debilitating and increasingly common diseases that affect humanity as well as on the strategies that are likely to be most effective for their prevention and treatment.

**Protein misfolding:** the conversion of a protein into a structure that differs from its native state

**Amyloid fibrils:** protein aggregates having a cross-β structure and other characeristics, e.g., specific dye-binding

**Protein deposition disease:** any pathological state associated with the formation of intracellular or extracellular protein deposits

## THE ROLE OF AMYLOID-LIKE STRUCTURES IN DISEASE AND IN NORMAL BIOLOGY

A broad range of human diseases arises from the failure of a specific peptide or protein to adopt, or remain in, its native functional conformational state. These pathological conditions are generally referred to as *protein misfolding* (*or protein conformational*) *diseases*. They include pathological states in which an impairment in the folding efficiency of a given protein results in a reduction in the quantity of the protein that is available to play its normal role. This reduction can arise as the result of one of several posttranslational processes, such as an increased probability of degradation via the quality control system of the endoplasmic reticulum, as occurs in cystic fibrosis (2), or the improper trafficking of a protein, as seen in early-onset emphysema (3). The largest group of misfolding diseases, however, is associated with the conversion of specific peptides or proteins from their soluble functional states ultimately into highly organized fibrillar aggregates. These structures are



Petitioner Merck, Ex. 1139, p. 335
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

**Amyloidosis:** any pathological state associated with the formation of extracellular amyloid deposits

**TEM:** transmission electron microscopy

**AFM:** atomic force microscopy

**ThT:** thioflavin T

**CR:** Congo red

**Protofilaments:** the constituent units of amyloid fibrils. They should not be confused with protofibrils

generally described as amyloid fibrils or plaques when they accumulate extracellularly, whereas the term "intracellular inclusions" has been suggested as more appropriate when fibrils morphologically and structurally related to extracellular amyloid form inside the cell (4). For simplicity, however, we shall describe all such species as amyloid fibrils in this article. It is also becoming clear that fibrillar species with amyloid characteristics can serve a number of biological functions in living organisms, provided they form under controlled conditions. Perhaps the most fascinating of these functions lies in the ability of such structures to serve as transmissible genetic traits distinct from DNA genes.

## Many Human Diseases Are Associated with Protein Aggregation

A list of known diseases that are associated with the formation of extracellular amyloid fibrils or intracellular inclusions with amyloid-like characteristics is given in **Table 1**, along with the specific proteins that in each case are the predominant components of the deposits. The diseases can be broadly grouped into neurodegenerative conditions, in which aggregation occurs in the brain, nonneuropathic localized amyloidoses, in which aggregation occurs in a single type of tissue other than the brain, and nonneuropathic systemic amyloidoses, in which aggregation occurs in multiple tissues (**Table 1**).

Some of these conditions, such as Alzheimer's and Parkinson's diseases, are predominantly sporadic (labeled [c] in **Table 1**), although hereditary forms are well documented. Other conditions, such as the lysozyme and fibrinogen amyloidoses, arise from specific mutations and are hereditary (labeled [d] in **Table 1**). In addition to sporadic (85%) and hereditary (10%) forms, spongiform encephalopathies can also be transmissible (5%) in humans as well as in other mammals. It has also been found that intravenous injection or oral administration of preformed fibrils from different sources can result in accelerated AA amyloidosis in mice subjected to an inflammatory stimulus (5, 6). It has therefore been postulated that an environment enriched with fibrillar material could act as a risk factor for amyloid diseases (6). Similarly, injection of the recombinant mouse prion protein in the form of amyloid-like fibrils has been reported to generate disease in mice that express the prion protein (7).

The extracellular proteinaceous deposits found in patients suffering from any of the amyloid diseases have a major protein component that forms the core and then additional associated species, including metal ions, glycosaminoglycans, the serum amyloid P component, apolipoprotein E, collagen, and many others (8, 9). Ex vivo fibrils, representing the amyloid core structures, can be isolated from patients, and closely similar fibrils can also be produced in vitro using natural or recombinant proteins; in this case, mildly denaturing conditions are generally required for their rapid formation, at least for proteins that normally adopt a well-defined folded structure (see below).

The fibrils can be imaged in vitro using transmission electron microscopy (TEM) or atomic force microscopy (AFM). These experiments reveal that the fibrils usually consist of a number (typically 2–6) of protofilaments, each about 2–5 nm in diameter (10). These protofilaments twist together to form rope-like fibrils that are typically 7–13 nm wide (10, 11) or associate laterally to form long ribbons that are 2–5 nm thick and up to 30 nm wide (12–14). X-ray fiber diffraction data have shown that in each individual protofilament the protein or peptide molecules are arranged so that the polypeptide chain forms β-strands that run perpendicular to the long axis of the fibril (11). The fibrils have the ability to bind specific dyes such as thioflavin T (ThT) and Congo red (CR) (15), although the specificity of binding of CR to amyloid fibrils and the resulting green birefringence under



Petitioner Merck, Ex. 1139, p. 336
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

**Table 1 Human diseases associated with formation of extracellular amyloid deposits or intracellular inclusions with amyloid-like characteristics**

| Disease | Aggregating protein or peptide | Number of residues[a] | Native structure of protein or peptide[b] |
|---|---|---|---|
| Neurodegenerative diseases | | | |
| Alzheimer's disease[c] | Amyloid β peptide | 40 or 42[f] | Natively unfolded |
| Spongiform encephalopathies[c,e] | Prion protein or fragments thereof | 253 | Natively unfolded (residues 1–120) and α-helical (residues 121–230) |
| Parkinson's disease[c] | α-Synuclein | 140 | Natively unfolded |
| Dementia with Lewy bodies[c] | α-Synuclein | 140 | Natively unfolded |
| Frontotemporal dementia with Parkinsonism[c] | Tau | 352–441[f] | Natively unfolded |
| Amyotrophic lateral sclerosis[c] | Superoxide dismutase 1 | 153 | All-β, Ig like |
| Huntington's disease[d] | Huntingtin with polyQ expansion | 3144[g] | Largely natively unfolded |
| Spinocerebellar ataxias[d] | Ataxins with polyQ expansion | 816[g,h] | All-β, AXH domain (residues 562–694); the rest are unknown |
| Spinocerebellar ataxia 17[d] | TATA box-binding protein with polyQ expansion | 339[g] | α+β, TBP like (residues 159–339); unknown (residues 1–158) |
| Spinal and bulbar muscular atrophy[d] | Androgen receptor with polyQ expansion | 919[g] | All-α, nuclear receptor ligand-binding domain (residues 669–919); the rest are unknown |
| Hereditary dentatorubral-pallidoluysian atrophy[d] | Atrophin-1 with polyQ expansion | 1185[g] | Unknown |
| Familial British dementia[d] | ABri | 23 | Natively unfolded |
| Familial Danish dementia[d] | ADan | 23 | Natively unfolded |
| Nonneuropathic systemic amyloidoses | | | |
| AL amyloidosis[c] | Immunoglobulin light chains or fragments | ~90[f] | All-β, Ig like |
| AA amyloidosis[c] | Fragments of serum amyloid A protein | 76–104[f] | All-α, unknown fold |
| Familial Mediterranean fever[c] | Fragments of serum amyloid A protein | 76–104[f] | All-α, unknown fold |
| Senile systemic amyloidosis[c] | Wild-type transthyretin | 127 | All-β, prealbumin like |
| Familial amyloidotic polyneuropathy[d] | Mutants of transthyretin | 127 | All-β, prealbumin like |
| Hemodialysis-related amyloidosis[c] | β2-microglobulin | 99 | All-β, Ig like |
| ApoAI amyloidosis[d] | N-terminal fragments of apolipoprotein AI | 80–93[f] | Natively unfolded |
| ApoAII amyloidosis[d] | N-terminal fragment of apolipoprotein AII | 98[i] | Unknown |
| ApoAIV amyloidosis[c] | N-terminal fragment of apolipoprotein AIV | ~70 | Unknown |
| Finnish hereditary amyloidosis[d] | Fragments of gelsolin mutants | 71 | Natively unfolded |
| Lysozyme amyloidosis[d] | Mutants of lysozyme | 130 | α+β, lysozyme fold |
| Fibrinogen amyloidosis[d] | Variants of fibrinogen α-chain | 27–81[f] | Unknown |
| Icelandic hereditary cerebral amyloid angiopathy[d] | Mutant of cystatin C | 120 | α+β, cystatin like |

(*Continued*)



Petitioner Merck, Ex. 1139, p. 337
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

**Table 1** (*Continued*)

| Disease | Aggregating protein or peptide | Number of residues[a] | Native structure of protein or peptide[b] |
|---|---|---|---|
| Nonneuropathic localized diseases | | | |
| Type II diabetes[c] | Amylin, also called islet amyloid polypeptide (IAPP) | 37 | Natively unfolded |
| Medullary carcinoma of the thyroid[c] | Calcitonin | 32 | Natively unfolded |
| Atrial amyloidosis[c] | Atrial natriuretic factor | 28 | Natively unfolded |
| Hereditary cerebral haemorrhage with amyloidosis[d] | Mutants of amyloid β peptide | 40 or 42[f] | Natively unfolded |
| Pituitary prolactinoma | Prolactin | 199 | All-α, 4-helical cytokines |
| Injection-localized amyloidosis[c] | Insulin | 21 + 30[j] | All-α, insulin like |
| Aortic medial amyloidosis[c] | Medin | 50[k] | Unknown |
| Hereditary lattice corneal dystrophy[d] | Mainly C-terminal fragments of kerato-epithelin | 50–200[f] | Unknown |
| Corneal amylodosis associated with trichiasis[c] | Lactoferrin | 692 | α+β, periplasmic-binding protein like II |
| Cataract[c] | γ-Crystallins | Variable | All-β, γ-crystallin like |
| Calcifying epithelial odontogenic tumors[c] | Unknown | ~46 | Unknown |
| Pulmonary alveolar proteinosis[d] | Lung surfactant protein C | 35 | Unknown |
| Inclusion-body myositis[c] | Amyloid β peptide | 40 or 42[f] | Natively unfolded |
| Cutaneous lichen amyloidosis[c] | Keratins | Variable | Unknown |

[a]Data refer to the number of residues of the processed polypeptide chains that deposit into aggregates, not of the precursor proteins.
[b]According to Structural Classification Of Proteins (SCOP), these are the structural class and fold of the native states of the processed peptides or proteins that deposit into aggregates prior to aggregation.
[c]Predominantly sporadic, although in some cases hereditary forms associated with specific mutations are well documented.
[d]Predominantly hereditary, although in some cases sporadic forms are documented.
[e]Five percent of the cases are transmitted (e.g., iatrogenic).
[f]Fragments of various lengths are generated and have been reported to be present in ex vivo fibrils.
[g]Lengths shown refer to the normal sequences with nonpathogenic traits of polyQ.
[h]Length shown is for ataxin-1.
[i]The pathogenic mutation converts the stop codon into a Gly codon, extending the 77-residue protein by 21 additional residues.
[j]Human insulin consists of two chains (A and B, with 21 and 30 residues, respectively) covalently linked by disulfide bridges.
[k]Medin is the 245–294 fragment of human lactadherin.

cross-polarized light has recently been questioned (16, 17).

The proteins found as intractable aggregates in pathological conditions do not share any obvious sequence identity or structural homology to each other. Considerable heterogeneity also exists as to secondary structure composition or chain length (**Table 1**). Interestingly, some amyloid deposits in vivo and fibrils generated in vitro have both been found to include higher-order assemblies, including highly organized species known as spherulites, which can be identified from a characteristic Maltese cross pattern when observed under cross-polarized light (18, 19). Such species are also observed in preparations of synthetic polymers, such as polyethylene, a finding consistent with the idea that amyloid fibrils have features analogous to those of classical polymers.



Petitioner Merck, Ex. 1139, p. 338
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

## Formation of Amyloid Fibrils Is Sometimes Exploited by Living Systems

An increasing number of proteins with no link to protein deposition diseases has been found to form, under some conditions in vitro, fibrillar aggregates that have the morphological, structural, and tinctorial properties that allow them to be classified as amyloid fibrils (20, 21). This finding has led to the idea that the ability to form the amyloid structure is an inherent or generic property of polypeptide chains, although, as we discuss below, the propensity to form such a structure can vary dramatically with sequence. This generic ability can increasingly be seen to have been exploited by living systems for specific purposes, as some organisms have been found to convert, during their normal physiological life cycle, one or more of their endogenous proteins into amyloid fibrils that have functional rather than disease-associated properties. A list of such proteins is reported in **Table 2**.

One particularly well-studied example of functional amyloid is that of the proteinaceous fibrils formed from the protein curlin that are used by *Escherichia coli* to colonize inert surfaces and mediate binding to host proteins. Consistent with the characteristics of other amyloid structures, these fibrils are 6–12 nm in diameter, possess extensive β-sheet structure, as revealed by circular dichroism (CD) spectroscopy, and bind to CR and ThT (22). A second example involves the filamentous bacterium *Streptomyces coelicolor* that produces aerial hyphae, which allow its spores to be dispersed efficiently; a class of secreted proteins called chaplins has been identified in the hyphae of this organism with the ability to form amyloid fibrils that act cooperatively to bring about aerial development (23). All these systems have extremely highly regulated assembly processes; generation of the bacterial *curli*, for example, involves several proteins, including one that nucleates a different protein to form fibrils.

As well as these examples from bacteria, the formation of functional amyloid-like structures has recently been observed in a mammalian system. The melanosomes, lysosome-related organelles that differentiate in melanocytes to allow the epidermal production of the melanin pigment, are characterized by intralumenal fibrous striations upon which melanin granules form. This fibrous material, sharing significant analogies with amyloid fibrils, is assembled from the intralumenal domain of the membrane protein Pmel17 that is proteolyzed by a proprotein convertase (24). This result is a direct indication that even in higher organisms amyloid formation can be physiologically useful for specific and specialized biological functions, provided it is regulated and allowed to take place under highly controlled conditions.

**Functional amyloid:** an amyloid structure found to have a beneficial function in living systems

**CD:** circular dichroism

## Amyloid Structures Can Serve as Nonchromosomal Genetic Elements

As we discussed in the previous paragraph, it is clear that living systems can utilize the amyloid structure as the functional state of some specific proteins. It is also clear, however, that nature has selected, or at the very least has not selected against, some proteins that can exist within normally functioning biological systems in both a soluble conformation and in an aggregated amyloid-like form. Remarkably, this phenomenon has resulted in the latter state being self-perpetuating, infectious, and inheritable as a non-Mendelian nonchromosomal genetic trait (25). Proteins with such behavior are called prions and are listed in **Table 2**. Although the only endogenous mammalian protein so far recognized to have such properties is associated with the group of invariably fatal and transmissible diseases, the heritable conformational changes of prion proteins from some other organisms have, in some cases, been found beneficial.

The prion proteins from *Saccharomyces cerevisiae*, including Ure2p and Sup35p, give rise to distinct phenotypes when adopting either one or the other forms of the



Petitioner Merck, Ex. 1139, p. 339
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

**Table 2 Proteins forming naturally nonpathological amyloid-like fibrils with specific functional roles**

| Protein | Organism | Function of the resulting amyloid-like fibrils | References |
|---|---|---|---|
| Curlin | *Escherichia coli* (bacterium) | To colonize inert surfaces and mediate binding to host proteins | 22 |
| Chaplins | *Streptomyces coelicolor* (bacterium) | To lower the water surface tension and allow the development of aerial hyphae | 23 |
| Hydrophobin[a] EAS | *Neurospora crassa* (fungus) | To lower the water surface tension and allow the development of aerial hyphae | 23a |
| Proteins of the chorion of the eggshell[b] | *Bombyx mori* (silkworm) | To protect the oocyte and the developing embryo from a wide range of environmental hazards | 23b |
| Spidroin | *Nephila edulis* (spider) | To form the silk fibers of the web | 23c |
| Intralumenal domain of Pmel17 | *Homo sapiens* | To form, inside melanosomes, fibrous striations upon which melanin granules form | 24 |
| Ure2p (prion) | *Saccharomyces cerevisiae* (yeast) | To promote the uptake of poor nitrogen sources ([URE3]) | 25 |
| Sup35p (prion) | *Saccharomyces cerevisiae* (yeast) | To confer new phenotypes ([PSI+]) by facilitating the readthrough of stop codons on mRNA | 26–28 |
| Rnq1p (prion) | *Saccharomyces cerevisiae* (yeast) | Not well understood ([RNQ+], also known as [PIN+], phenotype) | 28a |
| HET-s (prion) | *Podospora anserina* (fungus) | To trigger a complex programmed cell death phenomenon (heterokaryon incompatibility) | 31, 32 |
| Neuron-specific isoform of CPEB (prion) | *Aplisia californica* (marine snail) | To promote long-term maintenance of synaptic changes associated with memory storage | 30 |

[a]Other proteins from this class, collectively called hydrophobins, have been found to play similar roles in other species of filamentous fungi.
[b]Suggested to form amyloid-like fibrils in vivo, although amyloid formation has only been observed in vitro.

protein (soluble or fibrillar). These proteins are not related to each other, although they do have some characteristics in common, such as the presence of a globular domain and an unstructured portion of the sequence and the high occurrence of glutamine and asparagine residues in the unstructured domain. The polymerization-mediated inactivation of Sup35p, a protein involved in the termination of mRNA translation, confers a wide variety of novel phenotypes ([PSI+]) by facilitating the readthrough of stop codons (26–28). The aggregation of Ure2p destroys its ability to bind and sequester the transcription factor Gln3p; this results in the activation of a series of genes involved in the uptake of poor nitrogen sources (25). The resulting yeast cells [URE3] can grow on media that, for example, lack uracil but contain its precursor ureidosuccinate (25). Although the low natural occurrence of [URE3] and [PSI+] strains suggest that the corresponding phenotypes are not generally beneficial (29), they can still be advantageous under particular environmental circumstances.

In the marine snail *Aplysia californica*, a neuron-specific isoform of cytoplasmic polyadenylation element-binding protein (CPEB) has also been found to exist in a soluble and a self-perpetuating prion form (30). The prion form was found to be more active than the soluble form in stimulating translation of CPEB-regulated mRNA. From this finding, the suggestion was made that the polymerization of the protein could be essential for the long-term maintenance of synaptic changes associated with memory storage (30). Finally, the polymerization of the HET-s protein from *Podospora anserina* is involved in a controlled programmed cell death phenomenon termed heterokaryon incompatibility (31, 32).



Petitioner Merck, Ex. 1139, p. 340
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

It is evident, even from the relatively few examples that have been studied in detail so far, that the aggregation of proteins into amyloid-like structures can generate a number of extremely diverse biological functions. The presence of many other sequences in the genomes of different organisms with the characteristics of prions suggests that there may yet be surprises in store for us when their properties are investigated.

## THE STRUCTURES OF AMYLOID FIBRILS

For many years the only structural information about amyloid fibrils came from imaging techniques such as TEM, and more recently AFM, and from X-ray fiber diffraction (10, 11, 33). Despite the structural insight given by these techniques, as outlined above, one of the most common statements in the introductory sections of papers in this field until about three years ago was to the effect that "amyloid fibrils cannot be characterized in detail at the molecular level because they are not crystalline yet they are too large to be studied by solution NMR spectroscopy." The situation has changed dramatically recently as a result of major progress in the application of solid-state NMR (SSNMR) spectroscopy to preparations of amyloid fibrils (34–36) and of successes in growing nano- or microcrystals of small peptide fragments that have characteristics of amyloid fibrils yet are amenable to single crystal X-ray diffraction analysis (37, 38).

### High-Resolution Structural Studies Using Solid-State NMR

Using SSNMR, in conjunction with computational energy minimization procedures, Tycko and coworkers (34, 39, 40) have put forward a structure of the amyloid fibrils formed from the 40-residue form of the amyloid β peptide ($A\beta_{1-40}$) at pH 7.4 and 24°C under quiescent conditions. In this structure, each $A\beta_{1-40}$ molecule contributes a pair of β-strands, spanning approximately residues 12–24 and 30–40, to the core region of the fibrils (**Figure 1*a***). These strands, connected by the loop 25–29, are not part of the same β-sheet, however, but participate in the formation of two distinct β-sheets within the same protofilament (**Figure 1*a***). The different Aβ molecules are stacked on to each other, in a parallel arrangement and in register, at least from residue 9 to 39 (39, 40). By invoking additional experimental constraints, such as the diameter of the protofilaments observed using TEM, and the mass per unit length, measured by means of scanning transmission electron microscopy (STEM) (34, 41), it has been suggested that a single protofilament is composed of four β-sheets separated by distances of ~10 Å (**Figure 1*a***).

Support for key elements of this proposed structure comes from experiments of site-directed spin labeling coupled to electron paramagnetic resonance (SDSL-EPR) (42). The values of the inverse central line width in the EPR spectra for a series of labeled residues indicate that the segments of the $A\beta_{1-42}$ molecule corresponding to residues 13–21 and 30–39 are highly structured in the fibrils, parallel and in register. High flexibility and exposure to the solvent of the N-terminal region, in contrast to considerable structural rigidity detected for the remainder of the sequence, are also suggested by experimental strategies that use hydrogen-deuterium exchange methods in conjunction with mass spectrometry (43), limited proteolysis (44), and proline-scanning mutagenesis (45).

SSNMR, in conjunction with site-directed fluorescence labeling and an ingenious hydrogen/deuterium exchange protocol applied previously to probe the regions of β2-microglobulin fibrils that are involved in persistent structure (46), has led to identification of the regions of the C-terminal fragment of HET-s that are involved in the core of the fibril (36). In the proposed structure, each molecule contributes four β-strands, with strands one and three forming the same parallel β-sheet and with strands two and four

**SSNMR:** solid-state nuclear magnetic resonance

**STEM:** scanning transmission electron microscopy

**SDSL-EPR:** site-directed spin labeling coupled to electron paramagnetic resonance

**Aβ:** amyloid β peptide



Petitioner Merck, Ex. 1139, p. 341
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

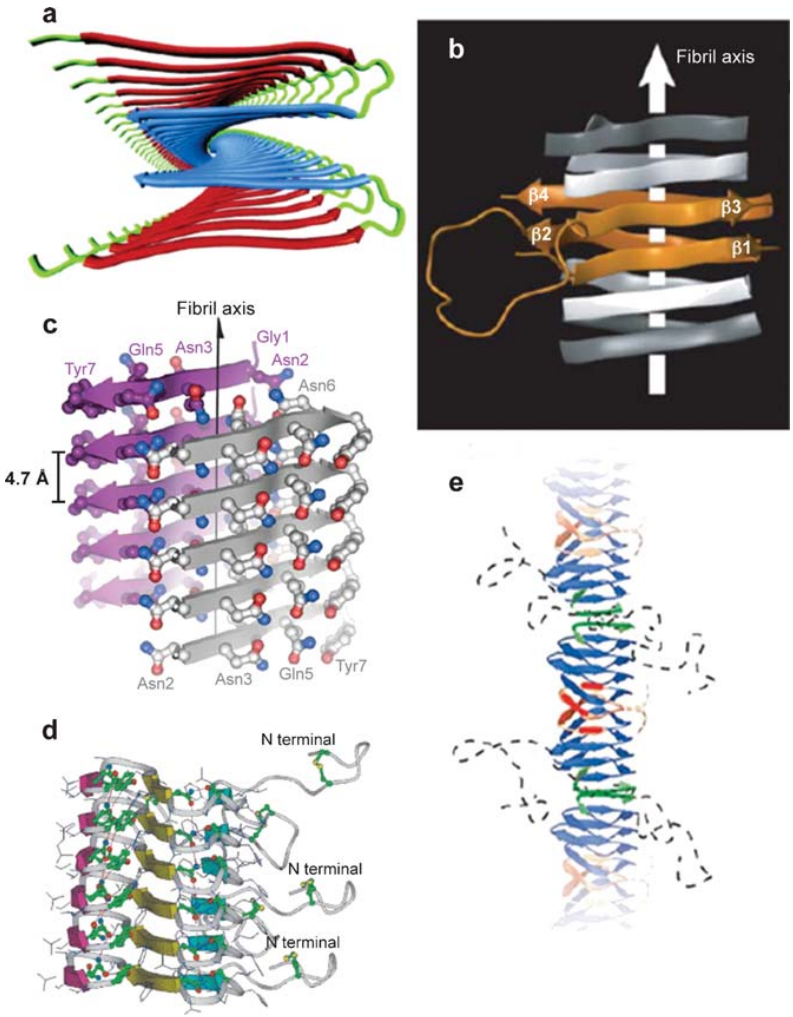


**Figure 1**

Recent three-dimensional structural models of fibrillar aggregates from different sources. (*a*) The protofilament of Aβ viewed down the long axis of the fibril. Reprinted with permission (177), copyright (2003) American Chemical Society. The segments 12–24 (*red*) and 30–40 (*blue*) are shown. (*b*) The fibril from the C-terminal domain 218–289 of the fungal prion protein HET-s [reproduced with permission (36)]. The ribbon diagram shows the four β-strands (*orange*) (residues 226–234, 237–245, 262–270, and 273–282) and the long loop between β2 and β3 from one molecule. Flanking molecules along the fibril axis (*gray*) are shown. (*c*) Atomic structure of the microcrystals assembled from the GNNQQNY peptide [reproduced with permission (38)]. Each β-strand is a peptide molecule. (*d*) The protofilament from amylin [reprinted with permission from Elsevier (51)]. Green, yellow, and pink β-strands indicate residues 12–17, 22–27, and 31–37, respectively. The unstructured N-terminal tail is shown on the right of the panel along with the disulfide bridge between Cys2 and Cys7. (*e*) The fibril from the NM region of Sup35p [reproduced with permission (52)]. The colored ribbons indicate residues 25–38 (*red*), 39–90 (*blue*), and 91–106 (*green*). The unstructured regions 1–20 (*red dashed lines*) and 158–250 (*black dashed lines*) are shown.



Petitioner Merck, Ex. 1139, p. 342
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

forming another parallel β-sheet ~10 Å away (**Figure 1*b***).

Advances in SSNMR techniques that enable specific internuclear distances and torsion angles to be measured have also allowed the structure of a 11-residue fragment of transthyretin within an amyloid-like fibril to be defined in atomic detail (35, 47). This study shows that the peptide adopts an extended β-strand within the fibrils. Most importantly, however, this pioneering study reveals that the molecules within the fibrils possess a degree of uniformity, even at the level of the side-chain torsion angles, that has previously only been associated with crystalline materials. Because this regularity is reflected in the very narrow resonance lines in the SSNMR spectra, we can anticipate that complete atomic-level structures will soon begin to emerge for a range of systems, transforming our understanding of this facet of the amyloid phenomenon.

## High-Resolution Structural Studies Using X-ray Crystallography

The remarkable achievement of inducing a peptide derived from Sup35p (GNNQQNY) and another with sequence KFFEAAAKKFFE, to form three-dimensional crystals that possess key characteristics of amyloid fibrils, has allowed both the structure of the peptides and the way the molecules could be packed together to be determined with unprecedented resolution (37, 38). In the case of the Sup35p fragment, the crystal consists of pairs of parallel β-sheets in which each individual peptide molecule contributes a single β-strand (**Figure 1*c***). The stacked β-strands are parallel and in register in both sheets. The two sheets interact with each other through the side chains of Asn2, Gln4, and Asn6 to such a degree that water is excluded from the region between them. The remaining side chains on the outer faces of the sheets are hydrated and more distant from the next pair of β-sheets, suggesting that this less intimate interaction could represent a crystal contact rather than a feature of the fibrillar state.

A particularly significant aspect of these structures determined with X-ray or SSNMR is that they are strikingly similar to proposals from cryo-electron microscopy (EM) analysis of the amyloid fibrils formed from an SH3 domain and from insulin, in which the electron density maps were interpreted as arising from pairs of relatively flat untwisted β-sheets (48, 49). Such similarities suggest that many amyloid fibrils could have core structures that have very similar features, which are primarily dictated by the intrinsic conformational preferences of polypeptide chains. The specific nature of the side-chain packing, including such characteristics as the alignment of adjacent strands and the separation of the sheets (50), however, provides an explanation for the occurrence of variations in the details of the structures for specific types of fibril. Hopefully, these pioneering X-ray and SSNMR studies may represent the first entries in a new database of structures similar to the current Protein Data Bank.

## Other Approaches to Defining the Structural Properties of Amyloid Fibrils

As discussed above, SSNMR and X-ray crystallography have recently made major contributions to our knowledge of the structures of amyloid fibrils. Considerable progress in this quest has also come from other approaches, typically involving the combination of data from a number of different biophysical experiments (13, 51, 52). One example is the three-dimensional structure proposed for amyloid fibrils from amylin (**Figure 1*d***) (51). The polypeptide chains were configured within the fibrils on the basis of a cross-β structure, deduced from X-ray diffraction data along with measurements of the protofilament diameter and mass per unit length, determined using TEM and STEM, respectively (53, 54). Additional constraints were provided by evidence of a parallel and in



Petitioner Merck, Ex. 1139, p. 343
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

register arrangement of the β-strands formed by adjacent molecules from SDSL-EPR data (55) and by evidence of the high propensity of various amylin segments to form fibrils when dissected from the rest of the sequence (51). In the resulting model, the N-terminal "tail" (residues 1–11) is unstructured, and residues 12–17, 22–27, and 31–37 form β-strands in a serpentine arrangement, contributing to different β-ribbons in the protofilament (**Figure1*d***).

In another particularly elegant example, detailed structural information on the fibrils, formed from the NM region (residues 1–250) of the yeast prion protein Sup35p, was also obtained by combining a variety of experimental strategies (52). Carefully chosen residues spaced along the fragment of the protein were mutated so as to generate 37 variants, each having a single cysteine residue at a desired position in the molecule. The variants were then labeled with fluorescent probes. The wavelength maximum and total emission intensity of the fluorescent probes were then used to provide information about the degree of burial from solvent of the various residues and about the distances between probes attached to different molecules within the fibrils. Dimeric constructs were also generated for each variant by covalently linking the free thiol group of one molecule to the same group in a second molecule, either directly by a disulfide bridge or by the insertion of a linker. The ability, or lack of ability, of such dimers to form fibrils was used to estimate the distances between corresponding regions of the sequence from adjacent molecules in the fibrils.

Taken together, these complementary sets of data allowed a model to be defined that describes the molecular structure of the filaments (52). In this structure (**Figure 1*e***), two segments of the N domain, corresponding to residues 25–38 and 91–106 (colored green and red in **Figure 1*e***, respectively), interact with the corresponding regions in other molecules to form a "head-to-head" and "tail-to-tail" arrangement. The large central region of the sequence between these two segments (blue in **Figure 1*e***) is folded in such a way that it forms only intramolecular interactions. The C-terminal region of the N domain and the proximal portion of the M domain (residues 107–157) are also structured within the fibrils, whereas the N-terminal region (residues 1–20) and the distal end of the M domain (residues 158–250) appear to be structurally heterogeneous and solvent exposed (dashed lines in **Figure 1*e***).

Finally, although detailed structural models have not yet been proposed, much has been learned about the characteristics of other types of fibrils through similar approaches. This has led, for example, to the identification of regions of the polypeptide chain that are associated with an ordered structure in α-synuclein and tau fibrils using SDSL-EPR (56, 57). It was also possible to determine the most structured regions in α-synuclein as well as in both straight and curly fibrils from β2-microglobulin using hydrogen-deuterium exchange (46, 58, 58a), limited proteolysis (59, 59a), and SSNMR (59b). In addition, from X-ray fiber diffraction studies a cylindrical β-sheet model for fibrils from a poly-Gln peptide and the exon-1 peptide of huntingtin has been proposed (60). The polyglutamine fibrils are of particular interest because of the possibility that the additional array of hydrogen-bonding interactions involving the side chains results in a structure significantly different from that of the classical amyloid fibrils. Evidence that this situation can arise comes from the absence of the 10 Å reflection in the X-ray fiber diffraction patterns of these systems.

## Similarities and Differences in Fibrillar Structures from Various Systems

Comparison of the information about the structural properties of various fibrillar systems, discussed in the previous three paragraphs, allows us to draw a number of tentative conclusions about their similarities and differences. Different fibrils clearly have many properties in common, including the



Petitioner Merck, Ex. 1139, p. 344
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

canonical cross-β structure and the frequent presence of repetitive hydrophobic or polar interactions along the fibrillar axis. The ubiquitous presence of a cross-β structure strongly supports the view that the physicochemical properties of the polypeptide chain are the major determinants of the fibrillar structure in each case. Moreover, several of the proposed structures, despite very different sequences of their component polypeptides, suggest that the core region is composed of two to four sheets that interact closely with each other. An interesting feature of these sheets is that they appear to be much less twisted than expected from the analysis of the short arrays of β-strands that form β-sheets in globular protein structures. This feature was first proposed from cryo-EM and has been supported by Fourier transform infrared (FTIR) analyses (48, 61).

Nevertheless, it is clear that there are significant differences in detail attributable to the influence of the side chains on the structures adopted by the various systems. These appear to include the lengths of the β-strands and whether they are arranged in a parallel or antiparallel arrangement within each sheet; the lengths and conformational properties of the loops, turns, and other regions that are not included within the core structure; and the number of β-sheets in the protofilament. It is clear that the fraction of the residues of a polypeptide chain that are incorporated in the core structure can vary substantially (e.g., from all the residues of the 7-mer peptide to only about 13% of the residues in the full-length HET-s) and that the exact spacing between the β-sheets varies with factors such as the steric bulk of the side chains that are packed together in the core (50). In addition, the presence of disulfide bonds in proteins such as insulin may perturb the way in which the sheets can stack together (49). In cases such as the polyglutamine sequences, other interactions between the side chains may generate larger perturbations of the structure to generate such motifs as β-helices (60), which are also seen under similar circumstances in the structures of globular proteins.

The structure that will normally be adopted in the fibrils will be the lowest in free energy and/or the most kinetically accessible. What is clear, therefore, is that the interactions of the various side chains with each other and with solvent are crucial in determining the variations in the fibrillar architecture even though the main-chain interactions determine the overall framework within which these variations can occur. In other words, the interactions and conditions (see below) involving the side chains in a given sequence can tip the balance between the alternative "variations on a common theme" arrangements of a polypeptide "polymer" chain in its fibrillar structure. Such a situation contrasts with that pertaining to the native structures of the highly selected protein molecules, which are able to fold to unique structures that are significantly more stable for a given sequence than any alternatives.

**FTIR:** Fourier transform infrared

## The Polymorphism of Amyloid Fibrils

Even before the molecular structures of amyloid fibrils began to emerge, it was clear that significant morphological variation can exist between different fibrils formed from the same peptide or protein (12, 48, 49, 54). Evidence is now accumulating that such variations in morphology is linked to heterogeneity in molecular structure, i.e., in the structural positioning of the polypeptide chains within the fibrils. One example of such heterogeneity involves the peptide hormone glucagon, wherein fibrils formed at different temperatures (25°C or 50°C) are morphologically distinct; measurements of CD and FTIR spectroscopy reveal differences in the secondary structure adopted by the constituent peptide molecules (14). A particularly important study in this regard addresses the origin of the marked differences in the morphology of $A\beta_{1-40}$ fibrils that can be observed



Petitioner Merck, Ex. 1139, p. 345
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

in TEM studies of samples prepared under agitation or quiescent conditions; differences in the SSNMR spectra recorded from the different preparations provide clear evidence that this polymorphism is linked to differences in molecular structure (62).

Another example of conformational variability involves fibrils formed from the yeast prion protein Ure2p, where two independent studies came to somewhat different conclusions about the fibril structure. Both studies find that the globular C-terminal domain maintains a largely native-like structure. However, in one case, it appears that the fibrils possess a cross-β core involving only the N-terminal domains, each arranged in a serpentine fashion and forming a series of consecutive strands and loops (63–65). The parallel and in register stacking of serpentines from different molecules then forms the cross-β core with the C-terminal globular units decorating it (64). In the other study, the C- and N-terminal domains of the protein appear to interact with each other, and these fibrils do not have the characteristic 4.7-Å reflection typical of a cross-β structure (17, 66, 67). These apparently conflicting reports are likely to reflect structural differences in the fibrils, probably caused by the slightly different conditions used to prepare them.

Conformational polymorphism has also been found in other yeast prion proteins and is of particular significance because of the light it sheds on the existence of "strains" of mammalian prions and on the nature of the crucial barriers to infectivity that limit transmissibility between species (68). Efficiency of interspecies prion transmission decreases as the sequences of the infectious prions diverge, probably because each prion sequence can give rise to a limited number of misfolded conformations, which have low cross-seeding efficiency. However, a strain conformation of Sup35p has recently been identified that allows transmission from *S. cerevisiae* to the highly divergent *Candida albicans* (68). Similarly, mammalian $PrP_{23-144}$ fibrils from different species vary in morphology and secondary structure, and these differences appear to be controlled by one or two residues in a critical region of the polypeptide sequence (69).

In all of these cases, preformed seeds can propagate their morphology and structure as well as overcome sequence- or condition-based structural preferences, resulting in fibrils that inherit the characteristics of the template (14, 62, 68, 69). These results show that each protein sequence can form a spectrum of structurally distinct fibrillar aggregates and that kinetic factors can dictate which of these alternatives is dominant under given circumstances. Of the many possible conformations that could be present in the amyloid core for a given protein, the specific ones that play this role will depend simply on the thermodynamic and, in many cases, the kinetic factors that are dominant under those circumstances. By contrast, natural globular proteins have been selected by evolution to fold into one specific three-dimensional structure, and the complex free-energy landscapes associated with their sequences have a single and well-defined minimum, under physiological conditions, corresponding to the native state.

## MECHANISMS OF AMYLOID FIBRIL FORMATION

The full elucidation of the aggregation process of a protein requires the identification of all the conformational states and oligomeric structures adopted by the polypeptide chain during the process and the determination of the thermodynamics and kinetics of all the conformational changes that link these different species. It also implies characterizing each of the transitions in molecular detail and identifying the residues or regions of the sequence that promote the various aggregation steps. The identification and characterization of oligomers preceding the formation of well-defined fibrils is of particular interest because of an increasing awareness that these species are likely to play a critical role in the pathogenesis of protein deposition diseases.



Petitioner Merck, Ex. 1139, p. 346
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

## Amyloid Formation Occurs via a Nucleated Growth Mechanism

It is widely established that amyloid fibril formation has many characteristics of a "nucleated growth" mechanism. The time course of the conversion of a peptide or protein into its fibrillar form (measured by ThT fluorescence, light scattering, or other techniques) typically includes a lag phase that is followed by a rapid exponential growth phase (70–73). The lag phase is assumed to be the time required for "nuclei" to form. Once a nucleus is formed, fibril growth is thought to proceed rapidly by further association of either monomers or oligomers with the nucleus.

Such a nucleated growth mechanism has been well studied both experimentally and theoretically in many other contexts, most notably for the process of crystallization of both large and small molecules (74). As with many other processes dependent on a nucleation step, including crystallization, addition of preformed fibrillar species to a sample of a protein under aggregation conditions ("seeding") causes the lag phase to be shortened and ultimately abolished when the rate of the aggregation process is no longer limited by the need for nucleation (70, 71). It has been shown also that changes in experimental conditions, or certain types of mutations, can also reduce or eliminate the length of the lag phase, again assumed to result from a situation wherein nucleation is no longer rate limiting (72, 73, 75). The absence of a lag phase, therefore, does not necessarily imply that a nucleated growth mechanism is not operating, but it may simply be that the time required for fibril growth is sufficiently slow relative to the nucleation process and that the latter is no longer the slowest step in the conversion of a soluble protein into the amyloid state. Although fibrils do not appear to a significant extent during the lag phase, it is increasingly clear that this stage in fibril formation is an important event in which a variety of oligomers form, including β-sheet-rich species that provide nuclei for the formation of mature fibrils.

The efficiency of preformed fibrils to promote further aggregation through a seeding mechanism decreases dramatically as the sequences diverge (68, 76, 76a). Using a number of immunoglobulin domains sharing different degrees of sequence identity, it was shown that coaggregation between different types of domain is not detectable if the sequence identity is lower than ~30% to 40% (76). A bioinformatics analysis of consecutive homologous domains in large multimodular proteins shows that such domains almost exclusively have sequence identities of less than 40%, suggesting that such low sequence identities could play a crucial role in safeguarding proteins against aggregation (76).

**Oligomers:** clusters of small numbers of protein or peptide molecules without a fibrillar appearance

**Protofibrils:** protein aggregates of isolated or clustered spherical beads 2–5 nm in diameter with β-sheet structure

## Oligomers Preceding Amyloid Fibril Formation: Structured Protofibrils

The past decade has seen very substantial efforts directed toward identifying, isolating, and characterizing the oligomeric species that are present in solution prior to the appearance of fibrils, both because of their likely role in the mechanism of fibril formation and because of their implication as the toxic species involved in neurodegenerative disorders. We focus initially on amyloid formation by the Aβ peptide because this has been widely studied owing to its links with Alzheimer's disease. Aggregation of this peptide is preceded by the formation of a series of metastable, nonfibrillar species that can be visualized using AFM and TEM (33, 77–79). Some appear to be spherical beads of 2–5 nm in diameter. Others appear to be beaded chains with the individual beads again having a diameter of 2–5 nm and seeming to assemble in linear and curly chains. Yet others appear as annular structures, apparently formed by the circularization of the beaded chains. All of these aggregates, which have been termed protofibrils by the authors who first observed them (33, 77–79), should not be confused with the



Petitioner Merck, Ex. 1139, p. 347
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

protofilaments that are the constituent units of mature fibrils. Protofibrils from Aβ can bind CR and ThT (79), contain an extensive β-sheet structure (79), and, in the form of the smaller spherical species, are made up of ~20 molecules (80). A first exciting attempt to determine the structure of Aβ protofibrils was published using proline-scanning mutagenesis (81).

Analogous spherical and chain-like protofibrillar structures have been observed for many other systems, including α-synuclein (82), amylin (80), the immunoglobulin light chain (83), transthyretin (84), polyQ-containing proteins (80), β2-microglobulin (85), equine lysozyme (86), the *Sulfolobus solfataricus* acylphosphatase (Sso AcP) (87), and an SH3 domain (87a). These species are generally characterized by extensive β-structure and sufficient structural regularity to bind ThT and CR. The exciting finding that a specific antibody can bind to protofibrillar species from different sources, but not to their corresponding monomeric or fibrillar states, suggests that such soluble amyloid oligomers have some important common structural elements (88).

Data have been reported showing that in some cases protofibrils can be on-pathway to fibrils (33, 71). In other cases, they appear to be off-pathway (85, 89). It has been reported that the transition from the protofibrillar to the fibrillar state of the peptide 109–122 of the Syrian hamster prion protein occurs concomitantly with the alignment of β-strands within sheets in which the strands are initially misaligned (89a). Such an aligment involves detachment and re-annealing of the strands, but may also occur through an internal structural reorganization within the sheets, depending on conditions (89b). Regardless of the precise role played by protofibrils in the overall process of fibril formation, the elucidation of their mechanism of formation and of their structures is extremely important, not least because these species could be the primary toxic agents involved in neurodegenerative disorders.

## Oligomers Preceding Fibril and Protofibril Formation: Unstructured Aggregates

Following the isolation and characterization of protofibrils, studies based on photo-induced cross-linking of unmodified proteins (PICUP) began to identify other oligomeric species that appeared to precede their formation (90, 91). Both the 40 and 42 residue forms of Aβ have been shown to exist as soluble oligomers in rapid equilibrium with the corresponding monomeric forms. These oligomers appear to be composed of 2–4 and 5–6 molecules for $A\beta_{1-40}$ and $A\beta_{1-42}$, respectively, and CD measurements suggest that they are relatively disorganized (91). Interest in these low-molecular-weight oligomers has been particularly intense as species of this type have also been detected in the brains of Alzheimer's disease patients (92) and in the lysates and conditioned media of cultured cells expressing the amyloid β protein precursor (93, 94).

The NM region of the yeast prion Sup35p has been shown to form "structurally fluid" oligomers rapidly, and these oligomers only later convert to species with extensive β-structures that are capable of nucleating fibril formation (71). Such a conversion has been found to be facilitated by the covalent dimerization of NM molecules when residues in the "head" region of N (residues 25–38) are cross-linked (52). Moreover, if the fluid oligomers are maintained under oxidizing conditions, intermolecular disulfide bridges are found to form more easily for variants in which cysteine residues are introduced into the head region of N rather than elsewhere. These results indicate that the interaction of the head regions of two N molecules nucleates the formation of an amyloid-like structure within the aggregates (52).

Similar behavior has been observed for the aggregation of denatured yeast phosphoglycerate kinase at low pH using dynamic light scattering and far-UV CD spectroscopy (95). β-sheet structure is increasingly stabilized as



Petitioner Merck, Ex. 1139, p. 348
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

the aggregates grow in size. When a critical mass is reached, the oligomers associate with each other to form short, curly protofibrils that are similar in appearance to those observed with Aβ and α-synuclein (95). Moreover, unfolding of the SH3 domain from the bovine phosphatidylinositol 3′ kinase at pH 3.6 results in the rapid formation of a broad distribution of unstructured oligomers that subsequently convert into thin, curly, ThT-binding protofibrils (87a). All these experimental results, along with computer simulations carried out using simple polyalanine peptides (96), suggest that structured protofibrillar species can form from the reorganization or assembly of small and relatively disorganized oligomers that are formed rapidly after the initiation of the aggregation process.

## Aggregation of Globular Proteins Can Occur via Partial Unfolding

So far we have discussed systems that are largely unstructured prior to the aggregation process. It is generally believed that globular proteins need to unfold, at least partially, to aggregate into amyloid fibrils (21, 97, 98). Evidence supporting this hypothesis comes from a large body of experimental data. It is clear, for example, that globular proteins have an increased propensity to aggregate under conditions that promote their partial unfolding, such as high temperature, high pressure, low pH, or moderate concentrations of organic solvents (85, 99–102). In addition, for some familial forms of disease in which the proteins involved in aggregation normally adopt folded conformations (see **Table 1**), there is clear evidence that a destabilization of the native structure, resulting in an increase in the population of nonnative states, is the primary mechanism through which natural mutations mediate their pathogenicity (103–105).

A strong correlation between a decreased conformational stability of the native state and an increased propensity to aggregate into amyloid-like structures has also been shown in vitro for nondisease-associated proteins (100, 106). Remarkably, aggregation of human lysozyme and HypF-N can be initiated by a population of less than 1% of a partially folded state that is in equilibrium with the native conformation (104, 107). Conversely, the binding of ligands and other species, such as antibodies, that stabilize the native state can decrease dramatically the propensity of proteins to aggregate (108–111). Such observations have inspired an extensive search of potential pharmaceutical compounds for the treatment of the diseases associated with transthyretin through specific binding to the tetrameric native state of the protein (109).

## Aggregation of Globular Proteins Can Occur via Formation of Native-Like Oligomers

Although the "conformational change hypothesis" is undoubtedly the most appropriate way to describe the formation of amyloid fibrils by many globular proteins, recent observations have suggested that in some cases the major conformational change associated with amyloid aggregation may not take place until after the initial aggregation step. Formation of amyloid fibrils by insulin at low pH, for example, is preceded by an oligomerization step in which a native-like content of α-helical structure is almost completely retained, and aggregates with a morphology reminiscent of amyloid protofibrils and with a high content of β-structure appear only later in the process (112). In addition, within a group of variants of the protein S6 from *Thermus thermophilus*, no significant correlation was found between the rate of fibril formation under conditions in which a quasi-native state was populated prior to aggregation and the unfolding rate or conformational stability (73). Similarly, the native state of the pathogenic variant of ataxin-3, the protein associated with spinocerebellar ataxia type-3, does not appear to be significantly destabilized, leading to the proposal that the pathway for fibril formation can be distinct from that of unfolding (113).



Petitioner Merck, Ex. 1139, p. 349
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

Details of the manner in which aggregation under these conditions can take place has come from studies of the aggregation of Sso AcP. These studies have shown that unfolding of the protein can be two orders of magnitude slower than the formation of amyloid protofibrils when the protein is placed under conditions in which the native state is thermodynamically more stable than the dominant partially unfolded state (87). The first event in the aggregation of Sso AcP under these conditions is the formation of oligomers that do not bind to ThT or CR and, remarkably, not only have a native-like topology but also retain enzymatic activity (114). These native-like oligomers then undergo structural reorganization to form amyloid protofibrils that have extensive β-structure, bind ThT and CR, but are not enzymatically active. The fact that protofibril formation is also faster than the rate of disaggregation of the initially formed oligomers shows that dissolution of the latter followed by renucleation cannot be the dominant process giving rise to the structural conversion.

In the case of Ure2p, a mechanism of the type observed for Sso AcP appears to give rise to a situation wherein a native-like conformation is even retained in the fibrils themselves under some conditions (17, 66, 67). The significant propensity of native or native-like structures to aggregate is not surprising if we consider that there is a multitude of conformers even in the native ensemble of a globular protein (115). Some of these conformers will be only transiently populated but could be significant for aggregation just as they are for the hydrogen exchange of their main-chain amide groups.

Finally in this section, despite their apparent differences, there are in fact substantial similarities between the fundamental mechanism of aggregation described here for folded proteins and that of natively unfolded systems, such as Aβ and Sup35p NM. In both cases, the polypeptide molecules assemble first into species that can have characteristics far from those of the final aggregates but similar to those of the precursor structures, whether natively unfolded or natively folded. The initial aggregates then transform into species that are not yet fibrillar in their morphologies but have other properties characteristic of amyloid-like structures, notably β-sheet structure and binding to CR and ThT. Clearly, fully or partially unfolded states of globular proteins are generally more susceptible to aggregation than the native states. Nevertheless, in some situations, particularly those close to physiological, the much higher populations of the latter can result in their playing an important role in initiating an aggregation process that could be significant on the very slow timescales of the amyloid disorders.

## A Multitude of Conformational States Is Accessible to Polypeptide Chains

The differing features of the aggregation processes, described in the previous paragraphs, reveal that polypeptide chains can adopt a multitude of conformational states and interconvert between them on a wide range of timescales. The network of equilibria, which link some of the most important of such states both inside and outside the cell, is schematically illustrated in **Figure 2**. Following biosynthesis on a ribosome, a polypeptide chain is initially unfolded. It can then populate a wide distribution of conformations, each of which contains little persistent structure, as in the case of natively unfolded proteins, or fold to a unique compact structure, often through one or more partly folded intermediates. In such a conformational state, the protein can remain as a monomer or associate to form oligomers or higher aggregates, some of which are functional with characteristics far from those of amyloid structures, such as in actin, myosin, and microtubules. Sooner or later, the vast majority of proteins will be degraded, usually under very carefully



Petitioner Merck, Ex. 1139, p. 350
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

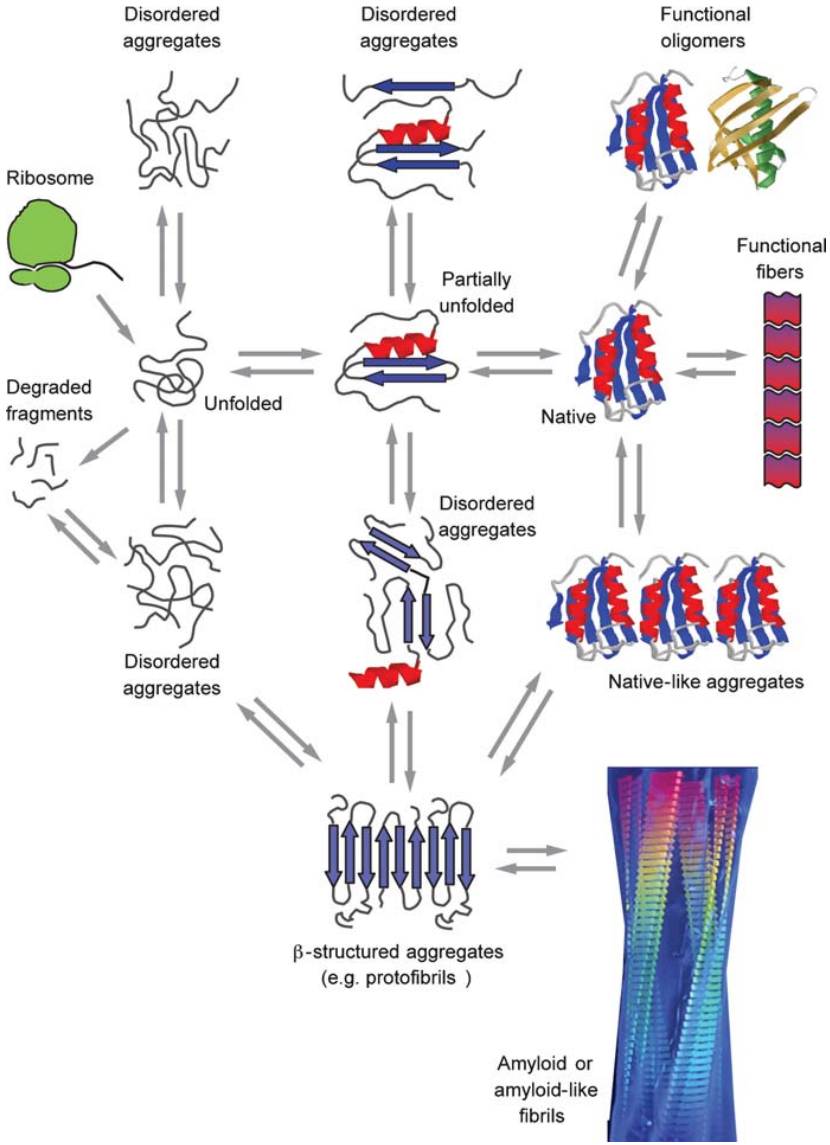


**Figure 2**

A schematic representation of some of the many conformational states that can be adopted by polypeptide chains and of the means by which they can be interconverted. The transition from β-structured aggregates to amyloid fibrils can occur by addition of either monomers or protofibrils (depending on protein) to preformed β-aggregates. All of these different conformational states and their interconversions are carefully regulated in the biological environment, much as enzymes regulate all the chemistry in cells, by using machinery such as molecular chaperones, degradatory systems, and quality control processes. Many of the various states of proteins are utilized functionally by biology, including unfolded proteins and amyloid fibrils, but conformational diseases will occur when such regulatory systems fail, just as metabolic diseases occur when the regulation of chemical processes becomes impaired.



Petitioner Merck, Ex. 1139, p. 351
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

controlled conditions and as a part of normal biochemical processes, with their amino acids often being recycled.

This description of normal functional behavior, honed by millions of years of evolution, is, however, only part of the story. Fully or partially unfolded ensembles on the pathways to their functional states (or generated as the result of stress, chemical modification, or genetic mutation) are particularly vulnerable to aggregation (**Figure 2**). Peptides and proteins that are natively unfolded, as well as fragments of proteins generated by proteolysis and unable to fold in the absence of the remainder of the polypeptide chain, can also aggregate under some circumstances, for example, if their concentrations become elevated. Some of the initial amorphous aggregates simply dissociate again, but others may reorganize to form oligomers with the germ of an amyloid structure, including the spherical, chain-like, and annular amyloid protofibrils observed for many systems. In order to generate long-range order in such structures, a critical number of molecules must be present such that the favorably enthaplic terms associated with their regular stacking can most effectively offset the accompanying loss of configurational entropy.

The structured polypeptide aggregates can then sometimes grow into mature fibrils by further self-association or through the repetitive addition of monomers. Proteins that adopt a folded structure under physiological conditions can also aggregate under some circumstances. This latter type of protein can either unfold, fully or partially, and aggregate through the mechanism described above or they can oligomerize prior to such a substantial conformational change. In the latter process, a structural reorganization to give amyloid-like assemblies occurs later and may in some cases be promoted by the existence of intermolecular contacts within native-like aggregates.

Every state of a polypeptide molecule, except the unique native state of globular proteins wherein the side chains pack together in a unique manner, is a broad ensemble of often diverse conformations. It is not surprising, therefore, that even the fibrillar end products of aggregation processes are characterized by morphological and structural diversity, representing variations on a common theme. Under most conditions in living systems, misfolding and aggregation of proteins are intrinsic side effects of the conformational transitions essential to the functioning of the organism. Formation of aggregates is normally inhibited by molecular chaperones and degradation processes as well as being disfavored by the amino acid sequences that are carefully selected by evolution to inhibit aggregation. But under some circumstances, as we discuss below, these aggregation processes can escape from the host of natural defenses and then give rise to pathogenic behavior.

## THE INFLUENCE OF SEQUENCE ON AMYLOID FORMATION

We have stressed that amyloid formation results primarily from the properties of the polypeptide chain that are common to all peptides and proteins. We have seen, however, that the sequence influences the relative stabilities of all the conformational states accessible to a given molecule, most notably the native state, and will thereby contribute to the susceptibility of a given polypeptide chain to convert into amyloid fibrils. Moreover, it is clear that polypeptide chains with different sequences can form amyloid fibrils at very different rates, even when these processes occur from fully or partially unfolded states. We start the exploration of this topic with a description of the determinants of the aggregation of those unfolded polypeptide chains that can broadly be described as unstructured, i.e., having no significant elements of persistent or cooperative structure. By considering these systems, we can examine how the properties of the sequence influence its intrinsic aggregation behavior rather than affect the stability of a given protein fold.



Petitioner Merck, Ex. 1139, p. 352
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

## Hydrophobicity, Charge, and Secondary Structure Propensities Strongly Influence Amyloid Formation

One important determinant of the aggregation of an unfolded polypeptide chain is the hydrophobicity of the side chains. Amino acid substitutions within regions of the sequence that play a crucial role (e.g., if they are in the region that nucleates aggregation) in the behavior of the whole sequence can reduce (or increase) the aggregation propensity of a sequence when they decrease (or increase) the hydrophobicity at the site of mutation (116–118). Moreover, there is evidence that protein sequences have evolved to avoid clusters of hydrophobic residues; for example, groups of three or more consecutive hydrophobic residues are less frequent in natural protein sequences than would be expected in the absence of evolutionary selection (119).

Another property likely to be a key factor in protein aggregation is charge, as a high net charge either globally or locally may hinder self-association (120, 121). For example, the effects of single amino acid substitutions were investigated on the propensity of AcP denatured in trifluoroethanol to aggregate (120). Although mutations decreasing the positive net charge of the protein resulted in an accelerated formation of β-sheet containing aggregates able to bind CR and ThT, mutations increasing the net charge resulted in the opposite effect. Further indications of the importance of charge in protein aggregation come from observations that aggregation of polypeptide chains can be facilitated by interactions with macromolecules, which exhibit a high compensatory charge (50, 122–125).

Comparison of large data sets of natively unfolded and natively folded proteins has shown that the former have a lower content of hydrophobic residues and a higher net charge than the latter (126). These properties undoubtedly contribute to maintaining the aggregation propensity of natively unfolded proteins sufficiently low to avoid the formation of aggregates under normal physiological conditions despite the fact that all, or at least the very large majority, of the side chains are accessible for intermolecular interactions.

In addition to charge and hydrophobicity, a low propensity to form α-helical structure and a high propensity to form β-sheet structure are also likely to be important factors encouraging amyloid formation (45, 50, 101, 117, 127, 128). Patterns of alternating hydrophilic and hydrophobic residues have been shown to be less frequent in natural proteins than expected on a random basis, suggesting that evolutionary selection has reduced the probability of such sequence patterns that favor β-sheet formation (127). Furthermore, it has been suggested that the high conservation of proline residues in a fibronectin type III superfamily and of glycine residues in AcPs can be rationalized on the grounds that such residues have a low propensity to form β-structure and hence inhibit aggregation (129, 130).

## The Amino Acid Sequence Affects Fibril Structure and Aggregation Rate

The demonstration that the various physicochemical factors described in the previous paragraph are important determinants of the formation of amyloid structure by unfolded polypeptide chains has proved to be of great value in understanding the mechanism of aggregation at a molecular level. For example, changes in the rate of aggregation of unfolded AcP following a series of mutations were used to generate a phenomenological equation, based on physicochemical principles, that is able to rationalize these rates in a robust manner (131). This expression was, remarkably, found to rationalize just as well similar data for a whole series of other unstructured peptides and proteins (**Figure 3*a***). This finding also provides compelling evidence for the close similarity of the principles underlying the aggregation behavior of different polypeptide molecules.



Petitioner Merck, Ex. 1139, p. 353
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

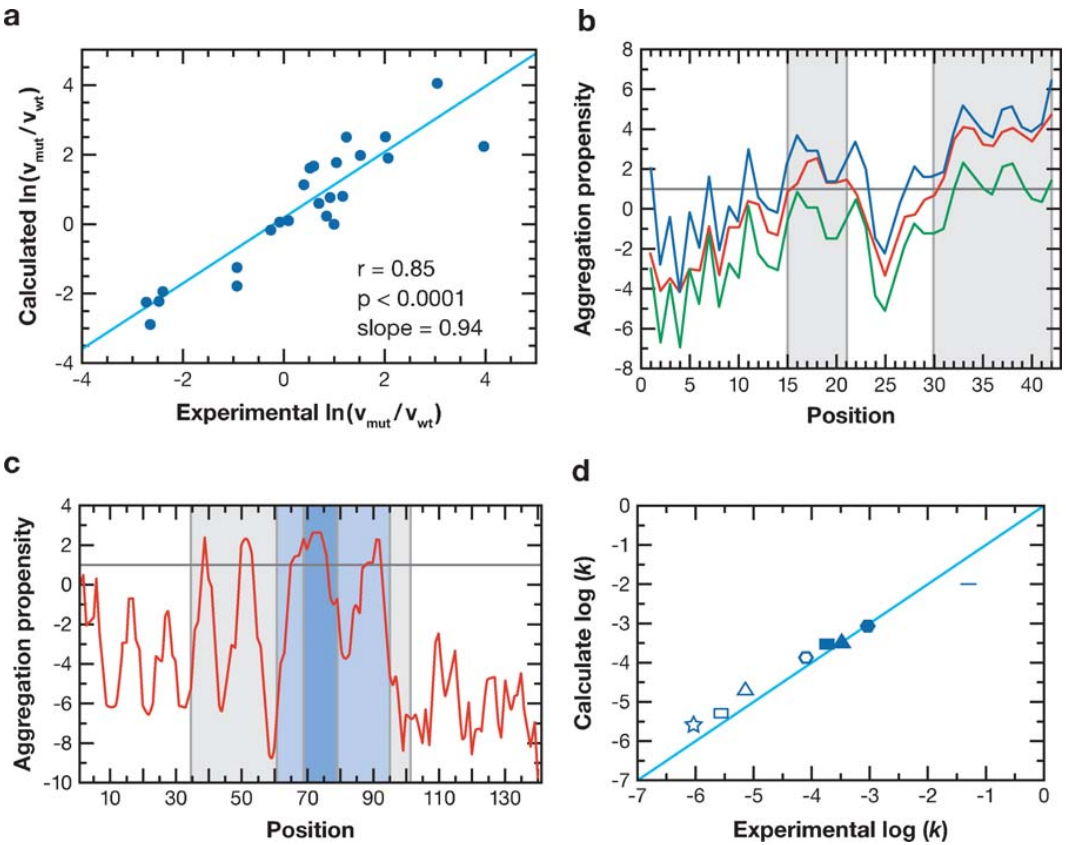


**Figure 3**

(*a*) Predicted versus experimental changes of aggregation rate following mutation. The mutations are in a variety of different unstructured polypeptide chains, such as Aβ, α-synuclein, amylin, and tau. Reproduced with permission (131). (*b*) Aggregation propensity profile (*red line*) for Aβ. The gray areas indicate regions of the sequence found experimentally to form and stabilize the amyloid cross-β core structure. A horizontal line at a propensity of 1 is drawn to highlight the aggregation-promoting regions that have values above this line. (*c*) Aggregation propensity profile (*red line*) for α-synuclein. The large region of the protein thought to be structured in the fibrils (*pale gray*) (56) is shown and includes all the peaks in the profile. The highly amyloidogenic NAC region (*light blue*) (178) and the 69–79 region (*dark blue*), found to be a particularly amyloidogenic segment within the NAC region (179) and containing the most prominent peak in the profile, are shown. The figures shown in panels *b* and *c* are reprinted with permission from Elsevier (134). (*d*) Predicted versus experimental aggregation rates (*k*) for a number of unstructured systems. The data points refer to $A\beta_{1-40}$ (Θ), ABri (□), denatured transthyretin (Δ), amylin (○), AChE (■), unfolded PrP (σ), unfolded human AcP (λ), and unfolded *E. coli* HypF-N (–). The straight line has a slope of 1 and indicates the ideal correlation between theory and experiment. Reprinted with permission from Elsevier (136).

A related approach, which considers additional factors such as the changes in the number of aromatic side chains, exposed surface area, and dipole moment upon mutation, has also been shown to predict the effect of mutations on the aggregation rates of a wide variety of polypeptide chains (132). The success of these rather straightforward



Petitioner Merck, Ex. 1139, p. 354
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

relationships between the rates of aggregation of unfolded polypeptide chains and simple physicochemical factors is strong support for the idea that such aggregation reflects the situation wherein a polypeptide chain behaves as a simple "polymer." Such behavior contrasts with the process of folding a globular protein for which the rates of folding are closely coupled to the specific structures of the highly evolved native states associated with individual sequences.

As described above, only a fraction of the residues of even the most highly amyloidogenic proteins are found in the core structure of the fibrils. In addition, conservative mutations have an effect on amyloid formation only when they are located in specific regions of the sequence (45, 117). Our increasing knowledge of such effects is beginning to lead to an understanding of the factors that cause specific segments of the sequence, rather than others, to form the characteristic cross-β structure. By extension of the analysis that led to the recognition of the links between aggregation propensities and the physicochemical characteristics of the constituent amino acid residues of a polypeptide chain, new algorithms have been developed to identify the regions of the sequence that are likely to promote aggregation within an unstructured polypeptide chain (133, 134). The outcome of both these approaches is a plot of aggregation propensity as a function of residue number, similar to the hydropathy profiles introduced to predict the regions of sequences that span the lipid bilayer in membrane proteins (135). The success of this type of approach is illustrated particularly well by the very good agreement between the regions of the sequence predicted to promote the aggregation of the Aβ peptide and α-synuclein and the regions found experimentally to form and stabilize the fibril core and/or to play a primary role in fibril formation (**Figure 3*b,c***) (134).

In a similar type of approach, multiple regression analysis has generated an equation that includes in the aggregation predictions the effects of extrinsic factors, such as the concentration of protein as well as the pH and ionic strength of the solution in which it is located, in addition to the intrinsic factors associated with the amino acid sequence (136). This equation reproduces the experimentally determined aggregation rates, which span five orders of magnitude, from the unstructured states of a set of nonhomologous protein sequences (**Figure 3*d***). In an exciting development, two reports have appeared recently that point to the validity of these concepts for in vivo situations, at least in bacteria (137, 138). The expenditure of effort and ingenuity in devising new methods to make quantitative analyses of such aggregation behavior in living organisms is therefore likely to bring rich rewards.

## Unfolded Regions Play Critical Roles in Promoting the Aggregation of Partially Folded States

Although the key regions of the sequence that promote fibril formation by an unfolded peptide or protein can now be broadly identified using relatively simple physicochemical parameters, the aggregation of polypeptides that contain significant levels of persistent secondary structure and long-range interactions will be influenced by additional factors. For example, the yeast prion proteins Sup35p, Ure2p, and HET-s all have unstructured and globular domains in their soluble states. In each of these three cases, the region forming the cross-β core and responsible for the prion properties has been found to be the domain that is unstructured in the soluble form of the protein (63, 139–141).

Further insights into this issue have again come from studies of proteins other than those found to form functional or pathogenic amyloid structures in vivo. An approach using limited proteolysis has shown that in the partially unfolded state adopted by AcP in the presence of moderate concentrations of trifluoroethanol, the regions of the sequence found to promote amyloid aggregation are flexible and/or solvent exposed in addition to having



Petitioner Merck, Ex. 1139, p. 355
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

an intrinsically high propensity to aggregate (142). Other regions that are not involved in aggregation despite having high propensities to aggregate were found to be at least partially buried in residual structure, whereas other solvent-exposed regions not involved in the aggregation process possess a low propensity to form amyloid fibrils. Similarly, the ease with which apomyoglobin converts to fibrils under different solution conditions correlates with the degree of denaturation, suggesting that fibrils assemble by association of unfolded polypeptide segments rather than by the docking of preformed structured elements (143).

Even α-synuclein, the protein associated with Parkinson's disease and assumed to have no significant structural preferences, has recently been shown to possess some significant long-range interactions between the negatively charged C-terminal region and the central amyloidogenic NAC region (144, 145). Structural perturbations that destabilize the interactions between these two portions of the protein molecule appear to increase the exposure of the amyloidogenic NAC region. Such perturbations include the presence of positively charged ions able to interact with the C terminus, a decrease in pH that reduces the net charge of the C-terminal region, and deletion of the C terminus; all result in a more rapid aggregation reaction (125, 146). Although it is clear that partial neutralization of the negative charge of α-synuclein will stimulate aggregation on a purely electrostatic argument, pairs of variants with a similar net charge but opposite signs (for example +3 and –3) aggregate more rapidly when the NAC region is unprotected (146).

## Variations in Fibrillar Structure Can Be Reconciled by Common Determinants of the Aggregation Process

Our ability to rationalize, and particularly to predict, important features of the process of amyloid assembly emphasizes in a dramatic manner that common traits are dominant in the aggregation behavior of different peptides and proteins. Although the structural analysis of fibrils at the level of specific residues (described above) highlights differences in the details of the manner in which individual molecules are incorporated into the fibrils, the fact remains that the generic cross-β structure and the frequent presence of stabilizing rows of hydrophobic interactions that run along the fibril axis (apart from important exceptions such as the fibrillar species associated with polyQ traits where additional side-chain hydrogen-bonding interactions are undoubtedly important) indicate the presence of common features in the aggregation of polypeptide chains. This commonality explains our ability to predict, often with a high degree of success, the regions involved in the formation of the amyloid core and the effect of mutations in this process.

Unlike the extreme dependence of the evolved native fold on protein sequence, it is unlikely that a single arrangement of a given chain in the amyloid core structure provides unique stability relative to all other arrangements. As noted above, this conclusion also means that the specific regions of a sequence found in such structures can vary with solution conditions, that there can be subtle differences in the manner in which a given polypeptide sequence is arranged in a cross-β core structure even under essentially identical conditions, and that the details of the resulting structures may be determined by kinetic rather than thermodynamic factors. This lack of a single unique structure, coupled with the extremely high degree of repetitive order within individual fibrils, may be the origin of the strain phenomena observed in both yeast and mammalian prions. Another important facet of this topic is that chemical modifications, for example those induced by physiologically formed metabolites (147, 148), or interactions with small molecules or metal ions (149) may play a much more important role in the aggregation process than might be imagined, e.g., by perturbing the thermodynamics of kinetics



Petitioner Merck, Ex. 1139, p. 356
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

sufficiently to alter the details of the resulting amyloid structure.

## THE PATHOGENESIS OF PROTEIN DEPOSITION DISEASES

The presence of highly organized and stable fibrillar deposits in the organs of patients suffering from protein deposition diseases led initially to the reasonable postulate that this material is the causative agent of the various disorders. This view was later reinforced by a number of observations; for example, amyloid fibrils formed from the Aβ peptide were found to be toxic to cultured neuronal cells (150, 151) and to cause both membrane depolarization and alterations in the frequency of their action potentials (152). Moreover, Aβ fibrils were shown to cause neuronal loss and microglial activation when injected into the cerebral cortex of aged rhesus monkeys (153). However, more recent findings have raised the possibility that precursors to amyloid fibrils, such as low-molecular-weight oligomers and/or structured protofibrils, are the real pathogenic species, at least in neuropathic diseases. Here we describe some of the most compelling evidence supporting this view, starting again from the well documented Aβ case.

### The Search Is on for the Causative Agents of Protein Aggregation Diseases

The severity of cognitive impairment in Alzheimer's disease correlates with the levels of low-molecular-weight species of Aβ, including small oligomers, rather than with the amyloid burden (154–156). In addition, transgenic mice show deficits in cognitive impairment, cell function, and synaptic plasticity well before the accumulation of significant quantities of amyloid plaques (157, 158). Similarly, phenotypic changes reminiscent of Alzheimer's disease precede amyloid plaque formation, or occur in their absence, in transgenic *Drosophila* expressing $A\beta_{1-42}$ and $A\beta_{1-40}$ (159, 160).

Further evidence comes from the finding that a single injection of a monoclonal anti-Aβ antibody does not reduce amyloid deposits in the brains of transgenic mice expressing $A\beta_{1-42}$, but it does reverse the associated memory loss, perhaps as a result of enhanced peripheral clearance and/or sequestration of soluble forms of the Aβ peptide (161). Genetic evidence also supports the theory that the precursor aggregates, as opposed to mature fibrils, are the pathogenic species: The aggressive "Arctic" (E693G) mutation of the amyloid β precursor protein, associated with a heritable early-onset manifestation of Alzheimer's disease, has been found in vitro to enhance protofibril, but not fibril, formation (162).

A similar scenario concerning the toxicity of early aggregates also holds for Parkinson's disease, a neurodegenerative condition associated with the formation of intracellular fibrillar deposits, notably Lewy bodies, in the dopaminergic neurons of the *substantia nigra*. In this disease, those dopaminergic neurons that survive, whether or not they contain Lewy bodies, show no quantifiable differences in viability (163, 164). Furthermore, mutations associated with juvenile Parkinson's disease or early-onset forms of Parkinsonism give rise to early neuronal degeneration in the absence of the accumulation of Lewy bodies (165). Overexpression of α-synuclein in transgenic flies or rats does not result in neuronal loss concomitant with the formation of detectable intracellular deposits (166, 167). By contrast, transgenic mice with nonfibrillar deposits of α-synuclein in various regions of the brain are characterized by substantial motor deficiencies and losses of dopaminergic neurons (168).

It is increasingly evident that prefibrillar aggregates from peptides and proteins other than Aβ and α-synuclein can either be toxic to cells or perturb their function. Early, nonfibrillar aggregates of transthyretin have been found toxic to neuronal cells under conditions



Petitioner Merck, Ex. 1139, p. 357
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

in which the native tetramer and the mature fibrils are not (169). Consistent with this finding, symptoms of familial amyloid polyneuropathy appear when transthyretin is deposited in an aggregated but nonfibrillar form that does not stain with CR (169). Reixach and coworkers (170) have found that such toxicity originates from low-molecular-weight oligomers of transthyretin of up to ~100 kDa in size.

The likelihood that such behavior is much more general is suggested by the finding that prefibrillar forms of the nondisease-related HypF-N from *E. coli*, the SH3 domain from bovine phosphatidylinositol 3′ kinase, lysozyme from horse, and apomyoglobin from sperm whale are also highly toxic to cultured fibroblasts and neurons, whereas the monomeric native states and the amyloid-like fibrils (all formed in vitro) displayed very little, if any, toxicity (171, 171a, 171b). Interestingly in this context, the most highly infective form of the mammalian prion protein has been identified as an oligomer of about 20 molecules, indicating that such small aggregates are the most effective initiators of transmissible spongiform encephalopathies (172).

## The Toxicity of Prefibrillar Aggregates Results from their Misfolded Nature

The reason why prefibrillar aggregates are toxic to cells, and hence appear to be the most likely culprits for the origins of at least some of the protein deposition diseases, is now at the front line of research in this field. A wide variety of biochemical, cytological, and physiological perturbations has been identified following the exposure of neurons to such species, both in vivo and in vitro. A detailed description of all of the reported effects is beyond the scope of the present review, and indeed, it is still too early to draw definitive conclusions about the similarities or differences of the effects of particular types of aggregates in different diseases.

Despite differences in the specific mechanisms of pathogenic behavior giving rise to distinct diseases, it is clear that the conversion of a protein from its soluble state into oligomeric forms will invariably generate a wide distribution of nonnative species, the populations of which will vary with sequence, time, and conditions. It seems likely that all of these inherently "misfolded" species will be toxic to some degree because they will inevitably expose on their surfaces an array of groups that are normally buried in globular proteins or dispersed in highly unfolded peptides or proteins. Small aggregates have a higher proportion of residues on their surfaces than larger aggregates, including mature amyloid fibrils, and therefore are likely in general to have a higher relative toxicity. In the crowded and highly organized environment of a living organism, the nonnative character of misfolded oligomers is particularly likely to trigger aberrant events resulting from their inappropriate interactions with cellular components, such as membranes, small metabolites, proteins, or other macromolecules. Such events will, in some situations, lead to the malfunctioning of crucial aspects of the cellular machinery, whether it is axonal transport, oxidative stress, ion balance, sequestration of essential proteins, or a combination of disparate factors, ultimately leading to apoptosis or other forms of cell death.

Although the natural defenses against misfolded proteins will act to sequester and neutralize such species, and/or inhibit their formation, it is inevitable that these mechanisms will sometimes be overwhelmed (98, 173). Such situations include mutations that dramatically increase aggregation rates, as in familial diseases; ingestion of preformed aggregates that are able to seed more extensive aggregation, as in prion diseases; or the age-related decline of chaperone and ubiquitin/proteasome responses, as in sporadic forms of diseases. Oligomer-mediated cytotoxicity is a key issue in neuropathic protein deposition diseases, although the question arises as to whether a similar mechanism is



Petitioner Merck, Ex. 1139, p. 358
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

central in the pathogenesis of nonneuropathic diseases. Systemic amyloidoses are often associated with accumulation of large quantities (even kilograms in some cases) of amyloid deposits in the affected tissues and organs (174). Undoubtedly, the impairment and disruption of tissue architecture, caused by these deposits in vital organs, are major features of these diseases and could well be the most important factors in the pathogenesis of at least some of these nonneuropathic degenerative conditions (174). Patients can have mechanical problems in carrying out even routine everyday tasks. Examples include difficulties in swallowing when amyloid accumulation occurs in the tongue and in moving because of extreme pain when amyloid accumulation occurs in joints. However, suggestions that early oligomeric species could have a more important role than fibril accumulation in the pathogenesis of nonneuropathic amyloidoses have been put forward (169, 174a). The elucidation of the mechanism of tissue damage by amyloid fibril proteins is undoubtedly an important issue in therapeutic approaches, although the optimum strategy must be to prevent aggregation or even production of the amyloidogenic protein before it can generate any potential damaging deposits.

## PERSPECTIVES

Despite the complexity of the protein aggregation process, the findings described above show that dramatic progress in its elucidation has been made in recent years. This progress relates particularly to our understanding of the nature and significance of amyloid formation and to how this process relates to the normal and aberrant behavior of living organisms. Increasingly sophisticated techniques are now being applied to elucidate the "amyloid phenomenon" in ever greater detail. Of special significance is the manner in which a wide variety of ideas from across the breadth of the biological, physical, and medical sciences is being brought together to probe important unifying principles. Much, of course, still remains to be discovered, but we are personally optimistic that the investigation of an increasing number of proteins both in vitro and in vivo will shed new light on the relationships between protein folding and misfolding as well as on the manner in which the multitude of different states accessible to proteins are regulated and interact with each other and with other cellular components. In addition, even our present understanding of the mechanism of amyloid formation is leading to more reliable methods of early diagnosis and to more rational therapeutic strategies that are either in clinical trials or approaching such trials (175, 176). Thus, despite the rapidity with which diseases of the type discussed here are increasingly afflicting the human populations of the modern world, there are grounds for optimism that present progress in understanding their nature and origins will lead, in the not too distant future, to the beginnings of widely applicable and effective means to combat their spread and their debilitating consequences.

### SUMMARY POINTS

1. A variety of human diseases is now thought to be associated with the formation of highly organized and generally intractable thread-like aggregates termed amyloid or amyloid-like fibrils.
2. Living organisms can take advantage of the inherent ability of proteins to form such structures to generate novel and diverse biological functions.
3. Dramatic advances have recently been made toward the elucidation of the structures of amyloid fibrils at a molecular level.



Petitioner Merck, Ex. 1139, p. 359
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

4. Amyloid fibril formation is preceded by formation of a wide range of aggregates such as unstructured oligomers and structured protofibrils.
5. We are now able to rationalize some of the issues regarding the molecular mechanism of amyloid formation, e.g., identify the regions of the sequence that form and stabilize the fibril core and/or play a primary role in fibril formation.
6. At least in some cases, prefibrillar aggregates, rather than the mature fibrils into which they convert, are the likely origins of pathological behavior. Despite obvious differences in detail, the pathogenic nature of these species lies in the exposure of groups that are normally buried in a folded protein or dispersed in an unfolded ensemble.

### FUTURE ISSUES TO BE RESOLVED

1. Although considerable progress has been made in the elucidation of amyloid fibril properties at a molecular level, very little is yet known about the structure of the amyloid protofibrils and unstructured aggregates that precede their formation and are likely to play a key role in the pathogenesis of protein deposition diseases.
2. Present research has been remarkably successful in providing a framework for understanding the fundamental nature of protein aggregation. The challenge now is to explore in more detail the links between these largely structural principles and the cellular and animal environments in which aggregation takes place.
3. The precise origin of the pathogenic nature of the amyloid deposits and their precursors remains elusive in each pathological condition associated with formation of these species.
4. The rational design of successful therapeutic strategies requires further characterization of the processes of amyloid formation occuring in vivo and of the interaction of the resulting aggregates with the various components of living organisms.

## LITERATURE CITED


1. Uversky VN, Fink AL. 2006. *Protein Misfolding, Aggregation and Conformational Diseases*, Vols. 1, 2. Kluwer Academic/Plenum
2. Amaral MD. 2004. *J. Mol. Neurosci.* 23:41–48
3. Lomas DA, Carrell RW. 2002. *Nat. Rev. Genet.* 3:759–68
4. Westermark P, Benson MD, Buxbaum JN, Cohen AS, Frangione B, et al. 2005. *Amyloid* 12:1–4
5. Lundmark K, Westermark GT, Nystrom S, Murphy CL, Solomon A, Westermark P. 2002. *Proc. Natl. Acad. Sci. USA* 99:6979–84
6. Lundmark K, Westermark GT, Olsen A, Westermark P. 2005. *Proc. Natl. Acad. Sci. USA* 102:6098–102
7. Legname G, Baskakov IV, Nguyen HO, Riesner D, Cohen FE, et al. 2004. *Science* 305:673–76
8. Hirschfield GM, Hawkins PN. 2003. *Int. J. Biochem. Cell. Biol.* 35:1608–13
9. Alexandrescu AT. 2005. *Protein Sci.* 14:1–12




Petitioner Merck, Ex. 1139, p. 360
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

10. Serpell LC, Sunde M, Benson MD, Tennent GA, Pepys MB, Fraser PE. 2000. *J. Mol. Biol.* 300:1033–39
11. Sunde M, Blake C. 1997. *Adv. Protein Chem.* 50:123–59
12. Bauer HH, Aebi U, Haner M, Hermann R, Muller M, Merkle HP. 1995. *J. Struct. Biol.* 115:1–15
13. Saiki M, Honda S, Kawasaki K, Zhou D, Kaito A, et al. 2005. *J. Mol. Biol.* 348:983–98
14. Pedersen JS, Dikov D, Flink JL, Hjuler HA, Christiansen G, Otzen D. 2005. *J. Mol. Biol.* 355:501–23
15. Nilsson MR. 2004. *Methods* 34:151–60
16. Khurana R, Uversky VN, Nielsen L, Fink AL. 2001. *J. Biol. Chem.* 276:22715–21
17. Bousset L, Redeker V, Decottignies P, Dubois S, Le Marechal P, Melki R. 2004. *Biochemistry* 43:5022–32
18. Jin LW, Claborn KA, Kurimoto M, Geday MA, Maezawa I, et al. 2003. *Proc. Natl. Acad. Sci. USA* 100:15294–98
19. Krebs MRH, MacPhee CE, Miller AF, Dunlop LE, Dobson CM, Donald AM. 2004. *Proc. Natl. Acad. Sci. USA* 101:14420–24
20. Stefani M, Dobson CM. 2003. *J. Mol. Med.* 81:678–99
21. Uversky VN, Fink AL. 2004. *Biochim. Biophys. Acta* 1698:131–53
22. Chapman MR, Robinson LS, Pinkner JS, Roth R, Heuser J, et al. 2002. *Science* 295:851–55
23. Claessen D, Rink R, de Jong W, Siebring J, de Vreugd P, et al. 2003. *Genes Dev.* 17:1714–26
23a. Mackay JP, Matthews JM, Winefield RD, Mackay LG, Haverkamp RG, et al. 2001. *Structure* 9:83–91
23b. Iconomidou VA, Vriend G, Hamodrakas SJ. 2000. *FEBS Lett.* 479:141–45
23c. Kenney JM, Knight D, Wise MJ, Vollrath F. 2002. *Eur. J. Biochem.* 269:4159–63
24. Berson JF, Theos AC, Harper DC, Tenza D, Raposo G, Marks MS. 2003. *J. Cell Biol.* 161:521–33
25. Chien P, Weissman JS, DePace AH. 2004. *Annu. Rev. Biochem.* 73:617–56
26. Eaglestone SS, Cox BS, Tuite MF. 1999. *EMBO J.* 18:1974–81
27. True HL, Lindquist SL. 2000. *Nature* 407:477–83
28. True HL, Berlin I, Lindquist SL. 2004. *Nature* 431:184–87
28a. Giaever G, Chu AM, Ni L, Connelly C, Riles L, et al. 2002. *Nature* 418:387–91
29. Nakayashiki T, Kurtzman CP, Edskes HK, Wickner RB. 2005. *Proc. Natl. Acad. Sci. USA* 102:10575–80
30. Si K, Lindquist S, Kandel ER. 2003. *Cell* 115:879–91
31. Coustou V, Deleu C, Saupe S, Begueret J. 1997. *Proc. Natl. Acad. Sci. USA* 94:9773–78
32. Saupe SJ. 2000. *Microbiol. Mol. Biol. Rev.* 64:489–502
33. Harper JD, Lieber CM, Lansbury PT Jr. 1997. *Chem. Biol.* 4:951–59
34. Petkova AT, Ishii Y, Balbach JJ, Antzutkin ON, Leapman RD, et al. 2002. *Proc. Natl. Acad. Sci. USA* 99:16742–47
35. Jaroniec CP, MacPhee CE, Astrof NS, Dobson CM, Griffin RG. 2002. *Proc. Natl. Acad. Sci. USA* 99:16748–53
36. Ritter C, Maddelein ML, Siemer AB, Luhrs T, Ernst M, et al. 2005. *Nature* 435:844–48
37. Makin OS, Atkins E, Sikorski P, Johansson J, Serpell LC. 2005. *Proc. Natl. Acad. Sci USA* 102:315–20
38. Nelson R, Sawaya MR, Balbirnie M, Madsen AO, Riekel C, et al. 2005. *Nature* 435:773–78



Petitioner Merck, Ex. 1139, p. 361
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

39. Antzutkin ON, Balbach JJ, Leapman RD, Rizzo NW, Reed J, Tycko R. 2000. *Proc. Natl. Acad. Sci. USA* 97:13045–50
40. Balbach JJ, Petkova AT, Oyler NA, Antzutkin ON, Gordon DJ, et al. 2002. *Biophys. J.* 83:1205–16
41. Goldsbury CS, Wirtz S, Muller SA, Sunderji S, Wicki P, et al. 2000. *J. Struct. Biol.* 130:217–31
42. Torok M, Milton S, Kayed R, Wu P, McIntire T, et al. 2000. *J. Biol. Chem.* 277:40810–15
43. Kheterpal I, Zhou S, Cook KD, Wetzel R. 2000. *Proc. Natl. Acad. Sci. USA* 97:13597–601
44. Kheterpal I, Williams A, Murphy C, Bledsoe B, Wetzel R. 2001. *Biochemistry* 40:11757–67
45. Williams AD, Portelius E, Kheterpal I, Guo JT, Cook KD, et al. 2004. *J. Mol. Biol.* 335:833–42
46. Hoshino M, Katou H, Hagihara Y, Hasegawa K, Naiki H, Goto Y. 2002. *Nat. Struct. Biol.* 9:332–36
47. Jaroniec CP, MacPhee CE, Bajaj VS, McMahon MT, Dobson CM, Griffin RG. 2004. *Proc. Natl. Acad. Sci. USA* 101:711–16
48. Jimenez JL, Guijarro JI, Orlova E, Zurdo J, Dobson CM, et al. 1999. *EMBO J.* 18:815–21
49. Jimenez JL, Nettleton EJ, Bouchard M, Robinson CV, Dobson CM, Saibil HR. 2002. *Proc. Natl. Acad. Sci. USA* 99:9196–201
50. Fandrich M, Dobson CM. 2002. *EMBO J.* 21:5682–90
51. Kajava AV, Aebi U, Steven AC. 2005. *J. Mol. Biol.* 348:247–52
52. Krishnan R, Lindquist SL. 2005. *Nature* 435:765–72
53. Makin O, Serpell LC. 2004. *J. Mol. Biol.* 335:1279–88
54. Goldsbury CS, Cooper GJ, Goldie KN, Muller SA, Saafi EL, et al. 1997. *J. Struct. Biol.* 119:17–27
55. Jayasinghe SA, Langen R. 2004. *J. Biol. Chem.* 279:48420–25
56. Der-Sarkissian A, Jao CC, Chen J, Langen R. 2003. *J. Biol. Chem.* 278:37530–35
57. Margittai M, Langen R. 2004. *Proc. Natl. Acad. Sci. USA* 101:10278–83
58. Yamaguchi K, Katou H, Hoshino M, Hasegawa K, Naiki H, Goto Y. 2004. *J. Mol. Biol.* 338:559–71
58a. Del Mar C, Greenbaum EA, Mayne L, Englander SW, Woods VL Jr. 2005. *Proc. Natl. Acad. Sci. USA* 102:15477–82
59. Monti M, Principe S, Giorgetti S, Mangione P, Merlini G, et al. 2002. *Protein Sci.* 11:2362–69
59a. Miake H, Mizusawa H, Iwatsubo T, Hasegawa M. 2002. *J. Biol. Chem.* 2002. 277:19213–19
59b. Heise H, Hoyer W, Becker S, Andronesi OC, Riedel D, Baldus M. 2005. *Proc. Natl. Acad. Sci. USA* 102:15871–76
60. Perutz MF, Finch JT, Berriman J, Lesk A. 2002. *Proc. Natl. Acad. Sci. USA* 99:5591–95
61. Zandomeneghi G, Krebs MR, McCammon MG, Fandrich M. 2004. *Protein Sci.* 13:3314–21
62. Petkova AT, Leapman RD, Guo Z, Yau WM, Mattson MP, Tycko R. 2005. *Science* 307:262–65
63. Baxa U, Taylor KL, Wall JS, Simon MN, Cheng NQ, et al. 2003. *J. Biol. Chem.* 278:43717–27
64. Kajava AV, Baxa U, Wickner RB, Steven AC. 2004. *Proc. Natl. Acad. Sci. USA* 101:7885–90
65. Baxa U, Cheng NQ, Winkler DC, Chiu TK, Davies DR, et al. 2005. *J. Struct. Biol.* 150:170–79



Petitioner Merck, Ex. 1139, p. 362
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

66. Bousset L, Thomson NH, Radford SE, Melki R. 2002. *EMBO J.* 17:2903–11
67. Bousset L, Briki F, Doucet J, Melki R. 2003. *J. Struct. Biol.* 141:132–42
68. Tanaka M, Chien P, Yonekura K, Weissman JS. 2005. *Cell* 121:49–62
69. Jones EM, Surewicz WK. 2005. *Cell* 121:63–72
70. Naiki H, Hashimoto N, Suzuki S, Kimura H, Nakakuki K, Gejyo F. 1997. *Amyloid* 4:223–32
71. Serio TR, Cashikar AG, Kowal AS, Sawicki GJ, Moslehi JJ, et al. 2000. *Science* 289:1317–21
72. Uversky VN, Li J, Souillac P, Millett IS, Doniach S, et al. 2002. *J. Biol. Chem.* 277:11970–78
73. Pedersen JS, Christensen G, Otzen DE. 2004. *J. Mol. Biol.* 341:575–88
74. Chayen NE. 2005. *Prog. Biophys. Mol. Biol.* 88:329–37
75. Fezoui Y, Teplow DB. 2002. *J. Biol. Chem.* 277:36948–54
76. Wright CF, Teichmann SA, Clarke J, Dobson CM. 2005. *Nature* 438:878–81
76a. Krebs MR, Morozova-Roche LA, Daniel K, Robinson CV, Dobson CM. 2004. *Protein Sci.* 13:1933–38
77. Harper JD, Wong SS, Lieber CM, Lansbury PT Jr. 1997. *Chem. Biol.* 4:119–25
78. Walsh DM, Lomakin A, Benedek GB, Condron MM, Teplow DB. 1997. *J. Biol. Chem.* 272:22364–72
79. Walsh DM, Hartley DM, Kusumoto Y, Fezoui Y, Condron MM, et al. 1999. *J. Biol. Chem.* 274:25945–52
80. Kayed R, Sokolov Y, Edmonds B, McIntire TM, Milton SC, et al. 2004. *J. Biol. Chem.* 279:46363–66
81. Williams AD, Sega M, Chen ML, Kheterpal I, Geva M, et al. 2005. *Proc. Natl. Acad. Sci. USA* 102:7115–20
82. Conway KA, Harper JD, Lansbury PT Jr. 2000. *Biochemistry* 39:2552–63
83. Ionescu–Zanetti C, Khurana R, Gillespie JR, Petrick JS, Trabachino LC, et al. 1999. *Proc. Natl. Acad. Sci. USA* 96:13175–79
84. Quintas A, Vaz DC, Cardoso I, Saraiva MJM, Brito RMM. 2001. *J. Biol. Chem.* 276:27207–13
85. Gosal WS, Morten IJ, Hewitt EW, Smith DA, Thomson NH, Radford SE. 2005. *J. Mol. Biol.* 351:850–64
86. Malisauskas M, Zamotin V, Jass J, Noppe W, Dobson CM, Morozova-Roche LA. 2003. *J. Mol. Biol.* 330:879–90
87. Plakoutsi G, Taddei N, Stefani M, Chiti F. 2004. *J. Biol. Chem.* 279:14111–19
87a. Bader R, Bamford R, Zurdo J, Luisi BF, Dobson CM. 2006. *J. Mol. Biol.* 356:189–208
88. Kayed R, Head E, Thompson JL, McIntire TM, Milton SC, et al. 2003. *Science* 300:486–89
89. Morozova-Roche LA, Zamotin V, Malisauskas M, Ohman A, Chertkova R, et al. 2004. *Biochemistry* 43:9610–19
89a. Petty SA, Adalsteinsson T, Decatur SM. 2005. *Biochemistry* 44:4720–26
89b. Petty SA, Decatur SM. 2005. *Proc. Natl. Acad. Sci. USA* 102:14272–77
90. Bitan G, Lomakin A, Teplow DB. 2001. *J. Biol. Chem.* 276:35176–84
91. Bitan G, Kirkitadze MD, Lomakin A, Vollers SS, Benedek GB, Teplow DB. 2003. *Proc. Natl. Acad. Sci. USA* 100:330–35
92. Roher AE, Chaney MO, Kuo YM, Webster SD, Stine WB, et al. 1996. *J. Biol. Chem.* 271:20631–35
93. Podlisny MB, Ostaszewski BL, Squazzo SL, Koo EH, Rydell RE, et al. 1995. *J. Biol. Chem.* 270:9564–70



Petitioner Merck, Ex. 1139, p. 363
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

94. Walsh DM, Tseng BP, Rydel RE, Podlisny MB, Selkoe DJ. 2000. *Biochemistry* 39:10831–39
95. Modler AJ, Gast K, Lutsch G, Damaschun G. 2003. *J. Mol. Biol.* 325:135–48
96. Nguyen HD, Hall CK. 2004. *Proc. Natl. Acad. Sci. USA* 101:16180–85
97. Kelly JW. 1998. *Curr. Opin. Struct. Biol.* 8:101–6
98. Dobson CM. 1999. *Trends Biochem. Sci.* 24:329–32
99. Guijarro JI, Sunde M, Jones JA, Campbell ID, Dobson CM. 1998. *Proc. Natl. Acad. Sci. USA* 95:4224–28
100. Chiti F, Taddei N, Bucciantini M, White P, Ramponi G, Dobson CM. 2000. *EMBO J.* 19:1441–49
101. Villegas V, Zurdo J, Filimonov VV, Aviles FX, Dobson CM, Serrano L. 2000. *Protein Sci.* 9:1700–8
102. Ferrao-Gonzales AD, Souto SO, Silva JL, Foguel D. 2000. *Proc. Natl. Acad. Sci. USA* 97:6445–50
103. Raffen R, Dieckman LJ, Szpunar M, Wunschl C, Pokkuluri PR, et al. 1999. *Protein Sci.* 8:509–17
104. Canet D, Last AM, Tito P, Sunde M, Spencer A, et al. 2002. *Nat. Struct. Biol.* 9:308–15
105. Hammarstrom P, Jiang X, Hurshman AR, Powers ET, Kelly JW. 2002. *Proc. Natl. Acad. Sci. USA* 99:16427–32
106. Ramirez-Alvarado M, Merkel JS, Regan L. 2000. *Proc. Natl. Acad. Sci. USA* 97:8979–84
107. Marcon G, Plakoutsi G, Canale C, Relini A, Taddei N, et al. 2005. *J. Mol. Biol.* 347:323–35
108. Chiti F, Taddei N, Stefani M, Dobson CM, Ramponi G. 2001. *Protein Sci.* 10:879–86
109. Sacchettini JC, Kelly JW. 2002. *Nat. Rev. Drug Discov.* 1:267–75
110. Dumoulin M, Last AM, Desmyter A, Decanniere K, Canet D, et al. 2003. *Nature* 424:783–88
111. Ray SS, Nowak RJ, Brown RH Jr, Lansbury PT Jr. 2005. *Proc. Natl. Acad. Sci. USA* 102:3639–44
112. Bouchard M, Zurdo J, Nettleton EJ, Dobson CM, Robinson CV. 2000. *Protein Sci.* 9:960–67
113. Chow MK, Ellisdon AM, Cabrita LD, Bottomley SP. 2005. *J. Biol. Chem.* 279:47643–51
114. Plakoutsi G, Bemporad F, Calamai M, Taddei N, Dobson CM, Chiti F. 2005. *J. Mol. Biol.* 351:910–22
115. Lindorff-Larsen K, Best RB, Depristo MA, Dobson CM, Vendruscolo M. 2005. *Nature* 433:128–32
116. Otzen DE, Kristensen O, Oliveberg M. 2000. *Proc. Natl. Acad. Sci. USA* 97:9907–12
117. Chiti F, Taddei N, Baroni F, Capanni C, Stefani M, et al. 2002. *Nat. Struct. Biol.* 9:137–43
118. Wurth C, Guimard NK, Hecht MH. 2002. *J. Mol. Biol.* 319:1279–90
119. Schwartz R, Istrail S, King J. 2001. *Protein Sci.* 10:1023–31
120. Chiti F, Calamai M, Taddei N, Stefani M, Ramponi G, Dobson CM. 2002. *Proc. Natl. Acad. Sci. USA* 99:16419–26
121. Schmittschmitt JP, Scholtz JM. 2003. *Protein Sci.* 12:2374–78
122. Konno T. 2001. *Biochemistry* 40:2148–54
123. Goers J, Uversky VN, Fink AL. 2003. *Protein Sci.* 12:702–7
124. Giasson BI, Forman MS, Higuchi M, Golbe LI, Graves CL, et al. 2003. *Science* 300:636–40
125. Fernandez CO, Hoyer W, Zweckstetter M, Jares-Erijman EA, Subramaniam V, et al. 2004. *EMBO J.* 23:2039–46



Petitioner Merck, Ex. 1139, p. 364
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

126. Uversky VN. 2002. *Protein Sci*. 11:739–56
127. Broome BM, Hecht MH. 2000. *J. Mol. Biol.* 296:961–68
128. Kallberg Y, Gustafsson M, Persson B, Thyberg J, Johansson J. 2001. *J. Biol. Chem.* 276:12945–50
129. Steward A, Adhya S, Clarke J. 2002. *J. Mol. Biol.* 318:935–40
130. Parrini C, Taddei N, Ramazzotti M, Degl'Innocenti D, Ramponi G, et al. 2005. *Structure* 13:1143–51
131. Chiti F, Stefani M, Taddei N, Ramponi G, Dobson CM. 2003. *Nature* 424:805–8
132. Tartaglia GG, Cavalli A, Pellarin R, Caflisch A. 2004. *Protein Sci.* 13:1939–41
133. Fernandez-Escamilla AM, Rousseau F, Schymkowitz J, Serrano L. 2004. *Nat. Biotechnol.* 22:1302–6
134. Pawar AP, DuBay KF, Zurdo J, Chiti F, Vendruscolo M, Dobson CM. 2005. *J. Mol. Biol.* 350:379–92
135. Kyte J, Doolittle RF. 1982. *J. Mol. Biol.* 157:105–32
136. DuBay KF, Pawar AP, Chiti F, Zurdo J, Dobson CM, Vendruscolo M. 2004. *J. Mol. Biol.* 341:1317–26
137. Ignatova Z, Gierasch LM. 2004. *Proc. Natl. Acad. Sci. USA* 101:523–28
138. Calloni G, Zoffoli S, Stefani M, Dobson CM, Chiti F. 2005. *J. Biol. Chem.* 280:10607–13
139. Paushkin SV, Kushnirov VV, Smirnov VN, Ter-Avanesyan MD. 1996. *EMBO J.* 15:3127–34
140. Masison DC, Maddelein ML, Wickner RB. 1997. *Proc. Natl. Acad. Sci. USA* 94:12503–8
141. Balguerie A, Dos Reis S, Ritter C, Chaignepain S, Coulary-Salin B, et al. 2003. *EMBO J.* 22:2071–81
142. Monti M, di Bard BLG, Calloni G, Chiti F, Amoresano A, et al. 2004. *J. Mol. Biol.* 336:253–62
143. Fandrich M, Forge V, Buder K, Kittler M, Dobson CM, Diekmann S. 2003. *Proc. Natl. Acad. Sci. USA* 100:15463–68
144. Bertoncini CW, Jung YS, Fernandez CO, Hoyer W, Griesinger C, et al. 2005. *Proc. Natl. Acad. Sci. USA* 102:1430–35
145. Dedmon MM, Lindorff-Larsen K, Christodoulou J, Vendruscolo M, Dobson CM. 2005. *J. Am. Chem. Soc.* 127:476–77
146. Hoyer W, Cherny D, Subramaniam V, Jovin TM. 2004. *Biochemistry* 43:16233–42
147. Nilsson MR, Driscoll M, Raleigh DP. 2002. *Protein Sci.* 11:342–49
148. Zhang Q, Powers ET, Nieva J, Huff ME, Dendle MA. 2004. *Proc. Natl. Acad. Sci. USA* 101:4752–57
149. Bush AI. 2003. *Trends Neurosci*. 26:207–14
150. Pike CJ, Walencewicz AJ, Glabe CG, Cotman CW. 1991. *Brain Res.* 563:311–14
151. Lorenzo A, Yankner BA. 1994. *Proc. Natl. Acad. Sci. USA* 91:12243–47
152. Hartley DM, Walsh DM, Ye CP, Diehl T, Vasquez S, et al. 1999. *J. Neurosci*. 19:8876–84
153. Geula C, Wu CK, Saroff D, Lorenzo A, Yuan M, Yankner BA. 1998. *Nat. Med*. 4:827–31
154. Lue LF, Kuo YM, Roher AE, Brachova L, Shen Y, et al. 1999. *Am. J. Pathol.* 155:853–62
155. McLean CA, Cherny RA, Fraser FW, Fuller SJ, Smith MJ, et al. 1999. *Ann. Neurol.* 46:860–66
156. Wang J, Dickson DW, Trojanowski JQ, Lee VM. 1999. *Exp. Neurol.* 158:328–37
157. Moechars D, Dewachter I, Lorent K, Reverse D, Baekelandt V, et al. 1999. *J. Biol. Chem.* 274:6483–92
158. Larson J, Lynch G, Games D, Seubert P. 1999. *Brain Res.* 840:23–35
159. Iijima K, Liu HP, Chiang AS, Hearn SA, Konsolaki M, Zhong Y. 2001. *Proc. Natl. Acad. Sci. USA* 101:6623–28



Petitioner Merck, Ex. 1139, p. 365
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

160. Crowther DC, Kinghorn KJ, Miranda E, Page R, Curry JA, et al. 2005. *Neuroscience* 132:123–35
161. Dodart JC, Bales KR, Gannon KS, Greene SJ, DeMattos RB, et al. 2002. *Nat. Neurosci.* 5:452–57
162. Nilsberth C, Westlind-Danielsson A, Eckman CB, Condron MM, Axelman K, et al. 2001. *Nat. Neurosci.* 4:887–93
163. Gertz HJ, Siegers A, Kuchinke J. 1994. *Brain Res.* 637:339–41
164. Bergeron C, Petrunka C, Weyer L. 1996. *Am. J. Pathol.* 148:273–79
165. Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, et al. 1998. *Nature* 392:605–8
166. Auluck PK, Chan HY, Trojanowski JQ, Lee VM, Bonini NM. 2002. *Science* 295:865–68
167. Lo Bianco C, Ridet JL, Schneider BL, Deglon N, Aebischer P. 2002. *Proc. Natl. Acad. Sci. USA* 99:10813–18
168. Masliah E, Rockenstein E, Veinbergs I, Mallory M, Hashimoto M, et al. 2000. *Science* 287:1265–69
169. Sousa MM, Cardoso I, Fernandes R, Guimaraes A, Saraiva MJ. 2001. *Am. J. Pathol.* 159:1993–2000
170. Reixach N, Deechongkit S, Jiang X, Kelly JW, Buxbaum JN. 2004. *Proc. Natl. Acad. Sci. USA* 101:2817–22
171. Bucciantini M, Giannoni E, Chiti F, Baroni F, Formigli L, et al. 2002. *Nature* 416:507–11
171a. Sirangelo I, Malmo C, Iannuzzi C, Mezzogiorno A, Bianco MR, et al. 2004. *J. Biol. Chem.* 279:13183–89
171b. Malisauskas M, Ostman J, Darinskas A, Zamotin V, Liutkevicius E, et al. 2005. *J. Biol. Chem.* 280:6269–75
172. Silveira JR, Raymond GJ, Hughson AG, Race RE, Sim VL, et al. 2005. *Nature* 437:257–61
173. Soti C, Pal C, Papp B, Csermely P. 2005. *Curr. Opin. Cell. Biol.* 17:210–15
174. Pepys MB. 2001. *Philos. Trans. R. Soc. London Ser. B* 356:203–10
174a. Merlini G, Bellotti V. 2003. *N. Engl. J. Med.* 349:583–96
175. Cohen FE, Kelly JW. 2003. *Nature* 426:905–9
176. Merlini G, Westermark P. 2004. *J. Intern. Med.* 255:159–78
177. Tycko R. 2002. *Biochemistry* 42:3151–59
178. Han HY, Weinreb PH, Lansbury PT Jr. 1995. *Chem. Biol.* 2:163–69
179. El-Agnaf OMA, Irvine GB. 2002. *Biochem. Soc. Trans.* 30:559–65



Petitioner Merck, Ex. 1139, p. 366
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

Annual Review
of Biochemistry

Volume 75, 2006

# Contents

Wanderings of a DNA Enzymologist: From DNA Polymerase to Viral Latency
*I. Robert Lehman* ..... 1

Signaling Pathways in Skeletal Muscle Remodeling
*Rhonda Bassel-Duby and Eric N. Olson* ..... 19

Biosynthesis and Assembly of Capsular Polysaccharides in *Escherichia coli*
*Chris Whitfield* ..... 39

Energy Converting NADH:Quinone Oxidoreductase (Complex I)
*Ulrich Brandt* ..... 69

Tyrphostins and Other Tyrosine Kinase Inhibitors
*Alexander Levitzki and Eyal Mishani* ..... 93

Break-Induced Replication and Recombinational Telomere Elongation in Yeast
*Michael J. McEachern and James E. Haber* ..... 111

LKB1-Dependent Signaling Pathways
*Dario R. Alessi, Kei Sakamoto, and Jose R. Bayascas* ..... 137

Energy Transduction: Proton Transfer Through the Respiratory Complexes
*Jonathan P. Hosler, Shelagh Ferguson-Miller, and Denise A. Mills* ..... 165

The Death-Associated Protein Kinases: Structure, Function, and Beyond
*Shani Bialik and Adi Kimchi* ..... 189

Mechanisms for Chromosome and Plasmid Segregation
*Santanu Kumar Ghosh, Sujata Hajra, Andrew Paek, and Makkuni Jayaram* ..... 211

Chromatin Modifications by Methylation and Ubiquitination: Implications in the Regulation of Gene Expression
*Ali Shilatifard* ..... 243



Petitioner Merck, Ex. 1139, p. 367
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

Structure and Mechanism of the Hsp90 Molecular Chaperone Machinery
*Laurence H. Pearl and Chrisostomos Prodromou* ..... 271

Biochemistry of Mammalian Peroxisomes Revisited
*Ronald J.A. Wanders and Hans R. Waterham* ..... 295

Protein Misfolding, Functional Amyloid, and Human Disease
*Fabrizio Chiti and Christopher M. Dobson* ..... 333

Obesity-Related Derangements in Metabolic Regulation
*Deborah M. Muoio and Christopher B. Newgard* ..... 367

Cold-Adapted Enzymes
*Khawar Sohail Siddiqui and Ricardo Cavicchioli* ..... 403

The Biochemistry of Sirtuins
*Anthony A. Sauve, Cynthia Wolberger, Vern L. Schramm, and Jef D. Boeke* ..... 435

Dynamic Filaments of the Bacterial Cytoskeleton
*Katharine A. Michie and Jan Löwe* ..... 467

The Structure and Function of Telomerase Reverse Transcriptase
*Chantal Autexier and Neal F. Lue* ..... 493

Relating Protein Motion to Catalysis
*Sharon Hammes-Schiffer and Stephen J. Benkovic* ..... 519

Animal Cytokinesis: From Parts List to Mechanisms
*Ulrike S. Eggert, Timothy J. Mitchison, and Christine M. Field* ..... 543

Mechanisms of Site-Specific Recombination
*Nigel D.F. Grindley, Katrine L. Whiteson, and Phoebe A. Rice* ..... 567

Axonal Transport and Alzheimer's Disease
*Gorazd B. Stokin and Lawrence S.B. Goldstein* ..... 607

Asparagine Synthetase Chemotherapy
*Nigel G.J. Richards and Michael S. Kilberg* ..... 629

Domains, Motifs, and Scaffolds: The Role of Modular Interactions in the Evolution and Wiring of Cell Signaling Circuits
*Roby P. Bhattacharyya, Attila Reményi, Brian J. Yeh, and Wendell A. Lim* ..... 655

Ribonucleotide Reductases
*Pär Nordlund and Peter Reichard* ..... 681

Introduction to the Membrane Protein Reviews: The Interplay of Structure, Dynamics, and Environment in Membrane Protein Function
*Jonathan N. Sachs and Donald M. Engelman* ..... 707



Petitioner Merck, Ex. 1139, p. 368
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

Relations Between Structure and Function of the Mitochondrial ADP/ATP Carrier
*H. Nury, C. Dahout-Gonzalez, V. Trézéguet, G.J.M. Lauquin, G. Brandolin, and E. Pebay-Peyroula* ..... 713

G Protein–Coupled Receptor Rhodopsin
*Krzysztof Palczewski* ..... 743

Transmembrane Traffic in the Cytochrome $b_6f$ Complex
*William A. Cramer, Huamin Zhang, Jiusheng Yan, Genji Kurisu, and Janet L. Smith* ..... 769

INDEXES

Subject Index ..... 791

Author Index ..... 825

ERRATA

An online log of corrections to *Annual Review of Biochemistry* chapters (if any, 1977 to the present) may be found at http://biochem.annualreviews.org/errata.shtml



Petitioner Merck, Ex. 1139, p. 369
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033