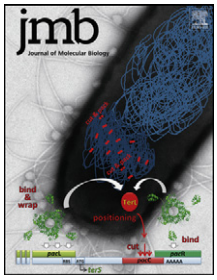

Amyloidogenicity at a Distance: How Distal Protein Regions Modulate Aggregation in Disease

Christina M. Lucato [1,2], Christopher J. Lupton [1,2], Michelle L. Halls [3] and Andrew M. Ellisdon [1,2]

*1 - **Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology,** Monash University, Clayton, Victoria 3800, Australia*
*2 - **Australian Research Council Centre of Excellence in Advanced Molecular Imaging,** Monash University, Clayton, Victoria 3800, Australia*
*3 - **Drug Discovery Biology Theme,** Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia*

***Correspondence to Andrew M. Ellisdon:** Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia. andrew.ellisdon@monash.edu*
***DOI of the original article:** http://dx.doi.org/10.1016/j.jmb.2015.10.029*
http://dx.doi.org/10.1016/j.jmb.2017.03.021
Edited by Ronald Wetzel

Abstract

The misfolding of proteins to form amyloid is a key pathological feature of several progressive, and currently incurable, diseases. A mechanistic understanding of the pathway from soluble, native protein to insoluble amyloid is crucial for therapeutic design, and recent efforts have helped to elucidate the key molecular events that trigger protein misfolding. Generally, either global or local structural perturbations occur early in amyloidogenesis to expose aggregation-prone regions of the protein that can then self-associate to form toxic oligomers. Surprisingly, these initiating structural changes are often caused or influenced by protein regions distal to the classically amyloidogenic sequences. Understanding the importance of these distal regions in the pathogenic process has highlighted many remaining knowledge gaps regarding the precise molecular events that occur in classic aggregation pathways. In this review, we discuss how these distal regions can influence aggregation in disease and the recent technical and conceptual advances that have allowed this insight.

Introduction

Aggregation to amyloid is the process by which proteins misfold from their native and functional state to one that is amenable to self-association. This self-association generally occurs concomitantly with a global rearrangement to a β-sheet-rich structure, held tightly together through extensive hydrogen bonding, to form the repeating units of long polymeric fibres known as “amyloid”. The common structural features of amyloid represent a stable state for a range of proteins of widely varying primary sequences. The presence and accumulation of amyloid fibrils are a classic hallmark of a range of amyloidoses that include the neurodegenerative diseases, Alzheimer's disease (AD) and Parkinson's disease (PD). This seminal observation led to an early focus of structural studies on end-stage amyloid fibrils; however, extensive research throughout the field has since implicated intermediate oligomers as a potential pathogenic species in these conditions. As such, the research emphasis has now shifted toward a structural characterization of these pathogenic intermediates and how they arise from the native proteins.

Until recently, classic structural techniques had limited application to the study of the earliest events in protein misfolding, as these early events occur over rapid time scales, and involve small and highly dynamic conformational changes. Therefore, this limited any insight to low-resolution structures of the early stage aggregation intermediates. As such, a structural understanding of the key molecular events that trigger aggregate nucleation has, for the most part, remained elusive. However, recent technological advances have facilitated more detailed examination of these phenomena. In particular, solid-state NMR (ssNMR) and cryo electron microscopy have been useful for determining the structural composition of fibrils, and theoretical techniques including molecular docking and dynamics simulations have allowed the modeling of protein

dynamics and interactions on very rapid time scales. In addition to the application of higher resolution experimental approaches, our understanding of the molecular basis of misfolding has been advanced by the recognition of the importance of regions distal to the classic amyloidogenic sequences in defining the aggregation potential of a protein. Although often not involved in the core amyloid fibril β-structures, these regions have proven influential in triggering aggregation.

In this review, we will focus on recent advances that have been made in characterizing the structural changes that occur during the earliest events in the misfolding pathway. In this context, we will revisit some classic examples of amyloid formation [including that of α-synuclein (α-syn), β_2-microglobulin (β2m), apolipoprotein A-I (apoA-I), polyglutamine (polyQ) proteins, amyloid β, and tau] to emphasize how the ability of each protein to aggregate can be initiated or enhanced by distal protein regions. These prototypical cases are redefining our view of the early molecular events in aggregation, whereby both global and local perturbations in structure beyond the amyloidogenic region itself appear to define pathological progression.

α-Syn: Aggregation Triggered by Global Deprotection of Amyloidogenic Regions

α-syn is a 140-residue protein that is localized to presynaptic nerve terminals and is proposed to play a key role in neurotransmitter release, dopamine synthesis, vesicle trafficking, and exocytosis [1]. α-syn amyloid was initially detected in Lewy bodies, the pathologic hallmark of PD. Subsequently, α-syn aggregates were detected in a number of other neurodegenerative diseases, collectively known as synucleinopathies, which are characterized by α-syn aggregation and inclusion formation [1]. α-syn is an intrinsically disordered protein (IDP) composed of three distinct regions: an N terminus (residues 1–60), which contains four imperfect repeats of the conserved KTKEGV hexamer motif; the NAC domain (residues 61–95), which contains the highly amyloidogenic NAC region and three imperfect KTKEGV repeats; and the C terminus (residues 96–140), which is highly acidic and proline-rich (Fig. 1a). The NAC domain is central to the ability of α-syn to form amyloid [2]; however, the N and C termini are critical modulators of nucleation, oligomer formation, and end-stage fibril structure.

Two recent technical feats, harnessing micro-electron diffraction and ssNMR, have provided high-resolution insights into the end-stage α-syn fibril structure [3,4]. Micro-electron diffraction revealed that short peptides, including part of the NAC domain, form similar steric-zipper protofilaments whereby the peptides pack in pairs to form in-register extended parallel β-sheets [3]. More recently, ssNMR studies provided a more complex model of the full-length α-syn fibril structure, whereby the NAC region forms the core of a complex β-serpentine arrangement [4]. Importantly, regions flanking the NAC domain are also incorporated into the amyloid core, which is composed of the N-terminal domain residues 46–54 in addition to the NAC domain residues 63–96 (Fig. 1c) [4].

Given this important structural role of the N terminus in the α-syn amyloid core, it is unsurprising that this region also plays a key role in modulating early steps of the aggregation pathway. The N terminus can interact with cell membranes, and while this is likely critical for its physiological function, it also promotes a conformational change in the protein, allowing it to sample a number of α-helical conformations (for example, an extended α-helix or broken α-helical states; Fig. 1b) [5–9]. These membrane-binding-

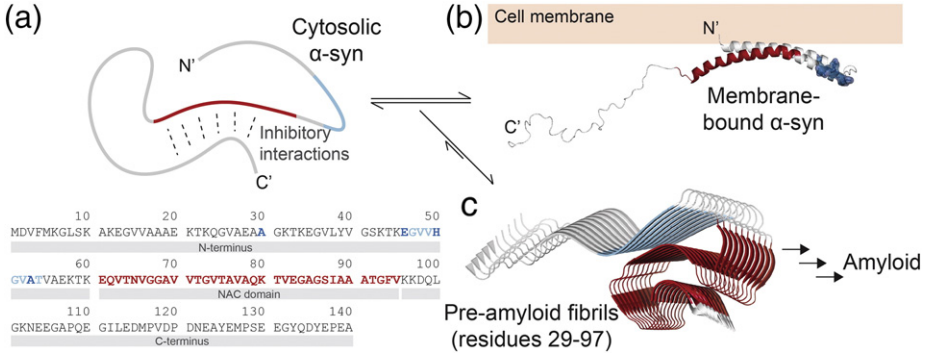

Fig. 1. Aggregation triggered by global deprotection of amyloidogenic regions. (a) α-syn is an IDP composed of an N terminus, the amyloidogenic NAC domain, and a C-terminal tail that forms inhibitory interactions with the NAC domain (upper left). The primary sequence of α-syn is shown (lower left), highlighting the amyloidogenic NAC domain (red), residues 46–54 that form part of the core of pre-amyloid fibrils (light blue), and PD-associated mutations (A30P, E46K, H50Q, A53T; dark blue). (b) α-syn acquires additional secondary structure upon interaction with the cell membrane (PDB ID 1XQ8 [143]), and (c) the equilibrium between the unstructured and partially structured forms is believed to be pivotal in the formation of amyloid fibrils (PDB ID 2N0A [4]; right).