

Osmolytes and macromolecular crowders diversely affect the aggregation of the cancer-related L106R mutant of the Axin RGS domain

Tommaso Garfagnini^a, Daniel Harries^{a,b*}, Assaf Friedler^{a*}

^a Institute of Chemistry, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem 9190401, Israel.

^b The Fritz Haber Center, Institute of Chemistry, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem 9190404, Israel.

* corresponding authors.

Abstract

Protein aggregation is involved in a variety of diseases, including neurodegenerative diseases and cancer. The cellular environment is crowded by a plethora of cosolutes comprising small molecules and biomacromolecules at high concentrations, which may influence the aggregation of proteins *in vivo*. To account for the effect of cosolutes on cancer-related protein aggregation, we studied their effect on the aggregation of the cancer-related L106R mutant of the Axin protein. Axin is a key player in the Wnt signaling pathway, and the L106R mutation in its RGS domain results in a native molten globule that tends to form native-like aggregates. This results in uncontrolled activation of the Wnt signaling pathway, leading to cancer. We monitored the aggregation process of Axin RGS L106R *in vitro* in the presence of a wide ensemble of cosolutes including polyols, amino acids, betaine and polyethylene glycol (PEG) crowders. Except *myo*-inositol, all polyols decreased RGS L106R aggregation, with carbohydrates exerting the strongest inhibition. Conversely, betaine and PEGs enhanced aggregation. These results are consistent with the reported effects of osmolytes and crowders on the stability of molten globular proteins and with both amorphous and amyloid aggregation mechanisms. We suggest a model of Axin L106R aggregation *in vivo*, whereby molecularly small osmolytes keep the protein as a free soluble molecule but the increased crowding of the bound state by macromolecules induces its aggregation at the nano-scale. Our study sheds light on the potential contribution of cosolutes to the onset of cancer as a protein misfolding disease, and on the relevance of aggregation in the molecular aetiology of cancer.

Introduction

Protein aggregation is a widespread phenomenon that results from an unbalance of proteostasis¹. Starvation, thermal or oxidative stresses, as well as ageing-related dysfunctions in the proteasome are known to trigger aggregative events¹. The mechanism of aggregation and the structure of the aggregates define the biological process outcomes. Several forms of aggregates have been described: (1) *Native aggregates*, which include proteins that retain their native fold and are stored within the native aggregates as a protective mechanism. The protein functionality is maintained within these aggregates or regained upon release²; (2) *Native-like aggregates*, which include proteins undergoing a limited aberration of their folding that can assemble into these native-like aggregates while retaining most of their native structure but with altered function³; (3) *Amorphous aggregates*, which include proteins that undergo larger modifications to native folding that can induce collapse into amorphous aggregates lacking a well-defined structure, with a complete loss of function⁴; (4) *Amyloid aggregation*, whereby proteins can undergo extensive structural rearrangements to β -sheet-rich structures that give rise to amyloid aggregates characterized by a cross- β structure and an unbranched fibrillar morphology⁵. The formation of mature intra/extra-cellular amyloid deposits is preceded and accompanied by soluble intermediates, such as transient amyloid oligomers, characterized by heterogeneous structure and inherent toxicity⁶.

The implications of protein aggregation for health are tremendous. The pathogenesis of a vast ensemble of neurodegenerative and systemic diseases correlates specifically with the appearance of amyloid aggregates⁵. Moreover, the aetiology of various types of cancer is traced back to either amyloid or native-like aggregation of proteins that regulate key processes for cell homeostasis. For instance, misfolded p53 mutants undergo amyloid fibrillation via prion-like propagation, causing related forms of cancer⁷. In another mechanism, destabilized p53 cancer mutants co-aggregate with its paralogue proteins p63 and p73 to form soluble amyloid-like assemblies that are unable to exert the tumour suppressor function⁸.

A newly reported case of protein aggregation in cancer is represented by Axin, the scaffold protein that orchestrates Wnt signaling in the cytoplasm⁹. The L106R cancer mutation destabilizes the RGS domain of Axin and leads to the aberrant assembly of a soluble, multi-protein nano-aggregate that is unable to carry out the Wnt-suppressive function of the WT protein¹⁰. The uncontrolled Wnt activation then promotes cell proliferation and induces the onset of a cancer phenotype¹⁰.

Protein-protein interaction (PPIs), including those involved in aggregation, are widely affected by the complexity of the cellular environment¹¹. The cytoplasm contains a variety of small molecules such as nutrients and metabolites at concentrations that are on the order of hundreds of millimolar or more. We collectively refer to these molecules here as osmolytes due to their involvement in the cellular response to osmotic stress¹². Proteins and other macromolecules also occupy a large volume fraction and contribute to creating a highly crowded environment in the cytoplasm. We use the term “macromolecular crowders” to designate the excluded-volume effect of this class of molecules¹³. Via their preferential exclusion from protein surfaces and other quinary interactions, osmolytes and crowders affect protein structure and dynamics, modulate PPIs, and drive association phenomena including aggregation^{14–18}.

Osmolytes typically exert an enthalpically-driven stabilization of compact folded conformations that minimizes the extent of the exposed protein surface and its hydration shell¹⁵. Thus, osmolytes are often able to oppose amyloid aggregation of structured and disordered proteins alike¹⁹. An investigation on the mechanism of inhibition using a model peptide revealed that polyols inhibit all stages of amyloid aggregation by favouring the adoption of a folded β -hairpin state over an ensemble of unfolded states²⁰. Yet, the extent of inhibition varies with the chemistry of the osmolyte and the protein. For instance, different polyols affect distinct stages of insulin fibrillation, and trehalose inhibits A β 40 oligomers and fibrils, but only suppresses A β 42 fibrils, and expedites the onset of α -synuclein early aggregation while decreasing the total amount of fibrillated protein^{21–23}. Polyols exert a similar protective effect on the heat or pH-induced amorphous aggregation of folded proteins like conalbumin and monoclonal antibodies^{24–27}. Amino acids display a varied effect. Proline suppresses the fibrillation of insulin, whereas it redirects huntingtin amyloidogenesis to amorphous aggregation^{28,29}. At low concentrations it enhances the thermal aggregation of a folded enzyme, while at high ones it has an inhibitory effect³⁰. Betaine also abolishes insulin fibrillation but, unlike proline, speeds up substantially that of huntingtin^{28,29}.

In contrast to osmolytes, the exclusion of macromolecular crowders from protein surfaces is typically dominated by entropy and favours the protein compact states so as to maximize the cosolutes free volume¹³. Crowders can destabilize folded states and enhance the aggregation of a large number of proteins, modulating diversely the molecular mechanism of the process³¹. For instance, they speed up the nucleation rate of the amyloid aggregation of a model peptide and promote the fragmentation of the mature fibrils, which seeds further aggregative events³². Moreover, PEG crowders trigger the liquid-liquid separation of α -

synuclein, that readily converts into an amyloid hydrogel composed of oligomers and fibrils³³. Yet, although increasing both nucleation and fibrillation rate of monomeric α -synuclein, crowders stabilize the native hexamers and tetramers of human insulin and bovine core histone, that are also compact states, thereby delaying their further aggregation³⁴. Crowders can assist functional self-assembly processes that contribute to correct proteostasis³⁵. For instance, they facilitate the complexation of the molecular chaperones GroEL, DnaK/ClpB and α -crystallin with unfolding and aggregating clients and induce the association of folded proteins into functional polymers or native aggregates^{36–40}. Based on these observations, it appears that the cosolutes from these two classes, osmolytes and crowders, tend to exert an effect on protein aggregation that is independent of the mechanism and of the specific type of the aggregate.

The effects of cosolutes on protein aggregation have so far been studied in processes that yield large, insoluble aggregates. How the cellular environment affects the assembly of soluble, aberrant complexes of misfolded proteins (or “nano-aggregates”) is still unknown despite their involvement in disease. To address this question, we assessed the impact of osmolytes and macromolecular crowders on the nano-aggregation of Axin RGS L106R. Some cancers also possess the characteristics of a protein aggregation disease, since cancer cells are hyper-crowded environments due to an accelerated metabolism⁴¹. We characterized the RGS L106R aggregation as a native-like process and found that polyols and amino acids generally inhibited it, with sugars and some amino acids acting as suppressors. To contrast, betaine and PEG crowders facilitated RGS L106R aggregation. Based on our results, we propose a mechanism for Axin L106R aggregation in vivo where osmolytes maintain the protein in the monomeric state as long as it is unbound. Upon recruitment into a multi-protein complex, the increased crowding within the complex triggers the aggregative events that turn it into a dysfunctional, yet soluble nano-aggregate.

Results

The L106R mutation induces partial unfolding of the RGS domain

Axin RGS domain WT and bearing the L106R cancer-related mutation were expressed as fusion proteins with glutathione S-transferase (GST). After cleavage of the GST-RGS constructs with TEV protease, RGS WT and L106R were purified as described in the Methods section. We performed a comprehensive biophysical characterization RGS L106R and its aggregation, in comparison to RGS WT. Circular dichroism (CD) spectra of RGS

L106R and RGS WT displayed minima at 208 nm and 222 nm that are typical of an α -helical secondary structure. The weaker intensity of the RGS L106R CD spectrum as compared to that of the WT protein correlated with a partial unfolding induced by the cancer-related L106R mutation (Fig. 1A). The melting temperature of the RGS domain dropped from 57°C for the wild type to 27°C for the L106R mutant, indicating a loss of thermodynamic stability upon mutation. The RGS L106R melting curve showed no lag phase and a longer denaturation phase, starting at a temperature as low as 5°C (Fig. 1B). This also reveals that the L106R mutation significantly destabilizes the protein, reflected in the more gradual unfolding compared with RGS WT, already at low temperatures. A weakening of RGS folding is consistent with L106 forming the structural pivot that holds together the α 1, α 2, and α 3 strands by interacting hydrophobically with F102, C111, L115 and T197⁴². The Leu-to-Arg mutation disrupts the hydrophobic core, resulting in partial unfolding.

We used the 8-anilinonaphthalene-1-sulfonic acid (ANS) fluorescent probe to follow the aggregation of the RGS WT and L106R. ANS is commonly used to assess the tertiary structure and the compactness of folding by detecting the exposure of hydrophobic patches on protein surfaces⁴³. ANS emission in the presence of RGS WT was indistinguishable from that of the free, unbound molecule, with a wavelength of maximal emission of 495 nm and no dose-dependence (Fig. 1C). In the presence of RGS L106R, ANS displayed a steep, dose-dependent increase of the emission intensity and a blue shift of the maximal emission wavelength from 495 to 470 nm (Fig. 1D). Such a change in the spectral properties is consistent with ANS becoming embedded into a hydrophobic chemical environment or bound to a protein hydrophobic pocket⁴³. Hence, we conclude that the L106R mutation induced a loosening of the folding of RGS and the formation of solvent-accessible hydrophobic patches capable of ANS binding.

Dynamic light scattering (DLS) at time zero of RGS L106R displayed a sharp peak at a hydrodynamic diameter of about 4-5 nm, compatible with the size of the monomeric protein, as previously reported¹⁰. A smaller, broad peak at about 12 nm (6.7% of the intensity) was indicative of the co-presence of association products of higher molecular weights (Fig. 2A). The DLS at time zero of RGS WT displayed a single, sharp peak at about 5.6 nm, accounting for a dimeric or trimeric complex (Fig. 3A).

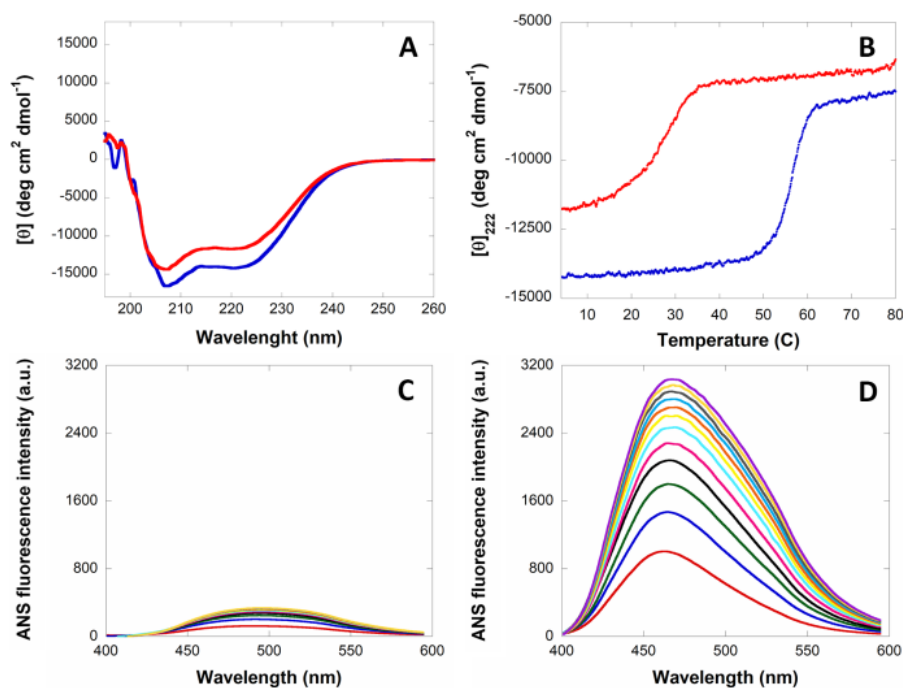


Figure 1: Biophysical characterization of RGS WT and RGS L106R. (A) CD spectra of RGS WT (blue) and RGS L106R (red) at 4 °C. (B) melting curves of RGS WT (blue) and RGS L106R (red) recorded by CD at 222 nm with a temperature gradient of 5 °C/min from 4 to 85 °C. ANS fluorescence spectra acquired by 12 sequential additions to (C) RGS WT and (D) RGS L106R. The RGS:ANS molar ratio varied from 1:1 to 1:12, with a starting protein concentration of 20 μ M.

RGS L106R aggregates are native-like

DLS measurements were used to monitor aggregation of RGS L106R over time. The DLS of RGS L106R recorded at different time points over 24 hours revealed that RGS L106R formed aggregates with increasing molecular weights. Within the first hour of incubation, the protein population shifted completely from the monomer to a 12-15 nm wide association product. Further aggregation occurred, as the hydrodynamic diameter increased up to about 38 nm in the following 23 hours of incubation. The peak broadening accounted for a wide distribution of aggregates of different sizes, with polydispersity increasing over time (Fig. 2A). Transmission electron microscopy (TEM) images acquired after 12 hours of incubation confirmed that RGS L106R forms aggregates characterized by a distinctive elongated, unbranched morphology. The aggregates were monodisperse in width (about 10 nm) but polydisperse in length (between 30 and 150 nm) (Fig. 2D).

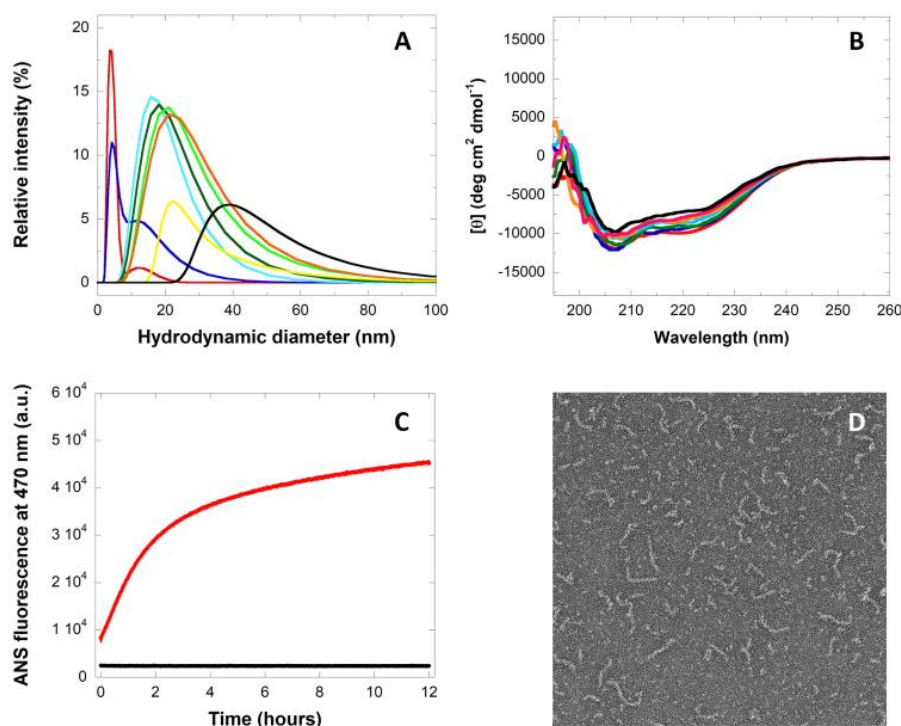


Figure 2: Biophysical characterization of RGS L106R aggregation. (A) DLS measurements and (B) CD spectra of RGS L106R acquired at fixed time points over 24 hours of incubation: time zero (red), 30 minutes (blue), 1 hour (light blue), 4 hours (orange), 12 hours (yellow), 24 hours (black). (C) time course kinetics of ANS fluorescence at 470 nm in the presence of RGS L106R (red) as opposed to the blank (black). (D) TEM images of RGS L106R after 12 hours of incubation.

Conversely, RGS WT did not undergo aggregation to any extent over time. The DLS displayed a single, sharp peak at about 4-5 nm hydrodynamic diameter over the whole incubation time and the TEM images revealed spheroidal oligomers with a narrow height distribution of around 10 nm (Fig. 3A,D). CD was then used to follow the changes in secondary structure during aggregation. The spectra, recorded over the same time window of DLS, showed that RGS L106R undergoes only partial unfolding during aggregation, with 75% of the native α -helical structure retained after 24 hours of incubation, as assessed by the CD at 222 nm (Fig. 2B). RGS WT, on the other hand, did not show any change in its CD spectrum throughout the incubation time (Fig. 3B). Taken together, our findings suggest that RGS L106R undergoes a process of native-like aggregation.

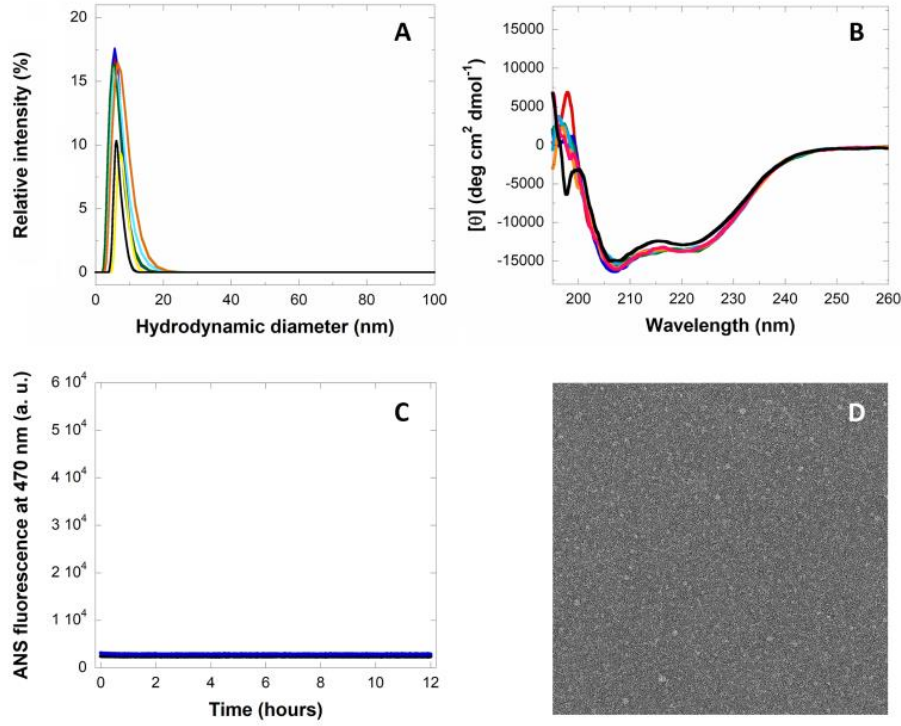


Figure 3: RGS WT does not display aggregation properties. (A) DLS measurements and (B) CD spectra of RGS WT acquired at fixed time points over 24 hours of incubation: time zero (red), 30 minutes (blue), 1 hour (light blue), 4 hours (orange), 12 hours (yellow), 24 hours (black). (C) time course kinetics of ANS fluorescence at 470 nm in the presence of RGS WT (blue) as opposed to the blank (black). (D) TEM images of RGS WT after 12 hours of incubation.

The spectral properties of ANS were then used to monitor how the hydrophobicity of RGS L106R surface evolves over aggregation. The experiments were performed at an RGS:ANS molar ratio of 1:3, corresponding to mid-saturation. The ANS time course kinetics in the presence of RGS L106R displayed a steep, time-dependent increase of the emission at 470 nm over time (Fig. 2C). The kinetic trace was fit to the following function,

$$(1) \quad y = m_{\infty} + m_1 \times t + m_2 \times \exp(-k_{\text{ANS}} \cdot t)$$

where m_{∞} is the ANS emission at the end of the kinetics at long times; m_1 is the slope of the linear term that corrects for the absence of a plateau at long times; m_2 is the difference between ANS emission at the end and at the beginning of the kinetics; k_{ANS} is the rate constant of the exponential growth term and t is time. From a physical point of view, k_{ANS} indicates the rate at which the number and/or the extent (area) of the hydrophobic patches increase on the RGS L106R surface⁴³. The exposure of hydrophobic patches correlates with the progression of protein aggregation, so k_{ANS} could be used to quantitatively characterize

the kinetics of the process⁴⁴. RGS WT maintained its structure and ANS-binding properties were unaltered over the whole incubation time (Fig. 3C).

Selection of osmolytes and crowders to be tested for their effect on RGS L106R aggregation

To mimic the effect of the cellular environment on the native-like aggregation of RGS L106R, we selected a pool of osmolytes that are known to populate the cytoplasm at high concentrations. These osmolytes differ in their molecular weight, geometry and structural constraints (Table 1). These properties are relevant, as size, shape and molecular dynamics may affect osmolyte solvation and hence may impact their effect on RGS L106R aggregation. The series of glycerol, erythritol, xylitol and sorbitol was selected since it represents a set of polyols with increasing size and number of hydroxyl groups. We also chose sucrose, maltose and trehalose that have the same molecular weight but different structures and connectivity. Sorbitol, *myo*-inositol and glucose have different structures and configurations due to the different constraints applied to their geometry. The effect of amino acids with different constraints on the torsion angle was investigated with glycine, proline and glycylglycine. To account for the hydrophobicity of the side chain we also tested the effect of the addition of a single hydroxyl group and picked serine as opposed to alanine and hydroxyproline as opposed to proline.

To understand how the macromolecules present in the cytoplasm affect the RGS L106R aggregation process *in vivo*, we were also interested in mimicking large, multi-domain hub proteins directly involved in regulative scaffolding activity and rich in intrinsically disordered regions (IDRs), such as Axin itself and its major protein partners. For this purpose, we selected polyethylene glycol (PEG) polymers, that are flexible macromolecules without a fixed shape or fold, representing extended IDRs. We tested PEG crowders of increasing size and molecular weight: PEG 400, PEG 3350 and PEG 6000.

Effect of osmolytes and crowders on RGS L106R aggregation

The aggregation of RGS L106R was followed by ANS time course kinetics in the presence of 750 mM of osmolyte or ethylene glycol (the PEG building block). This concentration is comparable to the actual osmolytes volume fraction in cellular environments and is often used as a proxy of the *in vivo* conditions for aggregation¹². With the exception of *myo*-inositol, all polyols slowed the rate of RGS L106R aggregation (Fig. 4A, Table 1). The k_{ANS} varied from $0.61 \pm 0.05 \text{ s}^{-1}$ in the absence of cosolutes to $0.13 \pm 0.03 \text{ s}^{-1}$ in the presence of sucrose, decreasing linearly as a function of the polyol molecular weight in the order: $k_{ANS,ref} \approx k_{ANS,Ins} > k_{ANS,Gyl} > k_{ANS,Eth} > k_{ANS,Xyl} > k_{ANS,Glc} > k_{ANS,Suc}$ (Table 1; Fig. 4B). With the

exception of *myo*-inositol, all the k_{ANS} values in the presence of the polyols were significantly different from that of the reference, as they display a p value below 0.05, corresponding to the minimum for statistical significance (Table 1). A strong divergence from the linear trend seen in figure 4B (blue line) was observed for polyols having the same molecular weight but different geometry and structural constraints (Fig. 4B).

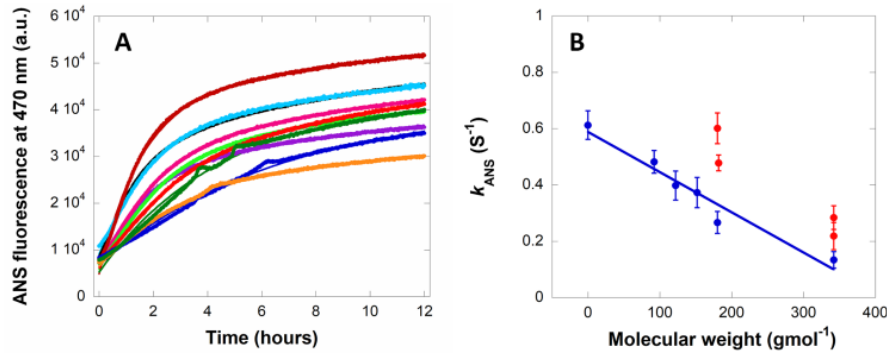


Figure 4: Effect of polyols on RGS L106R aggregation. (A) Time course ANS aggregation kinetics of RGS L106R in the absence (black) and in the presence of glycerol (purple), erythritol (pink), xylitol (light green), sorbitol (light blue) and *myo*-inositol (brown), glucose (red), sucrose (blue), maltose (orange) and trehalose (green). Each trace represents the average of at least three independent measurements. (B) Dependence of k_{ANS} of RGS L106R aggregation in the presence of glycerol, erythritol, xylitol, glucose and sucrose (blue line) and of sorbitol, *myo*-inositol, maltose and trehalose (red) on the molecular weight of the polyol osmolytes.

Despite having nearly the same molecular size, sorbitol, glucose and *myo*-inositol affected differently the k_{ANS} of RGS L106R aggregation, with $k_{ANS,Srb} = 0.48 \pm 0.03 \text{ s}^{-1}$, $k_{ANS,Glc} = 0.27 \pm 0.04 \text{ s}^{-1}$ and $k_{ANS,Ins} = 0.60 \pm 0.05 \text{ s}^{-1}$ (Table 1; Fig. 8). This means that only in the presence of glucose was k_{ANS} substantially decreased, while in the presence of sorbitol it was comparable to that in the presence of glycerol, and for *myo*-inositol it was unaltered as compared to that in the absence of any cosolute. The disaccharides sucrose, maltose and trehalose also had distinct effects on the k_{ANS} of the process, with $k_{ANS,Suc} = 0.13 \pm 0.03 \text{ s}^{-1}$, $k_{ANS,Mal} = 0.28 \pm 0.04 \text{ s}^{-1}$ and $k_{ANS,Thl} = 0.22 \pm 0.05 \text{ s}^{-1}$ (Table 1; Fig. 8). Among these disaccharides, only sucrose exerted a noticeable effect on RGS L106R aggregation while maltose and trehalose had a similar effect to that of glucose.

Amino acids also had an impact on RGS L106R aggregation. Except for glycine and alanine, amino acids generally decreased the rate of RGS L106R aggregation, with the k_{ANS} significantly different from that of the reference and with p value much below 0.05 (Fig. 5A; Table 1). We observed a non-linear size-dependence of the k_{ANS} of the aggregation process on the AA molecular weight (Fig. 5B).

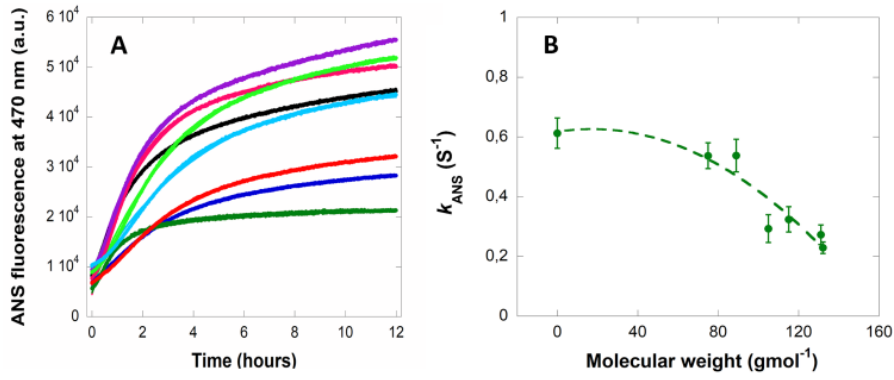


Figure 5: Effect of amino acids on RGS L106R aggregation. (A) Time course ANS aggregation kinetics of RGS L106R in the absence (black) and in the presence of glycine (pink), alanine (purple), serine (light green), proline (blue), glycylglycine (light blue), hydroxyproline (red) and betaine (green). Each trace represents the average of at least three independent measurements. (B) Dependence of the k_{ANS} of RGS L106R aggregation in the presence of glycine, alanine, serine, proline, glycylglycine, hydroxyproline and betaine on the molecular weight of the amino acid osmolytes. The dashed green line is a guide for the eye.

Glycine only slightly affected RGS L106R aggregation, with $k_{ANS,Gly} = 0.54 \pm 0.04 \text{ s}^{-1}$ (Table 1; Fig. 8). The presence of hydrophobic side chains of different sizes resulted in different outcomes: the effect of alanine was indistinguishable from that of glycine, with $k_{ANS,Ala} = 0.54 \pm 0.05 \text{ s}^{-1}$, while proline decreased it to $k_{ANS,Pro} = 0.32 \pm 0.04 \text{ s}^{-1}$ (Table 1; Fig. 8). The addition of a hydroxyl group to the side chain resulted in distinct effects: compared with alanine, the presence of serine induced a sizeable decrease to $k_{ANS,Ser} = 0.29 \pm 0.05 \text{ s}^{-1}$, while the effect of hydroxyproline was comparable to that of proline, $k_{ANS,HyPro} = 0.27 \pm 0.03 \text{ s}^{-1}$ (Table 1; Fig. 8). The strongest decrease was observed in the presence of the glycylglycine dipeptide, with $k_{ANS,GGI} = 0.23 \pm 0.02 \text{ s}^{-1}$ (Table 1; Fig. 8). Betaine showed a completely different impact to amino acids as it strongly enhanced RGS L106R aggregation, yielding $k_{ANS,Bet} = 0.92 \pm 0.03 \text{ s}^{-1}$ (Table 1; Fig. 8).

Finally, we tested PEG crowders and found that the presence of the smallest of the three, PEG 400, did not result in any effect on RGS L106R aggregation. Conversely, PEG 3350 and PEG 6000 crowders displayed an opposite behaviour compared to the osmolytes, as they increased the k_{ANS} of RGS L106R aggregation, with $k_{ANS,PEG3350} = 0.99 \pm 0.05 \text{ s}^{-1}$ and $k_{ANS,PEG6000} = 1.14 \pm 0.07 \text{ s}^{-1}$ (Table 1; Fig. 8). Both $k_{ANS,PEG3350}$ and $k_{ANS,PEG6000}$ were significantly different from the k_{ANS} in the absence of molecularly small cosolutes, with a p value much lower than 0.05 (Table 1). The PEG-induced enhancement of k_{ANS} displayed a positive near-linear dependence on the molecular weight of the crowder (Fig. 6B).

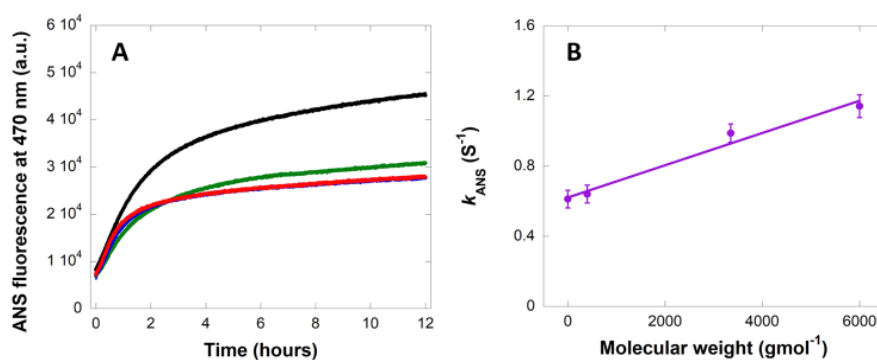


Figure 6: Effect of crowders on RGS L106R aggregation. (A) Time course ANS aggregation kinetics of RGS L106R in the absence (black) and in the presence of crowders: PEG 400 (green), PEG 3350 (blue) and PEG 6000 (red). Each trace represents the average of at least three independent measurements. (B) Dependence of the k_{ANS} of RGS L106R aggregation in the presence of PEG 400, PEG 3350 and PEG 6000 on the molecular weight of the crowders. The purple line is a guide for the eye.

TEM studies provided evidence that osmolytes that decreased k_{ANS} had a strong inhibitory effect on RGS L106R aggregation (Fig. 7).

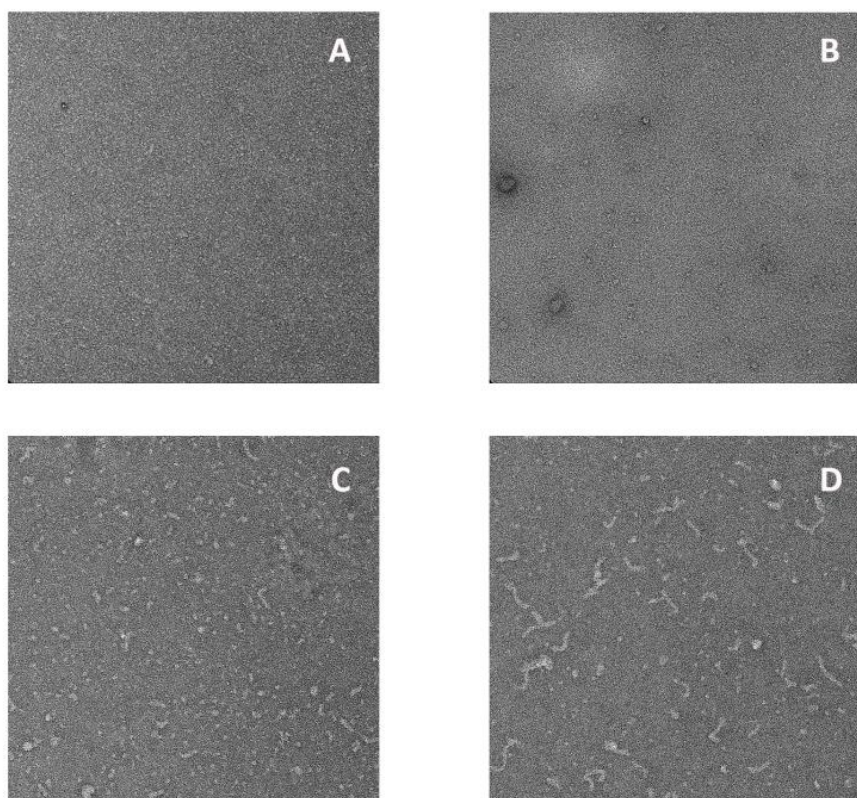


Figure 7: Effect of cosolutes on RGS L106R aggregation upon co-incubation. TEM images of RGS L106R after 12 hours of incubation in the presence of 750 mM of (A) glucose, (B) trehalose, (C) proline and (D) betaine.

No aggregates were observed when RGS L106R was incubated in the presence of glucose or trehalose (Fig. 7A,B). Proline also inhibited aggregation but to a lesser extent, with shorter, mostly round-shaped or truncated aggregates visible under TEM imaging (Fig. 7C). To contrast, co-incubation with the k_{ANS} -enhancer osmolyte betaine increased the number of longer worm-like RGS L106R aggregates (Fig. 7D).

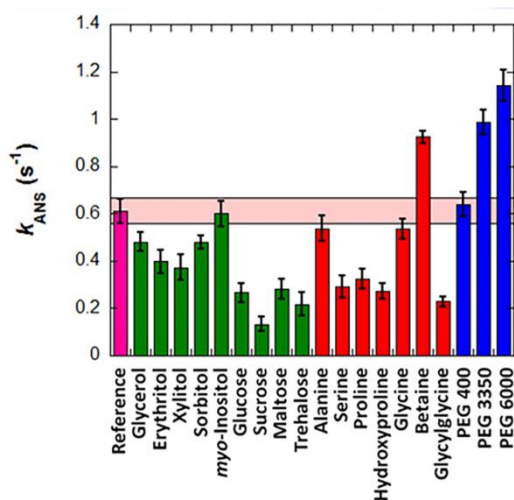


Figure 8: Apparent rate of RGS L106R aggregation. k_{ANS} of RGS L106R aggregation in the absence of cosolutes (pink) and in the presence of polyols (green), amino acids derivatives (red) and crowders (blue).

Table 1. Effect of osmolytes and crowders on RGS L106R aggregation.

Osmolyte	Class	Molecular weight (gmol^{-1})	k_{ANS} (s^{-1})
No osmolyte	-	-	0.61 ± 0.05
Glycerol (Gyl)	Polyol	92.10	0.48 ± 0.04 $p = 0.0094$
Erithrytol (Eth)	Polyol	122.12	0.40 ± 0.05 $p = 0.0004$
Xylitol (Xyl)	Polyol	152.15	0.37 ± 0.05 $p = 0.0002$
Sorbitol (Srb)	Polyol	182.17	0.48 ± 0.03 $p = 0.0060$
myo-Inositol (Ins)	Polyol	180.16	0.60 ± 0.05 $p = 0.7562$
Glucose (Glc)	Polyol	180.16	0.27 ± 0.04 $p = 0.0001$

Sucrose (Suc)	Polyol	342.30	0.13 ± 0.03 $p < 0.0001$
Maltose (Mal)	Polyol	342.30	0.28 ± 0.04 $p = 0.0001$
Trehalose (Thl)	Polyol	342.30	0.22 ± 0.05 $p < 0.0001$
Glycine (Gly)	Amino acid	75.07	0.54 ± 0.04 $p = 0.049$
Alanine (Ala)	Amino acid	89.09	0.54 ± 0.05 $p = 0.0945$
Serine (Ser)	Amino acid	105.09	0.29 ± 0.05 $p < 0.0001$
Proline (Pro)	Amino acid	115.13	0.32 ± 0.04 $p < 0.0001$
Hydroxyproline (HyPro)	Amino acid	131.12	0.27 ± 0.03 $p < 0.0001$
Glycylglycine (GGI)	Amino acid	132.12	0.23 ± 0.02 $p < 0.0001$
Betaine (Bet)	Trimethylamine	117.15	0.92 ± 0.03 $p = 0.0001$
PEG 400	Crowder	400.0(average)	0.64 ± 0.05 $p = 0.4353$
PEG 3350	Crowder	3350.0(average)	0.99 ± 0.05 $p = 0.0001$
PEG 6000	Crowder	6000.0(average)	1.14 ± 0.07 $p < 0.0001$

Discussion

ANS as a probe for studying protein aggregation

We followed the kinetics of RGS L106R native-like aggregation using ANS, a fluorescent probe used for detecting exposed hydrophobic patches on protein surfaces and accompanies aggregation with an increase of the emission over time⁴³. ANS has proved instrumental in monitoring aggregation processes of all kinds, irrespective of the mechanism and of the

internal structure of the aggregate⁴⁴. For instance, ANS was employed to follow the amorphous aggregation of carbonic anhydrase, the amyloid aggregation of transthyretin, A β , IAPP and HypF-N, and the heterogeneous amyloid/amorphous aggregation of concanavalin A, immunoglobulin light chain and β ₂-microglobulin^{4,45–51}. ANS was also used to characterize the native-like aggregation of β -lactoglobulin and interleukin-1 β in the absence of major structural rearrangements within the protein building blocks^{52,53}. Even though ANS cannot provide precise mechanistic nor structural information about an aggregation process, the increase of its emission is treated as proportional to the formation of new ANS binding sites on protein surface and to the progression of aggregation^{43,44}. The rate constant of the exponential growth of ANS emission k_{ANS} hence fully characterizes the native-like aggregation process of RGS L106R. The reliability of ANS as a probe is reinforced by its affinity to the native protein. ANS binds partially folded states like transient intermediates of protein folding and stabilized molten globules, while fully folded and unfolded states generally do not bind, or only weakly bind, ANS^{54,55}. The decrease in secondary structure and the intense ANS binding to monomeric RGS L106R are hence indicative of its molten globular nature. Consistently, R103, R106 and K107 at the centre of the hydrophobic core disrupted upon mutation are potential binding sites for the ANS sulfonate group via ion-pairing⁴⁴. This suggests that RGS L106R is already in the molten globule form in its native state and that the increase in ANS emission is solely due to aggregation.

The aggregation mechanism of RGS L106R

RGS L106R time course aggregation kinetics is fit to a single exponential curve. The absence of a lag phase is consistent with RGS L106R being a molten globule able to initiate aggregation even without first accessing an aggregation-competent monomeric state. This apparent simplicity can be deceiving, as even complex, multi-step kinetic processes may display a similar trend⁵⁶. RGS L106R forms assemblies having a worm-like, unbranched morphology. The width of the aggregate, as observed by TEM imaging, corresponds to the hydrodynamic diameter of the association products measured by DLS in the first hour of aggregation. This suggests that RGS L106R mature aggregates are formed by smaller aggregates that associate via geometrically constrained aggregation interfaces allowing for an ordered elongation (Fig. 9). This mechanism is reminiscent of the process of Barstar oligomerization followed by lateral association to form amyloid protofibrils, which is also characterized by a single exponential kinetics⁵⁶.

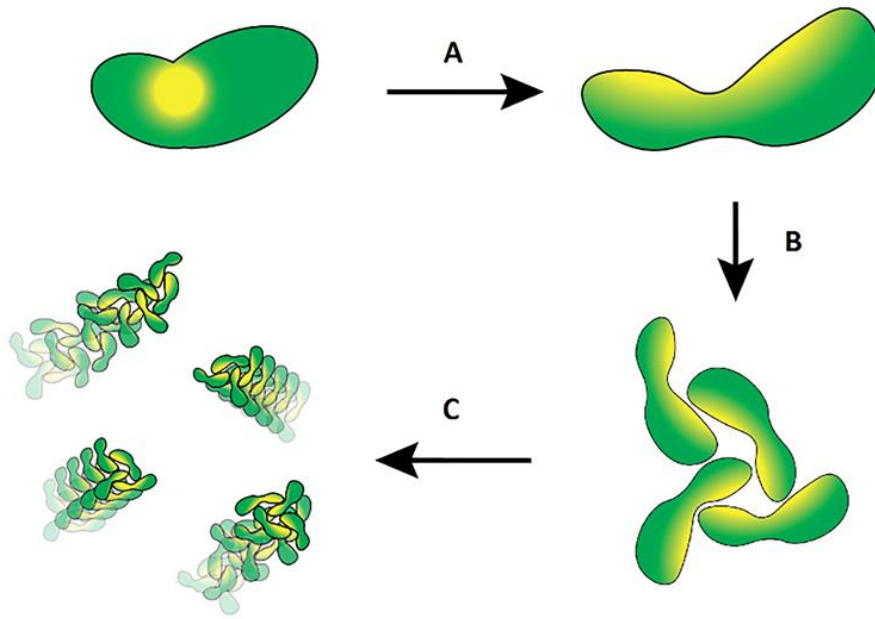


Figure 9: Suggested model of RGS L106R aggregation *in vitro*. RGS WT has a compact native fold (green) with a cluster of hydrophobic residues buried in the core of the protein (yellow). (A) The L106R cancer-related mutation induces partial unfolding and exposure of the core residues in extended hydrophobic patches on the protein surface. (B) During the first hour of incubation, tetrameric association products of RGS L106R are formed via hydrophobic interactions between the exposed patches. (C) Further aggregation proceeds by piling up of the tetrameric oligomers to form a worm-like, unbranched aggregate.

The distinctive worm-like appearance and the absence of structural conversion within the assembly make RGS L106R aggregates similar to those formed by γ D-crystallin P23T, which are also native-like³. Conversely, other known native-like aggregates display very different characteristics. The early native-like aggregates of SsoAcP, human/murine prion and β 2-microglobulin readily convert into proper amyloid fibrils, while RNase A aggregates combine native-like and amyloid nature, with folded domains decorating an inner amyloid fibril^{57–61}. The native-like aggregates formed by human pancreatitis-associated protein and yeast prion Ure2p still display a mature fibrillar morphology even in the absence of amyloid structure^{62,63}. RGS L106R native-like aggregates are distinctly different from all others also from a biological point of view, as they are the only ones to correlate with cancer¹⁰.

The effect of cosolutes on protein aggregation is independent of both the kinetic mechanism of aggregation and the structure of the aggregate

The effects of osmolytes and crowders on the native-like aggregation of RGS L106R are similar to those observed for other types of aggregation. Polyols such as alcohols and sugars can indeed oppose the amyloid and amorphous aggregation of a variety of proteins to

different extents, with the inhibitory effect also being proportional to the size of the molecule^{19–27}. Proline and hydroxyproline also inhibit the fibrillation and thermally-induced aggregation of a number of proteins^{19,28–30,64}. Betaine and PEG crowders are able to trigger some aggregative processes of either amyloid or amorphous nature^{19,28,31–35}. Less expectedly, *myo*-inositol, glycine and alanine do not display any effect, despite being reported to affect other processes of aberrant protein assembly^{15,19}. Overall, these findings are consistent with our observations on RGS L106R aggregation, as monitored through ANS fluorescence, a fluorescent probe unaffected by the presence of the cosolutes. ANS binding to proteins is indeed inhibited only by organic acids, anionic surfactants and tryptophan, that mimic ANS by having a large hydrophobic moiety with a negatively charged head⁴⁴. None of the cosolutes tested in the current study share such amphipathic molecular properties, so any decrease in k_{ANS} is unlikely due to ANS displacement during the kinetics.

The action of cosolutes hence seems to be independent of both the kinetic mechanism of aggregation and the structure of the aggregate, suggesting that they may impact a kinetic step that lies upstream of the aggregation itself. This has led us to suggest that cosolutes may exert their effect on RGS L106R native-like aggregation by increasing or decreasing the thermodynamic stability of the molten globular state and by modulating its structural properties. The stabilization of the molten globule would indeed disfavour the competing aggregation process, while a destabilization would accelerate its initiation. In support of this, according to several reports, cosolutes found to oppose or favour aggregation also either stabilize or destabilize molten globular proteins. For instance, sucrose and proline induced a contraction of molten globular ribonuclease A, thereby stabilizing a compact state⁶⁵. Carbohydrates and alcohols stabilized the molten globule state of equine ferricytochrome *c*^{66,67}. Glycerol stabilized an alkaline molten globular state of 5-aminolevulinate synthase, allowing for the retention of $\sim 1/5$ of the k_{cat} of the native enzyme⁶⁸. Our suggested mechanism for the osmolyte-mediated inhibition of aggregation is supported also by our results at the peptide level. Computer simulations show that sorbitol indeed shifts the conformational equilibrium of met16 peptide from an unfolded to a folded state with the aggregation-prone residues unexposed, decreasing the rate of nucleation of an amyloid aggregate⁶⁹. Conversely, PEG and ficoll crowders destabilize native myoglobin into a molten globule at physiological conditions^{70,71}, while ficoll and dextran destabilize the molten globular 5-aminolevulinate synthase and induce its association into amorphous aggregates⁶⁸. At the atomistic level, PEGs interact more favourably with aromatic C, amide N, and cationic N⁷². Such preferential interactions with portions of side chain or backbone

are likely to lead to attractive interactions for some exposed residues, inducing the destabilization of the native conformation and unfolding⁷².

A model for how the cellular environment affects Axin L106R aggregation

Based on the effects of osmolytes and crowders on L106R aggregation, we propose a general model for how the cellular environment affects Axin L106R aggregation *in vivo*. The osmolytes stabilize the molten globular state of RGS L106R and thereby oppose its structural collapse into a nano-aggregate, allowing Axin to reside in the cytoplasm as a monomer. The inhibition of aggregation may thus be a result of suppressed exposure of aggregation-prone regions of the protein in the compact state that could promote protein-protein associations. Upon recruitment into large complexes, such as the destruction complex or the signalosome, RGS L106R is exposed to higher local protein concentrations, without directly participating in PPIs. The crowding within the complexes outweighs the stabilizing effect of the osmolytes and triggers aggregation, potentially due to increased probabilities for protein-protein encounters (Fig. 10).

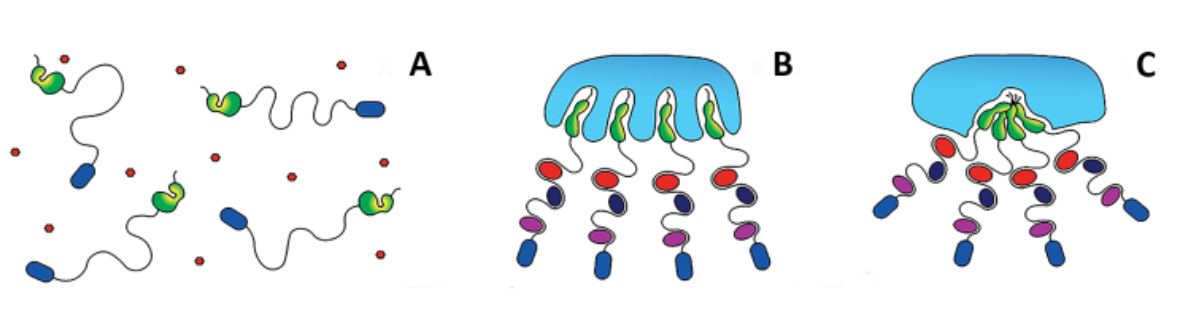


Figure 10: Model of Axin L106R aggregation *in vivo*. (A) In the cytoplasm Axin L106R fluctuates as a monomer, with the DIX (blue) and mutated RGS (green) domains free to move. In the presence of osmolytes (red), the hydrophobic patches (yellow) formed on RGS surface upon L106R-induced partial unfolding remain partially hidden from the solvent, thereby inhibiting aggregation. (B) Axin forms the destruction complex with GSK-3 β (red), β -catenin (violet), CK1 (purple) and the large, mostly intrinsically disordered scaffold protein APC (light blue) through the RGS domain. The high local concentration of disordered strands within the multi-protein complex triggers the destabilization of mutated RGS and the exposure of the hydrophobic patches. (C) Axin L106R aggregates via mutated RGS within the protein complex turning it non-functional.

This suggested mechanism may explain why cancer-related mutations are so abundant in hubs and scaffold proteins involved in wide networks of interactions, continuously switching from one complex to the other. Such proteins indeed experience elevated local concentration of crowders more than any other protein in the cytoplasm. Under acute crowding stress, even mildly destabilizing mutations may trigger the nano-aggregation of

structured domains and induce the cancer phenotype as a cascade event. The mechanism proposed here is potentially applicable to the cancer-related mutations of Axin, p53 and other protein hubs, with the crowding-assisted nano-aggregation of multi-protein complexes being a general mechanism for carcinogenesis in the absence of insoluble depositions.

Methods:

Site-directed mutagenesis and cloning of RGS WT and L106R

The modification of a pGEX-2T vector from Thrombin to TEV cleaving site was performed in a single step by multiple site-directed mutagenesis with the primers:

Forward ggcgaccatcctccaaaagagaatctgtattccaggatccaccgcc

Reverse ggcggtggatccctggaaatacagattctcttttgaggatggtcgcc

The genes encoding RGS WT and RGS L106R were amplified by PCR using primers with BamHI and EcoRI restriction sites at the 5' and 3' end, respectively. PCR products were purified by clean-up kits (Promega), cleaved by BamHI and EcoRI and ligated into the TEV-modified pGEX-2T vector. Bacterial colonies were screened and verified by DNA sequencing.

Expression and purification of RGS WT and L106R

BL21+RIL *E. coli* cells were transformed with the TEV-modified pGEX-2T plasmids containing RGS WT and RGS L106R genes. The cells were grown for 16 hours under vigorous agitation in LB broth media at 37 °C in the presence of 100 µM ampicillin and 35 µM chloramphenicol. The starter was diluted 1:200 (v/v) in fresh LB broth media and grown under vigorous agitation at 37 °C in the presence of 100 µM ampicillin and 35 µM chloramphenicol until [OD]₆₀₀ of 0.6. The expression was induced with 0.6 mM isopropyl β-D-1-thiogalactopyranoside for 16 hours at 16 °C under vigorous agitation. The bacteria were centrifuged at 4500 rpm for 15 minutes, harvested and stored at -80 °C. The pellet was resuspended in 50 mM Tris HCl pH 7.4, 300 mM NaCl, 5 mM β-mercaptoethanol, 10 µg/mL DNAase, 10 µg/mL lysozyme and 1 mM phenylmethanesulfonyl fluoride. The bacteria were lysed with a microfluidizer and the lysate separated from the pellet by centrifuging at 15000 rpm for 45 minutes. The lysate was purified using a Merck GST-bind 70541 resin and washed with 50 mM Tris HCl pH 7.4, 150 mM NaCl, 5 mM β-

mercaptoethanol. RGS WT and L106R were cleaved on the resin with 2.2 μM TEV protease in 5 mL buffer upon overnight incubation at 4 °C under mild agitation. The cleaved protein was collected and further purified by ÄKTA-assisted size exclusion chromatography using two Superdex75 200 mL columns. The purity of the protein was confirmed by electrophoresis with SDS-page gels. The protein was concentrated with Vivaspin 20 5 kDa MW cut-off centrifuge tubes (GE Healthcare), filtered with FILTSTAR hydrophilic nylon filters 0.22 μm cut-off and stored at -80 °C.

UV spectroscopy

UV spectra were recorded with a Shimadzu UV-1650PC spectrophotometer using a quartz cuvette of 0.1 cm path length for far-UV spectroscopy. The extinction coefficient at 280 nm (ϵ_{280}) of RGS WT and RGS L106R in TDW was 0.936 $\text{M}^{-1}\text{cm}^{-1}$, as computed by ExPASy Protparam. The extinction coefficient at 360 nm (ϵ_{360}) of ANS⁷³ in TDW was 5700 $\text{M}^{-1}\text{cm}^{-1}$.

Circular dichroism

CD spectra of RGS WT and RGS L106R were recorded with a J-810 spectropolarimeter (JASCO) equipped with a Peltier thermostat and with use of the supplied Spectra Manager software, in a 0.1 cm path length quartz cuvette for far-UV CD spectroscopy, in a spectral range comprised between 190 and 260 nm. The CD time course aggregation kinetics of RGS WT and RGS L106R was monitored by acquiring CD spectra at 25 °C, at time 0 or at the following time points: 15 min, 30 min, 45 min, 1 h, 1 h 30 min, 2 h, 3 h, 4 h, 6 h, 12 h and 24 h. The CD melting curves of RGS WT and RGS L106R were acquired by measuring the CD at 222 nm with a temperature gradient of 5 °C/min from 4 to 85 °C. In all conditions, the CD solutions were 8 μM RGS WT or RGS L106R in 50 mM Tris HCl pH 7.4, 300 mM NaCl, 5 mM β -mercaptoethanol.

ANS fluorescence, batch

300 μL of a solution 20 μM RGS WT or RGS L106R in 50 mM Tris HCl pH 7.4, 150 mM NaCl, 5 mM β -mercaptoethanol, 4 °C, were diluted with additions to an ANS solution in the same buffer conditions to reach protein:ANS molar ratios comprised between 1:1 and 1:12. Fluorescence spectra were recorded with a Perkin Helmer LS 55 fluorimeter using a quartz cuvette of 1 cm path length for far-UV fluorescence spectroscopy. The solution was excited at 380 nm and the emission spectra were recorded between 400 and 600 nm.

ANS fluorescence, high throughput

25 μ L of 32 μ M RGS WT or RGS L106R were added to 75 μ L of aggregating solution including the peptides at different concentrations. The solution was pipetted 10 times to obtain homogeneity. In all conditions, the solutions were 8 μ M RGS WT or RGS L106R in 50 mM Tris HCl pH 7.4, 150 mM NaCl, 5 mM β -mercaptoethanol, 24 μ M ANS or ThT, 25 °C. The aggregation kinetics were acquired with a BioTek Synergy H1 Hybrid Reader (Thermo Scientific) plate reader using 96 wells half-area plates (Costar 3696) for 100 μ L volume solutions. ANS was excited at 380 nm and the emission collected at 470 nm.

Dynamic light scattering

The DLS time course aggregation kinetics of RGS WT and RGS L106R was monitored by acquiring the DLS at 25 °C at time 0 or at the following time points: 15 min, 30 min, 45 min, 1 h, 1 h 30 min, 2 h, 3 h, 4 h, 6 h, 12 h and 24 h. In all conditions, the DLS solutions were 24 μ M RGS WT or RGS L106R in 50 mM Tris HCl pH 7.4, 300 mM NaCl, 5 mM β -mercaptoethanol. DLS measurement was performed with a Malvern Zetasizer Nano ZS instrument using disposable plastics cuvettes of 1 cm path length.

Transmission electron microscopy

RGS L106R was incubated at a concentration of 8 μ M in 50 mM Tris HCl pH 7.4, 150 mM NaCl, 5 mM β -mercaptoethanol at 25°C for 12 hours. The aggregation reaction solution was diluted 10x in the same buffer prior to sample preparation. A drop of 3-5 μ L sample was applied to a glow discharged TEM grid (carbon supported film on 300 mesh Cu grids, Ted Pella, Ltd.). After 30 sec the excess liquid was blotted, the grids were stained with 2% uranyl acetate for 30-60 sec, blotted and allowed to dry in air. The samples were examined using FEI Tecnai 12 G2 TWIN TEM operated at 120 kV. The images were recorded by a 4K x 4K FEI Eagle CCD camera.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We are grateful to Prof. Madelon Maurice of the Utrecht Medical Center and Dr. Yael Levi-Kalisman of the Center for Nanoscience and Nanotechnology at the Hebrew University of Jerusalem for their help and collaboration. This work was supported by the Innovative

Training Network 608180 “WntsApp” within by Marie-Curie Actions of the 7th Framework program of the EU.

Notes and references

- 1 M. S. Hipp, P. Kasturi and F. U. Hartl, *Nat. Rev. Mol. Cell Biol.*, 2019, **20**, 421–435.
- 2 E. W. J. Wallace, J. L. Kear-Scott, E. V. Pilipenko, M. H. Schwartz, P. R. Laskowski, A. E. Rojek, C. D. Katanski, J. A. Riback, M. F. Dion, A. M. Franks, E. M. Airoidi, T. Pan, B. A. Budnik and D. A. Drummond, *Cell*, 2015, **162**, 1286–1298.
- 3 J. C. Boatz, M. J. Whitley, M. Li, A. M. Gronenborn and P. C. A. Van Der Wel, *Nat. Commun.*, 2017, **8**, 1–10.
- 4 Y. Yoshimura, Y. Lin, H. Yagi, Y. H. Lee, H. Kitayama, K. Sakurai, M. So, H. Ogi, H. Naiki and Y. Goto, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 14446–14451.
- 5 F. Chiti and C. M. Dobson, *Annu. Rev. Biochem.*, 2017, **86**, 27–68.
- 6 F. Bemporad and F. Chiti, *Chem. Biol.*, 2012, **19**, 315–327.
- 7 A. P. D. Ano Bom, L. P. Rangel, D. C. F. Costa, G. A. P. De Oliveira, D. Sanches, C. A. Braga, L. M. Gava, C. H. I. Ramos, A. O. T. Cepeda, A. C. Stumbo, C. V. De Moura Gallo, Y. Cordeiros and J. L. Silva, *J. Biol. Chem.*, 2012, **287**, 28152–28162.
- 8 J. Xu, J. Reumers, J. R. Couceiro, F. De Smet, R. Gallardo, S. Rudyak, A. Cornelis, J. Rozenski, A. Zwolinska, J. C. Marine, D. Lambrechts, Y. A. Suh, F. Rousseau and J. Schymkowitz, *Nat. Chem. Biol.*, 2011, **7**, 285–295.
- 9 X. Song, S. Wang and L. Li, *Protein Cell*, 2014, **5**, 186–193.
- 10 Z. Anvarian, H. Nojima, E. C. Van Kappel, T. Madl, M. Spit, M. Viertler, I. Jordens, T. Y. Low, R. C. Van Scherpenzeel, I. Kuper, K. Richter, A. J. R. Heck, R. Boelens, J. P. Vincent, S. G. D. Rüdiger and M. M. Maurice, *Nat. Struct. Mol. Biol.*, 2016, **23**, 324–332.
- 11 R. J. Ellis and A. P. Minton, *Nature*, 2003, **425**, 27–28.
- 12 P. H. Yancey and J. F. Siebenaller, *J. Exp. Biol.*, 2015, **218**, 1880–1896.
- 13 S. B. Zimmerman and S. O. Trach, *J. Mol. Biol.*, 1991, **222**, 599–620.
- 14 M. C. Konopka, J. C. Weisshaar and M. T. Record, *Method. Enzymol.*, 2007, **428**, 487–504.
- 15 D. Harries and J. Rösger, *Methods Cell Biol.*, 2008, **84**, 679–735.
- 16 H.-X. Zhou, G. Rivas and A. P. Minton, *Annu. Rev. Biophys.*, 2008, **37**, 375–397.
- 17 M. Gruebele, K. Dave and S. Sukenik, *Annu. Rev. Biophys.*, 2016, **45**, 233–251.
- 18 R. D. Cohen and G. J. Pielak, *Protein Sci.*, 2017, **26**, 1698–1703.
- 19 M. Y. Bhat, L. R. Singh and T. A. Dar, eds. L. Rajendrakumar Singh and T. A. Dar, Springer

Singapore, Singapore, 2017, pp. 121–142.

- 20 S. Sukenik, R. Politi, L. Ziserman, D. Danino, A. Friedler and D. Harries, *PLoS One*, 2011, **6**, e15608.
- 21 S. Saha, A. Sharma and S. Deep, *RSC Adv.*, 2016, **6**, 28640–28652.
- 22 R. Liu, H. Barkhordarian, S. Emadi, B. P. Chan and M. R. Sierks, *Neurobiol. Dis.*, 2005, **20**, 74–81.
- 23 N. Katyal, M. Agarwal, R. Sen, V. Kumar and S. Deep, *ACS Chem. Neurosci.*, 2018, **9**, 1477–1491.
- 24 M. V. Khan, M. Ishtikhar, G. Rabbani, M. Zaman, A. S. Abdelhameed and R. H. Khan, *Int. J. Biol. Macromol.*, 2017, **94**, 290–300.
- 25 L. Nicoud, N. Cohrs, P. Arosio, E. Norrant and M. Morbidelli, *Biophys. Chem.*, 2015, **197**, 40–46.
- 26 O. O. Oyetayo, O. Méndez-Lucio, A. Bender and H. Kiefer, *Int. J. Pharm.*, 2017, **530**, 165–172.
- 27 C. M. Sudrik, T. Cloutier, N. Mody, H. A. Sathish and B. L. Trout, *Pharm. Res.*, 2019, **36**, 1–12.
- 28 T. Borwankar, C. Röthlein, G. Zhang, A. Tegen, C. Dosche and Z. Ignatova, *Biochemistry*, 2011, **50**, 2048–2060.
- 29 S. Choudhary, N. Kishore and R. V. Hosur, *Sci. Rep.*, 2015, **5**, 1–10.
- 30 T. B. Eronina, N. A. Chebotareva, S. G. Bazhina, V. F. Makeeva, S. Y. Kleymenov and B. I. Kurganov, *Biophys. Chem.*, 2009, **141**, 66–74.
- 31 S. Mittal, R. K. Chowhan and L. R. Singh, *Biochim. Biophys. Acta - Gen. Subj.*, 2015, **1850**, 1822–1831.
- 32 S. Sukenik and D. Harries, *Prion*, 2012, **6**, 26–31.
- 33 S. Ray, N. Singh, R. Kumar, K. Patel, S. Pandey, D. Datta, J. Mahato, R. Panigrahi, A. Navalkar, S. Mehra, L. Gadhe, D. Chatterjee, A. S. Sawner, S. Maiti, S. Bhatia, J. A. Gerez, A. Chowdhury, A. Kumar, R. Padinhateeri, R. Riek, G. Krishnamoorthy and S. K. Maji, *Nat. Chem.*, 2020, **12**, 705–716.
- 34 L. A. Munishkina, A. Ahmad, A. L. Fink and V. N. Uversky, *Biochemistry*, 2008, **47**, 8993–9006.
- 35 I. M. Kuznetsova, K. K. Turoverov and V. N. Uversky, *Int. J. Mol. Sci.*, 2014, **15**, 23090–23140.
- 36 J. Martin and F. U. Hartl, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 1107–1112.
- 37 G. Rivas, J. A. Fernández and A. P. Minton, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 3150–3155.

- 38 F. A. Ferrone and M. A. Rotter, *J. Mol. Recognit.*, 2004, **17**, 497–504.
- 39 S. G. Roman, N. A. Chebotareva, T. B. Eronina, S. Y. Kleymenov, V. F. Makeeva, N. B. Poliansky, K. O. Muranov and B. I. Kurganov, *Biochemistry*, 2011, **50**, 10607–10623.
- 40 I. Martín, G. Celaya, C. Alfonso, F. Moro, G. Rivas and A. Muga, *Biophys. J.*, 2014, **106**, 2017–2027.
- 41 M. V. Liberti and J. W. Locasale, *Trends Biochem. Sci.*, 2016, **41**, 211–218.
- 42 K. E. Spink, P. Polakis and W. I. Weis, *EMBO J.*, 2000, **19**, 2270–2279.
- 43 M. Cardamone and N. K. Puri, *Biochem. J.*, 1992, **282**, 589–593.
- 44 A. Hawe, M. Sutter and W. Jiskoot, *Pharm. Res.*, 2008, **25**, 1487–1499.
- 45 B. Kundu and P. Guptasarma, *Biochem. Biophys. Res. Commun.*, 2002, **293**, 572–577.
- 46 Z. Qin, D. Hu, M. Zhu and A. L. Fink, *Biochemistry*, 2007, **46**, 3521–3531.
- 47 V. Vetri, C. Canale, A. Relini, F. Librizzi, V. Militello, A. Gliozzi and M. Leone, *Biophys. Chem.*, 2007, **125**, 184–190.
- 48 M. Lindgren, K. Sörgjerd and P. Hammarström, *Biophys. J.*, 2005, **88**, 4200–4212.
- 49 C. A. Species, B. Bolognesi, J. R. Kumita, T. P. Barros, E. K. Esbjorner, L. M. Luheshi, D. C. Crowther, M. R. Wilson, C. M. Dobson, G. Favrin, J. J. Yerbury, E. K. Esbjörner and M. Leila, *ACS Chem. Biol.*, 2010, **5**, 735–740.
- 50 N. D. Younan and J. H. Viles, *Biochemistry*, 2015, **54**, 4297–4306.
- 51 B. Mannini, E. Mulvihill, C. Sgromo, R. Cascella, R. Khodarahmi, M. Ramazzotti, C. M. Dobson, C. Cecchi and F. Chiti, *ACS Chem. Biol.*, 2014, **9**, 2309–2317.
- 52 J. M. Finke and P. A. Jennings, *J. Biol. Phys.*, 2001, **27**, 119–131.
- 53 V. Vetri and V. Militello, *Biophys. Chem.*, 2005, **113**, 83–91.
- 54 G. V. Semisotnov, N. A. Rodionova, O. I. Razgulyaev, V. N. Uversky, A. F. Gripas' and R. I. Gilmanshin, *Biopolymers*, 1991, **31**, 119–128.
- 55 V. N. Uversky, S. Winter and G. Löber, *Biophys. Chem.*, 1996, **60**, 79–88.
- 56 S. Kumar, S. K. Mohanty and J. B. Udgaonkar, *J. Mol. Biol.*, 2007, **367**, 1186–1204.
- 57 C. M. Eakin, F. J. Attenello, C. J. Morgan and A. D. Miranker, *Biochemistry*, 2004, **43**, 7808–7815.
- 58 G. Plakoutsi, F. Bemporad, M. Calamai, N. Taddei, C. M. Dobson and F. Chiti, *J. Mol. Biol.*, 2005, **351**, 910–922.
- 59 S. Sambashivan, Y. Liu, M. R. Sawaya, M. Gingery and D. Eisenberg, *Nature*, 2005, **437**, 266–269.
- 60 N. J. Cobb, A. C. Apetri and W. K. Surewicz, *J. Biol. Chem.*, 2008, **283**, 34704–34711.
- 61 R. P. Honda, M. Xu, K. I. Yamaguchi, H. Roder and K. Kuwata, *Structure*, 2015, **23**, 1735–1742.

- 62 M. R. Ho, Y. C. Lou, W. C. Lin, P. C. Lyu, W. N. Huang and C. Chen, *J. Biol. Chem.*, 2006, **281**, 33566–33576.
- 63 L. Pieri, M. Bucciantini, D. Nosi, L. Formigli, J. Savistchenko, R. Melki and M. Stefani, *J. Biol. Chem.*, **281**, 15337–15344.
- 64 K. Kar and N. Kishore, *Biopolymers*, 2007, **87**, 339–351.
- 65 Y. Qu, C. L. Bolen and D. W. Bolen, *Proc. Natl. Acad. Sci.*, 2002, **95**, 9268–9273.
- 66 P. R. Davis-Searles, A. S. Morar, A. J. Saunders, D. A. Erie and G. J. Pielak, *Biochemistry*, 1998, **37**, 17048–17053.
- 67 T. Kamiyama, Y. Sadahide, Y. Nogusa and K. Gekko, *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.*, 1999, **1434**, 44–57.
- 68 B. M. Stojanovski, L. Breydo, V. N. Uversky and G. C. Ferreira, *RSC Adv.*, 2016, **6**, 114541–114552.
- 69 S. Sukenik, L. Sapir and D. Harries, *J. Chem. Theory Comput.*, 2015, **11**, 5918–5928.
- 70 Z. A. Parray, S. Shahid, F. Ahmad, M. I. Hassan and A. Islam, *Int. J. Biol. Macromol.*, 2017, **99**, 241–248.
- 71 K. Nasreen, S. Ahamad, F. Ahmad, M. I. Hassan and A. Islam, *Int. J. Biol. Macromol.*, 2018, **106**, 130–139.
- 72 D. B. Knowles, I. A. Shkel, N. M. Phan, M. Sternke, E. Lingeman, X. Cheng, L. Cheng, K. O'Connor and M. T. Record, *Biochemistry*, 2015, **54**, 3528–3542.
- 73 A. Stevens and R. C. Augusteyn, *Eur. J. Biochem.*, 1997, **243**, 792–797.