Molecular insights on the mechanism of α_1-antitrypsin condensate formation and maturation

Ignacio Sanchez-Burgos,[1] Andres R. Tejedor,[2,3] Rosana Collepardo-Guevara,[2] Jorge Bernardino de la Serna,[4] and Jorge R. Espinosa*[2,3,a)]

[1)] *Maxwell Centre, Cavendish Laboratory, Department of Physics, University of Cambridge, J J Thomson Avenue, Cambridge CB3 0HE, United Kingdom.*

[2)] *Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom.*

[3)] *Department of Physical-Chemistry, Universidad Complutense de Madrid, Av. Complutense s/n, 28040, Madrid, Spain.*

[4)] *Faculty of Medicine, Imperial College London, Ayrton Rd, South Kensington, London SW7 5NH, United Kingdom.*

(Dated: March 11, 2025)

The deficiency of α_1-antitrypsin protein is a genetic disorder characterized by the accumulation of misfolded protein aggregates within hepatocytes, leading to liver dysfunction. In the lung, it is found in macrophages, bronchial and epithelial alveolar cells type 2, leading to pulmonary emphysema. Despite extensive research, the precise mechanism underlying the formation of α_1-antitrypsin inclusion bodies remain elusive. In this study, we combine equilibrium and non-equilibrium molecular dynamics simulations to elucidate the intricate process of α_1-antitrypsin condensate formation and maturation. Our mechanistic model explains cluster accumulation—specifically the onset of this pathogenesis—through the emergence of phase-separated liquid-like protein droplets, which subsequently undergo inter-protein β-sheet transitions between misfolded variants, resulting in solid-like clusters. We find that this mechanism only applies to the misfolded variant, Z-α_1-antitrypsin, which phase-separates driven by its disordered C-terminus. In contrast, the native protein, M-α_1-antitrypsin, shows much lower propensity to phase-separate and form kinetically trapped aggregates. Furthermore, we explore how Z-α_1-antitrypsin exhibits an increased capacity to form condensates near external walls with different types of interactions. Such conditions can be similar to those found within the endoplasmic reticulum membrane, where phase separation and hardening take place. Overall, our results shed light on the molecular basis of α_1-antitrypsin-related disorders and provide valuable microscopic insights for the development of therapeutic strategies targeting protein misfolding and aggregation-related disorders.

I. INTRODUCTION

In living organisms, there is a crucial regulatory mechanism responsible for maintaining the correct balance of cellular proteins, known as proteostasis[1]. Disruptions in this system have been linked to the development of several common human diseases[2] such as cystic fibrosis[3] or type 2 diabetes[4]. At the heart of this regulatory network lies the intricate process of protein folding, where newly synthesized polypeptide sequences must fold correctly to attain their functional three-dimensional structure[5]. It is widely recognized that protein folding can occur during the synthesis process, which aids in the efficient achievement of functional structure. However, protein folding is also susceptible to errors, as evidenced by studies showing that misfolding can occur during protein synthesis, leading to degradation of aberrantly folded sequences[6–8].

α_1-antitrypsin, a glycoprotein crucial for regulating protease activity in the human body[9], is implicated in the pathology of α_1-antitrypsin deficiency (AATD), a genetic disorder[10] that leads to liver dysfunction and pulmonary emphysema. This protein is synthesized within hepatocytes and undergoes folding in the endoplasmic reticulum before being secreted into the bloodstream, where it primarily functions as an inhibitor of neutrophil elastase in the lungs[9,10]. While its correct folding has been shown to be essential for its functional activity[11], mutations in the α_1-antitrypsin gene lead to incorrect folding of the sequence, resulting in the formation of insoluble bodies within hepatocytes, which leads to AATD pathogenesis[10,12–14]. To elucidate the mechanisms underlying α_1-antitrypsin misfolding, researchers have conducted detailed investigations into the folding dynamics of nascent polypeptide sequences as they emerge from the ribosome during the synthesis. These studies have revealed that nascent α_1-antitrypsin can exhibit different folding behaviors in the ribosome, forming intermediates with molten-globule characteristics, where the C-terminus appears to be misfolded[14–16]. While structural analyses have provided major insights into the differences in folding kinetics between wild-type and mutant α_1-antitrypsin variants[10,17], shedding light on the precise origin of misfolding and their implications in health and disease remains a challenge.

The wild-type α_1-antitrypsin sequence, termed M-α_1-antitrypsin, rapidly folds and carries out the biological function of regulating the activity of proteases[11,14]. However, the mutated Z-α_1-antitrypsin variant contains a glutamic acid to lysine mutation in the C-terminal region (E342K), which impairs the protein's ability to at-

[a)] Electronic mail: jorgerene@ucm.es

tain its functional structure[16]. This variant has been identified as the principal cause of α_1-antitrypsin deficiency (AATD)[10,12,13,16]. In the original protein sequence, Glu_{342} forms a salt bridge to Lys_{290}, which is not present in the Z variant, and is thought to be the reason behind the delay in the sequence folding[13]. Since aberrant aggregates formed by the Z variant of α_1-antitrypsin resemble those observed in protein condensate-related neurodegenerative disorders[18–21], here we apply a computational approach developed by us to explore the protein phase behavior at submolecular level from individual sequences to higher-order assemblies[22–24].

We perform Molecular Dynamics (MD) simulations with a sequence-dependent coarse-grained protein model[25] to evaluate the phase diagrams of the M and Z variants, providing insights into their thermodynamic coexistence lines and intermolecular liquid network connectivity. Furthermore, to decipher the clustering mechanism of Z-α_1-antitrypsin, we model how different protein replicas can dynamically establish inter-protein β-sheet structures that lead to irreversible aggregation. This methodology relies on modeling inter-protein structural transitions, which give rise to kinetically arrested and highly insoluble phases similar to those observed in AATD. We additionally measure the viscosity of these protein assemblies, before and after the formation of intermolecular cross-β-sheet structures. We investigate the interplay between intra- and inter-protein structural changes, which is crucial for understanding the competition between correct folding and the formation of potentially insoluble bodies. Lastly, we simulate how the presence of a surface regulates the propensity of α_1-antitrypsin to form condensates via liquid-liquid phase separation (LLPS). Overall, our comprehensive computational approach provides valuable information on the phase behavior, viscosity, and maturation mechanisms of the M and Z-α_1-antitrypsin protein variants.

II. RESULTS

A. Phase separation propensity of M and Z-α_1-antitrypsin variants

To model the different α_1-antitrypsin variants, we employ the residue-level coarse-grained force field CALVADOS2[25]. Within this implicit-solvent model, proteins are simulated as polymers in which each bead represents a given amino acid with its own chemical identity. The CALVADOS2 model accurately predicts the conformational properties of intrinsically disordered proteins and propensities to undergo LLPS for diverse sequences and solution conditions. The intermolecular potential describing the interactions between all pairs of amino acids in the CALVADOS2 is detailed in the Supplementary Material (SM) Section I. To account for the structured domains of the protein, we preserve the relative positions of the residues (centered on the C_αs of each amino acid) that belong to the experimentally reported structured domains (PDB code: 3NE4 for both M and Z variants) with a rigid body integrator. In contrast, intrinsically disordered regions are simulated as flexible polymers using a harmonic bond potential. We use the LAMMPS MD package[26] to perform our simulations. Moreover, to mimic the emergence of inter-protein structured domains, we employ our previously developed aging algorithm[22,23,27], which dynamically evaluates high-density local fluctuations of protein domains (i.e., low-complexity aromatic-rich segments; LARKS[19,28]) which are prone to develop inter-protein β-sheet structures. Under these conditions, disorder-to-order structural transitions are susceptible to occur, and the intermolecular interactions of the involved residues are scaled, accounting for the stronger interaction associated with the formation of inter-protein β-sheets[24,29,30] (further numerical details on the algorithm and the precise simulation technical details and system sizes are described in the SM Section SII).

It has recently been suggested that the C-terminus of the Z-α_1-antitrypsin sequence can behave as an intrinsically disordered region[16] in contrast to the native M-α_1-antitrypsin protein, which contains multiple β-sheet structures across the C-terminus that are intramolecularly folded. In Figure 1(a), we sketch the different structural motifs of α_1-antitrypsin: (i) intrinsically disordered regions depicted by straight lines; (ii) α-helix structures by ovals; and (iii) β-sheet domains indicated by arrows. Dashed lines show the regions of the sequence that rapidly fold in the M variant, while remaining disordered in the Z variant. As can be seen, as a consequence of the (E342K) mutation, there is a segment of the sequence that, if not properly folded, may have the potential to promote phase-separation, and as a consequence, form inter-protein β-sheets instead of intramolecular folding. This domain (highlighted with dashed lines in Figure 1(a), and visibly disordered as indicated in Figure 1(b)) could be susceptible to promote LLPS among misfolded variants, and may explain the physicochemical impact of such mutation in the protein phase behavior leading to AATD.

To test this hypothesis, we build both variants using the high-resolution crystal structure (1.8 Å; PDB code[31]: 3NE4) resolved for M-α_1-antitrypsin. We employ the CALVADOS2[25] force field to model the proteins, and the LAMMPS rigid body integrator[26] to maintain the secondary and tertiary structures of the protein (see SM Section I for further details on the sequence implementation). The key difference between the Z and M variants lies in the substitution of glutamic acid to lysine (E342K mutation) which reportedly delays the correct folding of the C-terminal disordered region (Fig. 1(b)). We first perform Direct Coexistence (DC) simulations[32,33], in which a condensed and diluted protein phases cohabit within an elongated box. We employ DC simulations to compute the phase diagram in the temperature–density plane (see further details on DC

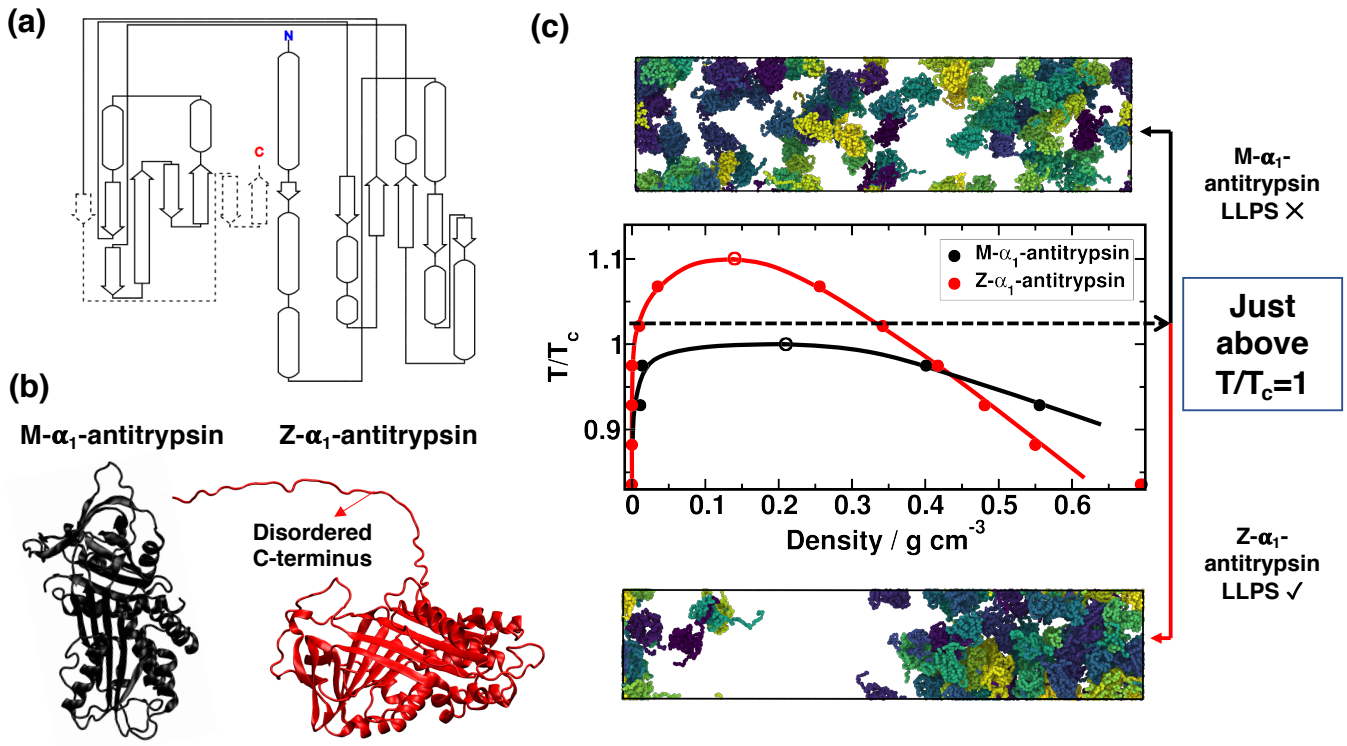

Figure 1: (a) Full sequence representation of α_1-antitrypsin illustrating the intrinsically disordered regions with straight lines, the α-helix structures with ovals, and β-sheets with arrows. The dashed region has no secondary structure in Z-α_1-antitrypsin. (b) Rendered images of the M and Z variants of α_1-antitrypsin employed in our simulations. (c) Phase diagram in the temperature (normalized by the critical temperature of the M variant) density plane for the M and Z α_1-antitrypsin sequences, along with representative snapshots at $T/T_c \sim 1.03$, where Z-α_1-antitrypsin displays a condensate coexisting with a diluted phase (Bottom), while M-α_1-antitrypsin cannot undergo phase-separation (Top).

calculations in Section SIII of the SM). Through this approach, we first determine the stability limits of both α_1-antitrypsin variants. In Figure 1(c), the coexistence lines of both α_1-antitrypsin sequences are shown, demonstrating the greater ability of Z-α_1-antitrypsin to undergo LLPS and form protein condensates. We also include in Figure 1(c) rendered images of DC simulations for the M and Z variants at a temperature just above the critical one (T_c) for the M variant, showing that the native form (Top panel) cannot form phase-separated condensates under these conditions while the Z variant (Bottom panel) undergoes LLPS. Importantly, higher critical solution temperatures have been proven to be associated with lower saturation concentrations in multiple *in silico* and *in vitro* studies of protein LLPS[34–39], which strongly indicates that Z-α_1-antitrypsin will form condensates (which could later potentially transition into pathological solid-like assemblies) at conditions at which the wild-type M variant would remain soluble. Moreover, the higher critical solution temperature for Z-α_1-antitrypsin is consistent with the fact that intrinsically disordered regions enhance the ability to form condensates in many other proteins[40–43] such as hnRNPA1, TDP-43 or FUS, where their low-complexity domains are crucial for enabling phase-separation[40,44–47]. In the case of α_1-antitrypsin, the mutated Z variant contains a disordered region (e.g., the C-terminus), that unless correctly folded into an intra-protein β-sheet structure[16], it promotes condensate formation under less favorable conditions (i.e., at higher temperatures or at lower protein saturation concentrations).

B. The disordered C-terminus of the Z variant enables greater intermolecular connectivity

We now investigate the molecular origin behind the greater propensity of Z-α_1-antitrypsin to form condensates by computing the contact frequency map among all possible residue-residue interactions across the sequence. From our calculations shown in Fig. 1(c), we perform an analysis of the frequency of pairwise intermolecular contacts established by both α_1-antitrypsin variants. To that goal, we examine the condensates formed at $T/T_c \simeq 0.97$, temperature at which the density of the condensates formed by both variants is similar (see Figure 1(c)). We consider an 'effective' intermolecular contact when two amino acids are closer than $1.2\sigma_{ij}$, being

$\sigma_{ij} = (\sigma_i + \sigma_j)/2$, where σ_i and σ_j represent the molecular diameter of the two residues involved in such interaction as proposed in Refs.[23,34].

In Figure 2 we show the intermolecular contact maps for both variants, where darker colors depict higher frequencies of pairwise residue-residue interactions. The average frequency of intermolecular contacts is expressed as a percentage, where 100% refers to a contact which persists across all the sampled configurations. Both variants show a similar pattern of intermolecular interactions until the 335th residue, where the C-terminus starts. Interestingly, domains from the 30th residue to the 45th, from the 102nd to the 110th, and from the 193rd to the 200th, which are enriched in charged and polar amino acids and remain intrinsically disordered, establish more frequent interactions than the rest of the sequence which is mostly structured. This is a common feature of multi-domain phase-separating proteins that combine both disordered and structured domains (through secondary and tertiary interactions) as previously reported for hnRNPA1[46,48,49] or TDP-43[50,51]. Remarkably, the major difference between the contact frequency maps is observed in the C-terminus of both variants, differentiated from the rest of the sequence with dashed lines (Figure 2). As detailed in Section II A, this region contains the E342K mutation in the Z variant which hinders intramolecular folding and thus, can establish much more frequent intermolecular contacts with other protein domains across the sequence; particularly between C-terminal domains, representing the key driver region inducing phase-separation.

C. Inter-protein β-sheet transitions promote condensate hardening

High-density local concentrations of disordered domains within protein condensates have previously been shown to induce the emergence of inter-protein β-sheets via structural transitions of low-complexity aromatic rich kinked segments[19,23,24,28,52]. While the folding of the M variant leads to the formation of intramolecular β-sheet structures, the disordered C-terminus of the Z intermediate variant has the potential to establish both intra- and inter-protein β-sheets within a condensed phase. Hence, after confirming the greater propensity of Z-α_1-antitrypsin to undergo LLPS compared to its native form (Fig. 1(c)), we now perform non-equilibrium simulations (using our dynamic algorithm[22–24]; see SM Section II for further technical details on the algorithm), so that inter-protein β-sheets in the relevant sequence segments (dashed arrows in Figure 1(a)) can be formed either at intramolecular level (i.e., accomplishing the correct folding of the protein), or between different protein replicas establishing inter-protein β-sheets. Through the original dynamic algorithm[22,53], we approximate β-sheet formation via disorder-to-order structural transitions by considering the atomistic implications (i.e., non-conservative strengthening of inter-protein binding, local protein rigidification, and changes in the intermolecular organization;[19,24,29]) of the gradual and irreversible accumulation of β-sheet structures in a time-dependent manner and as a function of the local protein density. Here, we adapt our dynamic algorithm[22,23] to describe both intra- and inter-protein β-sheet transitions within condensates for evaluating the interplay between both folding phenomena. We perform these calculations at T/T_c =1.02, conditions at which M-α_1-antitrypsin does not undergo LLPS, and hence, only condensate formation of Z-α_1-antitrypsin occurs as experimentally found[16].

In Figure 3(a) we plot the total number of β-sheets formed, distinguishing between those formed within the same protein (i.e., at intramolecular level; black solid line) and between different protein replicas (at intermolecular level; red solid line). Strikingly, we observe a significantly larger amount of the latter, giving rise to local high-density protein clusters in which the different proteins are strongly engaged to each other through cross-β-sheet assemblies (as illustrated in Figure 3(b)). In Figure 3(b)(right) we show a primitive path analysis (PPA)[23,24] of the formed cross-β-sheet network to visualize the intermolecular connectivity across the condensate. For this calculation, we enforce that: (1) β-sheet domains formed across the C-terminus are fixed in space; (2) the intramolecular excluded volume is set to zero; and (3) the bond interaction has an equilibrium bond length of 0 nm. In this manner, the PPA algorithm minimizes the contour length of the protein strands that connect the different β-sheet domains in the C-terminal region, while preserving the topology of the underlying inter-protein network[54], and allows for visualization of the connectivity generated by β-sheet clusters, which are mostly inter-protein arrays. Overall, our non-equilibrium simulations predict the ability of the Z-α_1-antitrypsin variant to form phase-separated condensates, and subsequently transition into highly insoluble states through the stabilization of cross-β-sheet structures. Importantly, our simulations are performed under conditions at which the native form, M-α_1-antitrypsin, does not form condensates, further validating our hypothesis since M-α_1-antitrypsin has not been found in aberrant solid-like assemblies of AATD patients[13,55].

Condensate maturation can lead to high viscous liquids or even solid-like states as the proteins within the condensates lose partial motility over time[40,56,57]. Such kinetically arrested dynamics is often associated with the emergence of condensate dysfunctional behavior, and has been widely linked to multiple neurological disorders[58,59]. Thus, we measure the viscosity of Z-α_1-antitrypsin condensates (just after their formation, and after an incubation time of 200 ns) in order to assess the change in the droplet's material properties over time due to the formation of inter-protein β-sheet structures. We note that the coarse-grained nature of our implicit-solvent model significantly underestimates the relaxation timescale of the proteins, and hence the viscosity of the condensates[34]; however, the observed trends and relative differences

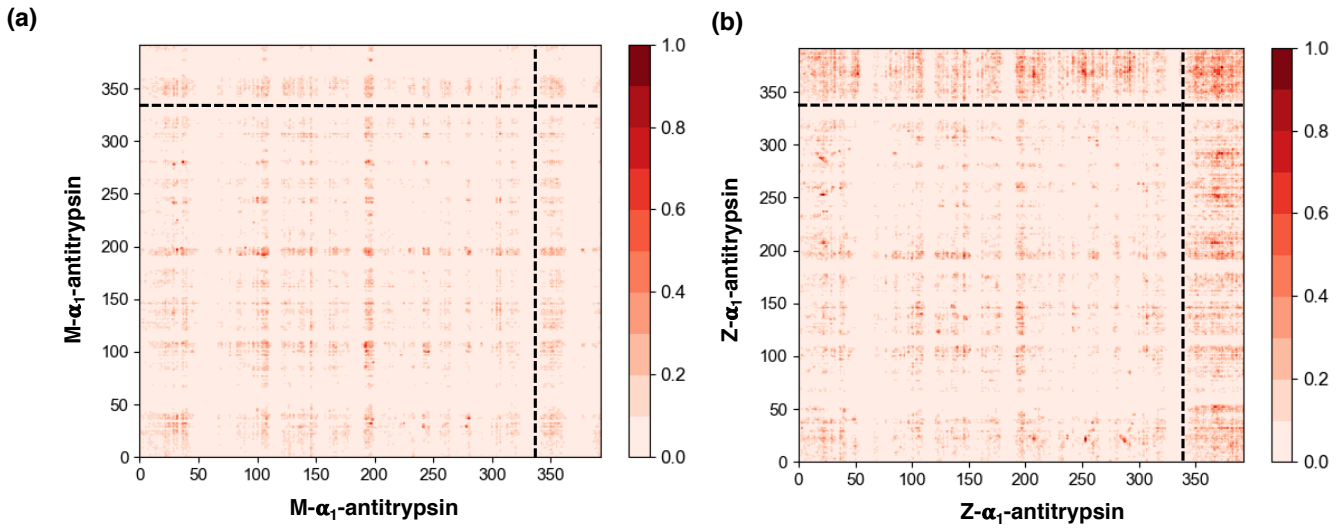

Figure 2: Average frequency of intermolecular contacts (expressed in %) per residue for (a) M-α_1-antitrypsin and (b) Z-α_1-antitrypsin. This analysis is performed over DC simulations (Figure 1(c)) at $T/T_c \simeq 0.97$. The black dashed line in both images indicates the region corresponding to the C-terminus, which remains disordered in the Z-α_1-antitrypsin mutant, and folded in the M-α_1-antitrypsin variant.

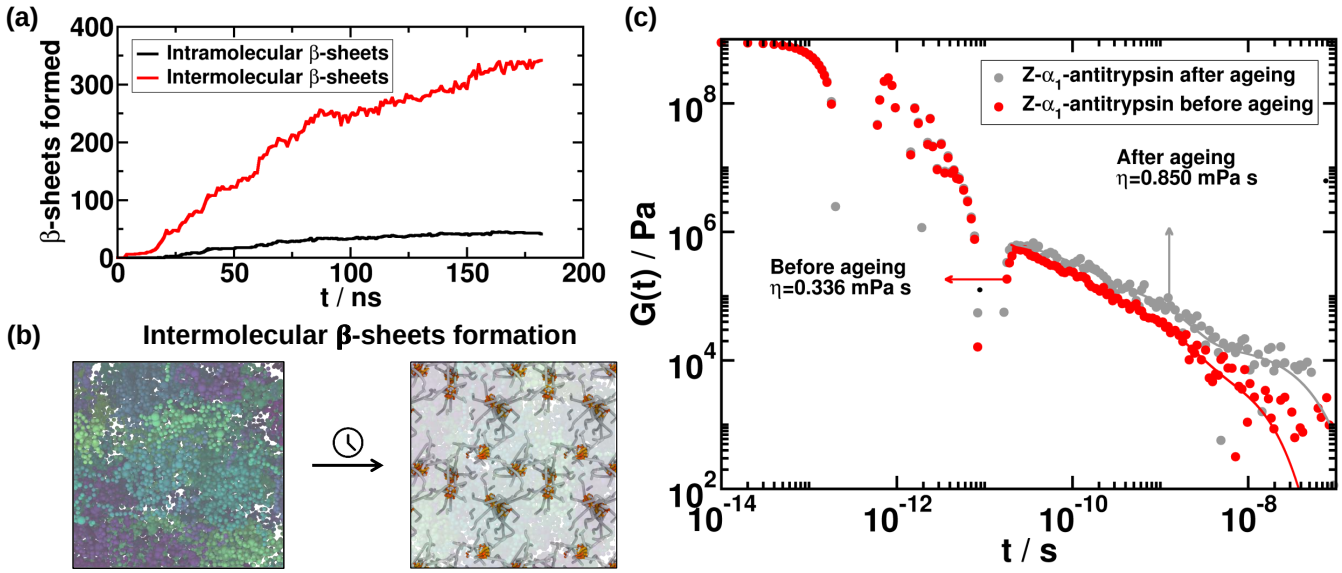

Figure 3: (a) Time-evolution of the different β-sheet structures formed (at inter- and intramolecular level) in Z-α_1-antitrypsin condensates at $T/T_c = 1.02$ (referred T_c to that of the M-α_1-antitrypsin variant) and the corresponding condensate equilibrium density at such temperature. (b) Primitive path analysis calculations showing the formation of inter-protein β-sheet structures among the C-terminus of the Z-α_1-antitrypsin variant in phase-separated condensates. (c) Shear stress relaxation modulus ($G(t)$) of the Z-α_1-antitrypsin bulk condensed phase at T/T_c=1.02, prior to the formation of any inter-protein structure (red points), and after condensate maturation where inter-protein β-sheets are formed (grey points). Continuous lines depict Maxwell mode fits to $G(t)$ data. From each set of points we obtain the viscosity for both systems (before and after condensate incubation time), by integrating in time the stress relaxation modulus.

in the viscoelastic properties between initially formed and incubated condensates are expected to hold despite the artificially faster dynamics of our residue-resolution simulations[23]. We compute the shear stress relaxation modulus ($G(t)$) in Z-α_1-antitrypsin bulk condensates prior (red curve) and post-maturation (gray curve in Fig. 3(c)) using separate NVT simulations, where we compute the auto-correlation function of any of the off-diagonal

components of the pressure tensor[60,61]. Since the system is isotropic, an accurate expression of $G(t)$ can be obtained by using the six independent components of the pressure tensor[62] (see section SIV in SM). The viscosity of the system is then computed by integrating the stress relaxation modulus along time, using the Green-Kubo formula[63]

$$\eta = \int_0^\infty dt\ G(t). \tag{1}$$

While at shorter timescales, $G(t)$ easily converges and can be integrated numerically, for longer timescales ($t > 10^{-10}$ s, Figure 3(c)) $G(t)$ needs to be fitted to a series of Maxwell modes ($G_i \exp(-t/\tau_i)$) equidistant in logarithmic time[60], resulting in a function that is integrated analytically to obtain the viscosity (see Ref.[61] or SM Section SIV for further details on these calculations).

In Figure 3(c), we show $G(t)$ as a function of time along with the fits to the Maxwell modes at long timescales (solid lines). From these curves, we obtain the viscosity, being 0.336 mPa·s for Z-α_1-antitrypsin before the formation of any inter-protein β-sheet structure, and 0.850 mPa·s for the incubated condensate. The shift in viscosity corresponds to the condensate maturation process of Z-α_1-antitrypsin, which gradually transforms into a high-viscous phase which is locally arrested and stabilized by inter-protein β-sheet clusters (Fig. 3(b)). The emergence of Z-α_1-antitrypsin foci has been consistently observed *in vivo* across multiple experimental studies[10,12,13,16]. In that respect, our simulations propose a mechanism that explains the experimentally observed formation of pathological Z-α_1-antitrypsin assemblies through: (1) condensate formation via LLPS; and (2) progressive accumulation of inter-protein β-sheets that, when forming an interconnected percolated network (Figure 3(b)), increase the viscosity of the condensates over time (Figure 3(c)) and prevents their spontaneous dissolution.

D. Heterogeneous nucleation of Z-α_1-antitrypsin condensates

Nucleation of Z-α_1-antitrypsin inclusion bodies naturally occurs within the confines of the endoplasmic reticulum (ER) membrane[16]. This observation suggests a possible mechanism by which the ER promotes heterogeneous condensate nucleation of Z-α_1-antitrypsin over its surface. In this section, we investigate condensate formation near a flat structureless surface, and therefore elucidate the effect of an external barrier on the regulation of α_1-antitrypsin LLPS.

As the specific intermolecular interactions responsible for protein self-assembly around a membrane are extremely challenging to describe with coarse-grained force fields[64,65], we adopt two different modeling approaches for mimicking the ER surface. First, we introduce a non-specific interaction between a flat structureless wall and α_1-antitrypsin using a soft attractive potential (see SM Section V for further details on the implementation). Essentially, the surface consists of a Lennard-Jones potential (with ϵ= 0.1 kcal mol^{-1} and σ= 15Å) inserted at the periphery of the simulation box (as illustrated in Figure 4(a)) which establishes a mild attractive region (of <0.5 k$_B$T) that mimics non-specific chemical adsorption onto the ER membrane. Conversely, we model a repulsive surface by employing a purely repulsive interaction described by a harmonic potential $E = K_H(r - r_c)^2$ for $r < r_c$. For the repulsive potential, the spring constant is set to K_H = 100 kcal mol^{-1} Å^{-2} and r_c = 4Å (see section SV of SM for further details). This results in a short-range repulsive structureless surface located at the edge of the simulation box, in an analogous arrangement to the attractive barrier depicted in Figure 4(a).

We perform DC simulations of the Z-α_1-antitrypsin variant well above its critical solution temperature (i.e., at $T/T_c = 1.1$ where T_c corresponds to the critical temperature for LLPS of the Z variant) in the presence of the moderately attractive and repulsive surfaces. At temperatures $T \geq T_c$, phase separation is unattainable, and therefore the emergence of condensates would only be a direct consequence of the presence of an external surface. We note that, due to the calculation setup, the simulation is not periodic along the elongated axis, as the surface remains uncrossable by the molecules. In Figure 4(b), we present density profiles along the long axis of the simulation box, starting from the edge where the attractive (or repulsive) surface is located. For the attractive surface (black curve), we observe a local density maximum induced by the external layer. Despite using a weak attractive interaction (i.e., lower than 0.5 k$_B$T), this density maximum indicates the presence of protein clusters with densities that correspond to those of condensates well below the critical solution temperature (please note that the critical density is ∼0.2 g/cm^3; Fig. 1(c)). These clusters exhibit significant local high-density fluctuations, which over time can induce the formation of inter-protein β-sheets, further increasing their stability and viscosity, and originating the characteristic features of insoluble condensates as described in Section II C and Fig. 3.

Additionally, for the simulation with a repulsive surface (depicted in Fig. 4(b) by the red curve), we observe a density maximum near the external layer, although less pronounced than for the moderately attractive wall (black curve). Although this maximum does not reach average densities exceeding 0.25 g cm^{-3}, our results suggest a significant propensity of Z-α_1-antitrypsin to form clusters on an external surface, despite lacking an explicit physicochemical affinity for the wall (i.e., attractive cross-interactions). This behavior can be attributed to the lower interfacial free energy that proteins may exhibit with the wall, compared to the surface tension between the condensate and the protein diluted phase. Under such circumstances, condensates would form on the surface through a classical heterogeneous nucleation mechanism[66,67]. In Ref.[68], a similar mechanism was re-

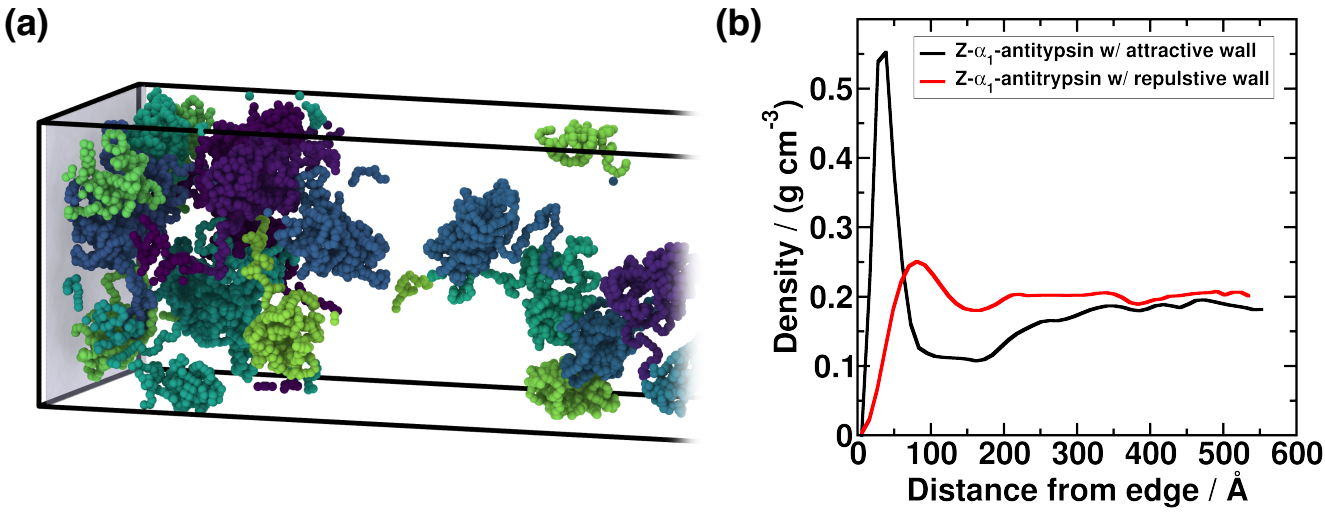

Figure 4: (a) Snapshot of a DC simulation of Z-α_1-antitrypsin at $T/T_c = 1.1$ (where T_c corresponds to the critical temperature for LLPS of the Z variant) in presence of an external surface indicated by the shaded plane at the edge of the simulation box. In the snapshot, each protein replica is represented in a different colour. (b) Averaged density profiles along the long axis of the DC simulation box, evaluated from the external surface. The nature of the inserted surface is indicated in the figure legend.

ported for hard-sphere colloids in the presence of purely repulsive, flat structureless walls, which favor the nucleation of crystalline clusters from the supersaturated fluid, leading to much higher nucleation rates compared to those observed in the absence of surfaces (i.e., homogeneous crystallization[69]). This phenomenon is driven by the lower wall-nucleus interfacial free energy promoting cluster wetting on the surface[70], and appears to be analogous to the behavior observed in our simulations for Z-α_1-antitrypsin. Remarkably, even in the case of the purely repulsive wall, the reported density maximum is of the order of condensate densities below the critical solution temperature (see phase diagram in Fig. 1(c)). Hence, protein surface coating may induce local high-density clusters that in turn promote inter-protein β-sheet transitions of Z-α_1-antitrypsin at conditions where LLPS is not expected to spontaneously occur, and subsequently trigger the formation of highly viscous condensates (Fig. 3). Overall, the results of Figure 4 reflect how the presence of either attractive or repulsive walls, enhances Z-α_1-antitrypsin propensity to form condensates, matching experimental observations from Ref.[16] where aberrant solid-like assemblies appear attached to ER membranes.

III. CONCLUSIONS

The results presented in this work shed light on the intricate phase behavior of the misfolded Z-α_1-antitrypsin variant and its implications in the development of AATD[10,12–14]. Through computational molecular modeling, we have elucidated the impact of the structural differences between the native M-α_1-antitrypsin sequence and the misfolded Z variant, particularly focusing on the phase behavior of the protein depending on the folding state of its C-terminus. The misfolded intermediate of Z-α_1-antitrypsin exhibits an intrinsically disordered C-terminus, contrasting with the structured intramolecular β-sheet domain found in the native M variant (Fig. 1(a) and 1(b)). We examine their phase diagram by means of Direct Coexistence simulations finding a greater propensity of the Z variant to undergo LLPS compared to the M native sequence (Fig. 1(c)). These results are consistent with previous experimental findings[13,16], and further validate our coarse-grained model in describing their coexistence lines[25]. The increased propensity for phase separation is unambiguously attributed to the presence of an intrinsically disordered region in the Z variant which is capable of establishing multiple intermolecular contacts with neighboring C-terminal domains of adjacent protein replicas (Figure 2(b)). Moreover, the disordered domain also facilitates intermolecular interactions with most regions of the sequence compared to the M variant in which such domain is practically inert. The enhanced ability of the Z variant to form condensates favors the progressive formation of inter-protein β-sheet clusters instead of individual intramolecular folding. Consequently, the emergence of interconnected strong-binding networks via inter-protein β-sheet accumulation significantly stabilizes the formation of condensates against their spontaneous dissolution.

We measure the viscosity of Z-α_1-antitrypsin condensates once they are formed (i.e., in the absence of inter-protein β-sheet clusters) and after incubation. The progressive emergence of cross-β-sheet clusters leads to an

increase of condensate viscosity due to the inter-protein assembly of local ordered structures that mutually bind via strongly π-π interactions and hydrogen bonding between backbone interactions[30]. These highly stable cross-β-sheet structures[24,28] lead to the progressive hardening of initially liquid-like condensates. The increased viscosity observed in aged Z-α_1-antitrypsin condensates provides a mechanistic explanation for the aberrant formation of misfolded Z-α_1-antitrypsin inclusion bodies found *in vivo*[10,12,13,16].

Furthermore, experimental evidence[16] suggests that the formation of aberrant Z variant clusters in the ER membrane may be promoted under confinement conditions, or nucleated over its surface. We carry out simulations of Z-α_1-antitrypsin in presence of both repulsive and moderately attractive flat structureless walls to elucidate their impact on protein phase-separation. Strikingly, we observe that both types of external surfaces (independently of their chemical adsorption affinity) induce local high-density fluctuations above the characteristic coexistence condensate densities to undergo LLPS (Fig. 4(b)). This suggests a heterogeneous nucleation mechanism by which Z-α_1-antitrypsin (or M-α_1-antitrypsin under higher concentrations) aided by an external surface could potentially achieve the critical saturation concentration to form small condensate nuclei. The initial nuclei can be further stabilized and grow through the formation of cross-β-sheet clusters which gradually increase the inter-protein interaction strength, and thus, the condensate viscosity and stability (Fig. 3(c)).

Altogether, our study advances our understanding of the intricate interplay between structure and phase behaviour in M- and Z-α_1-antitrypsin, as well as on their role in driving AATD pathogenesis. The proposed mechanism to explain its aberrant condensation is based on the formation of phase-separated droplets stabilized by C-terminus intermolecular interactions which over time develop inter-protein β-sheet transitions. This mechanism provides a strong hypothesis for understanding the onset of AATD. Furthermore, our findings at the molecular level can contribute to the development of therapeutic strategies[71] targeting the protein interactions involved in the accumulation of α_1-antitrypsin inclusion bodies at the ER. Potential routes could focus on the intramolecular folding of the C-terminus, or the selective engagement of external molecules with this domain to inhibit inter-protein β-sheet accumulation.

ACKNOWLEDGMENTS

I. S.-B. acknowledges funding from the Derek Brewer scholarship of Emmanuel College and EPSRC Doctoral Training Programme studentship, number EP/T517847/1, Ramon y Cajal fellowship (awarded to J.R.E.), as well as the UKRI EPSRC under the UK Government's guarantee scheme (EP/Z002028/1), following successful evaluation by the ERC (Consolidator Grant awarded to R.C.G.) under the European Union's Horizon Europe research and innovation programme. R.C.-G. acknowledges funding from the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme (grant agreement 803326). A. T. is funded by European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme (grant agreement 803326) and Ramon y Cajal fellowship (RYC2021-030937-I). J. R. E. also acknowledges funding from the Ramon y Cajal fellowship (RYC2021-030937-I) and the Spanish National Agency for Research (PID2022-136919NA-C33). This work has been performed using resources provided by the Cambridge Tier-2 system operated by the University of Cambridge Research Computing Service (http://www.hpc.cam.ac.uk) funded by EPSRC Tier-2 capital grant EP/P020259/1. The authors also thankfully acknowledge RES computational resources provided by MareNostrum 5 through the activities FI-2024-3-0001 and FI-2025-1-0008.

DATA AVAILABILITY

The data that supports the findings of this study are available within the article and its Supplementary Material.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

[1] D. Balchin, M. Hayer-Hartl, and F. U. Hartl, Science **353**, aac4354 (2016).
[2] F. Chiti and C. M. Dobson, Annual review of biochemistry **86**, 27 (2017).
[3] G. R. Cutting, Nature Reviews Genetics **16**, 45 (2015).
[4] J. R. Greenfield and L. V. Campbell, Current diabetes reviews **2**, 195 (2006).
[5] C. A. Waudby, C. M. Dobson, and J. Christodoulou, Trends in biochemical sciences **44**, 914 (2019).
[6] T. P. Knowles, M. Vendruscolo, and C. M. Dobson, Nature reviews Molecular cell biology **15**, 384 (2014).
[7] F. U. Hartl, Annual review of biochemistry **86**, 21 (2017).
[8] N. Louros, J. Schymkowitz, and F. Rousseau, Nature Reviews Molecular Cell Biology **24**, 912 (2023).
[9] L. T. Hunt and M. O. Dayhoff, Biochemical and biophysical research communications **95**, 864 (1980).
[10] B. Gooptu, J. A. Dickens, and D. A. Lomas, Trends in molecular medicine **20**, 116 (2014).
[11] S. P. Bottomley, Proceedings of the American Thoracic Society **7**, 404 (2010).
[12] S. V. Faull, E. L. Elliston, B. Gooptu, A. M. Jagger, I. Aldobiyan, A. Redzej, M. Badaoui, N. Heyer-Chauhan, S. T. Rashid, G. M. Reynolds, *et al.*, Science advances **6**, eabc1370 (2020).
[13] M.-H. Yu, K. N. Lee, and J. Kim, Nature structural biology **2**, 363 (1995).
[14] A. S. Knaupp, V. Levina, A. L. Robertson, M. C. Pearce, and S. P. Bottomley, Journal of molecular biology **396**, 375 (2010).

[15]P. Sivasothy, T. R. Dafforn, P. G. Gettins, and D. A. Lomas, Journal of Biological Chemistry **275**, 33663 (2000).
[16]J. E. Chambers, N. Zubkov, M. Kubánková, J. Nixon-Abell, I. Mela, S. Abreu, M. Schwiening, G. Lavarda, I. López-Duarte, J. A. Dickens, *et al.*, Science Advances **8**, eabm2094 (2022).
[17]E. Plessa, L. P. Chu, S. H. Chan, O. L. Thomas, A. M. Cassaignau, C. A. Waudby, J. Christodoulou, and L. D. Cabrita, Nature Communications **12**, 6447 (2021).
[18]X. Gui, F. Luo, Y. Li, H. Zhou, Z. Qin, Z. Liu, J. Gu, M. Xie, K. Zhao, B. Dai, *et al.*, Nature communications **10**, 2006 (2019).
[19]M. P. Hughes, M. R. Sawaya, D. R. Boyer, L. Goldschmidt, J. A. Rodriguez, D. Cascio, L. Chong, T. Gonen, and D. S. Eisenberg, Science **359**, 698 (2018).
[20]Y. Sun, K. Zhao, W. Xia, G. Feng, J. Gu, Y. Ma, X. Gui, X. Zhang, Y. Fang, B. Sun, *et al.*, Nature communications **11**, 6349 (2020).
[21]F. Luo, X. Gui, H. Zhou, J. Gu, Y. Li, X. Liu, M. Zhao, D. Li, X. Li, and C. Liu, Nature structural & molecular biology **25**, 341 (2018).
[22]A. Garaizar, J. R. Espinosa, J. A. Joseph, G. Krainer, Y. Shen, T. P. Knowles, and R. Collepardo-Guevara, Proceedings of the National Academy of Sciences **119**, e2119800119 (2022).
[23]A. R. Tejedor, I. Sanchez-Burgos, M. Estevez-Espinosa, A. Garaizar, R. Collepardo-Guevara, J. Ramirez, and J. R. Espinosa, Nature communications **13**, 5717 (2022).
[24]S. Blazquez, I. Sanchez-Burgos, J. Ramirez, T. Higginbotham, M. M. Conde, R. Collepardo-Guevara, A. R. Tejedor, and J. R. Espinosa, Advanced Science **10**, 2207742 (2023).
[25]G. Tesei and K. Lindorff-Larsen, bioRxiv , 2022 (2022).
[26]S. Plimpton, Journal of computational physics **117**, 1 (1995).
[27]I. Sanchez-Burgos, A. R. Tejedor, A. Feito, R. Collepardo-Guevara, and J. Rene Espinosa, bioRxiv , 2024 (2024).
[28]E. L. Guenther, Q. Cao, H. Trinh, J. Lu, M. R. Sawaya, D. Cascio, D. R. Boyer, J. A. Rodriguez, M. P. Hughes, and D. S. Eisenberg, Nature structural & molecular biology **25**, 463 (2018).
[29]A. Šarić, Y. C. Chebaro, T. P. Knowles, and D. Frenkel, Proceedings of the National Academy of Sciences **111**, 17869 (2014).
[30]A. Garaizar, J. R. Espinosa, J. A. Joseph, and R. Collepardo-Guevara, Scientific reports **12**, 4390 (2022).
[31]A. O. Patschull, L. Segu, M. P. Nyon, D. A. Lomas, I. Nobeli, T. E. Barrett, and B. Gooptu, Acta Crystallographica Section F: Structural Biology and Crystallization Communications **67**, 1492 (2011).
[32]A. J. Ladd and L. V. Woodcock, Chemical Physics Letters **51**, 155 (1977).
[33]J. R. Espinosa, A. Garaizar, C. Vega, D. Frenkel, and R. Collepardo-Guevara, The Journal of chemical physics **150**, 224510 (2019).
[34]A. R. Tejedor, A. Garaizar, J. Ramírez, and J. R. Espinosa, Biophysical Journal **120**, 5169 (2021).
[35]M. J. Maristany, A. Aguirre Gonzalez, R. Collepardo-Guevara, and J. A. Joseph, bioRxiv , 2023 (2023).
[36]A. Bremer, M. Farag, W. M. Borcherds, I. Peran, E. W. Martin, R. V. Pappu, and T. Mittag, Nature Chemistry **14**, 196 (2022).
[37]I. Alshareedah, W. M. Borcherds, S. R. Cohen, A. Singh, A. E. Posey, M. Farag, A. Bremer, G. W. Strout, D. T. Tomares, R. V. Pappu, *et al.*, bioRxiv , 2023 (2023).
[38]A. Feito, I. Sanchez-Burgos, I. Tejero, E. Sanz, A. Rey, R. Collepardo-Guevara, A. R. Tejedor, and J. R. Espinosa, bioRxiv , 2024 (2024).
[39]A. R. Tejedor, A. Aguirre Gonzalez, M. J. Maristany, P. Y. Chew, K. Russell, J. Ramirez, J. R. Espinosa, and R. Collepardo-Guevara, ACS Central Science **11**, 302 (2025).
[40]A. Molliex, J. Temirov, J. Lee, M. Coughlin, A. P. Kanagaraj, H. J. Kim, T. Mittag, and J. P. Taylor, Cell **163**, 123 (2015).
[41]S. Elbaum-Garfinkle, Y. Kim, K. Szczepaniak, C. C.-H. Chen, C. R. Eckmann, S. Myong, and C. P. Brangwynne, Proceedings of the National Academy of Sciences **112**, 7189 LP (2015).
[42]D. M. Mitrea, J. A. Cika, C. B. Stanley, A. Nourse, P. L. Onuchic, P. R. Banerjee, A. H. Phillips, C.-G. Park, A. A. Deniz, and R. W. Kriwacki, Nature communications **9**, 1 (2018).
[43]A. Garaizar, I. Sanchez-Burgos, R. Collepardo-Guevara, and J. R. Espinosa, Molecules **25**, 4705 (2020).
[44]S. Xiang, M. Kato, L. C. Wu, Y. Lin, M. Ding, Y. Zhang, Y. Yu, and S. L. McKnight, Cell **163**, 829 (2015).
[45]E. W. Martin, F. E. Thomasen, N. M. Milkovic, M. J. Cuneo, C. R. Grace, A. Nourse, K. Lindorff-Larsen, and T. Mittag, Nucleic acids research **49**, 2931 (2021).
[46]P. S. Tsoi, M. D. Quan, K.-J. Choi, K. M. Dao, J. C. Ferreon, and A. C. M. Ferreon, Protein Science **30**, 1408 (2021).
[47]P. S. Tsoi, J. C. Ferreon, and A. C. M. Ferreon, International Journal of Molecular Sciences **25**, 6825 (2024).
[48]M. Linsenmeier, L. Faltova, C. Morelli, U. Capasso Palmiero, C. Seiffert, A. M. Küffner, D. Pinotsi, J. Zhou, R. Mezzenga, and P. Arosio, Nature chemistry **15**, 1340 (2023).
[49]D. S. Protter, B. S. Rao, B. Van Treeck, Y. Lin, L. Mizoue, M. K. Rosen, and R. Parker, Cell reports **22**, 1401 (2018).
[50]S. Maharana, J. Wang, D. K. Papadopoulos, D. Richter, A. Pozniakovsky, I. Poser, M. Bickle, S. Rizk, J. Guillén-Boixet, T. M. Franzmann, M. Jahnel, L. Marrone, Y.-T. Chang, J. Sterneckert, P. Tomancak, A. A. Hyman, and S. Alberti, Science **360**, 918 (2018).
[51]A. Wang, A. E. Conicella, H. B. Schmidt, E. W. Martin, S. N. Rhoads, A. N. Reeb, A. Nourse, D. Ramirez Montero, V. H. Ryan, R. Rohatgi, *et al.*, The EMBO journal **37**, e97452 (2018).
[52]T. C. Michaels, A. Šarić, J. Habchi, S. Chia, G. Meisl, M. Vendruscolo, C. M. Dobson, and T. P. Knowles, Annual review of physical chemistry **69**, 273 (2018).
[53]A. R. Tejedor, I. Sanchez-Burgos, M. Estevez-Espinosa, A. Garaizar, R. Collepardo-Guevara, J. Ramirez, and J. R. Espinosa, bioRxiv (2022).
[54]S. K. Sukumaran, G. S. Grest, K. Kremer, and R. Everaers, Journal of Polymer Science Part B: Polymer Physics **43**, 917 (2005).
[55]C. M. Greene, S. J. Marciniak, J. Teckman, I. Ferrarotti, M. L. Brantly, D. A. Lomas, J. K. Stoller, and N. G. McElvaney, Nature reviews Disease primers **2**, 1 (2016).
[56]A. Patel, H. O. Lee, L. Jawerth, S. Maharana, M. Jahnel, M. Y. Hein, S. Stoynov, J. Mahamid, S. Saha, T. M. Franzmann, *et al.*, Cell **162**, 1066 (2015).
[57]S. Ray, N. Singh, R. Kumar, K. Patel, S. Pandey, D. Datta, J. Mahato, R. Panigrahi, A. Navalkar, S. Mehra, *et al.*, Nature Chemistry **12**, 705 (2020).
[58]B. Portz, B. L. Lee, and J. Shorter, Trends in biochemical sciences **46**, 550 (2021).
[59]S. Alberti and A. A. Hyman, Nature Reviews Molecular Cell Biology **22**, 196 (2021).
[60]M. Rubinstein, R. H. Colby, *et al.*, *Polymer physics*, Vol. 23 (Oxford university press New York, 2003).
[61]A. R. Tejedor, R. Collepardo-Guevara, J. Ramirez, and J. R. Espinosa, The Journal of Physical Chemistry B.
[62]J. Ramírez, S. K. Sukumaran, B. Vorselaars, and A. E. Likhtman, The Journal of chemical physics **133** (2010).
[63]D. Chandler, Mechanics. Oxford University Press, Oxford, UK **5**, 11 (1987).
[64]T. Mori, N. Miyashita, W. Im, M. Feig, and Y. Sugita, Biochimica et Biophysica Acta (BBA)-Biomembranes **1858**, 1635 (2016).
[65]Y. Rodríguez and M. Májeková, Biomolecules **10**, 214 (2020).
[66]N. E. Chayen, E. Saridakis, and R. P. Sear, Proceedings of the National Academy of Sciences **103**, 597 (2006).
[67]Y.-X. Liu, X.-J. Wang, J. Lu, and C.-B. Ching, The Journal of Physical Chemistry B **111**, 13971 (2007).
[68]J. R. Espinosa, C. Vega, C. Valeriani, D. Frenkel, and E. Sanz, Soft Matter **15**, 9625 (2019).
[69]J. R. Espinosa, C. Vega, C. Valeriani, and E. Sanz, The Journal of chemical physics **144**, 034501 (2016).
[70]M. Heni and H. Löwen, Physical Review E **60**, 7057 (1999).
[71]D. A. Lomas, J. A. Irving, C. Arico-Muendel, S. Belyanskaya, A. Brewster, M. Brown, C.-w. Chung, H. Dave, A. Denis, N. Dodic, *et al.*, EMBO Molecular Medicine **13**, e13167 (2021).