Regular Article - Soft Matter

# From isolated polyelectrolytes to star-like assemblies: the role of sequence heterogeneity on the statistical structure of the intrinsically disordered neurofilament-low tail domain

Mathar Kravikass[1,2], Gil Koren[1,2], Omar A. Saleh[3,4], and Roy Beck[1,2,a]

[1] School of Physics and Astronomy, The Center for Nanoscience and Nanotechnology, Tel Aviv University, Tel Aviv, Israel
[2] The Center of Physics and Chemistry of Living Systems, Tel Aviv University, Tel Aviv, Israel
[3] Materials Department, Biomolecular Sciences and Engineering Program, University of California, Santa Barbara, USA
[4] Physics Department, University of California, Santa Barbara, USA

**Abstract** Intrinsically disordered proteins (IDPs) are a subset of proteins that lack stable secondary structure. Given their polymeric nature, previous mean-field approximations have been used to describe the statistical structure of IDPs. However, the amino-acid sequence heterogeneity and complex intermolecular interaction network have significantly impeded the ability to get proper approximations. One such case is the intrinsically disordered tail domain of neurofilament low (NFLt), which comprises a 50 residue-long uncharged domain followed by a 96 residue-long negatively charged domain. Here, we measure two NFLt variants to identify the impact of the NFLt two main subdomains on its complex interactions and statistical structure. Using synchrotron small-angle x-ray scattering, we find that the uncharged domain of the NFLt induces attractive interactions that cause it to self-assemble into star-like polymer brushes. On the other hand, when the uncharged domain is truncated, the remaining charged N-terminal domains remain isolated in solution with typical polyelectrolyte characteristics. We further discuss how competing long- and short-ranged interactions within the polymer brushes dominate their ensemble structure and, in turn, their implications on previously observed phenomena in NFL native and diseased states.

## Introduction

Intrinsically disordered proteins (IDPs) are a subset of proteins that, instead of forming a rigid singular structure, fluctuate between different conformations in their native form [1,2]. Nonetheless, IDPs serve significant biological functions and account for about 44% of the human genome [3]. The lack of fixed structure provides IDPs many advantages in regulatory systems in which they often play a crucial role in mediating protein interaction [4,5]. These roles often come into play from intrinsically disordered regions (IDRs) of folded proteins interacting with other IDRs. For example, in the neurofilament proteins, tails emanating from the self-assembled filament backbone domains bind together and form a network of filaments [6–10].

The ensemble statistics of IDPs stem from their sequence composition and the surrounding solution [2]. For example, previous studies showed that IDPs comprising mostly negatively charged amino acids (polyelectrolytes) are locally stretched due to electrostatic repulsion between the monomers [11]. Moreover, different properties, such as hydrophobicity, were shown to be linked with local IDP domain collapse [12]. The complex interactions that arise from sequence heterogeneity allow IDPs to form specific complexes without losing their disordered properties [13]. For example, Khatun et al. recently showed how, under limited conditions, the human amylin protein self-assembles into fractal structures [14].

As IDPs are disordered chains, polymer theories are prime candidates to relate the measured structural statistics to known models, which can help link the sequence composition of the IDP to its conformations [15–18]. Specifically, polymer scaling theories allow us to derive the statistical structure of IDPs given sequence-derived parameters, such as charge density and hydrophobicity [11,12,19–21]. However, due to the heterogeneity of the IDP primary structure (i.e., the amino-acid sequence), some systems showed contradictions with the behavior theorized by standard heterogeneous polymer physics [17,19,22–24].

The unique biological properties of IDPs have given rise to numerous attempts to use them as building

blocks for self-assembled structures [25]. For example, IDPs were proposed as brush-like surface modifiers, due to their enhanced structural plasticity to environmental conditions [26,27]. Another example of an IDP brush system is the neurofilament (NF) protein system [6,28,29], described as interacting bottle-brushes. NF subunit proteins form mature filaments with protruding disordered C-terminus IDR known as 'tails.' NF tails were shown to mediate NF network formation and act as shock absorbers in high-stress conditions [29]. Moreover, NF aggregations are known to accumulate alongside other proteins in several neurodegenerative diseases, such as Alzheimer's and Parkinson's. [30].

The NF-low disordered tail domain (NFLt) sequence can be divided into two unique regions: an uncharged region (residues 1–50) starting from its N-terminal and a negatively charged region (residues 51–146). The NFLt can be described as a polyelectrolyte with a net charge per residue (NCPR) of $-0.24$. Furthermore, the statistical structures of segments within the NFLt are influenced by the amount, type, and disperse of the charged amino acid within a segment [22]. Nonetheless, other structural constraints, particularly long-range contacts, impact the local statistical structures. Additionally, NFLt was shown to have glassy dynamics with the response to tension [31]. Such dynamics were associated with multiple weakly interacting domains and structural heterogeneity.

In this paper, we revisit NFLt as a model system for charged IDP and focus on the contribution of its neutral and hydrophobic N-terminal domain. We will show that increased salt concentration causes NFLt to form star-like brushes with increased aggregation number ($Z$). Here, we are motivated by theoretical models, in particular the Pincus' model for salted polyelectrolytes [32], that capture key physical properties of IDPs, including the model system presented here [26,29,33]. We will further quantify the competition between hydrophobic attraction and electrostatic and steric repulsion in the formation of the structures of NFLt.

## Results

To study the N-terminal domain contribution to the structure of NFLt, we designed two variants and measured them at various buffer conditions. The first construct is the entire 146 residues of the NFLt chain, which we term as WT (NCPR = $-0.24$), and the second is isolating the 104 negatively charged residues from the C-terminal of NFLt (NCPR = $-0.33$), termed as $\Delta$N42. We expressed the variants in *E-coli* and purified it up to 96% (see methods).

We assessed the variants in solution using small-angle X-ray scattering (SAXS), a technique extensively used to characterize the statistical structures of IDPs [34]. From the raw SAXS data, measured at various salinities, we can already find high structural differences between the two variants (Fig. 1a). Dominantly at the low wave-vector ($q$) region, the WT variant scattering
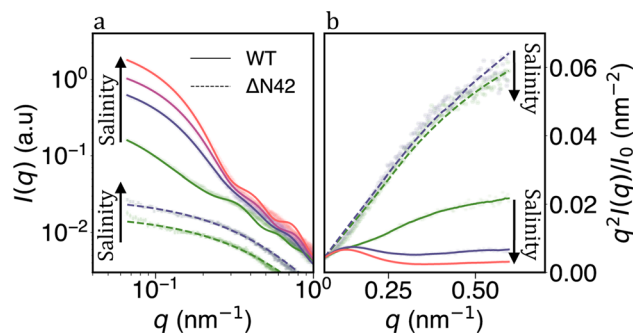


**Fig. 1** SAXS measurements of WT and $\Delta$N42 at different salinity ($C_s$). **a** For increasing $C_s$, the WT variant shows increased small-angle scattering, a signature for aggregation. In contrast, $\Delta$N42 remains structurally intrinsically disordered as $C_s$ vary. Data points are shifted for clarity. Lines are form-factor fittings, as described in the text. **b** Normalized Kratky plot of the same SAXS measurements. The $\Delta$N42 variant remains disordered and unchanged with salinity, while the WT variant shows a hump at low $q$, typical for a collapse region. With increasing $C_s$, the hump at the lower $q$ range becomes a sharper peak accompanied by a scattering rise at the higher $q$ range. Such behavior indicates that the aggregation coexists with the WT variant's highly dynamic and disordered regions. Both variants shown are at the highest measured concentration (Table S1, S3). WT measurements are in 20 mM Tris pH 8.0 with 0, 150, 250, and 500 mM added NaCl (from bottom to top). Likewise, for $\Delta$N42, measurements are in 20 mM Tris pH 8.0 with 0 and 150 mM added NaCl (bottom to top)

($I$) rises with added NaCl salt. Such an increase at low $q$ implies high molecular mass particles due to aggregation of the WT variant.

In contrast, $\Delta$N42 shows a separated Gaussian polymer profile (Figs. 1a, S1), nearly insensitive to total salinity ($C_s = 20 - 520$ mM). Similarly, the data presented in Kratky format ($qI^2$ vs. $q$, Fig. 1a) show the $\Delta$N42 has the signature of a disordered polymer. In contrast, the WT variant, in particular at high salinity, has a combination of a collapse domain (the peaks from below $q = 0.25$ nm$^{-1}$) and a disordered polymeric structure (the scattering rise at higher $q$ Fig. 1a).

Being completely disordered, $\Delta$N42 lacks a stable structure and can be described using a statistical ensemble of polymeric conformations [35] were:

$$I(q) = I_0 \exp\{-\frac{1}{3}(qR_{\mathrm{G}})^2 \\ + 0.0479(\nu - 0.212)(qR_{\mathrm{G}})^4\}. \tag{1}$$

Here, $I_0$ is the scattering at $q = 0$, $\nu$ is Flory scaling exponent, and $R_{\mathrm{G}}$ is the radius of gyration defined by:

$$R_{\mathrm{G}} = \sqrt{\frac{\gamma(\gamma + 1)}{2(\gamma + 2\nu)(\gamma + 2\nu + 1)}} bN^{\nu}, \tag{2}$$

where $\gamma = 1.615$ and $b = 0.55$ nm (see [35]) and the analysis is viable up to $qR_{\mathrm{G}} \sim 2$ (Fig. S2, S3). In all

ΔN42 cases, the scattering profile fits Eq. 1 and with $\nu$ ranging between 0.63–0.69 depending on the buffer salinity (Table S1). In 'infinite dilution' conditions (zero polymer concentration), we find $\nu$ to decrease monotonically from 0.73 to 0.62 with added salt (Table S2).

Given the noticeable aggregation for the WT variant, alternative form factors were considered to match the scattering profiles (lines in Fig. 1). The absence of structural motifs at high $q$ values ($q > 0.3$ nm$^{-1}$) indicates a disordered nature for WT at shorter length scales. Conversely, in the lower $q$ region ($q < 0.3$ nm$^{-1}$), the scattering suggests stable structural motifs or a larger molecular weight particles. Such SAXS resembles that of self-assembled decorated spherical micelles [36]. Variations of micelle models are shown to fit the data (Figs. 1, S4–S6). Sufficiently low aggregation number and core size distill the description of the spherical micelle into a 'star-like' brush. Alternative attempts to fit the scattering profiles to other form factors models, including vesicles and lamellar, were unsuccessful.

For the star-like model, the aggregated variants form a small spherical core of volume $V_{core}$ made out of $n \cdot Z$ monomers (comparison with different cores described in [37] and in Fig. S4), where $n$ denotes the peptide length per polypeptide within the core, and $Z$ is the aggregation number, i.e. the number of polypeptides per 'star.' The remainder of the WT variant then protrudes from the core as the star polymer brush (Figs. 2a, S4–S6).

The star-like scattering form factor is described as a combination of four terms [36]: the self-correlation term of the core $F_c$, the self-correlation term of the tails $F_t$, the cross-correlation term of the core and the tails $S_{ct}$ and the cross-correlation term of the tails $S_{tt}$:

$$F_{total}(q) = Z^2 \beta_c^2 F_c(q) + Z \beta_t^2 F_t(q) \\ + 2Z^2 \beta_c \beta_t S_{ct}(q) + Z(Z-1)\beta_t^2 S_{tt}(q). \quad (3)$$

Here, $\beta_c$ and $\beta_t$ are the excess scattering length of the core and the tails, respectively. From fitting the scattering data, we extracted the height of the tails $h = 2R_G$, the aggregation number $Z$, and the relevant core's parameters (e.g., core radius $R$ for a spherical core, cylinder radius $R$ and length $L$ for a cylindrical core [37]), schematically illustrated in Fig. 2a. All fitting parameters are found in Table S3.

To avoid misinterpretation and to minimize intermolecular interaction effects, we present the fitting results at the 'infinitely diluted regime' by extrapolating the relevant parameters measured at various protein concentrations to that at zero protein concentration (Fig. S7, Table S4). The parameters are mostly independent of the concentration unless explicitly mentioned.

At low salinity (20 mM), the aggregation number for the WT variant is of a dimer ($Z \approx 2$), and the core's shape is that of a cylinder (with a radius $R = 0.89$ nm and length $L = 1.19$ nm). At higher salt conditions (170–520 mM), the form factor fits spherical core aggregates with increasingly higher $Z$'s (Fig. 2a).

Given the relatively small core volume ($V_{core} \approx 1 - 2$nm$^3$, Fig. 2c), it is crucial to evaluate the 'grafting' distance between neighboring chains, $\rho$, on the core surface ($S = 4\pi R^2 = Z\rho^2$) and the brush extension, $h$, outside the core. As shown in Fig. 2d, in all cases, $h/\rho \gg 1$ indicates a 'brush regime' where neighboring chains repel each other while extending the tail's height [38].

The repulsion between the grafted tail is further emphasized when comparing $h/2$ for WT to the equivalent ΔN42 length-scale ($R_G$), showing a significant extension for WT (Fig. 2e). We notice that the WT tail's length ($h$) increases at low salt (during the transitions from a dimer to a trimer), followed by a steady mild decrease as the $C_s$, and following $Z$ increase. Similar compactness with increasing $C_s$ is shown for ΔN42 and is expected for polyelectrolyte due to the reduction in electrostatic repulsion [39]. To better compare the statistical structure of two variants of disordered regions, we followed the polymeric scaling notation $\nu$ that quantifies the compactness of the chain. For ΔN42, we extracted $\nu$ from Eqs. 1 and 2 and found a significant decrease in its value as 50 mM of NaCl is added to the 20 mM Tris buffer (Fig. 3a). The following monotonic decline is in line with polyelectrolytic models and electrostatic screening effects [40], shown in a solid red line in Fig. 3a. Interestingly, previous measurements of segments within the NFLt charged domain were shown to have similar $\nu$ values as in ΔN42 . However, the same decline in salinity was not observed (Fig. 3a) [22].

For the WT variant, the scaling factor ($\nu$) of the 'star-like polymer' brushes is extracted from Eq. 2. Here, we use $R_G = h/2$, where $h$ is obtained from Eq. 3. For $C_s = 20$ mM, we find that $\nu$ is of similar scale as for ΔN42 . This similarity can be attributed to the nature of the dimer, where the intramolecular electrostatic interactions dominate the expansion of each of the two tails. As $C_s$ increases by 150 mM, $\nu$ exhibits a considerable increase, presumably due to neighboring tail repulsion. Above $C_s = 170$ mM, $\nu$ shows a weak decrease. We attribute this weak decline to the salt-brush regime of polyelectrolyte brushes [41] shown in solid blue in Fig. 3a. In this regime, $h \propto C_s^{-1/3}$, and subsequently $\nu \propto -\frac{1}{3}log(C_s)$.

We note that the cores of the star-like polymers are relatively small and that each polypeptide aggregates through only a few, most likely hydrophobic, amino acids. From the tabulated amino-acid partial volume, $\langle\phi_{aa}\rangle$ [42], we estimate the comprising amino acids as spheres of volume $\langle\phi_{aa}\rangle$. From here, the average number of amino acids per polypeptide inside the core is estimated by the number of spheres that can fit within the core volume, divided by the aggregation number: $n = V_{core}/(\langle\phi_{aa}\rangle \cdot Z)$. Noticeably, our fit results with small $n$ values, ranging between $\sim 7 - 2$ residues on average within the aggregate ensemble and depending on the buffer salinity. Attempting to 'fix' $n$ to a larger constant residue per tail number results in a poorer fitting (Fig. S9). In Fig. 3a, we indeed see that the most significant change occurs at the low salt regime,
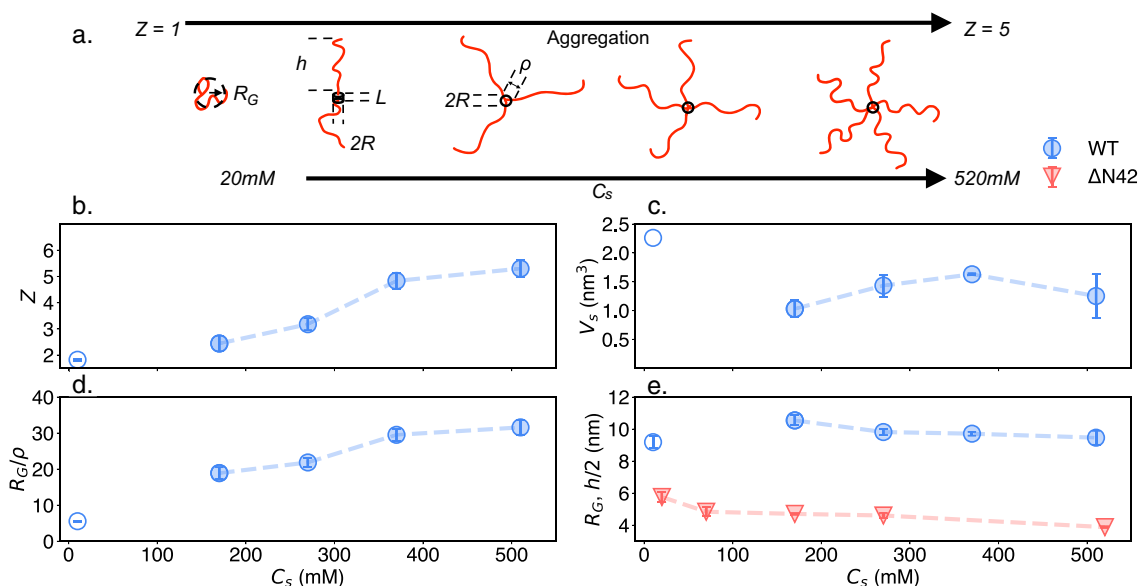
**Fig. 2** **a** Schematic of the system's structure variation with salinity ($C_s$). While $\Delta$N42 remains disordered and segregated, the WT variant aggregates to a star-like polymer with a higher aggregation number at higher $C_s$. **b–e** Structural parameters for WT (blue symbols) and $\Delta$N42 (red symbols) variants extracted from fitting the SAXS data. Full and hollow circles represent the spherical and cylindrical core fitted parameters, respectively. **d** In all cases, the brush heights ($h$) are much larger than the corresponding grafting length ($\rho$), indicative of a brush regime. **e** The structurally intrinsically disordered $\Delta$N42 variant compacts with higher $C_s$ values and remains more compacted from the projected brushes for the WT variant. All values are the extrapolated 'zero concentration' fitting parameters (see Fig. S7)
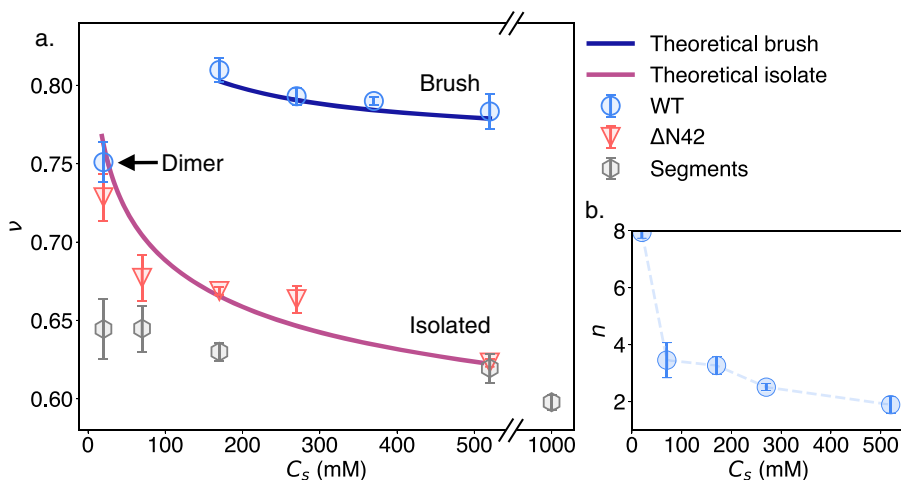


**Fig. 3** Deduced structural parameters from the SAXS data fitting. **a** Flory exponent ($\nu$) of WT tails and $\Delta$N42 variants showing extended disordered scaling. The red line refers to the theoretical brush model [41], and the blue line refers to the theoretical polyelectrolyte [40]. $\Delta$N42 shows a decrease in the protein extension due to the decline in intermolecular electrostatic repulsion (see also Fig. 4). WT shows an increase in the extension when shifting from a dimer to a trimer, followed by a slight decline with a further increase in salinity. In gray, average $\nu$ is obtained from measuring separate NFLt segments with an NCPR of $-0.3$ to $-0.6$ [22]. **b** The core (aggregated) peptide length per polypeptide as a function of salinity. At high salinity, each polypeptide aggregates via 2–3 amino acids that form the star-like polymer core. Both panels' values are the extrapolated 'zero concentration' parameters (supplementary Fig. S8)

where $n$ drops from an average of 7 to 3 amino acids ($C_s = 20, 170$ mM, respectively). Such behavior is known to occur within globular proteins [43] and was recently alluded to impact IDPs [44]. The following trend is a further decrease in $n$, albeit much weaker, which results in a final average $n$ of about two as the salinity reaches $C_s = 520$ mM.

Last, in Fig. 4, we quantify the intermolecular interactions by evaluating the second virial coefficient, $A_2$, using a Zimm analysis [45] (Table S5). Here, $A_2$ describes the deviation of the statistical ensemble from an ideal gas. In agreement with our previous data, we find that the inter-molecular interactions of $\Delta$N42 change from repulsive ($A_2 > 0$) to weakly attractive ($A_2 \leq 0$) as the salinity increases. In contrast, for WT, $A_2$ changes from a nearly neutral state of intermolecular interactions (i.e., ideal gas regime) to mildly attractive ($A_2 < 0$). These findings are reflected in the dependency of the variant Flory coefficient $\nu$ in concentration. While at the lowest salinity, $\Delta$N42 is shown to expand as protein concentration is decreased, for higher salinities and for the WT measurements, $\nu$ remain primarily unchanged (Fig. S8a).

Combining our results for both variants, we find an exemplary role of long-range electrostatic interactions tuning the statistical structure of IDPs. Without the uncharged N-terminal domain, the NFLt exhibited significant change as the electrostatic interactions were screened, causing them to condense further. In contrast, the presence of the uncharged domain incurred aggregation of the proteins, bringing the tails much closer to each other. The increase in proximity was reflected in a significant increase in the expansion compared to the truncated variant, which exhibited a much weaker contraction with salinity.

## Discussion and Conclusions

We investigated the effects of sequence heterogeneity on the interactions and structures of NFLt, an IDP model system. For NFLt, the N-terminal region consisting of the first $\sim 50$ residues is hydrophobic and charge neutral, while the remaining chain is highly charged. We found that the sequence heterogeneity differentiates between the structures of the entire WT NFLt and a variant lacking the N-terminal domain. In particular, the WT variant self-assembles into star-like structures, while the $\Delta$N42 one remains isolated in all measured cases.

Since $\Delta$N42 can be attributed as a charged polymer, weakly attractive interactions take center stage as the electrostatic repulsion diminishes with charge screening (Fig. 4). These interactions could be attributed to monomer–monomer attractions that arise from the sequence heterogeneity of the IDP, such as weak hydrophobic attraction from scattered hydropathic sites [22,28,29,46–48].

For the WT variant, the intermolecular interactions started from a near-neutral state and transitioned to

weakly attractive. However, as the WT measurements describe self-assembling complexes, the interpretation of these results differs from $\Delta$N42. As such, we interpret the intermolecular interactions as the 'aggregation propensity,' the protein complex's growing ability. The aggregation propensity grows as the attractivity between the complex and the other polypeptides in the solution increases. This behavior can be observed when examining the responsiveness of the aggregation number $Z$ to protein concentration $C$ (Fig. S7). In the lowest measured screening, $Z$ dependency on protein concentration was minimal. As we increase the screening effects, this dependency becomes more substantial. This characterization is also found in folded proteins, where intermolecular interactions were shown to indicate aggregation propensity [49]. The increased intermolecular attraction induced at increasing salinity is indicative of a salting-out phenomenon [50,51], although further investigation at higher salinity is needed.

The stability of the star-like polymer core should be evaluated by the participating residues per polypeptide ($n$). Indeed, while our fittings result with rather small $n$ values, the SAXS signal at low $q$ is dominated with aggregated structures under all salinity conditions. Within the occurring hydrophobic interactions, the release of bound water molecules and ions from the polypeptides is likely to contribute to the core's stability. Such entropic-based effects have been observed in similar processes such as protein flocculation [52,53] and in temperature-specific IDP binding modulation [54].

In our previous study [22], Flory exponents ($\nu$) of shorter segments from the same NFLt were measured independently and in the context of the whole NFLt using SAXS and time-resolved Förster resonance energy transfer (trFRET). There, regardless of the peptide sequence, in the context of the entire NFLt, the segments' structural statistics were more expanded (i.e., with larger $\nu$ values) than when measured independently. Similarly, these short segments measured with SAXS have smaller $\nu$ values (i.e., with a compacted statistical structure) than those of measured here for $\Delta$N42 in all salt conditions (Fig. 3a, gray symbols).

The expansion of segments in the context of a longer chain corroborates that long-range contacts contribute to the overall disordered ensemble [22]. Interestingly, at $C_s = 520$ mM salinity, we found similar $\nu$ values of the $\Delta$N42 and the previous short segment measurements, indicating a comparable expansion. We suggest that at higher salinities, the significance of electrostatic long-range contacts diminishes, aligning the expansion 'scaling laws' regardless of the chain length. Importantly, comparisons between our $\Delta$N42 variant results (and not to the WT variant) and the previous segments' measurements are more suitable as the chains did not aggregate in those cases.

Compared to $\Delta$N42, WT exhibits a mild contraction in salt, resembling the behavior of the 'salt-brush' regime observed in polyelectrolyte brushes, as demonstrated in Fig. 3. Similar salt-brush behavior was
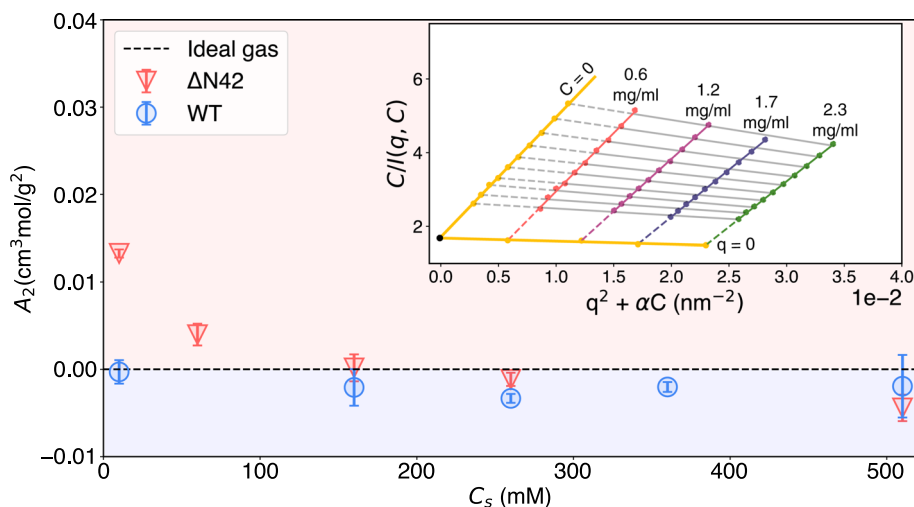
**Fig. 4** The osmotic second virial coefficient $A_2$ as a function of the two variants' salinity ($C_s$). ΔN42 intermolecular interactions transition from repulsive to attractive as $C_s$ increases. WT changes from a nearly neutral state of intermolecular interactions to attractive. Inset: A demonstration (WT variant, $20\,\mathrm{mM}$ Tris and $500\,\mathrm{mM}$ NaCl $p$H 8.0) for the Zimm analysis used to extract $A_2$ from SAXS data measured at various protein concentrations ($C$). Values shown in the graph are in mg/ml units. The dashed lines show the extrapolation from the measured data (colored lines) to the fitted $q \to 0$ and $C \to 0$ yellow lines, where $\alpha = 0.01$ is an arbitrary constant used in this analysis

previously observed in neurofilament-high tail domain brushes grafted onto a substrate [26], and in a recent polyelectrolytic brush scaling theory [55]. In the salt-brush regime, Pincus showed that brush mechanics resemble neutral brushes, determined by steric inter-chain interactions [32]. In this interpretation, the effective excluded volume per monomer enlarges and is proportional to $1/\kappa_s^2$, where $\kappa_s$ is the Debye length attributed to the added salt. Consequently, we suggest that the heightened charge screening in the WT solution allows steric interactions between brushes to play a more significant role in determining the brush ensemble. Additionally, we deduce that the increased prevalence of steric repulsion counteracts the attractive forces responsible for aggregation, thereby preventing brush collapse.

The NFLt contraction aligns with previous studies of native NFL hydrogel networks [28,29]. At high osmotic pressure, the NFL network showed weak responsiveness to salinity higher than $C_s = 100$ mM, in agreement with theory [55]. With the observed salt-brush behavior for WT, we suggest that weak salt response in NFL hydrogels coincides with the increase in steric repulsion shown for the star-like structures (Fig. 3a, blue line).

Additionally, our measurements show that the hydrophobic N-terminal regime of the NFLt domain aggregates. This result is consistent with the findings of Morgan et al. [31], where single-molecule pulling experiments were performed on WT NFLt, and slow aging effects were observed, likely due to collapse (and potential aggregation) of the neutral domain. Indeed, follow-up studies by Truong et al. [56] used single-molecule stretching to show that added denaturant led to a swelling of the chain (increased $\nu$), demonstrating that the WT chain has hydrophobic aggregation that can be disrupted by the denaturant. These observations suggest that at higher salt, the loss of repulsion may lead to attractive hydrophobic interactions growing more prominent in the NFL network. However, the steric repulsion from the remaining NFL tail may shield such an unwanted effect. Nonetheless, such effects may grow more prominent as the native filament assembly is disrupted.

In summary, we showed how the sequence composition of the NFLt IDP caused structural deviation from a disordered polyelectrolyte to a self-assembled star-like polymer brush. Together with the self-regulatory properties of the brushes, such behavior can be exploited to design structures that can resist specific environmental conditions. Additionally, our results showed possible implications on NFL aggregates that could shed light on the underlying correlations between the complex structure and the conditions driving it. While IDPs resemble polymers in many aspects, as we showed here, it is critical to assess their sequence to distinguish where and how to use the appropriate theoretical arguments to describe their statistical properties and structure.

## Methods

### Protein purification

Protein purification followed Koren et al. [22]. Variant ΔN42, included two cysteine residues at the C- and N terminals. After purification, ΔN42 variants were first reduced by $20\,\mathrm{mM}$ 2-Mercaptoethanol. Next, 2-Mercaptoethanol was dialyzed out with $1\,\mathrm{L}$ of $50\,\mathrm{mM}$ HEPES at pH 7.2. To block the cysteine sulfhydryl

group, we reacted $\Delta$N42 variants with 2-Iodoacetamide at a molar ratio of 1:20. At the reaction, the variants' concentrations were $\sim$2 mg/ml. The reaction solution was kept under dark and slow stirring for 5 hr and stopped by adding 50 mM 2-Mercaptoethanol followed by overnight dialysis against 1 L of 20 mM Tris at pH 8.0 with 0.1% 2-Mercaptoethanol. Final purity was >95% as determined by SDS-PAGE (Fig. S10).

### SAXS measurement and analysis

Protein samples were dialyzed overnight in the appropriate solution and measured with a Nanodrop 2000 spectrophotometer (Thermo Scientific) for concentration determination. Buffers were prepared with 1 mM of TCEP to reduce radiation damage and 0.2% of Sodium Azide to impair sample infection. The samples were prepared in a final concentration of 2 mg/ml, measured in a series of 4 dilutions. Preliminary measurements were measured at Tel-Aviv University with a Xenocs GeniX Low Divergence CuK$\alpha$ radiation source setup with scatterless slits [57] and a Pilatus 300K detector. All samples were measured at three synchrotron facilities: beamline B21, Diamond Light Source, Didcot, UK [58], beamline P12, EMBL, DESY, Hamburg, Germany [59], and beamline BM 29 ESRF, Grenoble, France [60]. Measurements at ESRF were taken using a robotic sample changer [61].

Integrated SAXS data were obtained from the beamline pipeline and 2D integration using the "pyFAI" Python library [62]. Extended Guinier analyses for the $\Delta$N42 variant were done with the "curve_fit" function from the "Scipy" Python library [63]. To extract $R_g$ and $\nu$, extended Guinier analysis was conducted for $0.7 < qR_g < 2$. Error calculation was done from the covariance of the fitting.

Model fittings for the WT variant were done using the "lmfit" Python library [64] using the model described in [36,37]. Due to the complexity of the model, cylindrical core fittings were done by binning the data in 100 logarithmic bins to reduce computation time. Within the same model, core parameters (cylinder radius $R$ and cylinder length $L$) were set constant, to offset fitting errors. Initial values of $R$ and $L$ were calculated with the highest measured concentration. Physical boundary conditions were imposed on the fitting, and scattering length (SL) values were set to be unchanged by the fitting process. SL values of both the core and the tail domains were determined by tabulated values of amino-acid SLD in 100% $H_2O$ [65] (Table S3). Fitting parameter error evaluation was done by finding the covariant of the returning fitting parameters. Error calculation of the volume was done using: $\frac{dV}{V} = \sqrt{3\left(\frac{dR}{R}\right)^2}$. In addition, $\nu$ values of WT were found by a recursive search of the corresponding tail height $h/2$ over Eq. 2. Errors of $\nu$ were then found by assuming a simple case of $R_g = bN^\nu$, from which $d\nu \sim \frac{\ln(1+dR/R)}{\ln N} \sim \ln(N)^{-1}\frac{dR}{R}$

### Zimm analysis

Zimm analysis was performed as described in [45]. Data normalization was done by first determining $I_0$ by fitting a linear curve over the Guinier plot ($\ln I(q)$ vs $q^2$). Normalized $1/I(q)$ linear fitting was done starting with the earliest possible data point until a deviation from the linear behavior occurs. Data points were then binned for visual clarity without impacting the result.

### Brush model fitting

Brush height model as described in [41] was fitted with a prefactor $c = 0.33$ to match data. Resulting heights were converted to $\nu$ by $h = bN^\nu$ where $b = 0.38$ nm and $N = 146$. To accommodate for the change in grafting density, a linear curve was fitted to the grafting density's change in salinity and was used to obtain a continuous plot.

### Polyelectrolyte fitting

The fitting model was used as described in [40] with a pre-factor $c = 1.24$ to match data.

## Author contribution statement

M.K., G.K., and R.B. designed the project. M.K. conducted experiments and analysis with G.K.'s and O.A.S.'s assistance. M.K., G.K., R.B., and O.A.S. wrote the paper.

**Data availability statement** The raw SAXS data is available in the Small-Angle Scattering Biological Data Bank (SASBDB) at: https://www.sasbdb.org/project/2190/.

**Declarations**

**Conflict of interest** The authors claim no conflicting interests.

# References

1. A.S. Holehouse, B.B. Kragelund, The molecular basis for cellular function of intrinsically disordered protein regions. Nat. Rev. Mol. Cell Biol. **2023**, 1–25 (2023)
2. A. Chowdhury, D. Nettels, B. Schuler, Interaction dynamics of intrinsically disordered proteins from single-molecule spectroscopy. Annu. Rev. Biophys. **52**, 433–462 (2023)
3. B. Xue, A.K. Dunker, V.N. Uversky, Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. J. Biomol. Struct. Dyn. **30**, 137–149 (2012)
4. V.N. Uversky, Intrinsic disorder-based protein interactions and their modulators. Curr. Pharm. Des. **19**, 4191–4213 (2013)
5. T. Ehm et al., Self-assembly of tunable intrinsically disordered peptide amphiphiles. Biomacromolecules **24**, 98–108 (2022)
6. A. Laser-Azogui, M. Kornreich, E. Malka-Gibor, R. Beck, Neurofilament assembly and function during neuronal development. Curr. Opin. Cell Biol. **32**, 92–101 (2015)
7. A.A. Chernyatina, S. Nicolet, U. Aebi, H. Herrmann, S.V. Strelkov, Atomic structure of the vimentin central $\alpha$-helical domain and its implications for intermediate filament assembly. Proc. Natl. Acad. Sci. **109**, 13620–13625 (2012)
8. E. Malka-Gibor et al., Phosphorylation-induced mechanical regulation of intrinsically disordered neurofilament proteins. Biophys. J . **112**, 892–900 (2017)
9. N. Hirokawa, M.A. Glicksman, M.B. Willard, Organization of mammalian neurofilament polypeptides within the neuronal cytoskeleton. J. Cell Biol. **98**, 1523–1536 (1984)
10. C.R. Safinya, J. Deek, R. Beck, J.B. Jones, Y. Li, Assembly of biological nanostructures: isotropic and liquid crystalline phases of neurofilament hydrogels. Annu. Rev. Condens. Matter Phys. **6**, 113–136 (2015)
11. S. Müller-Späth et al., Charge interactions can dominate the dimensions of intrinsically disordered proteins. Proc. Natl. Acad. Sci. **107**, 14609–14614 (2010)
12. S. Milles, E.A. Lemke, Single molecule study of the intrinsically disordered fg-repeat nucleoporin 153. Biophys. J . **102**, 10a (2012)
13. N. Sekiyama, R. Kobayashi, T.S. Kodama, Toward a high-resolution mechanism of intrinsically disordered protein self-assembly. J. Biochem. **174**, 391–398 (2023)
14. S. Khatun et al., Fractal self-assembly and aggregation of human amylin. Soft Matter **16**, 3143–3153 (2020)
15. J.-E. Shea, R.B. Best, J. Mittal, Physics-based computational and theoretical approaches to intrinsically disordered proteins. Curr. Opin. Struct. Biol. **67**, 219–225 (2021)
16. R. Van Der Lee et al., Classification of intrinsically disordered regions and proteins. Chem. Rev. **114**, 6589–6631 (2014)
17. U. Baul, D. Chakraborty, M.L. Mugnai, J.E. Straub, D. Thirumalai, Sequence effects on size, shape, and structural heterogeneity in intrinsically disordered proteins. J. Phys. Chem. B **123**, 3462–3474 (2019)
18. R.K. Das, R.V. Pappu, Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. Proc. Natl. Acad. Sci. **110**, 13392–13397 (2013)
19. H. Hofmann et al., Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. Proc. Natl. Acad. Sci. **109**, 16155–16160 (2012)
20. W. Zheng, G. Dignon, M. Brown, Y.C. Kim, J. Mittal, Hydropathy patterning complements charge patterning to describe conformational preferences of disordered proteins. J. Phys. Chem. Lett. **11**, 3408–3415 (2020)
21. D. Maltseva et al., Fibril formation and ordering of disordered fus lc driven by hydrophobic interactions. Nat. Chem. **15**, 1–9 (2023)
22. G. Koren et al., Intramolecular structural heterogeneity altered by long-range contacts in an intrinsically disordered protein. Proc. Natl. Acad. Sci. **120**, e2220180120 (2023)
23. J.A. Riback et al., Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water. Science **358**, 238–241 (2017)
24. X. Zeng, K.M. Ruff, R.V. Pappu, Competing interactions give rise to two-state behavior and switch-like transitions in charge-rich intrinsically disordered proteins. Proc. Natl. Acad. Sci. **119**, e2200559119 (2022)
25. P.G. Argudo, J.J. Giner-Casares, Folding and self-assembly of short intrinsically disordered peptides and protein regions. Nanoscale Adv. **3**, 1789–1812 (2021)
26. N. Srinivasan, M. Bhagawati, B. Ananthanarayanan, S. Kumar, Stimuli-sensitive intrinsically disordered protein brushes. Nat. Commun. **5**, 5145 (2014)
27. S. Pregent et al., Probing the interactions of intrinsically disordered proteins using nanoparticle tags. Nano Lett. **15**, 3080–3087 (2015)

28. R. Beck, J. Deek, J.B. Jones, C.R. Safinya, Gel-expanded to gel-condensed transition in neurofilament networks revealed by direct force measurements. Nat. Mater. **9**, 40–46 (2010)

29. M. Kornreich, E. Malka-Gibor, B. Zuker, A. Laser-Azogui, R. Beck, Neurofilaments function as shock absorbers: compression response arising from disordered proteins. Phys. Rev. Lett. **117**, 148101 (2016)

30. A. Didonna, P. Opal, The role of neurofilament aggregation in neurodegeneration: lessons from rare inherited neurological disorders. Mol. Neurodegener. **14**, 1–10 (2019)

31. I.L. Morgan, R. Avinery, G. Rahamim, R. Beck, O.A. Saleh, Glassy dynamics and memory effects in an intrinsically disordered protein construct. Phys. Rev. Lett. **125**, 058001 (2020)

32. P. Pincus, Colloid stabilization with grafted polyelectrolytes. Macromolecules **24**, 2912–2919 (1991)

33. E. Zhulina, F. Leermakers, On the polyelectrolyte brush model of neurofilaments. Soft Matter **5**, 2836–2840 (2009)

34. G. Tria, H.D. Mertens, M. Kachala, D.I. Svergun, Advanced ensemble modelling of flexible macromolecules using x-ray solution scattering. IUCrJ **2**, 207–217 (2015)

35. W. Zheng, R.B. Best, An extended guinier analysis for intrinsically disordered proteins. J. Mol. Biol. **430**, 2540–2553 (2018)

36. J.S. Pedersen, C. Svaneborg, Scattering from block copolymer micelles. Curr. Opin. Colloid Interface Sci. **7**, 158–166 (2002)

37. J.S. Pedersen, Form factors of block copolymer micelles with spherical, ellipsoidal and cylindrical cores. J. Appl. Crystallogr. **33**, 637–640 (2000)

38. W.-L. Chen, R. Cordero, H. Tran, C.K. Ober, 50th anniversary perspective: Polymer brushes: Novel surfaces for future materials. Macromolecules **50**, 4089–4113 (2017)

39. C.-H. Wang, M.-B. Luo, X. Xu, C. Wang, L.-Z. Sun, Effects of salt concentration on the polyelectrolyte translocation through a cylinder nanopore. Eur. Polym. J. **121**, 109332 (2019)

40. B.-Y. Ha, D. Thirumalai, Conformations of a polyelectrolyte chain. Phys. Rev. A **46**, R3012 (1992)

41. N.A. Kumar, C. Seidel, Polyelectrolyte brushes with added salt. Macromolecules **38**, 9341–9350 (2005)

42. A. Zamyatnin, Protein volume in solution. Prog. Biophys. Mol. Biol. **24**, 107–123 (1972)

43. H.I. Okur et al., Beyond the hofmeister series: Ion-specific effects on proteins and their biological functions. J. Phys. Chem. B **121**, 1997–2014 (2017)

44. S. Wohl, M. Jakubowski, W. Zheng, Salt-dependent conformational changes of intrinsically disordered proteins. J. Phys. Chem. Lett. **12**, 6684–6691 (2021)

45. B.H. Zimm, The scattering of light and the radial distribution function of high polymer solutions. J. Chem. Phys. **16**, 1093–1099 (1948)

46. V.N. Uversky et al., Natively unfolded human prothymosin α adopts partially folded collapsed conformation at acidic ph. Biochemistry **38**, 15009–15016 (1999)

47. A. Möglich, K. Joder, T. Kiefhaber, End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation. Proc. Natl. Acad. Sci. **103**, 12394–12399 (2006)

48. R.V. Pappu, R. Srinivasan, G.D. Rose, The flory isolated-pair hypothesis is not valid for polypeptide chains: implications for protein folding. Proc. Natl. Acad. Sci. **97**, 12565–12570 (2000)

49. A. Quigley, D. Williams, The second virial coefficient as a predictor of protein aggregation propensity: a self-interaction chromatography study. Eur. J. Pharm. Biopharm. **96**, 282–290 (2015)

50. P.K. Nandi, D.R. Robinson, Effects of salts on the free energies of nonpolar groups in model peptides. J. Am. Chem. Soc. **94**, 1308–1315 (1972)

51. P.K. Nandi, D.R. Robinson, Effects of salts on the free energy of the peptide group. J. Am. Chem. Soc. **94**, 1299–1308 (1972)

52. E. Dickinson, Strategies to control and inhibit the flocculation of protein-stabilized oil-in-water emulsions. Food Hydrocoll. **96**, 209–223 (2019)

53. T.B. Blijdenstein, C. Veerman, E. van der Linden, Depletion- flocculation in oil-in-water emulsions using fibrillar protein assemblies. Langmuir **20**, 4881–4884 (2004)

54. F. Zosel, A. Soranno, K.J. Buholzer, D. Nettels, B. Schuler, Depletion interactions modulate the binding between disordered proteins in crowded environments. Proc. Natl. Acad. Sci. **117**, 13480–13489 (2020)

55. E.B. Zhulina, O.V. Borisov, Cylindrical brushes with ionized side chains: scaling theory revisited. Soft Matter **19**, 8440–8452 (2023)

56. H.P. Truong et al., Pincus blob elasticity in an intrinsically disordered protein. Eur. Phys. J. E **46**, 100 (2023)

57. Y. Li, R. Beck, T. Huang, M.C. Choi, M. Divinagracia, Scatterless hybrid metal-single-crystal slit for small-angle x-ray scattering and high-resolution x-ray diffraction. J. Appl. Crystallogr. **41**, 1134–1139 (2008)

58. N.P. Cowieson et al., Beamline b21: high-throughput small-angle x-ray scattering at diamond light source. J. Synchrotron Radiat. **27**, 1438–1446 (2020)

59. C.E. Blanchet et al., Versatile sample environments and automation for biological solution x-ray scattering experiments at the p12 beamline (petra iii, desy). J. Appl. Crystallogr. **48**, 431–443 (2015)

60. P. Pernot et al., Upgraded esrf bm29 beamline for saxs on macromolecules in solution. J. Synchrotron Radiat. **20**, 660–664 (2013)

61. A. Round et al., Biosaxs sample changer: a robotic sample changer for rapid and reliable high-throughput x-ray solution scattering experiments. Acta Crystallogr. D Biol. Crystallogr. **71**, 67–75 (2015)

62. J. Kieffer, V. Valls, N. Blanc, C. Hennig, New tools for calibrating diffraction setups. J. Synchrotron Radiat. **27**, 558–566 (2020)

63. P. Virtanen et al., Scipy 1.0 fundamental algorithms for scientific computing in python. Nat. Methods **17**, 261–272 (2020)

64. M. Newville, et al. Lmfit: Non-linear least-square minimization and curve-fitting for python. Astrophys. Source Code Libr. ascl-1606 (2016)

65. B. Jacrot, The study of biological structures by neutron scattering from solution. Rep. Prog. Phys. **39**, 911 (1976)