Molecular basis of polyglutamine-modulated ELF3 phase change in Arabidopsis temperature response

Richard J. Lindsay[1], Philip A. Wigge[2,3], Sonya M. Hanson[1]

1. Center for Computational Biology and Center for Computational Mathematics, Flatiron Institute, Simons Foundation, 162 5th Ave, New York, NY 10010, USA
2. Leibniz-Institut für Gemüse- und Zierpflanzenbau, Großbeeren, Germany
3. Institute of Biochemistry and Biology, University of Potsdam, Potsdam, Germany

ABSTRACT

Temperature is a major environmental variable influencing the distribution and behavior of plants. Recent advances have led to the identification of a role for the circadian clock in sensing temperature in *Arabidopsis thaliana*. Elongation growth and flowering are accelerated at warmer temperatures, and these effects are mediated by the circadian clock gene *EARLY FLOWERING 3 (ELF3)*. ELF3 exists with a tripartite protein complex called the Evening Complex (EC) that functions as a DNA transcription repressor targeting growth-related genes. ELF3, a large scaffold protein with disordered domains, binds to the transcription factor LUX ARRYTHMO (LUX) and ELF4 to form the EC. A crucial feature of ELF3 is that it acts as a highly sensitive thermosensor that responds directly and rapidly to small increases of temperature of about 5 ºC and is rapidly reversible. At temperatures of about 22 ºC and below, the EC is active, binding and repressing the promoters of multiple growth promoting genes, reducing their expression and cell elongation. At around 27 ºC and above ELF3 undergoes rapid and reversible phase change and protein condensate formation. This temperature-dependent activity causes EC occupancy on target genes to decrease at 27 ºC, allowing their increased expression. A C-terminal prion-like domain (PrD) is sufficient for ELF3 phase change and temperature responsiveness. The PrD region contains a polyglutamine (polyQ) repeat of variable length, the size of which has been found to modulate the thermal responsiveness as measured by hypocotyl (stem) elongation and condensate formation. How the PrD is able to respond to temperature is however poorly understood. To understand the underlying biophysical basis for ELF3 thermal responsiveness, we use a polymer chain growth approach to build large ensembles and characterize monomeric ELF3-PrD at a range of polyQ lengths and temperatures. We then explore temperature-dependent dynamics of wild-type ELF3-PrD, ELF3-PrD with the variable polyQ tract removed, and a mutant (F527A) using chain growth structures as initial conformations for replica exchange (REST2) simulations. In addition to different mechanisms of temperature sensing with and without the variable polyQ tract, we find increased solvent accessibility of expanded polyQ tracts, promotion of temperature-sensitive helices adjacent to polyQ tracts, and exposure of a cluster of aromatic residues at increased temperature, all three of which promote inter-protein interaction. These results suggest a set of potential design principles for the engineering of temperature dependent molecular interactions. This has considerable potential for biotechnological application in medicine and agriculture.

INTRODUCTION

The ability to sense and respond to a changing environment is critical to the survival of life. All organisms employ a variety of mechanisms to mitigate the impact of potential stressors. These sensors are especially important in plants, as they need to respond to stressors in place. They possess, for instance, various mechanisms to modulate flowering in response to changing photoperiod and growth direction in

association with shade avoidance.[1,2] Many of these types of responses rely on the regulation of gene expression in response to extracellular signals. A classic example of an environmental response that regulates gene expression is the lac operon in prokaryotes,[3] in which a transcription inhibitor binds the lac promoter until it senses lactose, upon which it is released, allowing transcription and lactose metabolism to occur. Transcriptional repression systems are also important in eukaryotes, for example in *Arabidopsis thaliana*, the Evening Complex (EC), employs a similar mechanism to modify plant growth in response to ambient temperature.[4]

In *Arabidopsis thaliana*, the EC, a tripartite protein complex, provides a tunable temperature-sensing mechanism with implications for growth rate, flowering, and potentially senescence and circadian rhythm.[4–6] The EC is a circadian clock component of *Arabidopsis* that integrates temperature and temporal information from the environment to regulate seasonal growth and flowering.[7–9] The entire complex comprises a transcriptional repressor with three known components, ELF3, ELF4 and LUX ARRYTHMO (Fig. 1A). The largest and most functionally important component is ELF3. ELF3 is a 695 residue long intrinsically disordered protein (IDP) that inhibits expression of growth genes, like PIF4 and PIF5, by occluding the binding of transcription machinery. Another EC component, the transcription factor LUX ARRYTHMO, is a small protein responsible for EC DNA-binding specificity by guiding the complex to specific gene targets.[8,10] The third EC component, ELF4, is a small helical protein that stabilizes the ELF3/LUX/DNA complex and binds to the only region of ELF3 predicted to be structured.[7,8,11,12] While the precise function of ELF4 is not fully understood, its overexpression has been shown to enhance DNA binding and eliminate temperature sensitivity of the EC.[12] In wild-type *Arabidopsis thaliana* under normal cellular conditions, the EC binds and represses transcription of its targets until the temperature increases to about 27°C. At higher temperatures, ELF3 begins to form reversible nuclear condensates, sequestering ELF3, releasing the inhibitory complex, and allowing the expression of growth genes. An illustrated overview of the proposed temperature-sensing mechanism of the EC is shown in Fig. 1A.

ELF3 contains a C-terminal prion-like domain, first recognized by Jung *et. al,* which is necessary and sufficient for condensate formation and makes up a quarter of the ELF3 sequence (173 out of 695 residues).[4] Prion-like domains (PrD's) are intrinsically disordered regions of low sequence complexity known to promote liquid-liquid phase separation.[13] While the EC is conserved among land plants, its temperature sensing role appears to be under selection, since for example ELF3 in the Mediterranean grass species *Brachypodium distachyon* is functional in *Arabidopsis thaliana*, but has almost no PrD and does not undergo phase change or accelerated growth at higher temperatures. A central feature of ELF3-PrD in *Arabidopsis thaliana* is the presence of a poly-glutamine (polyQ) tract of variable length found to differ in *Arabidopsis* populations by geographic location.[9] Experiments by Jung *et. al* found the length of this polyQ tract modulates the sensitivity of the temperature response, observing that longer polyQ enhances condensate formation and hypocotyl growth rate at high temperature.[4] Previous studies have shown that polyQ tracts often enhance helical propensity in adjacent regions,[14] which is of particular interest because short regions of transient helicity called Short Linear Motifs (SLiMs) are often responsible for formation of intra- and inter-protein interactions necessary for biological function.[15]

In this study, we seek to understand the molecular basis of polyglutamine-mediated temperature sensitivity in the PrD of ELF3, and to articulate a physical mechanism by which polyglutamine tract length modulates condensate formation. We first build ensembles of ELF3-PrD with varying polyQ lengths at a range of representative temperatures using a fast and efficient chain-growth algorithm.[16,17] Next, we seek to better understand the role of the variable polyQ tract as well as a specific residue pair responsible for

the conformational variability of the protein by performing all-atom replica exchange with solute tempering (REST2) MD simulations on wild-type ELF3-PrD, a mutant of interest F527A, and ELF3-PrD with the variable polyQ tract removed.[18] Through this computational lens we are able to better understand the role and molecular mechanism of polyQ-adjacent helices and aromatic residues in the temperature-sensitive growth response of *Arabidopsis thaliana* mediated by the ELF3-PrD.

METHODS

Hierarchical Chain Growth Method and Simulation Details

The sequence for *Arabidopsis thaliana* ELF3 was obtained from UniProt. The PrD sequence (residues 432 to 604) was isolated and divided into 57 segments. Each segment was five amino acids in length including two residues that overlap with the previous segment and two residues that overlap with the subsequent segment. Each of the 57 segments was capped with an N-terminal acetyl group and a C-terminal N-methyl group to neutralize interactions of the charged ends of the protein. Segments were parametrized using the Amber03ws forcefield[19] and placed in a periodic rectangular box with 1nm of padding on each side of the protein and solvated with TIP4p2005 water molecules.[20]

Each segment was minimized until the maximum force reduction dipped below 1000 kJ/mol/nm per minimization step. For each of the 57 segments, 24 replicates were equilibrated to various temperatures ranging from 290K to 405K for 10ns. For this step, temperature effects were implemented using the v-rescale thermostat[21] and the Berendsen barostat.[22] Finally, replica exchange MD was run for each segment for 100ns using GROMACS version 2021.1.[23] For this production run, the v-rescale thermostat was used to maintain temperature and the Parrinello-Rahman barostat[24] was used to maintain a pressure of 1 bar.

Once the fragment simulations were finished, we used the hierarchical chain growth (HCG) application, developed by Pietrek et al, to construct a 7Q ELF3-PrD ensemble from these 57 segments consisting of 32,000 unique conformations.[16,17] Conformers were randomly chosen from each segment and joined in a pairwise fashion to create a set of fragment dimers. These were then joined into tetramers, octamers, and so on until the full 57-segment ELF3-PrD was reconstructed. This process was repeated until we had obtained ensembles with polyQ tracts of 0, 7, 13 and 19 residues in length each at temperatures of 290K, 300K, 320K and 415K for a total of 16 ensembles. Due to the two-residue overlap, we were required to run 19 segments made up of residues C-terminal to the polyQ region in order build the 0Q, 13Q and 19Q systems.

REST2 Simulation Details

The HCG method enabled us to quickly look at large ensembles of different polyQ conditions for potential structural differences and pointed us towards some regions and residues of interest. The usefulness of the HCG method, for our purposes, is limited by the absence of long-range interactions. In order to accurately study the dynamics of ELF3-PrD we turn to REST2 (Replica Exchange with Solute Tempering), an all-atom enhanced sampling technique capable of exploring longer effective time scales than traditional all-atom MD.[25] Much like temperature replica exchange (T-REMD), REST2 runs multiple replicas in parallel, each at a different effective temperature. Every few hundred MD steps, replicas attempt to swap with one another with the probability of acceptance tied to the energy difference between replicates.[18] Unlike T-REMD where temperatures values are set for each replica, REST2 modulates the strength of protein-protein and protein-water interactions to achieve an effective temperature, a feature

which reduces the number of replicas required compared with T-REMD. While REST2 can still be resource intensive, it is one of few methods able to provide Boltzmann-distributed ensembles of IDPs at biologically relevant timescales.

REST2 simulations[18] were performed of the full-length wild-type ELF3-PrD, the ELF3-PrD F527A mutant and the 0Q ELF3-PrD with the variable polyQ tract removed. For the initial structure of these simulations, we chose an especially helical conformer from our HCG wild-type ensemble created at 290K. For the F527A mutant, F527 was manually edited to alanine by removing side-chain atoms and changing the residue name in the structure file. Each system was inserted into a rectangular periodic box with 1nm of padding around each side of the protein and treated with the Amber03ws forcefield and the TIP4p2005 water model. The wild-type system contained 2606 protein atoms, 122623 TIP4P water molecules, 439 Na atoms and 342 Cl atoms for a total of 490489 atoms. The F527A mutant matched this setup but with 2596 protein atoms for a total of 493769 atoms. The 0Q system contained 2487 protein atoms, 210607 water molecules, 237 Na atoms and 240 Cl atoms for a total of 846088 atoms. A 50ns equilibration run was performed for each system using the v-rescale thermostat and Berendsen barostat. After equilibration, each protein atom was marked within the base topology file as a "hot atom" and replicas were created by scaling interactions involving these "hot atoms" by a lambda value corresponding to temperatures in the range of 290K to 405K. The wild-type and F527A systems each had 20 replicas while 0Q had 12 replicas in order to achieve an exchange probability between 35% and 45%. REST2 production runs of each system were performed for 500ns using the v-rescale thermostat and Parrinello-Rahman barostat. Convergence was verified using a split analysis of the radius of gyration values for each system at multiple temperatures (Fig. S1).

Statistical Analysis of Ensemble Contact Maps

We used two types of contact analysis to identify mechanistically important interactions and characterize differences in our ELF3-PrD chain-growth ensembles. Both approaches, E-PCA and I-PCA,[26] are part of the CAMERRA family of contact analyses developed by Shen *et. al*[26] and seek to explain sources of conformational variance of protein ensembles. In E-PCA, contact maps from each conformation of the ensemble are used to produce a contact covariance matrix of each possible residue pair, i and j, with all other residue pairs, k and l. PCA is performed using this covariance matrix as input and the top PCs are returned. This method has been used to uncover coordinated interactions underlying mechanisms of allostery, cooperative ligand binding, and general collective motions of a variety of proteins.[27–30] Due to the absence of long-range interactions in the HCG method used to generate our ensembles, E-PCA signals will be more prominent for local interactions. The dynamics of folded proteins tend to be dominated by the top few modes (PCs), however, intrinsically disordered proteins exhibit a slower eigen decay in which the first several PCs may contribute nearly equally (Fig. S2).

The second CAMERRA variant, I-PCA, provides information on the packing of residues and can identify regions of the sequence in which residues tend to interact more frequently. In this approach, individual residue pairs are considered against the rest of the protein, resulting in a covariance matrix of size $N \times N$ as opposed to $N^2 \times N^2$ for the E-PCA variant. This approach has been prominently used in studies of chromatin structure to differentiate transcriptionally active and inactive regions (euchromatin and heterochromatin, respectively) from Hi-C maps.[31] When applied to natively structured proteins I-PCA tends to identify the individual domains of the protein and while IDPs do not contain structured domains, I-PCA can nevertheless identify dynamically associated domains.[32]

RESULTS

Polyglutamine Tract Promotes Helicity in Proximal Residues

In order to understand how polyQ length and temperature change the behavior of ELF3-PrD, we examined the structural differences between PrD ensembles with varying polyQ tract lengths constructed at different representative temperatures. Using the hierarchical chain growth (HCG) method by Pietrek *et. al,*[16] we constructed ensembles of polyQ of 0, 7, 13 and 19 glutamine residues in length each at temperatures of 290K, 300K, 320K and 415K to produce a total of 16 ensembles with 32,000 conformations each, a total of 512,000 structures. For each residue we summed the conformers with alpha, 3_{10} or pi-helix character to obtain a total helical propensity (Fig. 1B). We found ELF3-PrD to be composed of short transient helices, similar to the 'SLiMs' (Short Linear Motifs) found in other disordered proteins, with most having helical propensity of around 5%. Though most of the SLiMs in ELF3-PrD had only light helical character, the helices adjacent to the variable-length polyQ tract had significantly higher propensities, up to 30% for the N-terminal residues and 10% for the C-terminal residues. These regions display some helicity with or without polyQ as seen in the 0Q system, but the polyQ tract seems to increase the portion of the ensemble with helical character. [14] An additional feature of the polyQ-adjacent SLiMs is a degree of temperature-sensitivity not seen elsewhere in the PrD. In the 7Q system, the N-terminal polyQ adjacent SLiM falls from ~30% at 290K to 20% at 300K and finally 10% propensity at 415K, while the same region in the 0Q system shows no appreciable reduction in propensity until 415K. The increase in temperature-responsiveness of these SLiMs in 7Q could be a by-product of the higher possible helical propensity enabled by the polyQ tract. Ensembles with polyQ tracts of 13 and 19 residues showed nearly identical results to the 7Q system (Fig. S3). This could be due to a limitation of the 5-residue fragments simulated to build libraries for the HCG method. Overall, it is clear the presence of a polyQ tract has a significant influence on the protein's local structural character and temperature responsiveness.

Contact Analysis of Chain-Growth Ensembles Reveals Spatial Segregation of ELF3-PrD

E-PCA[26] is a contact analysis method that isolates the top sources of conformational variance (or dynamics when used on MD trajectories) by identifying residue pairs that form and break in a concerted manner. We apply it here to our HCG-generated ELF3-PrD ensembles to identify regions of correlated contacts which may be of mechanistic importance. For each system, the biggest signal consistently comes from three SLiMs directly N-terminally adjacent to the polyQ tract, suggesting correlated contact formation or breaking of contacts in this region explains most of the conformational variance, at least partially due to helix formation. At 290K, PC1 of all polyQ-containing systems (Fig. 1C, left panel) contains a bimodal signal where each peak corresponds to interactions of residues within a polyQ-adjacent SLiM, suggesting some level of communication between these regions. The dominant mode of the 0Q system, on the other hand, exhibits a single peak, indicating communication between these SLiMs plays a less significant role in the contact dynamics of this system. At the maximum temperature, 415K, the contact dynamics of all systems converge, regardless of polyQ presence or length, as evidenced by the nearly identical PC1 values at this temperature (Fig. 1C, right panel). This peak at the high-temperature condition indicates formation of contacts that comprise the third polyQ N-terminal SLiM. However, one prominent negative value is consistently observed indicating interaction between F527 and Y530 strongly opposes contact formation within this SLiM. The persistence of this signal across ensembles and its location in the region identified as the primary source of conformational variability suggests a potential role for this contact pair in the general mechanism of ELF3-PrD. An overview of residues of interest in ELF3-PrD, including residue F527

and the N-terminal polyQ-adjacent SLiMs, is presented in Fig. 1D. E-PCA data for more modes, systems and temperatures is found in Fig. S4, and further contact analysis exploring the spatial organization of our ELF3-PrD HCG ensembles using the I-PCA approach can be found in Supplementary Information (Fig. S5).

Enhanced All-Atom MD Reveals Poly-glutamine Modulates Helical Propensity in a Temperature-Sensitive Manner

While the static conformers generated by the HCG approach provided a relatively quick and resource-inexpensive look at the effects of varying the polyQ tract length and temperature on the local structure of the ELF3-PrD, we turned to a more intensive MD simulation approach to gain a more complete understanding of dynamics underlying ELF3 function. Based on trends seen in the initial HCG results, we decided to run three separate all-atom MD simulations using the REST2 method: one of the wild-type ELF3-PrD to obtain a better understanding of its underlying dynamics, one of the F527A mutant to understand the contribution of F527 to condensate formation and temperature responsiveness, and one of the 0Q system to better grasp the mechanistic role of the variable polyQ tract. As with the HCG ensembles, helical propensity was calculated for both the WT and F527A systems at four separate effective temperatures (Fig. 2A and B, respectively). Helical propensity around the variable-length polyQ tract is similar in trend to the HCG ensemble but with significantly larger propensity values, in the range of 30-65% instead of the 10-30%, for the N-terminal adjacent helix (Fig. 2A). Additionally, some level of helical enhancement was observed adjacent to all three polyQ tracts of ELF3-PrD rather than only the first as in the HCG case (Fig. 2A/B).[14] The second polyQ tract (WT residues 568 to 572) exhibits enhancement in propensity and temperature-responsiveness in its N-terminal SLiM, and the third and last polyQ tract (WT residues 581 to 586) sees minor enhancement in its C-terminal SLiM, increasing to 10% or about double the baseline propensity. The concerted nature of the enhancement of helicity and temperature-responsiveness may be more than a coincidence as enhancement of potential helical propensity could increase the ability to sense temperatures by providing a greater range of signals with more differentiable outcomes. The enhancement of polyQ-adjacent helices does show a variety of structural outcomes dependent on sequence context, as factors like the propensity of helicity and which regions express propensity vary among the three tracts. The impact of sequence is not just local, as is apparent from the structural analysis of the F527A system. Relative to the WT, the F527A mutation attenuates helices, including those around each polyQ tract, despite being 15+ residues from the nearest tract. Variables like tract length,[14] sequence context,[33] tract location, and proximity to other polyQ tracts could all affect protein character and warrant further study.

Aromatic Interactions Play Major Role in Temperature-Sensitive Conformation Change

The REST2 simulations painted a more complete picture of the ELF3-PrD dynamics and revealed some mechanistically important features absent in the HCG ensembles, partly due to the inclusion of long-range interactions. The most prominent was a 10-residue helix of heavy aromatic character (residues 494 to 504, "H_{aro}") which was formed between 40-52% of the time in the REST2 WT simulations (Fig. 2A) but only present at baseline levels in the HCG ensembles. The F527A mutant sees a significant reduction in the size of this helix (Fig. 2B), suggesting long-range interactions with F527 may stabilize H_{aro}. Indeed, it seems this interaction may have consequences for the global conformation of ELF3-PrD. In the WT, both the regions around F527 and H_{aro} are largely buried, as shown by the low solvent-accessible surface area (SASA) in Fig. 2C, and their relative immobility as exhibited by RMSF values shown in Fig. 2D. A methionine residue of H_{aro}, M73, promotes compaction of this region through interaction of the sulfur atom with the ring of

multiple local aromatic (tyrosine) residues (Fig. S6), a common but often overlooked motif.[34] Upon mutation of F527 to alanine, both the F527 and H_{aro} regions become relatively solvent-exposed and flexible. This interaction influences the global shape of the protein by pushing the N-terminal region towards a globular conformation while the C-terminal region remains relatively extended. This is demonstrated by comparing the radius of gyration (RG) of the first 100 and last 73 residues for the WT and mutant systems, as in Fig. 2E. The first 100 residues of the WT (top panel) have sharp peak indicating a compactness not observed in the (mutant bottom panel) while the last 73 residues sample a wide range of RG values in each case, indicating a more expanded and flexible nature.

Upon discovering the global structural influence of the long-range interaction between F527 and H_{aro}, we wondered what role it may play in temperature-responsiveness. We examined the minimum distances between atoms of F527 and H_{aro} under different temperature conditions. We found a linear increase in F527 and H_{aro} distance (and decrease in interaction) with increasing temperature, as demonstrated by the histograms in Fig. 3A. To get a better sense for the role of this temperature-responsive property in the conformational landscape of the protein, we plotted the F527- H_{aro} distance vs the radius of gyration in Fig. 3B. At 290K, the WT is largely homogenous, mostly constrained to a single population of tightly interacting F527-H_{aro} and compact structures. However, with increasing temperature the ensemble shifts towards a second population that is slightly expanded with the F527-H_{aro} contacts broken. Representative structures of the low-temperature and emergent high-temperature populations are depicted in Fig. 3C and Fig. 3D, respectively. The F527A mutant reinforces the importance of F527 in the WT conformational landscape as F527A samples a diverse range of populations in this space. However, the most persistent F527A population has significant overlap with the WT high temperature population, suggesting the F527A mutant may mimic the phenotype of WT ELF3-PrD at high temperatures.

The temperature-dependent breaking of the F527/H_{aro} interaction is particularly interesting due to its potential relevance to ELF3-PrD condensate formation. A significant increase in total SASA is observed in the high-temperature population with a major source being aromatic residues (Fig. S7). As previously mentioned, H_{aro} and F527 are regions of enhanced aromatic residue density, as seen from the dashed red lines in Fig. 3E. Breaking the interaction between F527/H_{aro} frees up 5 aromatic residues (Y496, Y499, Y500, Y503 & F527). Indeed, our high-temperature population exhibits enhanced solvent accessible surface area for these residues and for the majority of the 18 aromatic amino acids present in the ELF3-PrD (Fig. 3E). Given the well-established role of aromatic residues in driving interactions in low complexity disordered protein regions,[35] it is reasonable to expect an increase in condensate formation propensity when these residues are accessible by other ELF3-PrD monomers. Indeed, previous work has established a link between single-chain dimensions and phase-separation propensity as a function of aromatic character.[36,37]

Polyglutamine Solvation Dynamics Contribute to ELF3-PrD Phase-separation Propensity

A key question to understand temperature sensing by ELF3 is how expansion of the polyQ tract enhances the phase-separation propensity of ELF3-PrD, thus increasing the temperature sensitivity of the molecule.[4] One driver is the increasing availability of glutamine on the protein surface with increasing tract length. Plotting the SASA values for our HCG ensembles (Fig. 4A) highlights all three polyQ tracts as regions of persistent high solvent accessibility. While the SASA value of polyQ regions is only moderately higher per residue than the average (indicated by a dashed black line), it persists for the length of the tract, creating patches of high solvent accessibility not seen elsewhere in the PrD. Solvent accessibility is

particularly consequential here because water is a poor solvent for glutamine, leading to self-interaction of polyQ regions.[38] SASA values for these HCG ensembles were temperature-invariant for each case. This data suggests polyQ tract expansion provides a direct and tunable means to enhance PrD phase separation in response to temperature.

Solvation dynamics serve as a determining factor in condensate formation behavior, in part by determining which regions remain accessible to the solvent and other proteins, and changes in temperature can alter these dynamics. To characterize how hydration differs across ELF3-PrD we obtained radial distribution functions (RDFs) of water oxygen atoms around ELF3-PrD as a whole and around each individual residue. Values for a per-residue solvation plot were obtained by using each residue as the reference point for an RDF calculation of water oxygen atoms and plotting the height of the first peak, representing the most tightly bound water molecules (Fig. 4B). At 405K, many residues exhibit a decrease in local solvation, with the most prominent and continuous decreases seen in the three polyQ tracts. Further investigation of ELF3-PrD polyQ interactions at high temperature show a tradeoff wherein a decrease in polyQ-water hydrogen bonding is compensated by an increase in protein-protein hydrogen bonds, even though the number of polyQ-proximal water molecules increases (Fig. S8). This indicates the tendency of polyQ tracts to interact intensifies as temperature rises. An illustration of water/protein contact dynamics is presented in Fig. 4C. Taken together that, at high temperature, polyQ tracts in this context 1) remain solvent accessible but 2) becomes less solvated while 3) becoming more extended (Fig. S9), we infer that protein-protein interactions between polyQ regions become more energetically favorable with increasing temperature.

Basis of Temperature-Responsiveness Differs in Absence of Variable PolyQ Tract

A key finding by Jung *et. al* observed limited temperature-responsiveness in the absence of the variable polyQ tract as measured by hypocotyl elongation. We sought to gain some insight into the mechanistic differences resulting from the absence of the variable polyQ tract in ELF3-PrD by running a REST2 simulation of the 0Q system, identical to the WT but with the first polyQ tract removed. A secondary structure analysis of this trajectory reveals drastic structural deviation due to the absence of the polyQ (Fig. 5A), including the abolition of the major temperature sensitive helix, H_{aro}, in the N-terminal end and the introduction of a new temperature-responsive helix near the C-terminus, pointing to fundamental differences in the temperature-sensing mechanism compared with WT.

While the 0Q system is reported to be significantly less sensitive to changes in temperature, ELF3-PrD retains some temperature responsive interactions and structural elements in the absence of this region. Our proposed mechanism for temperature responsiveness in the WT ELF3-PrD consisted of the H_{aro} region interacting with F527 in a temperature-sensitive manner, but in the 0Q system, H_{aro} and F527 interact less than 1/5th as frequently at 290K with little sensitivity to temperature at biologically relevant timescales, though an increase is seen at 405K (Fig. 5B). However, a new temperature-sensitive interaction is spotted between H_{aro} and the C-terminal polyQ tract, revealed by minimum distance distributions between these two regions (Fig. 5C). The extensive interaction at 290K is half as populated at 300K, making it particularly sensitive at biologically relevant temperature ranges. Each additional 10K increase in temperature leads to further reduction in contacts between this polyQ tract and H_{aro}.

In addition to the temperature-responsiveness of the H_{aro}/C-terminal polyQ interaction, another structural mechanism for temperature sensitivity arises in the form of a large helix between residues 554 and 566 (Fig. 5A). This helix exhibits impressive stability at 290K (77% helicity) and rapidly declines linearly

as temperature increases, bottoming out at 35% at 405K. This 0Q-native helix includes the first (second in WT) polyQ tract with the bulk of the helix forming N-terminal to polyQ. What was the most prominent helix in the WT is attenuated here in the absence of the variable polyQ tract, a characteristic also observed in the HCG ensembles. Between the 0Q-native helix and the temperature-responsive nature of the interaction between F527 and the third polyQ tract, it appears the 0Q system responds to temperature change by a mechanism vastly different from that of the WT.

DISCUSSION

Through this series of computational studies, we sought to better understand the molecular basis for the ELF3-PrD-based temperature response in plants such as *Arabidopsis thaliana*. In initial studies with the HCG method we saw evidence for temperature-dependent polyQ-adjacent helices and the importance of an aromatic residue. Our REST2 simulations confirmed these results and further helped elucidate the potential molecular basis of this temperature sensing mechanism as we observed increased solvent accessibility of expanded polyQ tracts, promotion of temperature-sensitive helices adjacent to polyQ tracts, and exposure of a cluster of aromatic residues at increased temperature, all three of which promote inter-protein interaction.

From these studies we propose a couple of ways in which the polyQ tract plays a significant role in ELF3-PrD aggregation. One is by promoting the formation of transient helices, SLiMs, in polyQ-proximal regions. SLiMs are known to be hotspots of protein-protein interaction[39], and in the systems studied here, SLiM helix propensities are attenuated by increasing temperature. This suggests a potential mechanism for destabilizing the EC and allowing ELF3 to dissociate. The polyQ tract also serves as a modulator of aggregation. PolyQ itself is capable of inter-protein interaction with other polyQ tracts[40], and increasing tract length enhances aggregation through a linear increase in polyQ solvent accessibility. It is notable that the polyQ tracts observed here lacked any persistent contact with other parts of the protein making them available for inter-protein interactions.

Another proposed contributor to ELF3 aggregation is the enhanced solvent exposure of aromatic residues with increasing temperature. At high temperature, the ELF3-PrD ensemble undergoes a conformational shift favoring conformations with several aromatic residues exposed to the solvent. The breaking of the interaction between F527 and H_{aro} alone drastically increases exposure of 5 aromatics, and in the population of structures which appears at high temperature, 8 additional aromatic residues see increases in SASA, representing 12/18 aromatic residues in the PrD. The total SASA of this population of ELF3-PrD sees a significant overall increase, presenting more opportunities for protein-protein interactions. The proposed mechanism of the ELF3-PrD-dependent temperature-sensing mechanism can be understood to be, in large part, a result of temperature-dependent conformational changes from the breaking of the F527-H_{aro}.

To make this study tractable we have studied only one monomer of the ELF3-PrD to make inferences about its temperature-dependent aggregation behavior in condensates. While we believe these monomer-level studies to have been informative, an obvious next step towards understanding the aggregation dynamics of this system is to expand our simulations to include multiple copies of the IDP in a simulation, as has previously been performed using a coarse-grain approach to study condensation of low complexity domains[41,42,43]. Additionally here we have chosen in our monomer modeling scheme, to use the HCG method combined with REST2, which we believe has given us adequate access to the temperature-dependent dynamics we were interested in for this system, given our choice of force field,[19]

compared to other popular methods for sampling IDP conformations.[44,45] Interestingly, while our HCG models of the ELF3-PrD exhibited helical character, they seemed to underestimate the secondary structure content compared to REST2 molecular dynamics. This illustrates that chain-growth Monte Carlo methods can be useful starting point to find and characterize short linear motifs in atomically-detailed ensembles of disordered proteins, but that an in-depth characterization by more in depth molecular dynamics and experiment where available is required to fully characterize local structural propensities.[17,46,47] Certainly there are caveats in our approach, most notably there is some evidence that the REST2 protocol favors condensed states of IDPs over more extended states[48] not to mention that MD force-fields are not typically tuned for optimal quantitative results at a wide range of temperatures. Nonetheless, with these limitations in mind, we believe our results to be relevant at biological temperature ranges and to have helped us develop a better understanding of this system.

Our atomistic approach has enabled the identification of molecular and thermodynamic contributions to the temperature sensitive properties of this disordered protein. While dissecting the entropic and enthalpic contributions to this process from molecular simulations is beyond the scope of this current work,[49] with atomistic molecular dynamics simulations we can be confident that we are sampling our statistical ensemble to the best approximation of our force field.[50] To this end our atomistic simulations are able to capture the enthalpic contributions of electrostatics and the contributions from conformational entropy with capabilities beyond those of coarse-grained or phenomenological simulations, such as just considering residues as either stickers or spacers. For example, we are able to observe that the interactions between the aromatic residues and the H_{aro} become less favorable at higher temperatures. Additionally, we are able to see that the waters creating close contacts with the polyQ tracts at low temperature becomes less favorable as entropic effects start to dominate our system at higher temperatures. Overall, our findings are a step toward understanding the key role disordered proteins play in biological temperature sensing, as they are dominated by different thermodynamic effects than ordered proteins, in which the hydrophobic effect, virtually absent in IDPs, plays a key role. Additionally, how the role of the large conformational entropy of IDPs, found to play a significant role in IDP protein-protein interactions,[51] and enthalpic forces balance each other out within the context of temperature-dependent phase separation remains to be fully understood. Most importantly, the statistical mechanical toolbox of molecular dynamics simulation has facilitated the identification of potential key residues in ELF3 to inform experimental assays *in planta* to validate and expand upon our findings.

Understanding the connection between polyQ length and aggregation behavior has important implications beyond temperature sensing in plants. The link between polyQ length and aggregation behavior of proteins containing PrD's is well documented,[52–54] with much of the work in the area focused on aberrant plaque formation associated with at least nine neurodegenerative diseases. More recently, polyQ-modulated reversible phase change of PrD's has been recognized as a ubiquitous biological mechanism promoting membrane-less organelle formation in eukaryotes, enabling a variety of regulatory functions in healthy cells.[13,55–57] Phenotypic variants arising from differences in polyQ tract length of aggregation-prone environment-sensing proteins have been observed in a genetically diverse range of species, from temperature sensing in fungi to influencing pacific salmon mating age and location.[53,58] Understanding the underlying principles governing the behavior of thermosensory proteins offers the potential to design unique temperature responsive networks that may be of considerable value in medicine and agricultural biotechnology.

DATA AVAILABILITY

All code used and molecular trajectories generated in this study are available via: https://github.com/flatironinstitute/ELF3.

ACKNOWLEDGEMENTS

We are grateful to the help of Lisa M. Pietrek, Lukas S. Stelzl, and Gerhard Hummer for their support in the use of the HCG method during the early stages of this project. Thanks to Lucy Reading-Ikkanda for contributions to figure illustrations and design. S.M.H was supported by a Humboldt Research Fellowship for Experienced Researchers from the Alexander von Humboldt Foundation during the inception and preliminary stages of this research. The Flatiron Institute is a division of the Simons Foundation.

(1) Morelli, G.; Ruberti, I. Light and Shade in the Photocontrol of Arabidopsis Growth. *Trends Plant Sci.* **2002**, *7* (9), 399–404. https://doi.org/10.1016/S1360-1385(02)02314-2.
(2) Hayama, R.; Coupland, G. The Molecular Basis of Diversity in the Photoperiodic Flowering Responses of Arabidopsis and Rice. *Plant Physiol.* **2004**, *135* (2), 677–684. https://doi.org/10.1104/pp.104.042614.
(3) Oehler, S.; Eismann, E. R.; Krämer, H.; Müller-Hill, B. The Three Operators of the Lac Operon Cooperate in Repression. *EMBO J.* **1990**, *9* (4), 973–979. https://doi.org/10.1002/j.1460-2075.1990.tb08199.x.
(4) Jung, J.-H.; Barbosa, A. D.; Hutin, S.; Kumita, J. R.; Gao, M.; Derwort, D.; Silva, C. S.; Lai, X.; Pierre, E.; Geng, F.; Kim, S.-B.; Baek, S.; Zubieta, C.; Jaeger, K. E.; Wigge, P. A. A Prion-like Domain in ELF3 Functions as a Thermosensor in Arabidopsis. *Nature* **2020**, *585* (7824), 256–260. https://doi.org/10.1038/s41586-020-2644-7.
(5) Box, M. S.; Huang, B. E.; Domijan, M.; Jaeger, K. E.; Khattak, A. K.; Yoo, S. J.; Sedivy, E. L.; Jones, D. M.; Hearn, T. J.; Webb, A. A. R.; Grant, A.; Locke, J. C. W.; Wigge, P. A. ELF3 Controls Thermoresponsive Growth in Arabidopsis. *Curr. Biol.* **2015**, *25* (2), 194–199. https://doi.org/10.1016/j.cub.2014.10.076.
(6) Zagotta, M. T.; Hicks, K. A.; Jacobs, C. I.; Young, J. C.; Hangarter, R. P.; Meeks-Wagner, D. R. The Arabidopsis ELF3 Gene Regulates Vegetative Photomorphogenesis and the Photoperiodic Induction of Flowering. *Plant J* **1996**, *10* (4), 691–702. https://doi.org/10.1046/j.1365-313x.1996.10040691.x.
(7) Zhao, H.; Xu, D.; Tian, T.; Kong, F.; Lin, K.; Gan, S.; Zhang, H.; Li, G. Molecular and Functional Dissection of EARLY-FLOWERING 3 (ELF3) and ELF4 in Arabidopsis. *Plant Sci.* **2021**, *303*, 110786. https://doi.org/10.1016/j.plantsci.2020.110786.
(8) Nusinow, D. A.; Helfer, A.; Hamilton, E. E.; King, J. J.; Imaizumi, T.; Schultz, T. F.; Farré, E. M.; Kay, S. A. The ELF4-ELF3-LUX Complex Links the Circadian Clock to Diurnal Control of Hypocotyl Growth. *Nature* **2011**, *475* (7356), 398–402. https://doi.org/10.1038/nature10182.
(9) Undurraga Soledad, F.; Press Maximilian, O.; Legendre, M.; Bujdoso, N.; Bale, J.; Wang, H.; Davis Seth, J.; Verstrepen Kevin, J.; Queitsch, C. Background-Dependent Effects of Polyglutamine Variation in the Arabidopsis Thaliana Gene ELF3. *Proc. Natl. Acad. Sci.* **2012**, *109* (47), 19363–19367. https://doi.org/10.1073/pnas.1211021109.
(10) Zhang, C.; Gao, M.; Seitz, N. C.; Angel, W.; Hallworth, A.; Wiratan, L.; Darwish, O.; Alkharouf, N.; Dawit, T.; Lin, D.; Egoshi, R.; Wang, X.; McClung, C. R.; Lu, H. LUX ARRHYTHMO Mediates Crosstalk between the Circadian Clock and Defense in Arabidopsis. *Nat. Commun.* **2019**, *10* (1), 2543. https://doi.org/10.1038/s41467-019-10485-6.
(11) Huang, H.; Nusinow, D. A. Into the Evening: Complex Interactions in the Arabidopsis Circadian Clock. *Trends Genet.* **2016**, *32* (10), 674–686. https://doi.org/10.1016/j.tig.2016.08.002.
(12) Silva Catarina, S.; Nayak, A.; Lai, X.; Hutin, S.; Hugouvieux, V.; Jung, J.-H.; López-Vidriero, I.; Franco-Zorrilla Jose, M.; Panigrahi Kishore, C. S.; Nanao Max, H.; Wigge Philip, A.; Zubieta, C. Molecular Mechanisms of Evening Complex Activity in Arabidopsis. *Proc. Natl. Acad. Sci.* **2020**, *117* (12), 6901–6909. https://doi.org/10.1073/pnas.1920972117.
(13) Hennig, S.; Kong, G.; Mannen, T.; Sadowska, A.; Kobelke, S.; Blythe, A.; Knott, G. J.; Iyer, K. S.; Ho, D.; Newcombe, E. A.; Hosoki, K.; Goshima, N.; Kawaguchi, T.; Hatters, D.; Trinkle-Mulcahy, L.; Hirose, T.; Bond, C. S.; Fox, A. H. Prion-like Domains in RNA Binding Proteins Are Essential for Building Subnuclear Paraspeckles. *J. Cell Biol.* **2015**, *210* (4), 529–539. https://doi.org/10.1083/jcb.201504117.

(14) Totzeck, F.; Andrade-Navarro, M. A.; Mier, P. The Protein Structure Context of PolyQ Regions. *PloS One* **2017**, *12* (1), e0170801. https://doi.org/10.1371/journal.pone.0170801.
(15) Davey, N. E.; Van Roey, K.; Weatheritt, R. J.; Toedt, G.; Uyar, B.; Altenberg, B.; Budd, A.; Diella, F.; Dinkel, H.; Gibson, T. J. Attributes of Short Linear Motifs. *Mol. Biosyst.* **2012**, *8* (1), 268–281. https://doi.org/10.1039/C1MB05231D.
(16) Pietrek, L. M.; Stelzl, L. S.; Hummer, G. Hierarchical Ensembles of Intrinsically Disordered Proteins at Atomic Resolution in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2020**, *16* (1), 725–737. https://doi.org/10.1021/acs.jctc.9b00809.
(17) Stelzl, L. S.; Pietrek, L. M.; Holla, A.; Oroz, J.; Sikora, M.; Köfinger, J.; Schuler, B.; Zweckstetter, M.; Hummer, G. Global Structure of the Intrinsically Disordered Protein Tau Emerges from Its Local Structure. *JACS Au* **2022**, *2* (3), 673–686. https://doi.org/10.1021/jacsau.1c00536.
(18) Wang, L.; Friesner, R. A.; Berne, B. J. Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2). *J. Phys. Chem. B* **2011**, *115* (30), 9431–9438. https://doi.org/10.1021/jp204407d.
(19) Best, R. B.; Zheng, W.; Mittal, J. Balanced Protein-Water Interactions Improve Properties of Disordered Proteins and Non-Specific Protein Association. *J. Chem. Theory Comput.* **2014**, *10* (11), 5113–5124. https://doi.org/10.1021/ct500569b.
(20) Abascal, J. L. F.; Vega, C. A General Purpose Model for the Condensed Phases of Water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123* (23), 234505. https://doi.org/10.1063/1.2121687.
(21) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), 014101. https://doi.org/10.1063/1.2408420.
(22) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684–3690. https://doi.org/10.1063/1.448118.
(23) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001.
(24) Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Phys. Rev. Lett.* **1980**, *45* (14), 1196–1199. https://doi.org/10.1103/PhysRevLett.45.1196.
(25) Jo, S.; Jiang, W. A Generic Implementation of Replica Exchange with Solute Tempering (REST2) Algorithm in NAMD for Complex Biophysical Simulations. *Comput. Phys. Commun.* **2015**, *197*, 304–311. https://doi.org/10.1016/j.cpc.2015.08.030.
(26) Johnson, Q. R.; Lindsay, R. J.; Shen, T. CAMERRA: An Analysis Tool for the Computation of Conformational Dynamics by Evaluating Residue-Residue Associations. *J. Comput. Chem.* **2018**, *39* (20), 1568–1578. https://doi.org/10.1002/jcc.25192.
(27) Lindsay, R. J.; Pham, B.; Shen, T.; McCord, R. P. Characterizing the 3D Structure and Dynamics of Chromosomes and Proteins in a Common Contact Matrix Framework. *Nucleic Acids Res.* **2018**, *46* (16), 8143–8152. https://doi.org/10.1093/nar/gky604.
(28) Clark, A. K.; Wilder, J. H.; Grayson, A. W.; Johnson, Q. R.; Lindsay, R. J.; Nellas, R. B.; Fernandez, E. J.; Shen, T. The Promiscuity of Allosteric Regulation of Nuclear Receptors by Retinoid X Receptor. *J. Phys. Chem. B* **2016**, *120* (33), 8338–8345. https://doi.org/10.1021/acs.jpcb.6b02057.
(29) Johnson, Q. R.; Lindsay, R. J.; Nellas, R. B.; Fernandez, E. J.; Shen, T. Mapping Allostery through Computational Glycine Scanning and Correlation Analysis of Residue–Residue Contacts. *Biochemistry* **2015**, *54* (7), 1534–1541. https://doi.org/10.1021/bi501152d.
(30) Johnson, Q. R.; Lindsay, R. J.; Petridis, L.; Shen, T. Investigation of Carbohydrate Recognition via Computer Simulation. *Mol. Basel Switz.* **2015**, *20* (5), 7700–7718. https://doi.org/10.3390/molecules20057700.

(31) Lajoie, B. R.; Dekker, J.; Kaplan, N. The Hitchhiker's Guide to Hi-C Analysis: Practical Guidelines. *Methods San Diego Calif* **2015**, *72*, 65–75. https://doi.org/10.1016/j.ymeth.2014.10.031.
(32) Lindsay, R. J.; Mansbach, R. A.; Gnanakaran, S.; Shen, T. Effects of PH on an IDP Conformational Ensemble Explored by Molecular Dynamics Simulation. *Biophys Chem* **2021**, *271*, 106552. https://doi.org/10.1016/j.bpc.2021.106552.
(33) Ramazzotti, M.; Monsellier, E.; Kamoun, C.; Degl'Innocenti, D.; Melki, R. Polyglutamine Repeats Are Associated to Specific Sequence Biases That Are Conserved among Eukaryotes. *PloS One* **2012**, *7* (2), e30824. https://doi.org/10.1371/journal.pone.0030824.
(34) Valley, C. C.; Cembran, A.; Perlmutter, J. D.; Lewis, A. K.; Labello, N. P.; Gao, J.; Sachs, J. N. The Methionine-Aromatic Motif Plays a Unique Role in Stabilizing Protein Structure. *J. Biol. Chem.* **2012**, *287* (42), 34979–34991. https://doi.org/10.1074/jbc.M112.374504.
(35) Martin, E. W.; Mittag, T. Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions. *Biochemistry* **2018**, *57* (17), 2478–2487. https://doi.org/10.1021/acs.biochem.8b00008.
(36) Dignon, G. L.; Zheng, W.; Best, R. B.; Kim, Y. C.; Mittal, J. Relation between Single-Molecule Properties and Phase Behavior of Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci.* **2018**, *115* (40), 9929–9934. https://doi.org/10.1073/pnas.1804177115.
(37) Martin, E. W.; Holehouse, A. S.; Peran, I.; Farag, M.; Incicco, J. J.; Bremer, A.; Grace, C. R.; Soranno, A.; Pappu, R. V.; Mittag, T. Valence and Patterning of Aromatic Residues Determine the Phase Behavior of Prion-like Domains. *Science* **2020**, *367* (6478), 694–699. https://doi.org/10.1126/science.aaw8653.
(38) Khare, S. D.; Ding, F.; Gwanmesia, K. N.; Dokholyan, N. V. Molecular Origin of Polyglutamine Aggregation in Neurodegenerative Diseases. *PLOS Comput. Biol.* **2005**, *1* (3), e30. https://doi.org/10.1371/journal.pcbi.0010030.
(39) Pancsa, R.; Fuxreiter, M. Interactions via Intrinsically Disordered Regions: What Kind of Motifs? *IUBMB Life* **2012**, *64* (6), 513–520. https://doi.org/10.1002/iub.1034.
(40) Wen, J.; Scoles, D. R.; Facelli, J. C. Molecular Dynamics Analysis of the Aggregation Propensity of Polyglutamine Segments. *PLOS ONE* **2017**, *12* (5), e0178333. https://doi.org/10.1371/journal.pone.0178333.
(41) Benayad, Z.; von Bülow, S.; Stelzl, L. S.; Hummer, G. Simulation of FUS Protein Condensates with an Adapted Coarse-Grained Model. *J. Chem. Theory Comput.* **2021**, *17* (1), 525–537. https://doi.org/10.1021/acs.jctc.0c01064.
(42) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111* (27), 7812–7824. https://doi.org/10.1021/jp071097f.
(43) Dignon, G. L.; Zheng, W.; Kim, Y. C.; Best, R. B.; Mittal, J. Sequence Determinants of Protein Phase Behavior from a Coarse-Grained Model. *PLoS Comput. Biol.* **2018**, *14* (1), e1005941. https://doi.org/10.1371/journal.pcbi.1005941.
(44) Vitalis, A.; Pappu, R. V. ABSINTH: A New Continuum Solvation Model for Simulations of Polypeptides in Aqueous Solutions. *J. Comput. Chem.* **2009**, *30* (5), 673–699. https://doi.org/10.1002/jcc.21005.
(45) Metskas, L. A.; Rhoades, E. Conformation and Dynamics of the Troponin I C-Terminal Domain: Combining Single-Molecule and Computational Approaches for a Disordered Protein Region. *J. Am. Chem. Soc.* **2015**, *137* (37), 11962–11969. https://doi.org/10.1021/jacs.5b04471.
(46) Pesce, F.; Lindorff-Larsen, K. Combining Experiments and Simulations to Examine the Temperature-Dependent Behaviour of a Disordered Protein. *bioRxiv* **2023**, 2023.03.04.531094. https://doi.org/10.1101/2023.03.04.531094.

(47) Thomasen, F. E.; Lindorff-Larsen, K. Conformational Ensembles of Intrinsically Disordered Proteins and Flexible Multidomain Proteins. *Biochem. Soc. Trans.* **2022**, *50* (1), 541–554. https://doi.org/10.1042/BST20210499.
(48) Zhang, Y.; Chen, J. Recalibrating Replica Exchange Solute Tempering for Sampling Disordered Protein Conformations. *Biophys. J.* **2022**, *121* (3), 196a. https://doi.org/10.1016/j.bpj.2021.11.1730.
(49) Wereszczynski, J.; McCammon, J. A. Statistical Mechanics and Molecular Dynamics in Evaluating Thermodynamic Properties of Biomolecular Recognition. *Q. Rev. Biophys.* **2012**, *45* (1), 1–25. https://doi.org/10.1017/S0033583511000096.
(50) Alavi, S. Statistical Mechanics: Theory and Molecular Simulation. By Mark E. Tuckerman. *Angew. Chem. Int. Ed.* **2011**, *50* (51), 12138–12139. https://doi.org/10.1002/anie.201105752.
(51) Lindström, I.; Dogan, J. Dynamics, Conformational Entropy, and Frustration in Protein–Protein Interactions Involving an Intrinsically Disordered Protein Domain. *ACS Chem. Biol.* **2018**, *13* (5), 1218–1227. https://doi.org/10.1021/acschembio.7b01105.
(52) Kokona, B.; Rosenthal, Z. P.; Fairman, R. Role of the Coiled-Coil Structural Motif in Polyglutamine Aggregation. *Biochemistry* **2014**, *53* (43), 6738–6746. https://doi.org/10.1021/bi500449a.
(53) Michael, T. P.; Park, S.; Kim, T.-S.; Booth, J.; Byer, A.; Sun, Q.; Chory, J.; Lee, K. Simple Sequence Repeats Provide a Substrate for Phenotypic Variation in the Neurospora Crassa Circadian Clock. *PLOS ONE* **2007**, *2* (8), e795. https://doi.org/10.1371/journal.pone.0000795.
(54) Michalik, A.; Van Broeckhoven, C. Pathogenesis of Polyglutamine Disorders: Aggregation Revisited. *Hum. Mol. Genet.* **2003**, *12* (suppl_2), R173–R186. https://doi.org/10.1093/hmg/ddg295.
(55) Gomes, E.; Shorter, J. The Molecular Language of Membraneless Organelles. *J. Biol. Chem.* **2019**, *294* (18), 7115–7127. https://doi.org/10.1074/jbc.TM118.001192.
(56) Van Treeck, B.; Protter, D. S. W.; Matheny, T.; Khong, A.; Link, C. D.; Parker, R. RNA Self-Assembly Contributes to Stress Granule Formation and Defining the Stress Granule Transcriptome. *Proc. Natl. Acad. Sci.* **2018**, *115* (11), 2734. https://doi.org/10.1073/pnas.1800038115.
(57) Wang, J.; Choi, J.-M.; Holehouse, A. S.; Lee, H. O.; Zhang, X.; Jahnel, M.; Maharana, S.; Lemaitre, R.; Pozniakovsky, A.; Drechsel, D.; Poser, I.; Pappu, R. V.; Alberti, S.; Hyman, A. A. A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* **2018**, *174* (3), 688-699.e16. https://doi.org/10.1016/j.cell.2018.06.006.
(58) O'Malley, K. G.; Ford, M. J.; Hard, J. J. Clock Polymorphism in Pacific Salmon: Evidence for Variable Selection along a Latitudinal Gradient. *Proc. R. Soc. B Biol. Sci.* **2010**, *277* (1701), 3703–3714. https://doi.org/10.1098/rspb.2010.0762.

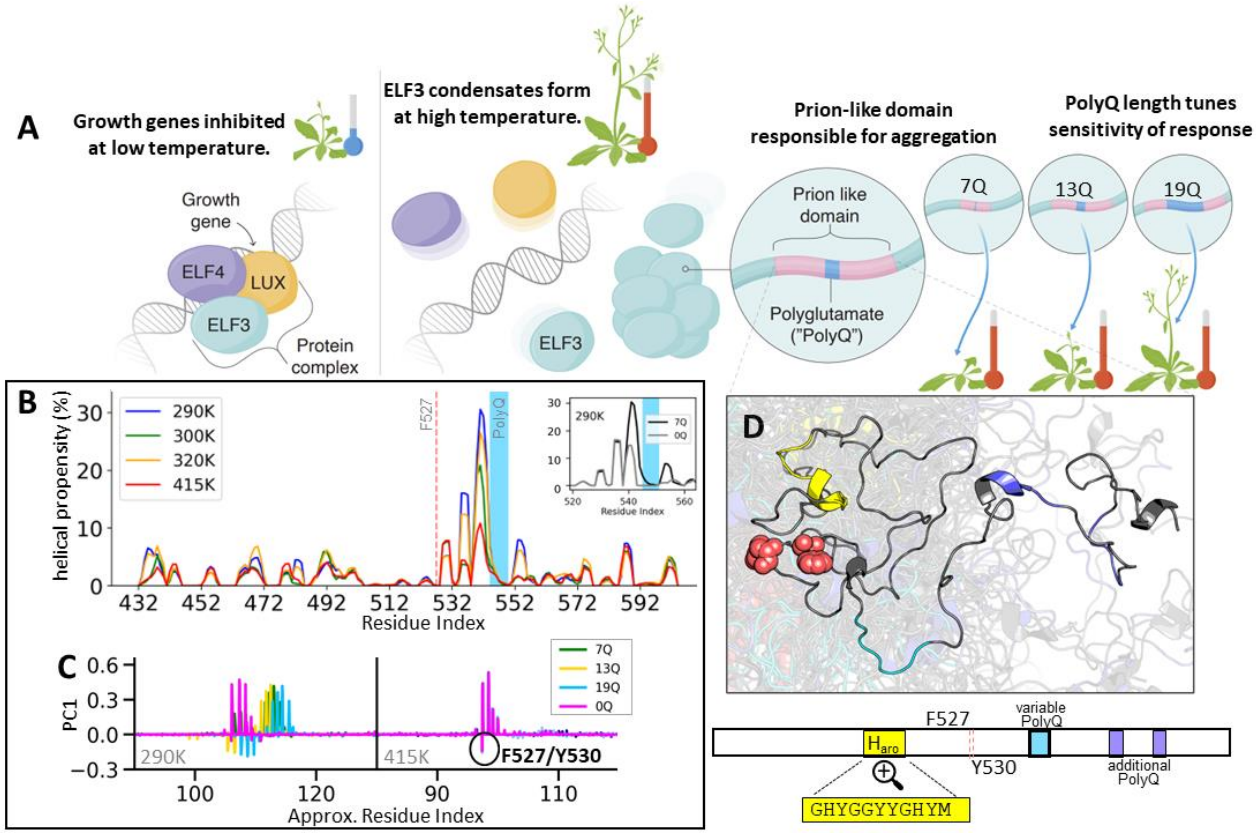

Figure 1. Temperature sensing by the Evening Complex is driven by a prion-like domain and tunable by a variable length polyQ tract. A. An illustration of the temperature-responsive mechanism of the evening complex (EC). At low temperature the EC is bound to growth genes. ELF3 dissociates at higher temperature and forms nuclear condensates. A prion-like domain is responsible for ELF3 aggregation and the length of a variable polyQ tract tunes the sensitivity of the aggregation response. **B.** Percentage of hierarchical chain growth (HCG) ensemble in a helical state by residue at four temperatures. The 7Q (wild-type) system is shown in the main panel with a region of interest of 7Q and 0Q shown in the inset. The light blue region marks the location of the variable poly-glutamine tract. **C.** PC1 of the E-PCA analysis of the ELF3-PrD HCG ensembles is shown with the lowest temperature studied (290K) in the left panel and the maximum temperature (415K) in red. **D**. Top: A structure of wild-type ELF3-PrD taken from a REST2 simulation. Bottom: A bar representing the ELF3-PrD protein sequence with regions of interest highlighted, including H_{aro} in yellow, F527 and Y530 marked with red dashed lines, variable polyQ in light blue and two additional polyQ tracts in darker blue.

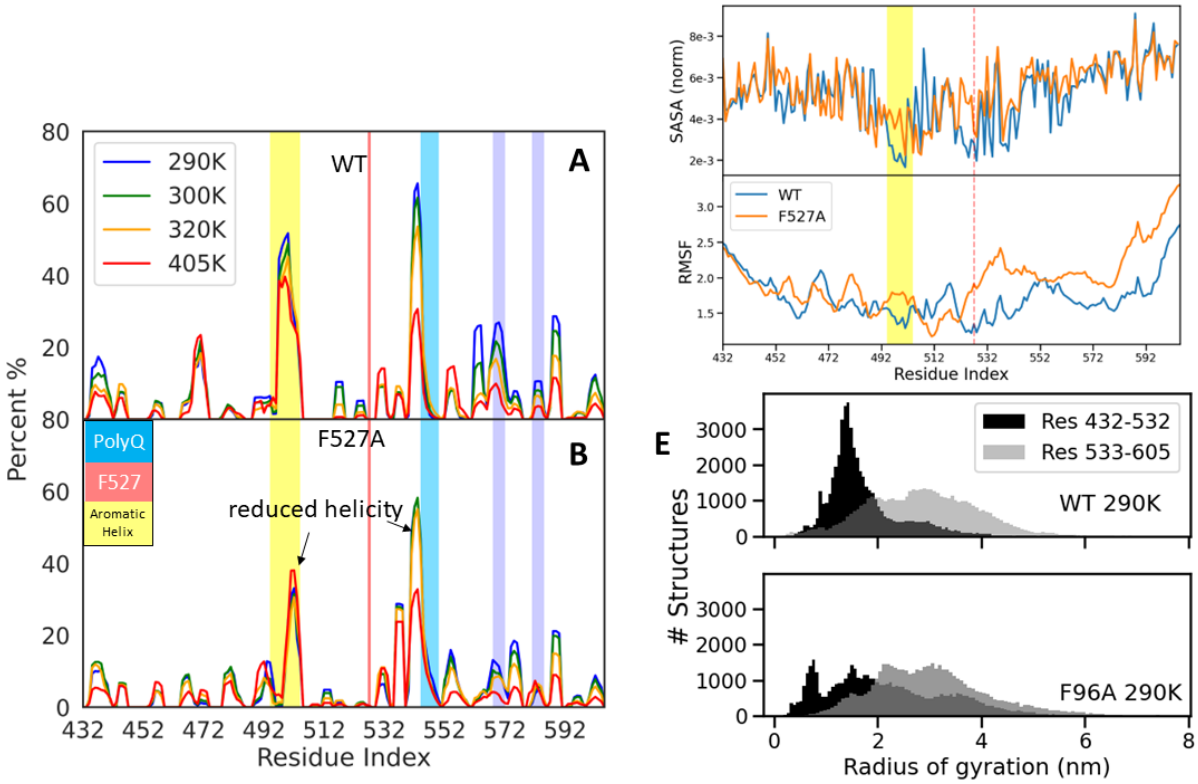

Figure 2. REST2 simulations of ELF3-PrD reveal a prominent role for an aromatic helix. A. Percentage of WT REST2 simulation in which residues were in a helical conformation, shown at a range of temperatures**.** **B.** Helical percentages of residues from the F527A REST2 simulation at a range of temperatures. **C**. Top: Solvent-accessible surface area of WT and F527A systems taken from REST2 trajectories. H_{aro} is highlighted in yellow and residue 527 is indicated by a dashed red line. **D.** The radius of gyration of the first 100 residues (black) and last 73 residues (grey) of ELF3-PrD. Top panel depicts and the WT and the bottom F527A, both at 290K.

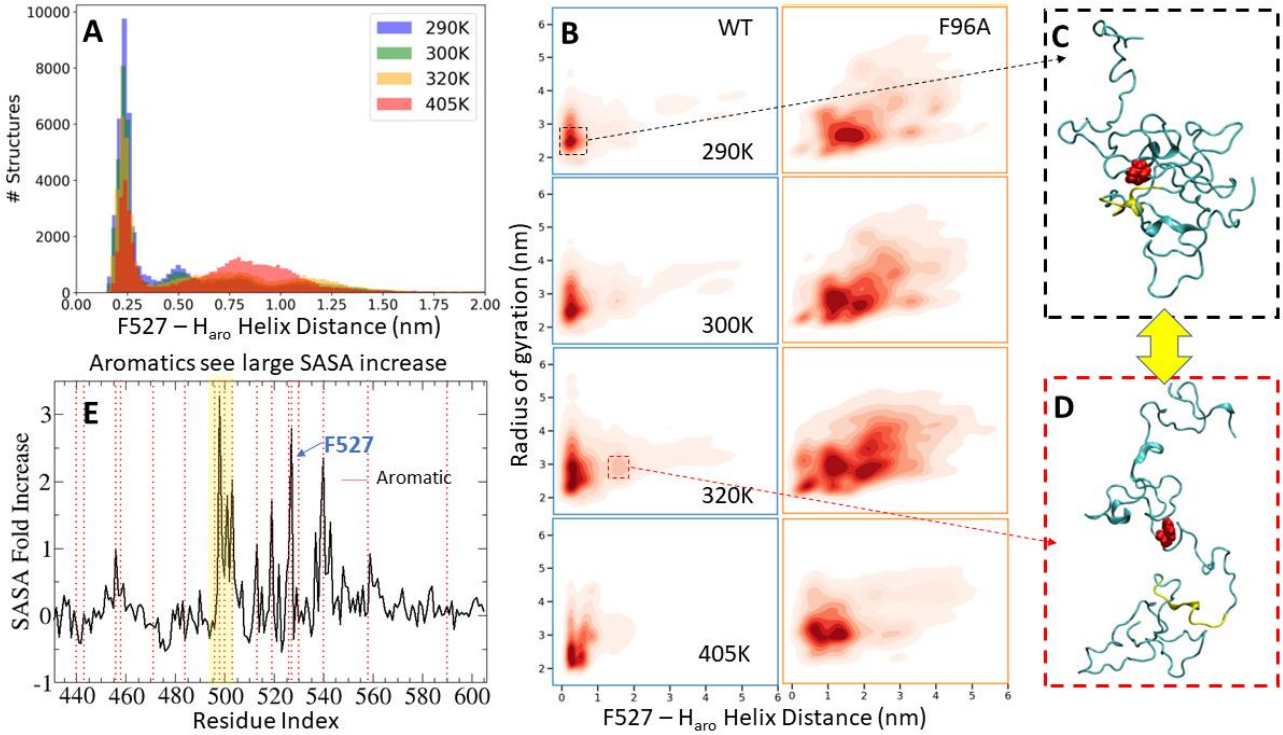

Figure 3. Conformational change uncovers 'sticky' aromatic residues. A. The minimum distance between residue F527 and H_{aro} is by temperature from the WT REST2 simulation. **B**. Free energy landscapes of WT (left) and F527A (right) with the X-axis defined as the minimum distance between residue 96 and H_{aro} and the y-axis defined as the radius of gyration. Each ensemble is represented at four temperatures. The most prominent low-temperature population is demarcated by a dashed black box. The emergent high-temperature population is bounded by a dashed red box. **C.** A representative structure of the dominant low-temperature population. This structure illustrates the interaction of F527 and H_{aro} at low temperature. **D**. A representative structure of the emergent high-temperature population illustrating separation between F527 and H_{aro}. **E**. Fold increase in SASA between the identified low- and high-temperature populations. Dotted red lines represent aromatic residues and the yellow-shaded region are residues of H_{aro}.

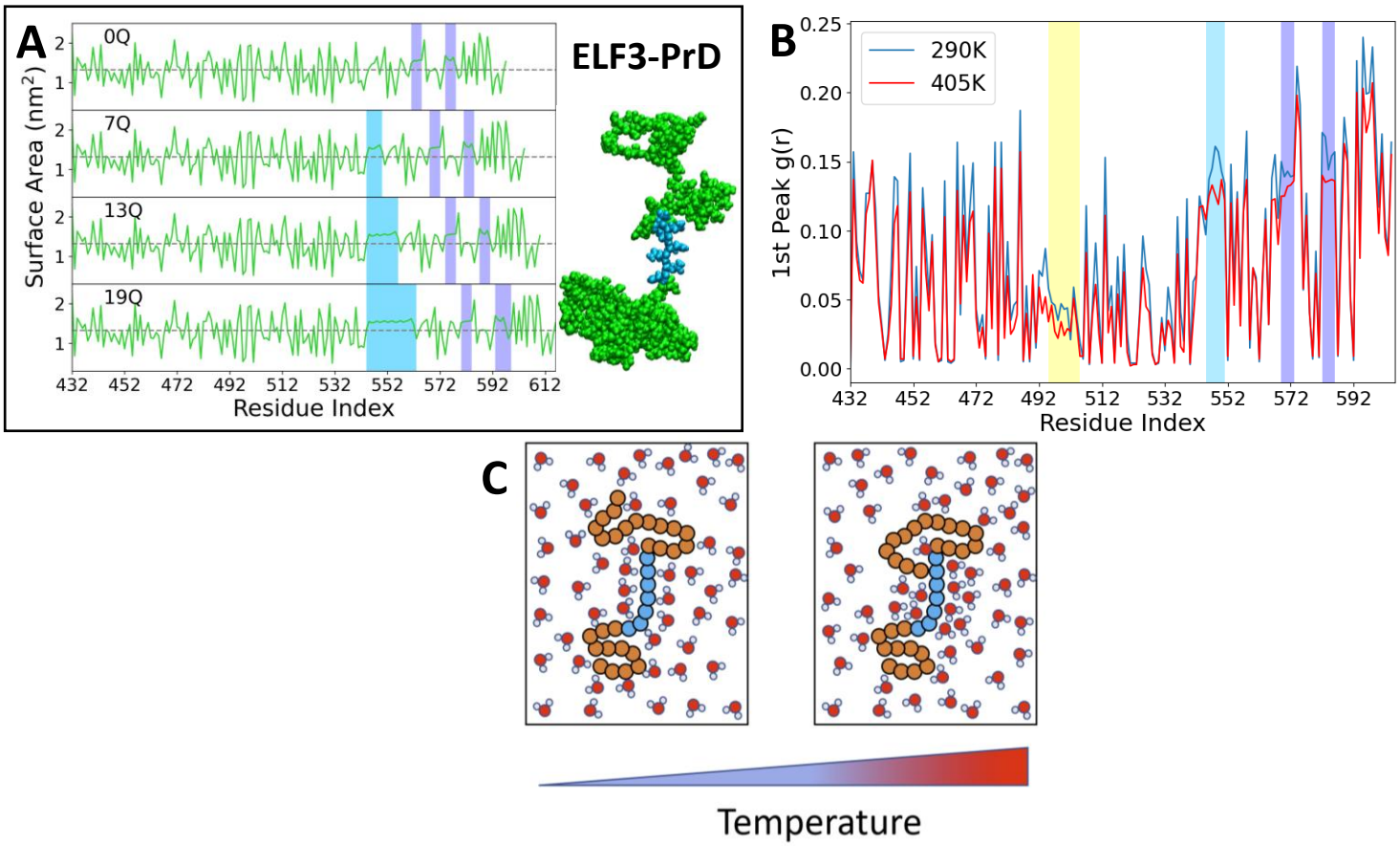

Figure 4. Solvation dynamics promote protein-protein interactions in polyQ tracts at high temperatures. A. The solvent-accessible surface area (SASA) is shown for each HCG polyQ condition at 290K. The variable polyQ tract is highlighted in sky blue and the second and third polyQ tracts in a lighter blue. The black dashed line demarcates the average SASA value for the system**. B.** Per-residue hydration plot of ELF3-PrD WT at 290K (blue) and 405K (red). Each value represents the maximum height of the first peak of an RDF between the given residue and solvent oxygen atoms. The aromatic helix, H_{aro}, is highlighted in yellow and the polyQ tracts in sky blue. **C**. A cartoon of the hydration dynamics of ELF3-PrD at low and high temperature. Blue circles represent the polyQ tract and orange circles are other residues.

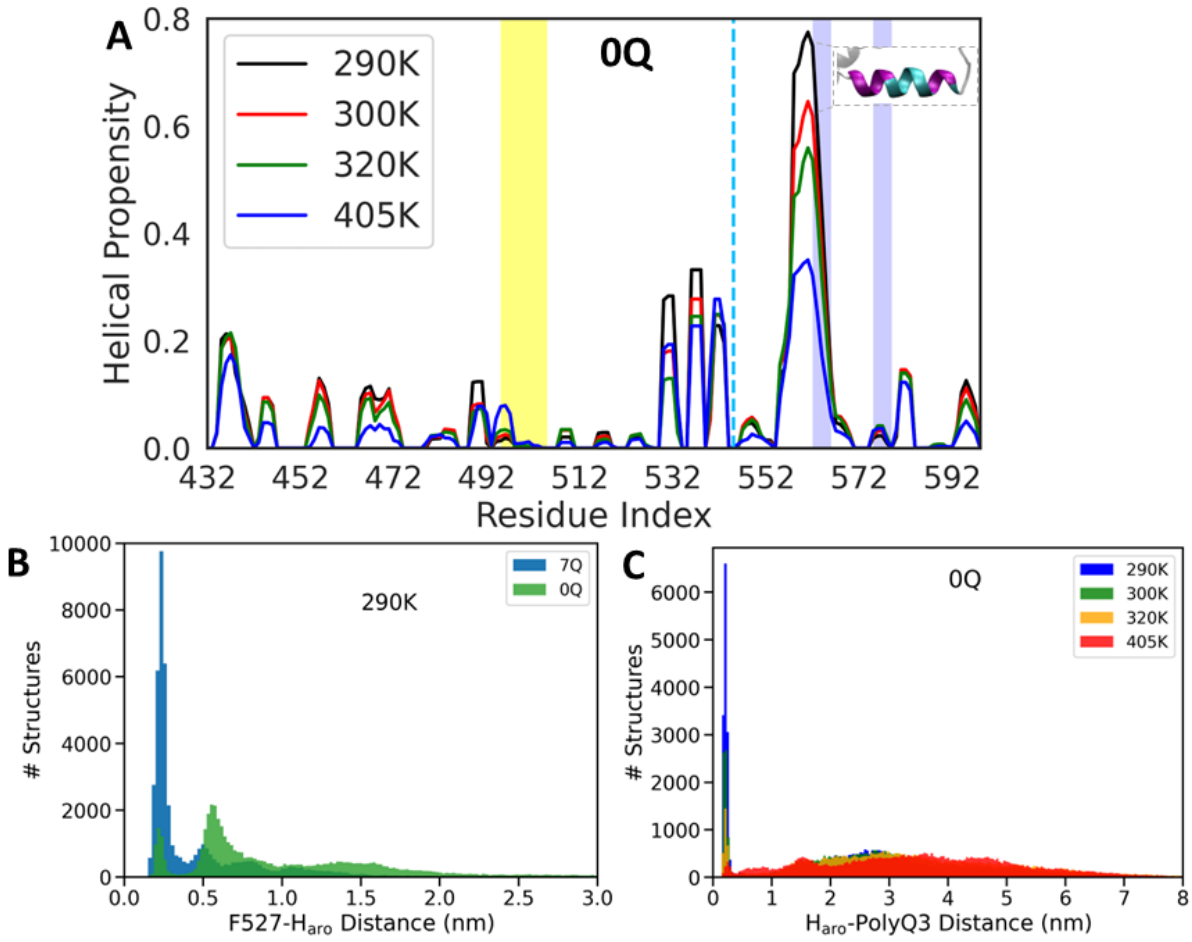

Figure 5. Absence of variable polyQ tract alters mechanism of temperature sensing. A. The helical propensity of the 0Q REST2 trajectory at a range of temperatures. H_{aro} is highlighted in yellow, the location of the removed polyQ tract is demarcated by a dashed blue line and the remaining polyQ tracts are highlighted in purple. **B**. F527-Aromatic Helix minimum distance at 290K with 7Q REST2 results in black and 0Q in gray. **C.** F527-polyQ3 minimum distance plotted for a range of temperatures.