

# Alternative low-populated conformations prompt phase transitions in polyalanine repeat expansions

Received: 16 June 2023

Accepted: 19 February 2024

Published online: 02 March 2024

Check for updates

Rosa Antón<sup>1</sup>, Miguel Á. Treviño <sup>1</sup>, David Pantoja-Uceda<sup>1</sup>, Sara Félix <sup>2,3</sup>, María Babu <sup>4</sup>, Eurico J. Cabrita <sup>2,3</sup>, Markus Zweckstetter <sup>4,5</sup>, Philip Tinnefeld<sup>6</sup>, Andrés M. Vera <sup>6</sup> & Javier Oroz <sup>1</sup> ✉

Abnormal trinucleotide repeat expansions alter protein conformation causing malfunction and contribute to a significant number of incurable human diseases. Scarce structural insights available on disease-related homorepeat expansions hinder the design of effective therapeutics. Here, we present the dynamic structure of human PHOX2B C-terminal fragment, which contains the longest polyalanine segment known in mammals. The major  $\alpha$ -helical conformation of the polyalanine tract is solely extended by polyalanine expansions in PHOX2B, which are responsible for most congenital central hypoventilation syndrome cases. However, polyalanine expansions in PHOX2B additionally promote nascent homorepeat conformations that trigger length-dependent phase transitions into solid condensates that capture wild-type PHOX2B. Remarkably, HSP70 and HSP90 chaperones specifically seize PHOX2B alternative conformations preventing phase transitions. The precise observation of emerging polymorphs in expanded PHOX2B postulates unbalanced phase transitions as distinct pathophysiological mechanisms in homorepeat expansion diseases, paving the way towards the search of therapeutics modulating biomolecular condensates in central hypoventilation syndrome.

Eukaryotic proteomes contain abundant stretches of sequentially repeated aminoacids or homorepeats, being polyglutamine, polyasparagine and polyalanine the most prevalent<sup>1</sup>. Although the function of homorepeats is still unclear, they are presumed to play roles in protein structure and interactomes, acting as polymorphic adaptors in a length- and boundaries-dependent manner<sup>2</sup>. Recombination between two mispaired alleles can cause expansions as well as contractions in homorepeats<sup>3</sup>. Indeed, the aberrant increments of consecutive glutamine residues by trinucleotide repeat expansion (TRE)

mutations were linked to multiple human diseases more than 30 years ago<sup>4</sup>. Significantly less is known about polyalanine-containing proteins, which are enriched in transcription factors and RNA-binding proteins<sup>2</sup>. Expansions of the polyalanine segments above certain thresholds are involved in at least nine different human diseases, including developmental defects and neurological disorders<sup>5</sup>. Several mechanisms were proposed to explain the pathogenicity of TREs, such as protein deregulation or loss-of-function, cellular mislocalization, dominant-negative effects or a toxic gain of function<sup>4-6</sup>. Considering that the

<sup>1</sup>Instituto de Química Física Blas Cabrera (IQF), CSIC, E-28006 Madrid, Spain. <sup>2</sup>Associate Laboratory i4HB – Institute for Health and Bioeconomy, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2819-516 Caparica, Portugal. <sup>3</sup>UCIBIO, Department of Chemistry, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2819-516 Caparica, Portugal. <sup>4</sup>German Center for Neurodegenerative Diseases (DZNE), 37075

Göttingen, Germany. <sup>5</sup>Department for NMR-based Structural Biology, Max Planck Institute for Multidisciplinary Sciences, 37077 Göttingen, Germany.

<sup>6</sup>Department of Chemistry and Center for NanoScience, Ludwig-Maximilians-Universität München, München 81377, Germany. ✉e-mail: [joroz@iqf.csic.es](mailto:joroz@iqf.csic.es)

structural conversions induced by the TREs have a profound impact in protein malfunction, intense research was dedicated to describe the structure of TRE mutants<sup>7–10</sup>. However, due to their challenging nature<sup>10</sup>, structural descriptions of homorepeat-containing proteins are currently scarce. Yet, it is essential to understand the molecular mechanisms of TRE-based diseases to develop efficient therapeutics<sup>2</sup>.

Congenital Central Hypoventilation Syndrome (CCHS or Ondine's curse) is a rare neuropathy characterized by a decreased sensitivity to hypoxia and hypercapnia, which is paralleled by the inadequate control of breathing, particularly during sleep<sup>11</sup>. Additionally, CCHS shows impaired heart rate, blood pressure and body temperature regulation. CCHS is a chronic disorder affecting mainly children, with autosomal dominant transmission and has no effective treatment<sup>12</sup>. 90% of CCHS patients show mutations in the vital developmental transcription factor PHOX2B (Paired mesoderm homeobox protein 2B)<sup>13,14</sup>. PHOX2B is a master gene in charge of the development of autonomic nervous system regions responsible for medullary control reflexes of autonomic functions<sup>12,15</sup>. In particular, heterozygous in-frame duplications within a 20 alanine stretch located in the C-terminus of PHOX2B, leading to expansions from +4 to +13 alanine residues, being +5, +6 and +7 the most common, were associated with CCHS<sup>13,14,16</sup>. Remarkably, the length of the polyalanine expansions (also called polyalanine repeat mutations or PARMs) in PHOX2B correlates with the severity of the disease phenotype in CCHS<sup>3,14,17</sup>. In parallel, PHOX2B showed a polyalanine (polyAla) length-dependent cytoplasmic aggregation when overexpressed in culture cells<sup>18–20</sup>. Consequently, loss of PHOX2B transcription function due to its cytosolic aggregation has been postulated as a possible pathogenic mechanism in CCHS<sup>5,12,18</sup>. However, the role of PHOX2B aggregates in neuronal failure in CCHS remains controversial, since cytoplasmic deposits are only observed upon high overexpression of PHOX2B carrying the longer PARMs<sup>18,21</sup> and is not observed in a knock-in mouse carrying a +7 PARM<sup>22</sup>. Alternatively, polyAla tracts were shown to form abnormal soluble oligomers *in vitro* and in cells<sup>23</sup>, and a recent study demonstrated that PARMs in the paralogue protein HOXD13 cause hereditary synpolydactyly in humans due to altered liquid-liquid phase separation (LLPS)<sup>24</sup>. Interestingly, protein misfolding within membraneless compartments formed by phase separation and percolation<sup>25</sup> can overpower the refolding capacity of molecular chaperones, triggering liquid-to-solid transitions in aged condensates which could be the basis for protein loss-of-function in degenerative diseases<sup>26</sup>. Therefore, unbalanced phase transitions due to proteostasis misregulation could be a key molecular process triggering PHOX2B malfunction in CCHS, by means of a conformational pathway in PHOX2B that would be preferentially routed by the longer PARMs. Hence, the structural polymorphism induced by the PARMs could promote nascent misfolded conformations in PHOX2B that prompt liquid-to-solid transitions in the cytosol<sup>27,28</sup>. How the PARMs, or even the TREs in general, promote protein misfolded conformations that elude chaperone control, remains a mystery.

The highly repetitive nature of homorepeats makes them particularly challenging for structural studies. Still, Nuclear Magnetic Resonance (NMR) spectroscopy is the technique of choice to characterize the structure of homorepeats, due to their inherent dynamics. However, only limited structural information has been obtained so far for polyglutamine peptides by NMR<sup>9,10</sup>. Besides a central DNA-binding homeobox domain, sequence analysis reveals flanking transcriptional activation regions composed of disordered segments and putative structured motifs in PHOX2B. Therefore, the large disordered regions and the particular content of aromatic residues could contribute to PHOX2B de-mixing in the cell<sup>25</sup>. Liquid-liquid and liquid-solid phase transitions are expected to account with protein structural rearrangements towards  $\beta$ -sheeted structural motifs which self-assemble through steric zipper-type stickers<sup>27–32</sup>. However, proof of structural polymorphism within condensates is scarce<sup>28</sup>, and several lines of

evidence proposed that proteins remained mainly disordered and dynamic in the dense phases<sup>33–35</sup>.

Here, we present the solution structure of the C-terminal domain of human PHOX2B containing the longest polyAla fragment known in mammals (20 consecutive alanines). Pathogenic PARMs relevant in CCHS promote nascent structural conformations that trigger fast phase transitions into solid condensates that arrest wild-type PHOX2B. Remarkably, these incipient conformations are precluded upon chaperone recognition, which blocks PHOX2B de-mixing. We propose that fast phase transitions promoted by PARMs disrupt PHOX2B function, and advance that preventing the nascent polymorphism could restore the normal function of expanded PHOX2B in CCHS.

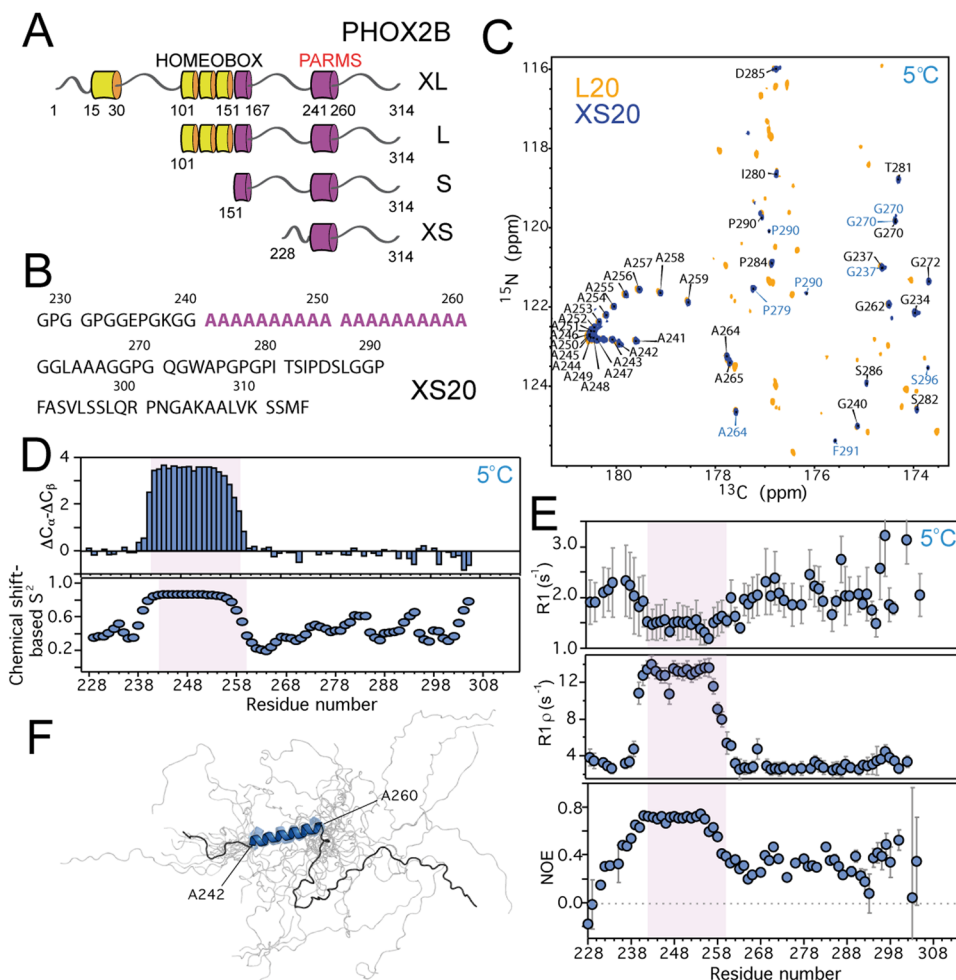
## Results

### PHOX2B polyAla tract populates a rigid $\alpha$ -helix in solution

The structural changes induced by the PARMs in PHOX2B may alter PHOX2B's interactome and transcription function, leading to protein malfunction. We employed solution NMR spectroscopy to understand the structural impact of PARMs in PHOX2B. The polyAla tract present in the C-terminal domain of human PHOX2B is singularly conserved among mammals, ranging from 16 consecutive alanines in horse to 21 in cat. Sequence analysis and structural prediction reveal short structured elements flanked by long disordered regions in human PHOX2B (Fig. 1A, B). Several constructs of PHOX2B were generated containing different protein regions to enhance solubility and minimize spectral complexity (Fig. 1A). We focused on the fragment 228–314 (termed XS20, for the number of consecutive alanines contained in the tract) of PHOX2B because its NMR spectra mirrored that of the longer PHOX2B fragments (Fig. 1C). Production of high-purity, highly soluble samples enabled us to acquire 3D NMR experiments on XS20 which follow backbone connectivities with high spectral resolution (Methods), circumventing the significant signal overlap from the polyAla repetitive region and enabling us to obtain unambiguous assignments (Fig. 1C, Supplementary Fig. 1A).

The polyAla region (residues 241–260, Fig. 1B) displayed a remarkable arrangement in the 2D CON NMR spectra (Fig. 1C), where the consecutive alanine moieties mostly follow an N-to-C termini trail. NMR secondary chemical shift analysis and chemical shift-derived  $S^2$  order parameters calculated by TALOS-N<sup>36</sup> indicate that the polyAla tract particularly populates  $\alpha$ -helical conformations, while the rest of the protein remains largely disordered (Fig. 1D). J-coupling analysis showed values close to 4 Hz for the polyAla tract (Supplementary Fig. 1B), which is in agreement with a strong tendency to adopt  $\alpha$ -helical structures<sup>26</sup>. Close inspection of the spectra reveals multiple minor conformations in the disordered regions due to significant *cis-trans* isomerization of the Pro residues in the numerous GPG motifs of PHOX2B (Fig. 1B, Supplementary Fig. 1A, C). <sup>15</sup>N spin relaxation experiments for XS20 at 5 °C show a significant deviation of the longitudinal ( $R_1$ ) and rotating-frame ( $R_{1\rho}$ ) relaxation rates and the heteronuclear nuclear Overhauser effect (NOE) values for the polyAla region, indicating local rigidity due to the formation of secondary structure (Fig. 1E). The C-terminal region of the polyAla fragment appears less rigid, which would agree with the lower secondary chemical shift values observed (Fig. 1D). Overall correlation time ( $\tau_C$ ) obtained from polyAla relaxation parameters<sup>37</sup> (6.7 ns, Supplementary Table 1) indicates that XS20 may establish intermolecular associations through the polyAla segment at 5 °C.

A significant set of <sup>1</sup>H-<sup>15</sup>N and <sup>1</sup>H-<sup>13</sup>C heteronuclear NOESY experiments which, in combination with chemical shifts and torsion angle restraints, defined distance restraints for XS20 structure calculation (Supplementary Table 2). A set of 100 conformers were calculated by CYANA, and the 20 conformers with the lowest target function were selected and assessed by PRO-CHECKNMR (Fig. 1F). Almost the whole polyAla segment is contained



**Fig. 1 | PHOX2B polyAla forms a stable  $\alpha$ -helix in solution.** **A** Different constructs containing PHOX2B domains.  $\alpha$ -helical regions are represented by cylinders, and the polyAla regions colored in purple. PARMs are located in the polyAla segment covering residues 241–260. **B** XS20 primary structure. The polyAla is highlighted in purple. **C** Overlaid CON NMR spectra for L20 (orange) and XS20 (blue) at 5°C. Assignments for XS20 are shown. Blue assignments correspond to alternative isomers. **D** Secondary chemical shifts (top) and chemical shift-based  $S^2$  order

parameters (bottom) for XS20 at 5°C. **E**  $^{15}\text{N}$  spin relaxation parameters for XS20 at 5°C. Error bars were calculated by fitting an exponential decay to the data.

**F** Structural ensemble of the 20 XS20 conformers with the lowest conformational energy at 5°C. Structures were aligned onto the structured elements. The boundaries of the  $\alpha$ -helix (in blue) are indicated. The purple shadows in the plots in **D** and **E** limit the polyAla region. Protein concentration was 0.5 mM. Source data are provided as a Source Data file.

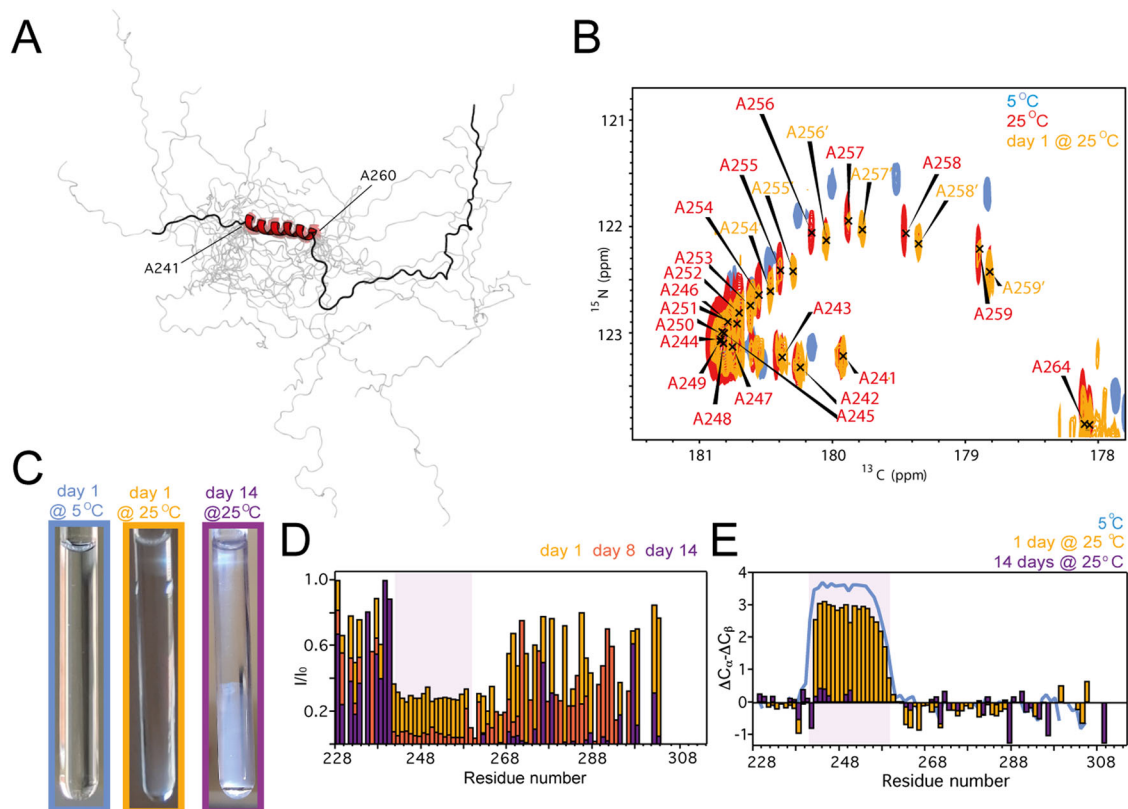
in the  $\alpha$ -helix (residues 242–260), which shows a significant structural convergence within all the conformers assessed (RMSD of 1.41 Å, Supplementary Table 2). The rest of XS20 is highly dynamic, in close agreement with the chemical shifts, J-couplings and relaxation data (Fig. 1D, E, Supplementary Fig. 1B).

### PHOX2B undergoes a temperature-dependent de-mixing

The solution structure of XS20 at 25°C was also calculated on the basis of experimental restraints (Fig. 2A, Supplementary Table 3). The structure presents a well-delimited  $\alpha$ -helix, covering the entire polyAla segment, flanked by highly disordered regions. Remarkably, the CON NMR spectra revealed the appearance of secondary conformations in XS20 samples upon incubation for 1 day at 25°C, particularly evident in the C-terminal region of the polyAla  $\alpha$ -helix (Ala255–Ala259; Fig. 2B). These conformations showed lower values in the CO dimension, indicative of decreased helical content<sup>38</sup>. In addition, continued incubation of XS20 at 25°C promoted turbidity in the NMR samples (Fig. 2C). Analysis of the  $^{15}\text{N}$  HSQC NMR spectra for the incubated XS20 samples indicated significant signal decay, with further intensity decay in the polyAla  $\alpha$ -helix (Fig. 2D, Supplementary Fig. 2). Because an increase in turbidity is associated with protein de-mixing or aggregation processes<sup>39</sup>, we wondered whether the NMR signal decay

observed in the polyAla  $\alpha$ -helical segment was due to intermolecular associations involved in PHOX2B multimerization or rather to structural conversions coupled to phase transitions<sup>28</sup>. Analysis of the secondary chemical shifts indeed revealed that incubated XS20 showed a time-dependent decrease in the  $\alpha$ -helical content of the polyAla region (Fig. 2E), which suggests that PHOX2B displays metamorphism towards disordered structures upon de-mixing or aggregation. Line broadening analysis and diffusion NMR revealed that the observable NMR moieties (Fig. 2B) belong to species of similar assembly order, since large oligomer assembly lay NMR signals beyond detection (Supplementary Fig. 3A, C). Importantly, XS20 showed no significant degradation during long incubations at 25°C (Supplementary Figs. 3 and 4). Therefore, increase in order parameters or protein degradation are not responsible for the deepened signal decay detected in the polyAla  $\alpha$ -helix upon incubation (Fig. 2D), but it is rather due to slow conformational exchange towards disordered conformations (Fig. 2B, E).

J-coupling assessment and chemical shift-based order parameters confirmed that the structure of XS20 is largely identical at 5°C and 25°C (Supplementary Fig. 5). Moreover, the  $^{15}\text{N}$  spin relaxation parameters for XS20 at 25°C indicate that the protein maintains the main rigid element covering the polyAla segment, with increased flexibility



**Fig. 2 | PHOX2B polyAla is coupled to temperature-dependent phase transitions.** **A** Structural ensemble of XS20 20 lowest energy conformers at 25°C. The lowest energy structure is displayed on top of the ensemble. Boundaries of the  $\alpha$ -helix (in red) are indicated. **B** Overlaid CON spectra for XS20  $\alpha$ -helical region at 5°C (blue), 25°C (red) and after 1 day at 25°C (orange). Alternative conformations appearing upon incubation are labeled in orange. This region disappears with

longer incubation times. **C** XS20 NMR samples following the indicated incubations. **D** Decay in intensity plot for XS20 at day 1 (orange), 8 (dark orange) and 14 (purple) of incubation at 25°C. **E** Secondary chemical shift plots for XS20 at 5°C (blue line), and after 1 day (orange bars) and 14 days (purple bars) of incubation at 25°C. Purple shadows in **D** and **E** limit the  $\alpha$ -helix. Protein concentration was 0.5 mM. Source data are provided as a Source Data file.

in the C-terminal region of the  $\alpha$ -helix (Fig. 3A, B, Supplementary Fig. 6).  $\tau_c$  calculations based on experimental relaxation data indicates that XS20 may also self-assemble through the polyAla tract at 25°C (Supplementary Table 1). Remarkably, these self-assembly and structured patterns are maintained at considerably lower protein concentrations, (0.05 mM, Supplementary Fig. 6), revealing that XS20 in solution is not monomeric and displays time-dependent assembly into large oligomers at high protein concentrations (Supplementary Fig. 3). Considering that the major structures of XS20 at 5°C and 25°C are fairly identical (Supplementary Fig. 5C, D), the remarkable propensity of PHOX2B to self-associate into large oligomers or aggregates upon incubation at higher temperatures could be attributed to the increased contact potential of hydrophobic Ala sidechains at 25°C<sup>40</sup> or to temperature-dependent structural rearrangements that trigger multimerization.

### PARMs extend the major polyAla $\alpha$ -helical conformation

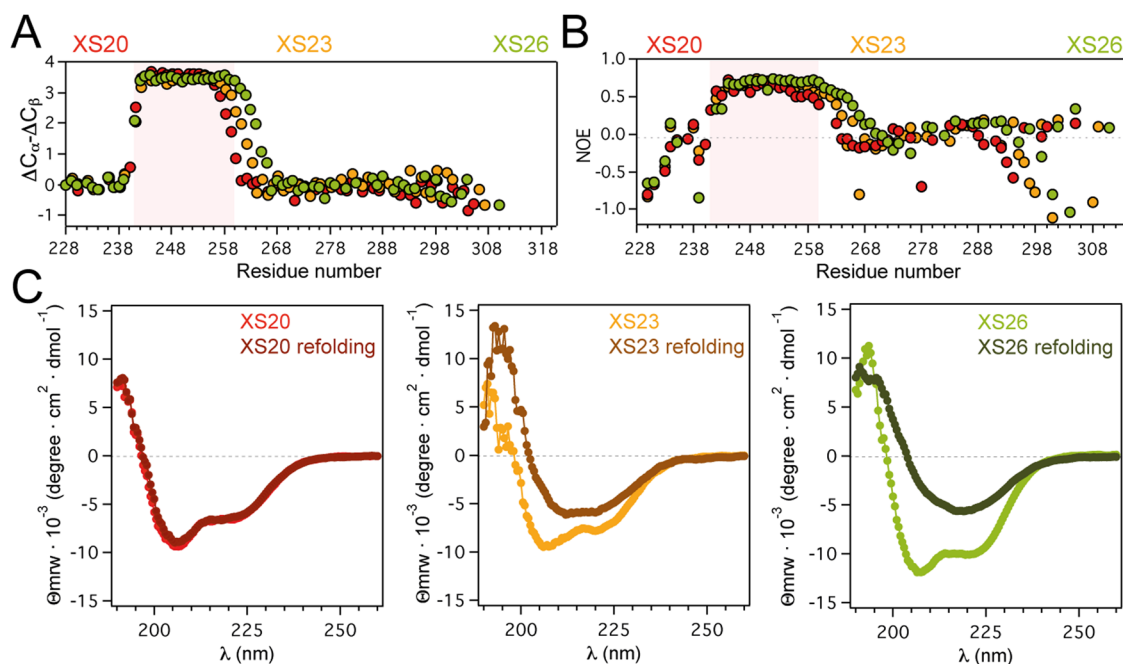
Pathogenic PARMs are expected to alter protein conformation impacting interactomes<sup>5</sup>. To determine the effect of PARMs on PHOX2B structure, we generated a +3 Ala variant (adding 3 alanine residues to the 20 polyAla tract, named XS23) and a +6 Ala variant (named XS26). While a +3 Ala expansion would not be strongly correlated to CCHS, XS26 would be representative of the most abundant PARMs observed in the disease<sup>13,14,16</sup>. NMR secondary chemical shifts showed that the Ala inclusions in XS23 and XS26 extend the  $\alpha$ -helix covering the polyAla segment (Fig. 3A). In addition, <sup>15</sup>N spin relaxation data indicated that the expanded  $\alpha$ -helices from XS23 and XS26 are also rigid (Fig. 3B) and show a stronger propensity to establish length-dependent intermolecular associations at 25°C, particularly evident

for XS26 (Supplementary Figs. 3, 7, Supplementary Table 1). Therefore, PARMs do not alter the structural and dynamic properties of the major  $\alpha$ -helical conformation of the polyAla tract but increase its self-assembly potential.

At low temperatures, polyAla repeats can form coiled-coil structures that increase in stability with polyAla length<sup>7</sup>. However, PARMs were also reported to decrease PHOX2B thermal stability favoring amyloid aggregation<sup>6</sup>. Circular dichroism (CD) measurements determined that both XS20 and the expanded variants display a high content in  $\alpha$ -helical structure (Fig. 3C). However, spectral analysis revealed no indication of coiled-coil formation for any PHOX2B variant at 25°C, in agreement with previous data obtained on model polyAla sequences<sup>7</sup>. Strikingly, although XS20 showed a fully reversible refolding after heat unfolding, both XS23 and XS26 showed irreversible unfolding. Indeed, the spectra revealed an  $\alpha$ -to- $\beta$  switch similar to that reported elsewhere for designed peptides<sup>41</sup>. Therefore, we wondered whether PARMs modulate the equilibrium between conformational states giving rise to nascent conformations which could promote associations into higher order species.

### PARMs promote alternative minor conformations

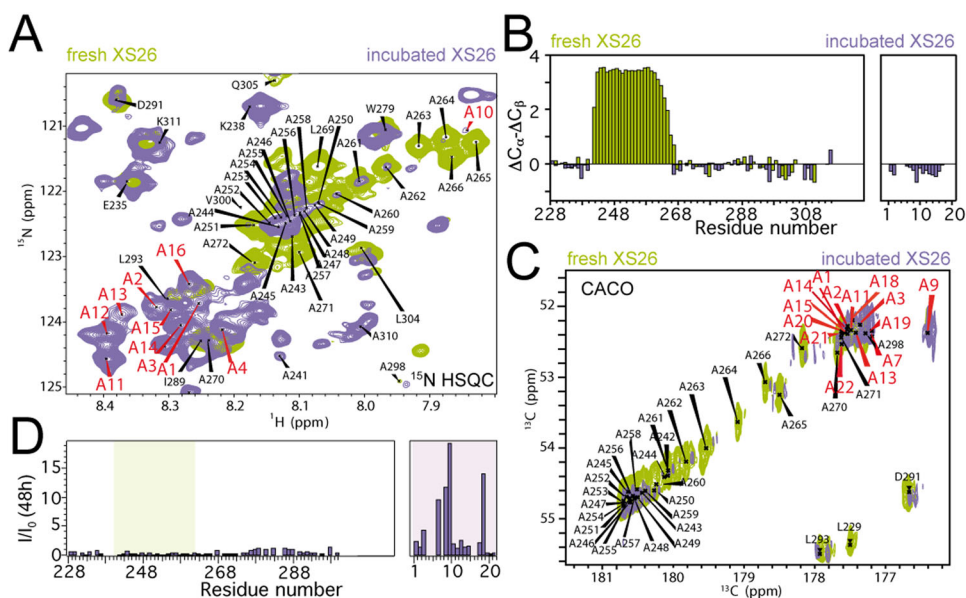
The 2D CON spectra showed shifts towards higher values in the CO dimension for the expanded polyAla fragments, which is indicative of a stronger tendency to populate  $\alpha$ -helical conformations (Supplementary Fig. 8). However, incubation of XS26 and XS23 at 25°C promoted considerable sample turbidity, triggering significant changes in the <sup>15</sup>N HSQC NMR spectra (Fig. 4A). Besides the disappearance of the polyAla  $\alpha$ -helical region, as described for XS20 after prolonged incubations (Supplementary Fig. 2), new crosspeaks appeared in the spectra which



**Fig. 3 | PARMs extend the  $\alpha$ -helix promoting irreversible unfolding.**

**A** Secondary chemical shifts for XS20 (red), XS23 (orange) and XS26 (green) at 25 °C. **B**  $^{15}\text{N}$  heteronuclear NOEs for XS20 (red), XS23 (orange) and XS26 (green) at 25 °C. Errors bars (<0.4) are not included for simplicity. The red shadows in the

plots in **A** and **B** limit the  $\alpha$ -helix in XS20. **C** CD spectra at 25 °C for XS20 (red), XS23 (orange) and XS26 (green) before and after heat denaturing at 95 °C (darker spectra in the plots). Protein concentration was 0.5 mM (**A**, **B**) and 0.017 mM (for XS23) and 0.036 mM (for XS20 and XS26) in **C**. Source data are provided as a Source Data file.



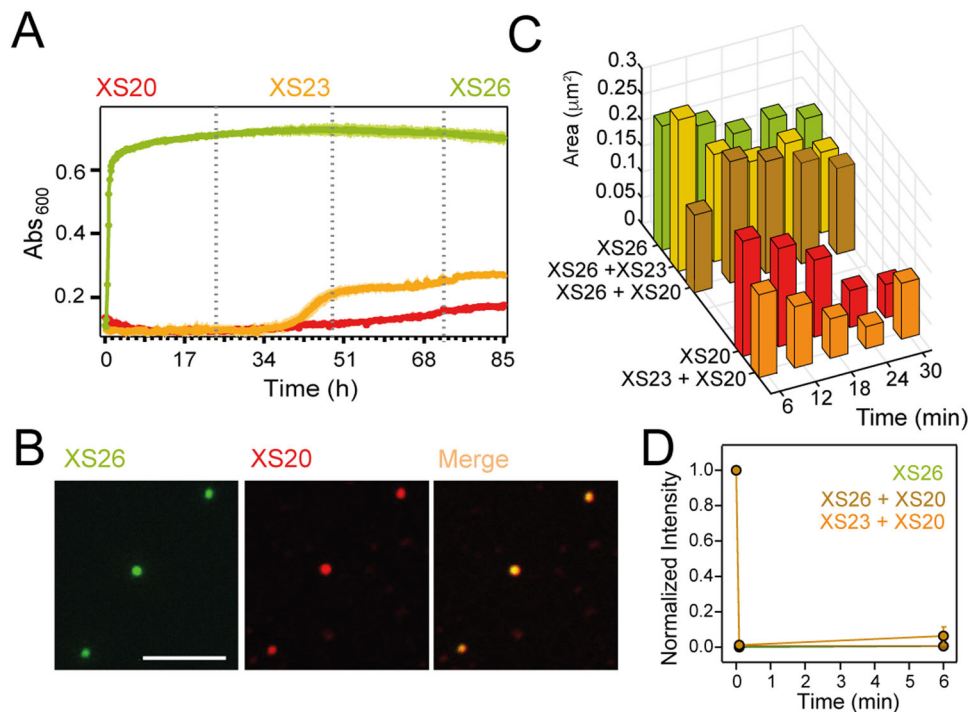
**Fig. 4 | PARMs promote nascent conformations.** **A** Detailed  $^{15}\text{N}$ -HSQC spectra at 25 °C for fresh XS26 (green) and XS26 incubated 7 days at 25 °C (navy). **B** Secondary chemical shifts for fresh XS26 (green) and XS26 incubated 7 days at 25 °C (navy). Values for new alanine moieties are plotted on the right. **C** Detailed CACO spectra at 25 °C for fresh XS26 (green) and XS26 incubated for 2 days at 25 °C (navy). **D** CACO

signal intensity for XS26 after 2 days at 25 °C. Green shadow limits XS26 polyAla  $\alpha$ -helix, while purple shadow limits nascent disordered alanine moieties (plotted on the right). Nascent alanine moieties (red in **A** and **C**) follow identical numbering and are plotted correlatively in **B** and **D**. Protein concentration was 0.5 mM in **A** and **B** and 0.225 mM in **C** and **D**. Source data are provided as a Source Data file.

were not attributable to protein degradation or large oligomer assembly (Supplementary Figs. 3, 4), but rather to slow conformational exchange in low assembly order species. Sequential assignments corroborated that the new crosspeaks belong to Ala residues which display disordered conformations, clearly identified by the characteristic chemical shifts for alanine's C $\alpha$  and C $\beta$  moieties<sup>38</sup> (Fig. 4A, B). The signal intensity of these new Ala resonances increase with the incubation time, and is highly reproducible among different protein

batches (Supplementary Fig. 9).  $^{15}\text{N}$  spin relaxation measurements for the incubated samples confirmed that the new polyAla crosspeaks are highly dynamic, in stark contrast with the major, rigid  $\alpha$ -helical conformation for the polyAla segment in fresh samples (Supplementary Fig. 10, Supplementary Table 1).

In addition, comparison of the 3D NMR CBCA(CO)NH spectra for XS20, XS23 and XS26 showed that the region where central polyAla crosspeaks appear additionally contained emerging crosspeaks with



**Fig. 5 | Pathogenic PARMs promote rapid phase transitions.** **A** Turbidity at 25 °C for 0.225 mM XS20 (red), XS23 (orange) and XS26 (green). Broken lines represent consecutive days of incubation. **B** Confocal fluorescent microscopy images showing co-localization of XS20 in XS26 condensates. XS26 was labeled with ATTO-488 (green) and XS20 with ATTO-565 (red). Scale bar: 5 µm. Data were biologically reproduced at least three times (see Supplementary Fig. 13). **C** Histogram distribution of the area (in µm<sup>2</sup>) of the condensates formed by the protein mixtures with respect to the time spent for the microscopic observations. Error bars

(<0.3 in µm<sup>2</sup>) are not included for clarity. **D** TIRF-FRAP data for XS26 condensates (green), and for mixtures of XS26 + XS20 (brown) and XS23 + XS20 (orange) incubated for 6-18 minutes at 25 °C. FRAP data obtained at different incubation times was combined. Condensates showed marginal fluidity in all conditions tested and no fluorescence was recovered after bleaching.  $n_{\text{XS26}} = 64$ ;  $n_{\text{XS26+XS20}} = 50$ ;  $n_{\text{XS23+XS20}} = 16$  ( $n$ : number of condensates). Data are presented as mean values  $\pm$  SD. Samples in **B–D** contained 0.1 mM XS23 and XS26 and 0.15-0.25 mM XS20. Source data are provided as a Source Data file.

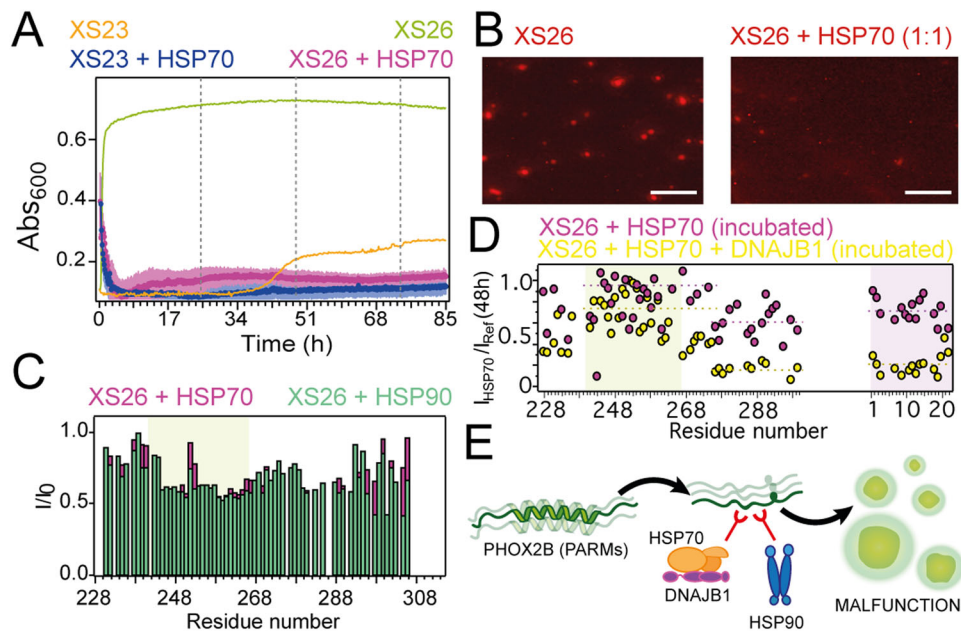
chemical shifts which are characteristic for alanines in disordered (or random coil; 52,67 $\pm$ 1.76 ppm) or  $\beta$ -strand (50,86 $\pm$ 1.28 ppm) conformations<sup>38</sup> (Supplementary Fig. 11). Remarkably, only the expanded XS23 and XS26 variants showed the nascent crosspeaks, with higher signal intensity for XS26. However, strong signal overlap in this region of the spectra could lead to dubious spectral processing. To discard misassignment due to artifactual spectral processing, the presence of the alternative conformations in the PARMs was validated in 2D NMR spectra with direct <sup>13</sup>C detection (CACO spectra), in addition to 2D <sup>13</sup>C HSQC spectra (Fig. 4C, Supplementary Fig. 12). CACO spectra provided also direct evidence of the emergence of polyAla disordered conformations upon incubation (Fig. 4D). Therefore, PARMs promote alternative minor conformations in PHOX2B, which could be the basis for their reported aggregation propensity<sup>18–20,42</sup>. However, a direct correlation between the abundance of the alternative polyAla conformations in PHOX2B and its propensity to aggregate, is missing.

### Pathogenic PARMs trigger rapid phase transitions

Light dispersion measurements showed a length-dependent increase in turbidity for PHOX2B, with a rapid response for XS26 (Fig. 5A). Confocal microscopy confirmed that XS26 promoted phase de-mixing rather than fibrillar aggregation<sup>42</sup> (Fig. 5B). Interestingly, the condensates triggered by XS26 de-mixing incorporated XS20, which correlates with the dominant negative effect of the pathogenic PARMs in PHOX2B<sup>18,19,21</sup>. Fluorescence microscopy validated the minimal de-mixing for XS20 (Supplementary Fig. 13, Supplementary Fig. 14A). In addition, XS20 is arrested into XS26 condensates at a slower rate (Supplementary Fig. 13), in agreement with the reported slower

kinetics of aggregation for the shortest PHOX2B variants<sup>6</sup> (Fig. 5A, Supplementary Fig. 15).

A recent study reported that PARMs in the protein HOXD13 alter its de-mixing properties and its capacity to co-condense with transcriptional co-activators, presenting dysregulated phase separation as the potential basis for human synpolydactyly<sup>24</sup>. To determine the phase transition properties of XS26, we performed fluorescent microscopy experiments on individual condensates (Fig. 5C, D). Recruitment of XS20 into heterotypic condensates promoted a time-dependent reduction in condensate size and fluorescence intensity (Fig. 5C, Supplementary Fig. 14B), which is indicative of lower protein content due to diffusion in the aged condensates containing XS20. The considerable small size of PHOX2B condensates (<1 µm, Fig. 5B, C, Supplementary Fig. 13, Supplementary Fig. 14C) makes the acquisition of fluorescent recovery after photobleaching (FRAP) measurements by confocal microscopy highly challenging (Supplementary Fig. 16). Thus, FRAP measurements were performed on total internal reflection fluorescence (TIRF). TIRF-FRAP rates revealed minimal mobility inside the homotypic and heterotypic PHOX2B condensates irrespective of the incubation time (Fig. 5D), consistent with confocal-based FRAP measurements (Supplementary Fig. 16). Therefore, PARMs promote a rapid transition into solid condensates, which can retain wild type PHOX2B. Interestingly, concentration-dependent PHOX2B phase transitions are not affected by increased ionic strength (Supplementary Figs. 15, 17), suggesting that, while hydrophobicity may promote small oligomers via polyAla coiled coils at low temperatures<sup>7</sup>, phase de-mixing is not driven by polyAla's hydrophobicity. However, increasing temperatures accelerated phase transition kinetics (Supplementary Figs. 15, 17B), which may be explained by enhanced thermal



**Fig. 6 | Chaperones suppress PHOX2B de-mixing arresting nascent conformations.** **A** Turbidity at 25 °C for 0.225 mM XS23 or XS26 in presence of 0.1 mM HSP70 (blue and magenta, respectively). Plots for the unaccompanied PHOX2B constructs are included. **B** Fluorescence microscopy images for XS26 (0.1 mM, labeled with ATTO-565) in the absence and presence of HSP70. Scale bars = 10  $\mu$ m. Data were biologically reproduced three times. **C**  $^{15}$ N-HSQC-based signal intensity plots for XS26 major conformation (0.07 mM) in complex with HSP70 (magenta) and HSP90 (green), all in 1:2 molar ratios. Green shadow limits polyAla  $\alpha$ -helix. **D** XS26

(0.225 mM) CACO signal intensity decay upon incubation (2 days at 25 °C) in presence of HSP70 (1:2 equivalents, in magenta) and HSP70:DNAJB1 (1:2:1 equivalents, in yellow). Green shadow limits polyAla  $\alpha$ -helix, while purple shadow limits nascent disordered alanine moieties (plotted on the right). Broken lines mark the average intensity for the different segments. **E** PARMs promote disorder in PHOX2B polyAla segment triggering fast phase transitions. Chaperones arrest nascent disordered conformations blocking phase separation. Source data are provided as a Source Data file.

destabilization of polyAla major helical conformation<sup>6</sup>. Therefore, we wondered whether the structural disorder induced by PARMs is coupled to phase transitions<sup>28</sup>.

### Chaperones arrest PHOX2B nascent conformations

Molecular chaperones are essential for neutralizing protein misfolding and malfunction. Heat shock proteins HSP70 and HSP40 co-localize with cellular PHOX2B aggregates, and are directly involved in their clearance<sup>19,43,44</sup>. Turbidimetry and fluorescence microscopy revealed that HSP70 minimized XS26 phase transitions (Fig. 6A, B). Because phase transitions are coupled to rearrangements within the polyAla  $\alpha$ -helix (Figs. 2–4), we reasoned that chaperones diminish PHOX2B homo-associations by binding the polyAla  $\alpha$ -helix in high affinity. This hypothesis would agree with the chaperone capacity to recognize hydrophobic, structured motifs in disordered clients<sup>26</sup>. Surprisingly, NMR titrations showed that HSP70 and HSP90 molecular chaperones are not preferentially bound to PHOX2B polyAla  $\alpha$ -helix (Fig. 6C). However, since there is a patent correlation between the emergence of polyAla alternative conformations and the propensity to phase separate by PHOX2B (Figs. 4 and 5), we considered that chaperones could target specifically polyAla alternative conformations rather than the major  $\alpha$ -helical structure. 2D  $^{13}$ C HSQC and CACO spectra provide convenient templates to confirm this specific recognition because they expose the whole diversity of conformations in the polyAla tract and are short time consuming. Moreover, because chaperone binding will result in signal broadening beyond detection<sup>26</sup>, increased difference in signal intensity in presence of chaperones is an indication of preferential chaperone targeting. Analysis of signal intensity of the different polyAla conformations in XS26 revealed that HSP70 and HSP90 molecular chaperones predominantly target disordered conformations (Fig. 6D, Supplementary Fig. 18). In brief, the signal intensity of XS26 polyAla tract in  $\alpha$ -helical conformation diminishes up to 20% in presence of freshly added HSP70 and 40% for HSP90, while the

intensity of the polyAla moieties in disordered conformation are reduced above 50% in the presence of chaperones (Supplementary Fig. 18). In addition, the population of alternative disordered alanines that emerge upon incubation are selectively targeted by HSP70, and diminish about 30 % after 48 h of incubation (Fig. 6D). Remarkably, addition of the HSP40 co-chaperone DNAJB1, which is known to target disordered clients<sup>26</sup>, significantly blocked 80% of the population of XS26 disordered polyAla conformations (Fig. 6D). To confirm that polyAla disorder promote phase transitions, we generated a Gly-Pro mutant to induce increased disorder in PHOX2B's polyAla  $\alpha$ -helix (Supplementary Fig. 19). The Gly-Pro mutant (termed XS<sub>GP</sub>) disrupts polyAla  $\alpha$ -helix increasing disordered conformations, which significantly accelerates phase transition kinetics. Therefore, the data clearly indicate that chaperones target disordered nascent conformations induced by PARMs, which are the template structures priming phase transitions<sup>28</sup>(Fig. 6E).

### Discussion

About 15% of eukaryotic proteins contain homorepeats<sup>2</sup>. Polyalanine, polyserine and polyglutamine segments primarily adopt  $\alpha$ -helical conformations<sup>2</sup>, reportedly promoting coiled coil associations triggering length-dependent aggregation<sup>7</sup>. Interestingly, polyAla coiled coils showed a higher tendency to aggregate compared to polyQ and polyS<sup>45</sup>, explaining why polyAla homorepeats are usually shorter and less polymorphic than polyQ repeats<sup>46</sup> and display lower thresholds for disease-causing TREs<sup>47</sup>. PHOX2B contains the longest polyAla segment known in mammals and one of the shortest thresholds for disease-causing expansions<sup>13,19</sup>. Long PARMs in PHOX2B cause shifts in localization and cytoplasmic aggregation<sup>18–21</sup>. Therefore, a loss-of-function mechanism triggered by aggregation has been postulated as the basis for mutated PHOX2B malfunction in CCHS<sup>5,12,18</sup>. Similar mechanisms were proposed for the impact of PARMs in SOX3, RUNX2, FOXL2, HOXA13 and HOXD13<sup>43,48,49</sup>. However, clear connections

between expanded homorepeats aggregation and cellular toxicity are still missing<sup>3</sup>. In addition, the nature of PARM aggregates remained controversial, since both amyloid<sup>42</sup> and  $\alpha$ -helical-rich aggregates<sup>23</sup> have been described for polyAla fragments. Most of the PARMs are related to the origin of human developmental diseases, which would be in agreement with PARM-induced protein chronic dysfunction, rather than with amyloid aggregation<sup>23</sup>. Indeed, lack of detection of PHOX2B aggregates in mouse models<sup>22</sup> is consistent with the assumption that PARMs may cause toxicity triggering aberrant oligomers<sup>23,50</sup>. Frameshift mutations in PHOX2B promote nucleolar mispartitioning and cytotoxicity by aberrant phase transitions<sup>31</sup>. In addition, PARMs in HOXD13 dysregulate LLPS, favoring homotypic phase separation and reducing condensate fluidity<sup>24</sup>. This aberrant demixing lessens the blending of HOXD13 with transcription factors in heterotypic condensates, potentially inhibiting transactivation and ultimately promoting human synpolydactyly. Our data indicate that a similar mechanism could underlie expanded PHOX2B loss-of-function in CCHS (Supplementary Fig. 20). Pathogenic PARMs in PHOX2B trigger fast phase transitions into small, solid condensates which seclude the wild-type form. Loss of PHOX2B function is therefore previous to aggregation into large protein deposits, consistent with the lack of protein aggregation in mice models of human CCHS<sup>22</sup>. Still, PHOX2B condensates evolve into irregular clusters (Supplementary Fig. 13, Supplementary Fig. 17C, D) rather than fibrillar aggregates<sup>23</sup>, explaining why expanded PHOX2B is deposited in the cell cytoplasm upon over-expression<sup>18,21</sup>. These features are also present in other PARMs, in agreement with a common mechanism underlying polyAla human diseases<sup>5,23,24</sup>. Interestingly, a recent report revealed cytosolic condensates triggered by polyS motifs which recruit and promote tau aggregation<sup>52</sup>. Therefore, condensate formation could represent distinct targets for therapeutic intervention in TRE human diseases, towards the discovery of compounds that modify the properties or prevent biomolecular condensates<sup>24,51</sup>.

Basu and colleagues showed that condensate fluidity decreased dramatically with the length of the expanded PARM<sup>24</sup>. In particular, 23-alanine HOXD13 showed minimal fluidity, and 25-alanine HOXA13 and 27-alanine RUNX2 showed absence of FRAP, indicating the solid properties of the condensates formed. Here, we observed fast solid condensate formation in PHOX2B promoted by a +6 alanine expansion (26 alanines). This reduced condensate fluidity is not alleviated by the recruitment of wt PHOX2B (20 alanines) into the condensates. Interestingly, the solid condensates formed by a +3 alanine expansion showed a remarkable delayed phase transition. Because a +3 PARM in PHOX2B is not significantly related with CCHS, we could postulate that fast phase transitions triggered by long PARMs would promote protein malfunction, and solid condensates should therefore be considered the relevant therapeutic targets. Similar to HOXD13<sup>24</sup>, malignancy of PHOX2B condensates could result from the absence of additional partners in heterotypic condensates due to the fast de-mixing of expanded PHOX2B. Indeed, because of its ability to directly detect the abundance of alternative conformations responsible for phase transitions, NMR spectroscopy is a useful screening platform for structure-based design of phase transitions inhibitors. Importantly, because imbalanced phase transitions could well be the basis for protein malfunction in PARMs<sup>24,51</sup>, the structural approach presented here could be useful to target several human diseases. Yet, revealing the protein structures and interactions that drive aberrant phase transitions is indeed highly advantageous to design compounds that decrease the population of these active states.

Here, we report PHOX2B C-terminal atomic structures including the polyAla fragments, and reveal that PARMs alter the conformational transitions promoting nascent structures. Previous studies detected conformational changes induced by homorepeat expansions<sup>33</sup> which could be modulated by the flanking regions and protonation states<sup>8</sup>. Indeed, flanking sequences modulate conformational transitions and

aggregation propensities in homorepeats<sup>29</sup>. However, a direct observation of the co-existence of different conformations in expanded homorepeats in steady state was lacking. Taking advantage of the increased NMR signal intensity of the core of the polyAla tract moieties, in addition to the characteristic alanine chemical shifts, we observed the emergence of alternative conformations in PARMs coupled to phase transitions. In brief, we detected abundant conversions from  $\alpha$ -helical to disordered conformations for the polyAla segment, in agreement with measurements of distant proteins in the dense phases<sup>33–35</sup>. This increased disorder in de-mixed phases is also in agreement with a recent study that described the different compaction of protein chains in the dilute and dense phases<sup>28</sup>. We additionally observed  $\beta$ -strand conformations, which appear exclusive for expanded PHOX2B. This  $\alpha$ - to  $\beta$ -sheet interconversion by peptide plane flipping, postulated as a key step in amyloid formation<sup>54</sup>, has never been observed in condensed milieu to date. Intriguingly, despite the established relevance of  $\beta$ -sheet structures in protein deposits<sup>29–32,55</sup>, formation of polyAla  $\beta$ -sheet aggregates have only been observed under extreme experimental conditions<sup>1,23,42,56</sup>. Indeed, whereas the abundance of  $\beta$ -strand conformations correlates with the affluence and persistence of expanded PHOX2B condensates, remains to be established.

Phase separation predictors suggest that the regions flanking PHOX2B polyAla tract (rich in GPG motifs) show high tendency to phase separate (stickers<sup>25</sup>), while the hydrophobic polyAla tract would act as a spacer<sup>25</sup>. Therefore, it follows that conformational changes and increased disorder in the hydrophobic polyAla spacer would facilitate contacts within the stickers in the dense phases<sup>28</sup>. This is somehow in disagreement with the premise stating that polyAla expansions trigger protein aggregation through coiled coil stabilization<sup>7</sup>, but is in accordance with the role of disordered polyQ segments supporting Whi3 phase separation<sup>57</sup>. While coiled coils are prevalent in proteins that phase separate, no evidence for coiled coil formation in the polyAla tract was inferred from our CD data. Interestingly, NMR relaxation data showed that polyAla segments trigger small oligomeric assembly regardless of the concentration (probably through coiled-coil formation). Although this property may ultimately lead to protein aggregation, it is not sufficient to explain the temperature-dependent phase transitions observed. Indeed, increasing the ionic strength have no effect on phase separation, but rather promote aberrant aggregation by hydrophobic collapse<sup>40</sup>. Relaxation data also showed a stronger tendency of the PARMs to assemble into small oligomers, which could not be explained on the basis of the major  $\alpha$ -helical conformation of the extended polyAla tract. In addition, significant  $\alpha$ -helix to disordered conversions are observed coupled to phase transitions in expanded PHOX2B, in accordance with the decrease in thermal stability induced by PARMs<sup>6</sup>. Disruption of the  $\alpha$ -helix indeed accelerated phase transitions. Disordered to  $\alpha$ -helix conversions are established as early steps in amyloid aggregation<sup>58</sup>. However, the switch from  $\alpha$ -helix to disordered conformations observed for PARMs coupled to phase transitions could represent the polymorphism that is expected to occur in the condensate interface<sup>27,28</sup>. Therefore, the remarkable disorder in addition to the reduced condensate fluidity indicates that PHOX2B phase transitions are triggered by phase separation coupled to networking or percolation transitions<sup>25</sup>. Considering that the polyAla-flanking stickers in PHOX2B display significant polymorphism due to proline *cis-trans* isomerization, enriched disorder in the polyAla tract would increase networking during phase transitions. Interestingly, there is a significant abundance of Gly and Pro residues flanking polyAla tracts in humans<sup>59</sup>, suggesting that the sticking mechanism proposed here could be general in PARMs. Therefore, our data indicate that the search of compounds that stabilize the major  $\alpha$ -helical structures and deter PARM-promoted nascent conformations would represent a reasonable rationale for the search of therapeutics for CCHS and other polyAla disorders. In particular, our data imply that



**Table 1 | List of oligonucleotides used for cloning**

	Forward	Reverse
XL20 insert	5'ATTTCCAGGGATCCATGTATAAAATGGAATATTC3'	5'GTGGTGTCTCGAGTTAGAACAATACTGCTC3'
XL20 vector	5'CTCGAGCACCACCACCACCACCCTG3'	5'GACATGGATCCCTGGAAATACAGGTTTTTC3'
L20 insert	5'ATTTCCAGGGATCCATCCGCACCACCTTCAC3'	5'GTGGTGTCTCGAGTTAGAACAATACTGCTC3'
S20 insert	5'ATTTCCAGGGATCCGCTAAGTTTCGCAAGC3'	5'GTGGTGTCTCGAGTTAGAACAATACTGCTC3'
S20 vector	5'CTCGAGCACCACCACCACCACCCTG3'	5'CTTAGCGGATCCCTGGAAATACAGGTTTTTC3'
XS20 insert	5'ATTTCCAGGGATCCGCGCCCGGGGGCCCCGGG3'	5'GTGGTGTCTCGAGTTAGAACAATACTGCTC3'
XS20 vector	5'CTCGAGCACCACCACCACCACCCTG3'	5'GGGCCGGATCCCTGGAAATACAGGTTTTTC3'
XS23	5'GCAAGGGCGGTGCAGCAGCAGCGGCTGCGGCTG3'	5'CAGCCGCAGCCGCTGCTGCTGCACCGCCCTTGC3'
XS26	5'GCAAGGGCGGTGCTGCAGCAGCAGCAGCAGCG3'	5'CGTGTCTGCTGCTGCTGCAGCACCGCCCTTGC3'
C315	5'GAGCAGTATGTTCTGTTAACTCGAGCACCACC3'	5'GGTGTGTCTCGAGTTAACAACAATACTGCTC3'

the C-terminal region of polyAla  $\alpha$ -helix is significantly more dynamic, and could be the region to target the design of stabilizing compounds.

HSP40, HSP70 and HSP90 chaperones primarily identify nascent conformations and impede PHOX2B phase transitions, providing a direct link between conformational transitions and phase de-mixing. We recently identified a chaperone tendency to recognize short structured motifs in disordered clients<sup>60</sup>. Interestingly, here we show the selective chaperone arrest of minor disordered conformations in an otherwise structured segment of mutated PHOX2B. In particular, chaperones seize disordered conformations which could serve as secondary-structure templates priming fast phase transitions<sup>28,61</sup>. This selective identification is in agreement with the co-localization of HSP70 with the cytoplasmic aggregates formed by expanded PHOX2B, while it remained disperse in the cytosol in the presence of transfected wild-type PHOX2B<sup>19</sup>. In addition, up-regulation of HSP70-HSP40 machinery promotes the clearance of PARM aggregates<sup>43,44,49</sup>. Liquid-solid phase transitions are usually driven by the acquisition of  $\beta$ -sheet structure<sup>29–32,55</sup>. Remarkably, the minor  $\beta$  metamorphic population observed in expanded PHOX2B is consistent with ThT negative,  $\alpha$ -helical PHOX2B aggregates<sup>23</sup> which are adhered by minimal  $\beta$ -sheet steric zippers<sup>7,45</sup>. This peculiar sticking procedure would not be shared by other proteins undergoing phase separation and promoting ThT positive amyloid aggregates<sup>26</sup>. For instance, chaperones promote TDP-43 LLPS and further amyloid aggregation to protect the cell against the prion-like behavior of soluble TDP-43<sup>62</sup>, but inhibit PHOX2B phase transitions since it is the basis of its malfunction. Although we still do not fully understand the underlying mechanistic basis, realizing how chaperones display this variety of roles towards cellular protection is indeed fascinating.

## Methods

### Sample preparation

All the different human PHOX2B (UniProt KB Q99453, Origene) cDNA constructs were subcloned into modified pET28a vectors (Novagen) containing Thioredoxin (TXA) as a fusion protein followed by a six-histidine tag for Ni<sup>2+</sup> affinity purification and a Tobacco Etch Virus (TEV) protease cleavage site. Construct nomenclature refers to the length of the contained sequence (XL, L, S and XS; being XL the full-length protein) followed by the number of alanines in the polyAla tract (20, 23 and 26). All cloning procedures were performed following the Gibson Assembly method (New England Biolabs). PolyAla expanded mutants were generated by mutagenic PCR. XS20 and XS26 were subsequently used as templates for the addition of a C-terminal cysteine residue (C315) for the covalent attachment of a fluorescent tag (detailed below). XS26 was used as template for the generation of the XSGP variant (GenScript). Human HSP72, HSP90 $\beta$  (named HSP70 and HSP90, respectively, for simplicity) and DNAJB1 cDNA sequences were cloned into respective pET28a vectors (GenScript). All sequences were verified by DNA sequencing. Cloning procedures were performed with

the following oligonucleotides in both DH5 $\alpha$  and XL1 *E. coli* strains (Table 1).

PHOX2B fragments (including WT, PARMs and XSGP variant) were produced in BL21 star (DE3) *E. coli* strain. Cells were cultured in rich media until OD<sub>600nm</sub> = 0.6. For <sup>15</sup>N/<sup>13</sup>C isotopic labeling, cultures were then harvested by a 20-min centrifugation at 5,300 g and the pellets resuspended in ¼ of the initial volume in M9 medium supplemented with 4 g/l of <sup>13</sup>C-glucose and 1 g/l of <sup>15</sup>NH<sub>4</sub>Cl as the sole sources of C and N, respectively<sup>63</sup>. Cultures were incubated at 37 °C during 1–1.5 h for the adaption to the new media until the optical density increased one more unit than the initial value after resuspension. Protein expression (both in rich and minimal media) was induced by the addition of 0.5 mM (0.1 mM for L20 production) isopropyl  $\beta$ -d-1-thiogalactopyranoside (IPTG) for 16–20 h at 25 °C. Cells were harvested by centrifugation at 5,300 g during 25 minutes at 4 °C and resuspended in 30 ml/L of lysis buffer (50 mM KPi/500 mM NaCl/10 mM imidazole/2 mM  $\beta$ -mercaptoethanol [pH 8]), including 2  $\mu$ l of protease inhibitors (ThermoFisher Scientific), 0.08 mg/ml of DNase I (Sigma-Aldrich) and 1 mg/ml lysozyme (Sigma-Aldrich) and sonicated. Subsequently, a clarification step at 4000 g for 20 minutes at 4 °C in an Optima XPN-90 Ultracentrifuge (Beckman Coulter) was performed to remove the insoluble debris.

The clarified lysates were loaded onto pre-equilibrated Ni<sup>2+</sup> affinity columns (Cytiva), adding 500 mM imidazole to the buffer for the elution (under the application of gradient elutions). TEV digestion was performed in 5 mM KPi/10 mM NaCl/10 mM imidazole/2 mM  $\beta$ -mercaptoethanol [pH 8], by an overnight incubation at 4 °C with 0.5–1 mg/ml of TEV protein (produced in house) per 15 mg of fusion protein. Cleaved PHOX2B protein constructs were directly cleared with a Ni<sup>2+</sup> affinity purification as previously detailed, where PHOX2B was present in the flowthrough fractions. Samples were further purified by a serial anion and cation exchange chromatography using Hitrap Q HP and SP HP columns (Cytiva), respectively, using 5 mM KPi/10 mM NaCl/2 mM  $\beta$ -mercaptoethanol [pH 6.8]. Up to 1 M NaCl was added in gradient to the buffer for sample elution. Highly pure samples were concentrated below 0.06 mM using Vivaspin® Turbo 15 (3 MWCO) ultrafiltration tubes (Sartorius) in 5 mM KPi/10 mM NaCl/2 mM  $\beta$ -mercaptoethanol/0.03 % sodium azide [pH 6.8] (NMR buffer) and snap-frozen at –80 °C.

DNAJB1, HSP70 and HSP90 molecular chaperones were produced in BL21 (DE3) and Rosseta 2 (DE3) star *E. coli* strains, using 1 mM IPTG for induction at OD<sub>600</sub> = 0.8–0.9 during 4 h at 37 °C. Cells were harvested and lysed in 20 mM Tris-HCl/500 mM NaCl/10 mM imidazole [pH 8] as explained previously. Recombinant proteins in the soluble fraction were purified by Ni<sup>2+</sup> affinity chromatography using high density NiSO<sub>4</sub> agarose beads (ABT), 20 mM Tris-HCl/500 mM NaCl/10 mM imidazole [pH 8] as binding buffer and binding buffer including 500 mM imidazole for elution. Proteins were further purified by size exclusion chromatography using HiLoad™ 26/60 Superdex™ 75 pg columns (Cytiva) in 10 mM Hepes/500 mM KCl/5 mM DTT [pH 7.5].

Fractions containing pure monomeric protein were pooled, concentrated using Vivaspin® 20 ultrafiltration tubes (10 MWCO, Sartorius) and stored at  $-80^{\circ}\text{C}$ . Chaperone and co-chaperone samples were buffer exchanged before the experiments into 20 mM Hepes/5 mM  $\text{MgCl}_2$ /10 mM KCl/1 mM DTT [pH 6.8] using Zeba spin desalting columns (ThermoFisher).

XS20 and XS26 C315 mutants were tagged with ATTO-488 and ATTO-565 (ATTO-TEC GmbH) following the manufacturer protocol for fluorescent microscopy assays. After tagging, the excess of dye was removed with a size exclusion chromatography, using a Superdex™ 75 10/300 GL column (Cytiva). An average 45% tagging efficiency was confirmed spectrophotometrically.

### Turbidimetry assays

Protein samples of pure XS20, XS23, XS26 and L20 were prepared at 0.05, 0.125, 0.225 and 0.5 mM concentration in 5 mM KPi/10 mM NaCl/1 mM DTT/1 mM benzamidine/0.03% sodium azide [pH 6.8] (incubation buffer).  $\text{X}_{\text{S}_{\text{GP}}}$  was measured at 0.165 mM. To monitor turbidity in presence of HSP70 and HSP90, 0.225 mM of PHOX2B proteins were incubated with 0.1 mM of the chaperones in 20 mM Hepes/10 mM NaCl/5 mM  $\text{MgCl}_2$ /2 mM ATP/1 mM DTT/0.03% sodium azide [pH 6.8]. Turbidity was measured at  $25^{\circ}\text{C}$  using  $\text{Abs}_{600}$  in a FLUOstar Omega (BMG LABTECH) multiwell plate reader using 96-well flat-bottom plates (Porvair Sciences). Measurements at  $37^{\circ}\text{C}$ , including 10, 150 and 300 mM NaCl and 10% (w/v) PEG 4000 (Fisher Scientific) were performed using the same buffer and protein concentration as mentioned above. Samples were incubated for 84 h with 60 s of agitation at 100 rpm before each measurement point (10 minutes between measurement points). All data was technically and biologically replicated, and presented as the average between replicates with the standard deviation.

### Fluorescent microscopy

Fluorescent microscopy images were obtained in a Nikon Eclipse TE2000-U inverted microscope, a Nikon Ti2e-AXR-DUX-ST super-resolution confocal microscope, a Leica AF6000 LX and a Leica TCS SP8 STED confocal microscope. Protein samples were incubated in NMR buffer ranging from 0.1-0.25 mM protein concentration. Samples were first incubated at  $37^{\circ}\text{C}$  for 30 seconds, placed for 3 minutes at  $4^{\circ}\text{C}$  and finally incubated again at  $37^{\circ}\text{C}$  for the corresponding period of time. Fluorescent labeled XS20 and XS26 (both labeled with ATTO-488 and ATTO-565) were mixed with unlabeled samples at a molar ratio 1:400-1:600 before exchange into NMR buffer. Colocalization experiments were performed using 0.1 mM (1:400) XS26-ATTO-488 and 0.15 mM (1:600) XS20-ATTO-565 incubated at  $37^{\circ}\text{C}$  for 3.5 min. To study PHOX2B condensates in presence of chaperones, 0.1 mM PHOX2B proteins were incubated with 0.1 mM chaperones (1:1 ratio) in 20 mM Hepes/10 mM NaCl/5 mM  $\text{MgCl}_2$ /2 mM ATP/1 mM DTT/0.03 % sodium azide [pH 6.8] buffer. Samples were spotted onto glass slides (ThermoFisher scientific) and imaged upright.

Confocal FRAP experiments were performed on a home-built confocal microscope based on an Olympus IX-71 inverted microscope. The ATTO-488 labeled samples were excited by a 532 nm pulsed laser (LDH-P-FA-530B, PicoQuant GmbH) at a repetition rate of 40 MHz. The laser is coupled into a single mode fiber (P3-488PM-FC, Thorlabs GmbH) to obtain a Gaussian beam profile. A linear polarizer (LPVISE100-A, Thorlabs GmbH) and a quarter-wave plate (AQWP05M-600, Thorlabs GmbH) is used to obtain Circular polarized light. The light is focused onto the sample using an oil-immersion objective (UPLSAPO100XO, NA 1.40, Olympus Deutschland GmbH) and the position of the sample is adjusted using a piezo stage (P-517.3CD, Physik Instrumente (PI) GmbH & Co. KG) and controller (E-727.3CDA, Physik Instrumente GmbH & Co. KG). The emitted light is separated from the excitation beam by a dichroic beamsplitter (zt532/640rpc, Chroma) and focused onto a 50  $\mu\text{m}$  diameter pinhole (Thorlabs

GmbH). After the pinhole, the emission signal is directed by a dichroic beamsplitter (640 LPXR, Chroma) into a green (Brightline HC582/75, AHF; RazorEdge LP 532, Semrock) detection channel. Emission is focused onto avalanche photodiodes (SPCMAQRH-14-TR, Excelitas) and the signals are registered by a time-correlated single photon counting (TCSPC) unit (HydraHarp400, PicoQuant). The setup is controlled by a commercial software package (SymPhoTime64, Picoquant GmbH). The surface was scan using a power of  $1\ \mu\text{W}$  (as measured at the entrance of the microscope body) to detect the condensates. To bleach part of the condensate, the confocal spot was focused nearby the condensate and the power increased to  $175\ \mu\text{W}$  for 2 seconds. For the preparation of the microscopy chamber, SecureSeal Hybridization chambers (2.6 mm depth, Grace Bio-Labs, USA) were glued on UV-Ozone cleaned (PSD-UV4, Novascan Technologies, USA) microscope coverslips. The surface was further cleaned with 1 M KOH for 15 min. After that, the chambers were washed three times with water and three times with PBS buffer. The surface was finally passivated with BSA-biotin for at least 1 h. Right before the experiments, the BSA was removed and the chamber washed three times with NMR buffer. Protein samples were incubated as mentioned below.

TIRF-based FRAP measurements and time evolution observation of the condensates were performed on a commercial Nanoimager S microscope (ONI Ltd., UK) under TIRF illumination. Blue excitation was applied at 488 nm to excite the ATTO-488 labeled proteins. The auto focus of the setup was used to maintain the imaged samples on focus. Protein sample concentration, mixtures and incubation protocols were as mentioned above. Incubated samples were diluted 1/30 with NMR buffer in the microscopy chamber. Measurements were performed at  $30^{\circ}\text{C}$ . In order to standardize the measurements, data acquisition was started 6 min after the last  $37^{\circ}\text{C}$  incubation step. To study the time evolution of the condensates, observations (two frame video with 100 ms integration time per frame and 0.2 mW blue excitation) were performed at 6, 12, 18, 24 and 30 min. In each observation, 25 field of views were acquired in a rectangular pattern ( $5 \times 5$ ) covering  $117\ \mu\text{m} \times 117\ \mu\text{m}$ . Duplicates were done for all the conditions tested, and averaged data is presented including standard deviation. To avoid being biased by diffraction limit, only condensates equal or larger than 2 pixels (each pixel being 117 nm) were included in the analysis. The following light program was set for FRAP measurements: i) Pre-bleaching observation (two frame video with 100 ms integration time per frame and 0.2 mW of blue excitation). ii) Bleaching (2.5 s of blue excitation at 169 mW power). iii) After-bleaching observation (two frame video with 100 ms integration time per frame and 0.2 mW blue excitation). iv) 360 s of recovery step without light excitation. v) Recovery after photobleaching observation (two frame video with 100 ms integration time per frame and 0.2 mW blue excitation).

TIRF videos were analyzed using Imajj<sup>64</sup>. Only the first frame of the video was analyzed to avoid significant photobleaching of the field of view during the observation. The background of the images was subtracted using a rolling ball of 50 pixels radius. Condensates were analyzed using the Analyze Particles routine, which automatically identifies the condensates and calculates their intensity and size. For the quantification, the images were binarised (16-bit) and thresholded using the Otsu algorithm<sup>64</sup>. Only properly focused condensates were included in the analyses, which contained information on particle area and fluorescence intensity.

### NMR assignments, $^{15}\text{N}$ relaxation and structure calculation

All the NMR spectra were acquired in an 800 MHz ( $^1\text{H}$ ) Bruker AVNEO spectrometer equipped with a Z-gradient cryoprobe.  $^1\text{H}$  chemical shifts were referenced to the internal reference sodium trimethylsilylpropanesulfonate (DSS), and  $^{15}\text{N}$  and  $^{13}\text{C}$  chemical shifts were referenced indirectly to  $^1\text{H}$  using the corresponding gyromagnetic ratios<sup>65</sup>. NMR assignments for XS20 were obtained at  $5^{\circ}\text{C}$  and  $25^{\circ}\text{C}$  using 0.5 mM of  $^{15}\text{N}/^{13}\text{C}$ -labeled XS20 in NMR buffer. The following NMR

spectra were acquired to obtain unambiguous assignments at both temperatures: 2D  $^{15}\text{N}$ - and  $^{13}\text{C}$ -HSQCs, 2D CON, 3D HNCO, HN(CA)CO, HNCA, CBCA(CO)NH, hCC(CO)NH (15 ms mixing time) and HBHA(CO)NH. Sequential connectivities were obtained by high-resolution 3D HNNH spectra (using >90 points in F1 dimension), which enabled the unambiguous assignment of the polyAla segments even for the highly overlapped regions of the spectra. The C-terminal residues from the XS constructions (namely residues 307–314) were not assigned in the spectra. HNHA and  $^{15}\text{N}$ -NOESY-HSQC and  $^{13}\text{C}$ -NOESY-HSQC (using 100 and 120 ms mixing time, respectively) were used to obtain J-couplings and NOEs for structure calculation. Assignments for XS20 at both temperatures were deposited in the BMRB database (codes 51978 and 51979, respectively). Additional XS20 samples were incubated up to 14 days at 25 °C with no agitation in incubation buffer before the NMR experiments. Only minor sample degradation within NMR incubation times was detected by mass spectrometry and gel electrophoresis. Assignment of 0.5 mM  $^{15}\text{N}/^{13}\text{C}$ -labeled XS23 and XS26 in NMR buffer were obtained at 25 °C by means of 2D  $^{15}\text{N}$ - and  $^{13}\text{C}$ -HSQCs, 2D CON, 3D HNCO, HN(CA)CO, HNCA, CBCA(CO)NH and HNNH spectra. Additional samples of XS23 and XS26 were incubated for 14 and 7 days at 25 °C in incubation buffer with no agitation, respectively. Sequential connectivities were lost for the new crosspeaks appearing in the  $^{15}\text{N}$ -HSQC XS23 and XS26 spectra upon incubation. However, 3D CBCA(CO)NH spectra revealed that the large majority of those peaks belonged to residues preceded by Ala residues in disordered conformation, due to their characteristic C $\alpha$  and C $\beta$  chemical shifts<sup>38</sup>. Therefore, the new crosspeaks in the  $^{15}\text{N}$ -HSQC were assigned following arbitrary numbering, which was maintained in the nascent crosspeaks detected in  $^{13}\text{C}$ -HSQC and CACO spectra. In addition, alanine residues from the central section of the polyAla  $\alpha$ -helix (residues 253–258 in particular) showed strong signal overlap in the H-N correlations, which enabled to observe nascent crosspeaks in the 3D CBCA(CO)NH spectra with chemical shifts corresponding to alanines in alternative conformations. These disordered polyAla crosspeaks were also observed in 2D  $^{13}\text{C}$ -HSQC spectra, which displays the whole palette of conformation-dependent C-H correlations. The nascent alanine moieties were reproduced in different protein batches upon incubation. 0.5 mM  $^{15}\text{N}/^{13}\text{C}$ -labeled L20 was used to validate the structure of XS20 using 2D  $^{15}\text{N}$ - and  $^{13}\text{C}$ -HSQCs, 2D CON, 3D HNCO and CBCA(CO)NH spectra at 5 °C. 0.5 mM  $^{15}\text{N}/^{13}\text{C}$ -labeled XS<sub>CP</sub> was assigned at 25 °C by means of 2D  $^{15}\text{N}$ - and  $^{13}\text{C}$ -HSQCs, 2D CON, 3D CBCA(CO)NH and HNNH spectra.

Loss of NMR signal intensity was calculated comparing the  $^{15}\text{N}$ -HSQC NMR signal intensities for freshly prepared XS20 samples at 0.5 mM and 25 °C *vs.* the NMR signal intensity for XS20 samples which were incubated at 25 °C. For CACO spectra, signal intensity for 0.225 mM XS26 incubated in incubation buffer for 48 h was compared to that for fresh XS26. Secondary chemical shifts were calculated as  $\Delta\text{C}\alpha$ - $\Delta\text{C}\beta$ , where  $\Delta\text{C}\alpha$  and  $\Delta\text{C}\beta$  are the differences between experimentally obtained C $\alpha$  and C $\beta$  chemical shifts at the specific protein concentrations and computed shifts for disordered PHOX2B constructs at the same temperature and pH ([https://spin.niddk.nih.gov/bax/nmrserver/Poulsen\\_rc\\_CS/](https://spin.niddk.nih.gov/bax/nmrserver/Poulsen_rc_CS/))<sup>66,67</sup>.  $^3J_{\text{HNHA}}$  coupling constants were calculated on the basis of 3D HNHA experiments<sup>68</sup>.  $^{15}\text{N}$  longitudinal ( $R_1$ ), rotating frame ( $R_{1\rho}$ ) and ( $^1\text{H}$ )- $^{15}\text{N}$  NOE relaxation data for 0.05 and 0.5 mM XS20, 0.5 mM XS23 and 0.5 mM XS26 samples in NMR buffer were obtained using standard Bruker pulse sequences acquired at 25 °C (or 5 °C and 25 °C for XS20) in the 800 MHz spectrometer. In all relaxation experiments the spectral width was 12500 Hz for  $^1\text{H}$  and 2027 Hz for  $^{15}\text{N}$  dimensions. Eight relaxation delays (20; 60; 100; 240; 600; 1000; 1600 and 2400 ms) were used for  $^{15}\text{N}$   $R_1$  measurements, as well as to measure  $^{15}\text{N}$   $R_{1\rho}$  values (using 8.00; 35.99; 75.99; 100.00; 156.00; 200.00; 347.99 and 699.99 ms). Relaxation values and uncertainties were calculated by fitting an exponential decay to the data. Het-NOEs were calculated from the ratio of cross peak intensities in

spectra collected with and without amide proton saturation during the recycle delay. Uncertainties in peak heights were determined from the standard deviation of the intensity distribution in signal-less spectral regions. All NMR spectra were processed in Topspin 4.1.1 and analyzed in Sparky 3.19<sup>69</sup>.

The  $S^2$  order parameter was derived from the chemical shifts obtained by TALOS-N<sup>70</sup>. The overall correlation time ( $\tau_c$ ) was estimated from the ratios of the mean values of  $T_1$  and  $T_2$  as  $\tau_c \approx 1/(4 \pi \cdot ^{15}\text{N}$  frequency in Hz) \*  $((6 \cdot R_2/R_1) - 7)^{1/2}$  which was derived from eqn. 8 of Kay et al.<sup>37</sup> that considers  $J(0)$  and  $J(\omega)$  spectral densities and discounts terms of higher frequencies from a subset of residues with little internal motion and no significant exchange broadening. This subset excluded residues with  $T_2$  values lower than the average minus one standard deviation, unless their corresponding  $T_1$  values were larger than the average plus one standard deviation<sup>71</sup>. Experimental  $\tau_c$  values for XS20 at 5 °C and 25 °C were compared to the correlation time calculated on the basis of its size and shape in solution, using HydroNMR<sup>72</sup>. The atomic structures for the fragment 241–260 of PHOX2B at both temperatures (pure  $\alpha$ -helix) were considered in the calculations, and  $\tau_c$  values were averaged for the ensemble structures. Temperature-dependent solvent viscosity was considered in the calculations, applying 3.0 Å of atomic element radius. 3D coordinates of canonical  $\alpha$ -helices for 20, 23 and 26 alanines were also generated using the builder module in PyMOL (PyMOL Molecular Graphics System, Version 1.8.x Schrödinger, LLC.). HydroNMR was applied to the generated  $\alpha$ -helices, which reported the following  $\tau_c$  averaged values (in ns): 1.85 (XS20 at 5 °C), 1.05 (XS20 at 25 °C), 1.24 (XS23 at 25 °C) and 1.43 (XS26 at 25 °C). These values indicate an increase of 0.19 ns in  $\tau_c$  every 3 additional alanines, which is significantly smaller than the experimental values obtained for the extended variants. These discrepancies, in addition to the difference between experimental and calculated  $\tau_c$  for XS20 structure, indicate that PHOX2B has a tendency to assemble in solution through the polyAla helix, and that PARMs increase these associations. Relaxation and correlation time values are shown in Supplementary Table 1.

The NMR structure of XS20 at 5 °C and 25 °C was calculated with the program CYANA v3.98.13<sup>73</sup> based on experimental NOE-derived distance constraints and TalosN-derived dihedral constraints<sup>36</sup> following the standard 7-cycle iterative process and a final annealing using the list of restraints obtained in the last cycle. One hundred structures were generated using the mentioned procedure. Structural statistics for the calculated structures are summarized in Supplementary Tables 2 and 3.  $\alpha$ -helical boundaries were experimentally validated by the observation of  $i+3$  and  $i+4$  H-H NOEs. The 20 conformers with the lowest target function values were selected and deposited in the Protein Data Bank under the accession numbers 8PTL and 8PUI. The structural ensembles were visualized and examined using MolMol<sup>74</sup> and Pymol v2.0 (PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.). PROCHECK-NMR<sup>75</sup> version 3.4.4 was used to analyze the quality of the refined structures. Ramachandran statistics showed 77.5% most favored regions and 22.5% for the additionally allowed favored regions for the structure at 5 °C, while the statistics were 76.3% of most favored regions and 23.7% of additionally allowed favored regions for the structure at 25 °C.

NMR titrations were based on 2D  $^{15}\text{N}$ -HSQC,  $^{13}\text{C}$ -HSQC and CACO spectra measured at 25 °C in the 800 MHz spectrometer in 20 mM Hepes/10 mM NaCl/5 mM MgCl<sub>2</sub>/2 mM ATP/1 mM DTT/0.03 % sodium azide [pH 6.8] buffer. 0.07 mM (for the  $^{15}\text{N}$ -HSQC spectra), 0.1 mM (for the  $^{13}\text{C}$ -HSQC spectra) and 0.225 mM (for the CACO spectra)  $^{15}\text{N}/^{13}\text{C}$ -labeled XS26 were mixed with fresh 0.14 mM unlabeled (therefore, NMR invisible) HSP70 and HSP90 (for the  $^{15}\text{N}$ -HSQC-based titrations) or 0.1 mM fresh HSP70 (for the  $^{13}\text{C}$ -HSQC-based titrations), while CACO samples were incubated 48 h with no agitation at 25 °C with 0.45 mM HSP70. CACO-based titrations with the co-chaperone DNAJBI were done mixing 0.14 mM of  $^{15}\text{N}/^{13}\text{C}$  XS26, 0.28 mM of

unlabeled HSP70 and 0.14 mM of unlabeled DNAJB1 in the above-mentioned buffer and incubating for 48 h with no agitation at 25 °C. The data shown are the signal intensity decay for the interactions, which is obtained comparing the normalized NMR signal intensity for the complex ratios and the intensity for XS26 alone at the same concentration. For the <sup>13</sup>C-HSQC-based titrations, values shown are averaged between the C $\alpha$  and C $\beta$  crosspeak intensities, and include alternative conformations both for polyAla and nascent alanine moieties.

Diffusion experiments were acquired at 25 °C in NMR buffer in the 800 MHz spectrometer using a pseudo 2D <sup>1</sup>H NMR pulse sequence using stimulated echo with bipolar gradient pulses. Diffusion time (160 ms) and length of DOSY gradient pulse (1.4 ms) were optimized to obtain a correct signal decay using an exponential gradient ramp. Both XS20 and XS26 were measured at 0.05, 0.1 and 0.5 mM at different incubation times (7 days at 25 °C). Gradients were calibrated before measuring DOSY experiments using a “doped water” sample (1 % H<sub>2</sub>O + 0.1 mg GdCl<sub>3</sub>/ml D<sub>2</sub>O + 0.1 % <sup>13</sup>CH<sub>3</sub>OH) following the Bruker protocol. The pseudo 2D DOSY dataset was processed in Topspin 4.3.0 and analyzed with Dynamics Center (Bruker). Diffusion coefficient in 8 M urea was corrected applying a viscosity coefficient of 1.663 compared to water<sup>76</sup>.

### Circular Dichroism (CD)

CD spectra were recorded at 25 °C before and after thermal denaturation at 95 °C in a J-810 Jasco polarimeter using a 0.1 cm cuvette and 0.265 mg/ml (for XS20 and XS26) and 0.132 mg/ml (for XS23) protein concentration. The sample buffer was 5 mM KPi/10 mM NaCl /1 mM DTT/0.01% azide [pH 6.8]. Samples were serially diluted in water to minimize signal intensity. Thermal unfolding was induced by heating the samples at 95 °C for 15 min followed by a 6 min cool down to 25 °C. For each measurement, ten scans were averaged and, after correction for the buffer contribution, transformed into mean residue weight ellipticity ( $\Theta \cdot 10^{-3}$  (degree  $\cdot$  cm<sup>2</sup>  $\cdot$  dmol<sup>-1</sup>)) as follows (Eq. (1)):

$$\text{Molar Ellipticity} = \frac{\theta / 10 \times \text{Optical path length (cm)} \times \text{Concentration (M)}}{\text{residue number} - 1} \quad (1)$$

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

All materials are readily available from the corresponding author upon request. Source data are provided with this paper. The structural data generated in this study have been deposited in the PDB database under the accession codes 8PTL and 8PUJ and the chemical shifts data in the BMRB database under the accession codes 51978 and 51979. Source data are provided as a Source Data file. Source data are provided with this paper.

### References

- Bernacki, J. P. & Murphy, R. M. Length-dependent aggregation of uninterrupted polyalanine peptides. *Biochemistry* **50**, 9200–9211 (2011).
- Chavali, S., Singh, A. K., Santhanam, B. & Babu, M. M. Amino acid homorepeats in proteins. *Nat. Rev. Chem.* **4**, 420–434 (2020).
- Matera, I. et al. PHOX2B mutations and polyalanine expansions correlate with the severity of the respiratory phenotype and associated symptoms in both congenital and late onset Central Hypoventilation syndrome. *J. Med. Genet.* **41**, 373–380 (2004).
- Cummings, C. J. & Zoghbi, H. Y. Fourteen and counting: unraveling trinucleotide repeat diseases. *Hum. Mol. Genet.* **9**, 909–916 (2000).
- Messaed, C. & Rouleau, G. A. Molecular mechanisms underlying polyalanine diseases. *Neurobiol. Dis.* **34**, 397–405 (2009).
- Pirone, L. et al. Molecular insights into the role of the polyalanine region in mediating PHOX2B aggregation. *FEBS J.* **286**, 2505–2521 (2019).
- Pelassa, I. et al. Association of polyalanine and polyglutamine coiled coils mediates expansion disease-related protein aggregation and dysfunction. *Hum. Mol. Genet.* **23**, 3402–3420 (2014).
- Baias, M. et al. Structure and Dynamics of the Huntingtin Exon-1 N-Terminus: A Solution NMR Perspective. *J. Am. Chem. Soc.* **139**, 1168–1176 (2017).
- Escobedo, A. et al. Side chain to main chain hydrogen bonds stabilize a polyglutamine helix in a transcription factor. *Nat. Commun.* **10**, 2034 (2019).
- Elena-Real, C. A. et al. The structure of pathogenic huntingtin exon 1 defines the bases of its aggregation propensity. *Nat. Struct. Mol. Biol.* **30**, 309–320 (2023).
- Coleman, M., Boros, S. J., Huseby, T. L. & Brennom, W. S. Congenital central hypoventilation syndrome. A report of successful experience with bilateral diaphragmatic pacing. *Arch. Dis. Child* **55**, 901–903 (1980).
- Di Lascio, S., Benfante, R., Cardani, S. & Fornasari, D. Research Advances on Therapeutic Approaches to Congenital Central Hypoventilation Syndrome (CCHS). *Front. Neurosci.* **14**, 615666 (2020).
- Amiel, J. et al. Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central hypoventilation syndrome. *Nat. Genet.* **33**, 459–461 (2003).
- Di Lascio, S. et al. Structural and functional differences in PHOX2B frameshift mutations underlie isolated or syndromic congenital central hypoventilation syndrome. *Hum. Mutat.* **39**, 219–236 (2018).
- Brunet, J. F. & Pattyn, A. Phox2 genes - from patterning to connectivity. *Curr. Opin. Genet. Dev.* **12**, 435–440 (2002).
- Sasaki, A. et al. Molecular analysis of congenital central hypoventilation syndrome. *Hum. Genet.* **114**, 22–26 (2003).
- Patwari, P. P. et al. Congenital central hypoventilation syndrome and the PHOX2B gene: a model of respiratory and autonomic dysregulation. *Respir. Physiol. Neurobiol.* **173**, 322–335 (2010).
- Bachetti, T. et al. Distinct pathogenetic mechanisms for PHOX2B associated polyalanine expansions and frameshift mutations in congenital central hypoventilation syndrome. *Hum. Mol. Genet.* **14**, 1815–1824 (2005).
- Trochet, D. et al. Molecular consequences of PHOX2B missense, frameshift and alanine expansion mutations leading to autonomic dysfunction. *Hum. Mol. Genet.* **14**, 3697–3708 (2005).
- Di Zanni, E. et al. In vitro drug treatments reduce the deleterious effects of aggregates containing polyAla expanded PHOX2B proteins. *Neurobiol. Dis.* **45**, 508–518 (2012).
- Di Lascio, S. et al. Transcriptional dysregulation and impairment of PHOX2B auto-regulatory mechanism induced by polyalanine expansion mutations associated with congenital central hypoventilation syndrome. *Neurobiol. Dis.* **50**, 187–200 (2013).
- Dubreuil, V. et al. A human mutation in Phox2b causes lack of CO<sub>2</sub> chemosensitivity, fatal central apnea, and specific loss of parafacial neurons. *Proc. Natl Acad. Sci. USA* **105**, 1067–1072 (2008).
- Polling, S. et al. Polyalanine expansions drive a shift into  $\alpha$ -helical clusters without amyloid-fibril formation. *Nat. Struct. Mol. Biol.* **22**, 1008–1015 (2015).
- Basu, S. et al. Unblending of Transcriptional Condensates in Human Repeat Expansion Disease. *Cell* **181**, 1062–1079.e1030 (2020).
- Mittag, T. & Pappu, R. V. A conceptual framework for understanding phase separation and addressing open questions and challenges. *Mol. Cell* **82**, 2201–2214 (2022).
- Carrasco, J. et al. Metamorphism in TDP-43 prion-like domain determines chaperone recognition. *Nat. Commun.* **14**, 466 (2023).

27. Ma, B. & Nussinov, R. Molecular dynamics simulations of alanine rich beta-sheet oligomers: Insight into amyloid formation. *Protein Sci.* **11**, 2335–2350 (2002).
28. Farag, M. et al. Condensates formed by prion-like low-complexity domains have small-world network structures and interfaces defined by expanded conformations. *Nat. Commun.* **13**, 7722 (2022).
29. Murray, D. T. et al. Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell* **171**, 615–627.e616 (2017).
30. Hughes, M. P. et al. Atomic structures of low-complexity protein segments reveal kinked  $\beta$  sheets that assemble networks. *Science* **359**, 698–701 (2018).
31. Luo, F. et al. Atomic structures of FUS LC domain segments reveal bases for reversible amyloid fibril formation. *Nat. Struct. Mol. Biol.* **25**, 341–346 (2018).
32. Gui, X. et al. Structural basis for reversible amyloids of hnRNPA1 elucidates their role in stress granule assembly. *Nat. Commun.* **10**, 2006 (2019).
33. Burke, K. A., Janke, A. M., Rhine, C. L. & Fawzi, N. L. Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Mol. Cell* **60**, 231–241 (2015).
34. Brady, J. P. et al. Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. *Proc. Natl Acad. Sci. USA* **114**, E8194–E8203 (2017).
35. Galvanetto, N. et al. Extreme dynamics in a biomolecular condensate. *Nature* **619**, 876–883 (2023).
36. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
37. Kay, L. E., Torchia, D. A. & Bax, A. Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972–8979 (1989).
38. Wang, Y. & Jardetzky, O. Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Sci.* **11**, 852–861 (2002).
39. Alberti, S., Gladfelter, A. & Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419–434 (2019).
40. van Dijk, E., Hoogveen, A. & Abeln, S. The hydrophobic temperature dependence of amino acids directly calculated from protein structures. *PLoS Comput. Biol.* **11**, e1004277 (2015).
41. Ciani, B., Hutchinson, E. G., Sessions, R. B. & Woolfson, D. N. A designed system for assessing how sequence affects alpha to beta conformational transitions in proteins. *J. Biol. Chem.* **277**, 10150–10155 (2002).
42. Shinchuk, L. M. et al. Poly-(L-alanine) expansions form core beta-sheets that nucleate amyloid assembly. *Proteins* **61**, 579–589 (2005).
43. Albrecht, A. N. et al. A molecular pathogenesis for transcription factor associated poly-alanine tract expansions. *Hum. Mol. Genet.* **13**, 2351–2359 (2004).
44. Bachetti, T. et al. Geldanamycin promotes nuclear localisation and clearance of PHOX2B misfolded proteins containing polyalanine expansions. *Int J. Biochem. Cell Biol.* **39**, 327–339 (2007).
45. Lilliu, E. et al. Polyserine repeats promote coiled coil-mediated fibril formation and length-dependent protein aggregation. *J. Struct. Biol.* **204**, 572–584 (2018).
46. Gojobori, J. & Ueda, S. Elevated evolutionary rate in genes with homopolymeric amino acid repeats constituting nondisordered structure. *Mol. Biol. Evol.* **28**, 543–550 (2011).
47. Orr, H. T. & Zoghbi, H. Y. Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* **30**, 575–621 (2007).
48. Caburet, S. et al. A recurrent polyalanine expansion in the transcription factor FOXL2 induces extensive nuclear and cytoplasmic protein aggregation. *J. Med. Genet.* **41**, 932–936 (2004).
49. Utsch, B. et al. Molecular characterization of HOXA13 polyalanine expansion proteins in hand-foot-genital syndrome. *Am. J. Med. Genet. A* **143A**, 3161–3168 (2007).
50. Iizuka, Y. et al. Toxicity of internalized polyalanine to cells depends on aggregation. *Sci. Rep.* **11**, 23441 (2021).
51. Mensah, M. A. et al. Aberrant phase separation and nucleolar dysfunction in rare genetic diseases. *Nature* **614**, 564–571 (2023).
52. Lester, E. et al. Cytosolic condensates rich in polyserine define subcellular sites of tau aggregation. *Proc. Natl Acad. Sci. USA* **120**, e2217759120 (2023).
53. Bravo-Arredondo, J. M. et al. The folding equilibrium of huntingtin exon 1 monomer depends on its polyglutamine tract. *J. Biol. Chem.* **293**, 19613–19623 (2018).
54. Milner-White, J. E., Watson, J. D., Qi, G. & Hayward, S. Amyloid formation may involve alpha- to beta sheet interconversion via peptide plane flipping. *Structure* **14**, 1369–1376 (2006).
55. Lu, J. et al. CryoEM structure of the low-complexity domain of hnRNPA2 and its conversion to pathogenic amyloid. *Nat. Commun.* **11**, 4090 (2020).
56. Blondelle, S. E., Forood, B., Houghten, R. A. & Pérez-Payá, E. Polyalanine-based peptides as models for self-associated beta-pleated-sheet complexes. *Biochemistry* **36**, 8393–8400 (1997).
57. Zhang, H. et al. RNA Controls PolyQ Protein Phase Transitions. *Mol. Cell* **60**, 220–230 (2015).
58. Abedini, A. & Raleigh, D. P. A role for helical intermediates in amyloid formation by natively unfolded polypeptides? *Phys. Biol.* **6**, 015005 (2009).
59. Mier, P., Elena-Real, C. A., Cortés, J., Bernadó, P. & Andrade-Navarro, M. A. The sequence context in poly-alanine regions: structure, function and conservation. *Bioinformatics* **38**, 4851–4858 (2022).
60. Hervás, R. & Oroz, J. Mechanistic Insights into the Role of Molecular Chaperones in Protein Misfolding Diseases: From Molecular Recognition to Amyloid Disassembly. *Int J. Mol. Sci.* **21**, 9186 (2020).
61. Kameda, T. & Takada, S. Secondary structure provides a template for the folding of nearby polypeptides. *Proc. Natl Acad. Sci. USA* **103**, 17765–17770 (2006).
62. Bolognesi, B. et al. The mutational landscape of a prion-like domain. *Nat. Commun.* **10**, 4162 (2019).
63. Marley, J., Lu, M. & Bracken, C. A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* **20**, 71–75 (2001).
64. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
65. Markley, J. L. et al. Recommendations for the presentation of NMR structures of proteins and nucleic acids. IUPAC-IUBMB-IUPAB Inter-Union Task Group on the Standardization of Data Bases of Protein and Nucleic Acid Structures Determined by NMR Spectroscopy. *J. Biomol. NMR* **12**, 1–23 (1998).
66. Kjaergaard, M. & Poulsen, F. M. Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *J. Biomol. NMR* **50**, 157–165 (2011).
67. Kjaergaard, M., Brander, S. & Poulsen, F. M. Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *J. Biomol. NMR* **49**, 139–149 (2011).
68. Vuister, G. W., Delaglio, F. & Bax, A. The use of 1JC alpha H alpha coupling constants as a probe for protein backbone conformation. *J. Biomol. NMR* **3**, 67–80 (1993).
69. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).

70. Wang, I. F. et al. The self-interaction of native TDP-43 C terminus inhibits its degradation and contributes to early proteinopathies. *Nat. Commun.* **3**, 766 (2012).
71. Pawley, N. H., Wang, C., Koide, S. & Nicholson, L. K. An improved method for distinguishing between anisotropic tumbling and chemical exchange in analysis of <sup>15</sup>N relaxation parameters. *J. Biomol. NMR* **20**, 149–165 (2001).
72. García de la Torre, J., Huertas, M. L. & Carrasco, B. HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.* **147**, 138–146 (2000).
73. Güntert, P. Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* **278**, 353–378 (2004).
74. Koradi, R., Billeter, M. & Wüthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 29–32 (1996).
75. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486 (1996).
76. Kawahara, K. & Tanford, C. Viscosity and density of aqueous solutions of urea and guanidine hydrochloride. *J. Biol. Chem.* **241**, 3228–3232 (1966).

## Acknowledgements

This work was supported by the grants PID2019-109276RA-I00 and PID2022-142382OB-I00 funded by MCIN/AEI/10.13039/501100011033 and by FEDER A way to make Europe (to J.O.), and by donations from the Spanish and Mexican Ondine Associations. J.O. was a recipient of a Leonardo Grant from the Spanish BBVA Foundation (BBM\_TRA\_0203) and is a Ramón y Cajal Fellow of the Spanish AEI-Ministry of Science and Innovation (RYC2018-026042-I funded by MCIN/AEI/10.13039/501100011033 and by “ESF Investing in your future”). E.C. acknowledges the funding by Fundação para a Ciência e a Tecnologia (FCT-Portugal) of the UCIBIO project (UIDP/04378/2020 and UIDB/04378/2020) and Associate Laboratory Institute for Health and Bioeconomy – i4HB Project (LA/P/0140/2020). S.F. is a PhD student under the PTNMRPhD Program from FCT-Portugal (PD/BD/148028/2019). M.Z. was supported by the European Research Council (ERC) under the EU Horizon 2020 research and innovation programme (grant agreement No. 787679). P.T. acknowledges funding from the Free State of Bavaria through the ONE MUNICH Project Munich Multiscale Biofabrication and A.M.V. is grateful for the support from the Center for NanoScience Munich. N.M.R. experiments were performed in the “Manuel Rico” NMR Laboratory (LMR) of the Spanish National Research Council (CSIC), a node of the Spanish Large-Scale National Facility (ICTS R-LRB). Authors would like to acknowledge Dr. José Manuel Pérez Cañadillas (IQF/CSIC) for providing the TXA fusion plasmid, and for critical reading of the manuscript. Authors acknowledge the use of the Servicio de Microscopía Láser Confocal y Multidimensional in vivo (CIB-CSIC) and Servicio de Proteómica y Genómica (CIB-CSIC).

## Author contributions

Conceptualization, J.O. Sample preparation, data acquisition and analysis, R. A., M. A. T., A. M. V., S. F., D. P.U., M. B., E. J. C., P. T., M. Z., J. O. NMR structure calculation, D. P. U., J. O. Writing – Original Draft, J. O. Writing, Review and Editing, all authors. Funding acquisition, M.Z., J. O. Supervision, J. O.

## Competing interests

R.A., M.A.T., D.P.U. and J.O. are co-inventors of the patent application EP23383165 filed on 15/11/2023 by Consejo Superior de Investigaciones Científicas, that includes the production and structural features of PHOX2B upon phase separation. S.F., M.B., E.J.C., M.Z., P.T. and A.M.V. declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-46236-5>.

**Correspondence** and requests for materials should be addressed to Javier Oroz.

**Peer review information** *Nature Communications* thanks Xavier Salvatella and the other, anonymous, reviewer for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024