Biomolecular condensates sustain pH gradients at equilibrium driven by charge neutralisation

Hannes Ausserwöger[1,*], Rob Scrutton[1,*], Tomas Sneideris[1,*], Charlotte M. Fischer[1,*], Daoyuan Qian[1], Ella de Csilléry[1], Kadi L. Saar[1], Alan Z. Białek[1], Marc Oeller[2], Georg Krainer[1,3], Titus M. Franzmann[4], Sina Wittmann[5], Juan M. Iglesias-Artola[6], Gaetano Invernizzi[7], Anthony A. Hyman[6], Simon Alberti[4], Nikolai Lorenzen[7], Tuomas P. J. Knowles[1,8,+]

[1] Centre for Misfolding Diseases, Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

[2] Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Martinsried, Germany

[3] Institute of Molecular Biosciences (IMB), University of Graz, Humboldtstraße 50, 8010 Graz, Austria

[4] Biotechnology Center (BIOTEC), Center for Molecular and Cellular Bioengineering (CMCB), Technische Universität Dresden, Tatzberg 47/49, Dresden, Germany

[5] Institute of Molecular Biology (IMB), Ackermannweg 4, 55128 Mainz, Germany

[6] Max Planck Institute of Cell Biology and Genetics (MPI-CBG), 01307 Dresden, Germany

[7] Global Research Technologies, Novo Nordisk A/S, 2760 Måløv, Denmark

[8] Cavendish Laboratory, Department of Physics, University of Cambridge, JJ Thomson Road, Cambridge CB3 0HE, United Kingdom

* contributed equally

[+] correspondence should be addressed to tpjk2@cam.ac.uk

Abstract

Electrochemical gradients are essential to the functioning of cells and are typically formed across membranes using active transporters and require energy input to maintain them. Here, we show by contrast that biomolecular condensates are able to sustain significant pH gradients without any external energy input. We explore the thermodynamic driving forces that establish this gradient using a microfluidics-based droplet platform that allows us to sample in a continuous manner both the stability and composition of the condensates across a wide pH range. These results reveal that condensed biomolecular systems adjust the pH of the dense phase towards the isoelectric point (pI) of the component polypeptide chains. We demonstrate, on the basis of two representative systems, FUS and PGL3, that condensates can create both alkaline and acidic gradients with a magnitude exceeding one pH unit. Investigations of multicomponent protein/nucleic acid systems further show that heterotypic interactions can modulate condensate pH gradients. We further investigate using a bioinformatics approach the diversity of electrochemical properties of complex condensates by studying a large set of human condensate networks, showing that these span a wide range of mixture pIs and pH-response behaviours. In summary, our results reveal that protein condensation may present a fundamental physico-chemical mechanism for the effective segregation and optimisation of functional processes through changes in the emergent electrochemical microenvironment.

Introduction

The proton (H^+) concentration, typically represented as the pH, is a uniquely important environmental factor in biology. It affects all fundamental processes underlying viability and functioning of cells including electrochemical potential gradients[1], chemical reaction rates[2] or macromolecular conformation and assembly[3,4]. pH mediated changes in the functional behaviour of proteins, for example, are driven by the inherent ability of functional groups to display different protonation states, shifting their physico-chemical properties. As a consequence, the pH is tightly regulated within the intracellular environment, with deviations from the cytosolic pH only common in membrane-bound organelles[5,6]. Through energy input using active transporters the organellar pH is regulated to optimise for the specific reactions and processes performed in the organelle[7–12]. Precise control over the intracellular pH is further aided by a high buffering capacity of the cytosol with the latter originating primarily from smaller ionic species such as phosphates and carbonates[5,6]. In comparison the concentration of protein associated ionisable side groups is typically far too low to contribute, accounting to only about 1% of the cytosolic buffering capacity in yeast cells[13].

Remarkably, this low-protein-concentration picture can be drastically altered as a result of biomolecular phase separation leading to the formation of a protein-rich dense phase and a protein-poor dilute phase. Here, collective interactions drive local concentration increases of biomolecules in the condensed state to give rise to unique chemical microenvironments at equilibrium[14,15]. These unique microenvironments, often referred to as condensates, are believed to play critical roles in cellular organisation and pathology[14,16,17]. They can yield distinct material properties or differential molecular partitioning[18], which has even been shown to extend to the formation of elemental ion gradients[19,20]. Reports even suggest that condensates could drive pH gradients intracellularly[21–26]. Hence, we sought a generalisable understanding of how pH influences and drives the collective assembly of biomolecules as well as how the internal environment of these assemblies can modulate external pH conditions and establish pH gradients.

In this work, we demonstrate that biomolecular condensates sustain pH gradients at equilibrium by minimising intermolecular electrostatic repulsion. Experiments using a high-throughput, microfluidics-based experiments capable of mapping pH responsive phase boundaries highlight that phase behaviour is governed by macromolecular charge neutrality and that polypeptide chains buffer pH towards their pI. The change of the dense phase pH appears to be aided by an exclusion of the buffer molecules, which, combined with the local protein concentration, allows the proteinaceous amino acids (AAs) to affect pH. Investigations of representative phase-separation-prone protein systems further show that condensates can create pH gradients, either alkaline or acidic in nature, exceeding a unit in pH under a range of phase separation triggers.

Moreover, we show that phase separation-prone proteins generally exhibit a charge bias, suggesting condensates could establish diverse range of electrochemical properties comparable to that found in membrane bound organelles. To capture the complexity of relevant condensation systems we apply a bioinformatics approach capable of correlating the physico-chemical properties of condensate proteomes with observed intracellular pH gradients. Investigations on a set of 85 predicted human condensate networks highlight that condensate proteomes directly encode for a diversity of different electrochemical properties and response behaviours. Together, our findings suggest that condensates create a spectrum of pH microenvironments by equilibrium phase separation potentially expanding the range of accessible intracellular functional chemical microenvironments.

Results

Mapping pH responsive phase boundaries in a continuous manner

To quantify pH responsive phase behaviour, we apply an established Histidine and Succinic acid based buffering system (H/S buffer) to exercise precise control of pH between 3.5 to 9,[27] and integrate it with a combinatorial droplet microfluidic platform[28]. Specifically, we prepare equimolar Histidine and Succinic acid solutions at two different pH values (pH: 9 and 3.5) and mix these stocks to give a final pH that depends linearly on the stock volume fraction (see SI Fig. 1). This enables the calculation of pH from measuring the volume fraction of the individual buffers without requiring direct pH determination. Applying this buffering system to water-in-oil droplets generated using microfluidic technology then makes it possible to generate large numbers of individual chemical environments with distinct pH while also ensuring rapid and efficient mixing (Fig. 1a, b). These individual aqueous droplets are then subjected to image analysis to characterise each for the presence of phase separation by labelling the peptide with a covalent fluorescent NHS-ester dye (1:100 molar ratio) to detect dense phase features. We first subject this platform to a simple insulin peptide model system, comparing human insulin with a modified glargine variant (see SI Table 1 for sequences), which are known for their highly pH responsive phase behaviour and massive shifts in the physico-chemical properties with only few sequence differences. For insulin glargine this yields a continuous map of over 75000 data points with varying pH and insulin glargine variant concentrations enabling detailed quantification of the phase separation probability in the range of the probed conditions (Fig. 1c).

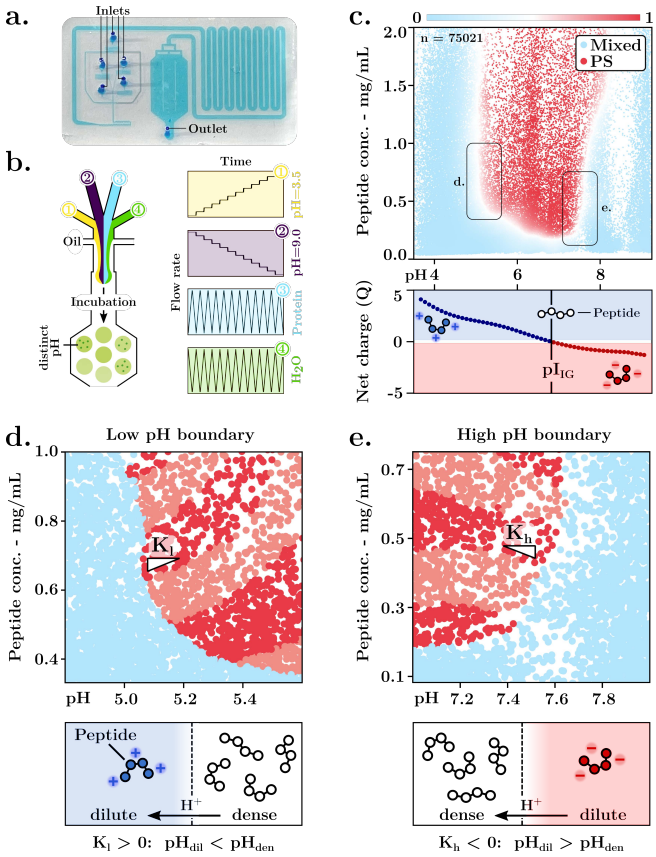

Figure 1: Deciphering the biophysical mechanisms underlying pH responsive phase behaviour. (a) Image of the microfluidic chip, filled with a coloured solution to highlight the device features, used to generate continuous pH phase diagrams. **(b)** Schematic of the chip design and flow profiles used to generate continuous pH phase boundaries. **(c)** pH responsive phase boundary of an insulin variant peptide, correlated with the molecular net charge (Q). The individual data points are shaded for their phase separation probability (P_{PS} = 1: red, P_{PS} = 0.5: white, P_{PS} = 0: blue) as determined by averaging over neighbouring data points (n_{tot} = 75021, see Supporting Information). **(d, e)** Visualisation of the dilute phase band gradients (K) at the low pH (d) and high pH (e) phase boundary. Dilute phase bands are displayed using red shading of phase separated points as generated by sectioning for the dilute phase concentration obtained in each individual microdroplet[19,29,30]. Lower panels: Schematic representations visualising the proton partitioning as inferred from the reduced tie line gradients.

The resulting pH responsive phase boundary displays a close correlation to the peptide net charge (Q) as phase separation is more prone closer to the peptide isoelectric point (pI) at 6.8. No phase separation was observed at extreme pHs and at low peptide concentrations (see Fig. 1c), suggesting molecular charge neutrality is a requirement for the intermolecular interactions to become favourable, which is in accordance with previous pH responsive phase behaviour preditions[31]. It is worth noting that the phase boundary is not perfectly symmetrical suggesting that both individual ionisation changes of specific amino acids as well as the effective charge experienced by the polypeptide chain can modify pH responsive phase boundary. Human insulin (pI_{Ins} = 5.4; glargine variant: N21G mutation, 2x R additions in the β-chain) displays a similar correlation between the phase boundary and the sequence charge, where in both cases the minimum in the critical concentration coincides with the pI as also confirmed by a support vector machine (SVM) fitting to the data points (see SI Fig. 2). Here, the dilute phase concentration (c_{dil}) is also minimal around the pI, meaning that phase separation propensity attains a maximum at the pI (see SI Fig. 3). Interestingly, a larger width of the pH phase boundary at a fixed peptide concentration of 1 mg/ml was observed for human insulin compared to the glargine variant (see SI Fig. 2), indicating that phase separation is still favourable at more extreme pH's. This coincides with an increase in the overall sequence hydrophilicity ($H_{pH=7,\,I}$ = 14, $H_{pH=7,\,IG}$ = 21; Kyte-Doolittle scale[32]) of the variant compared to insulin given the addition of two charged arginine residues. Hence, insulin is less hydrophilic, therefore exerting stronger van der Waals interactions, indicating a trade-off between intermolecular charge repulsion and attractive interactions.

Intermolecular electrostatic repulsion drives dense phase pH gradients

Next, we set out to characterise the proton partitioning between the dense and dilute phase by applying a previously published partitioning analysis approach[19,29,33]. For this analysis, we quantify the peptide dilute phase concentrations in each individual water-in-oil droplet obtained from our pH scanning microfluidics approach. Then, by simply sectioning all phase separated droplets for their dilute phase concentration in set thresholds of a < c_{dil} < b, we obtain so-called dilute phase bands (see Fig. 1d, e; light and dark red

shaded bands). The edge between two bands, which we refer to as reduced tie line gradients, corresponds to a set of conditions which have the same protein dilute phase concentrations but different total concentrations. This characterisation of the demixing behaviour of the system is connected to the higher dimensional tie lines[19,29,33] (see SI for interpretation in pH chemical space) enabling us to infer the directionality of the solute partitioning. When triggering insulin glargine phase separation at low pH (~ 5) the dilute phase band shows a positive gradient (see Fig. 1d, $K_l > 0$), indicating that the pH in the dilute (pH_{dil}) and dense phase (pH_{den}) are not equivalent. Specifically, this suggests that $pH_{den} > pH_{dil}$, i.e. protons are being excluded from the dense phase. This would bring the dense phase pH closer to the peptide pI at 6.8 than the starting pH condition. At the high pH phase boundary (around pH ~ 8) we find a negative reduced tie line gradient ($K_h < 0$), indicating $pH_{den} < pH_{dil}$, suggesting partitioning of protons into the dense phase under these conditions (see Fig. 1e). Our measurements suggest that pH_{den} is expected to be closer to the peptide pI than pH_{dil} independently of the pH starting condition. This would allow a reduction of the peptide net charge in the dense phase, therefore, decreasing charge repulsion. A similar behaviour is observed for insulin at the high pH boundary where $K_{pH} < 0$ (see, SI Fig. 4). The low pH phase boundary could not be resolved with the available buffer system.

We next set out to confirm the observation of pH gradients between the dense and dilute phase. To achieve quantitative pH readouts, we applied the ratiometric pH sensing dye SNARF-4F, which changes its emission spectrum as a function of the environmental pH (see SI Fig. 5 and 6). Using confocal imaging the SNARF dye signal can then be read out in a spatially resolved manner, making it possible to determine the dense and dilute phase pH (SI Fig. 7). Using this approach, we determine dilute and dense phase pH of insulin glargine samples at varying initial pH conditions to confirm the presence of significant pH gradients (see Fig. 2a). Crucially, while the dilute phase pH changes substantially with the starting pH, the dense phase pH stays roughly constant between pH ~ 6.5 and 7. Therefore, independent of the exact starting pH the dense phase pH for all samples coincides with the isoelectric point of the peptide at 6.8. Hence, the dense phase displays a distinct buffering capacity allowing it to regulate pH. This leads to a reduction in electrostatic repulsion and favours attractive intermolecular interactions. When applying the SNARF-4F dye to insulin samples a similar trend is observed where the dilute phase pH changes massively, yet the dense phase pH is constant independent of the starting pH (see SI Fig. 8).

To further understand the driving force behind the observed pH gradients, we next set out to characterise the dense phase repulsion free energy change with pH. Using a scaling argument, we suggest that the dense phase repulsion free energy scales with the square of sequence charge density (q) of the polypeptide chain, as the repulsion free energy of assembling a sphere of fixed radius is proportional to the square of its charge density[34]. Although factors including the dense phase concentration (c_{den}) will modulate the repulsion free energy, q^2 can serve as a proxy for the relative energetic change in repulsion with pH. We then calculated q at a range of pHs by normalising the peptide net charge, as obtained from tabled pKa values of the individual amino acids, by the sequence length N ($q = Q/N$). For insulin glargine the pH dependent q^2 profile yields a minimum around the pI, with a sharp increase at around pH 5 and 8 (Fig. 2b), corroborating that the observed pH gradients will help mitigate electrostatic repulsion in the dense phase. Accordingly, the repulsion profile also resembles the overall phase boundary, which is similarly observed for human insulin (SI Fig. 9), further supporting the generality of this simple concept.

We next sought to understand the mechanism driving the change in the dense phase pH by varying the amount of H/S buffer in the system. We quantified the insulin c_{dil} at varying concentrations of H/S buffer but fixed pH = 6.4 ($pI_{Ins} = 5.4$) and total peptide concentration (Fig. 2c). This gives a decrease in c_{dil} with decreasing H/S buffer strength, meaning that more protein is converted into the dense phase at lower buffer strength. This is expected as the formation of a dense phase necessitates a pH gradient towards the peptide pI to minimise repulsion and so, competes with the environmental buffer. By combining data sets at 2 and 1.5 mg/mL total insulin concentration we can then also perform partitioning analysis[19,30]. We find a negative reduced tie line gradient between insulin and the buffer (Fig. 2c), suggesting that the buffer concentration is lower in the dense phase than the dilute phase. Hence, the necessity to regulate pH to favour intermolecular interactions appears to even drive exclusion of the buffer from the dense phase to drive the buffering of protein associated side groups.

Effect of electrostatic modulation on pH responsive collective interactions

Given the observed importance of macromolecular charge to pH responsive assembly, we next seek to investigate the effect of charge screening via addition of potassium chloride (KCl), which generally weakens electrostatic interactions. The introduction of fixed KCl concentrations of 75 and 150 mM to the insulin glargine pH scan led to an increase in the width of the pH responsive phase boundary (see Fig. 2d, see SI Fig 10 for effect of KCl on H/S buffer system). Hence, phase separation can occur at more extreme peptide net charge in presence of ions due to the added charge shielding. It is noteworthy that the width of the phase boundary shifts more drastically in the high pH regime compared to the low pH regime for a given fixed peptide c_{tot} (e.g. 0.5 mg/mL). This is likely associated with differential affinities of the individual

counterions with the polypeptide chain: although the phase boundary is still expanding away from the pI, ions can differentially affect the charge state of individual amino acid residues, which is expected based on the large chemical differences in the amino acids as well as the counter ions.

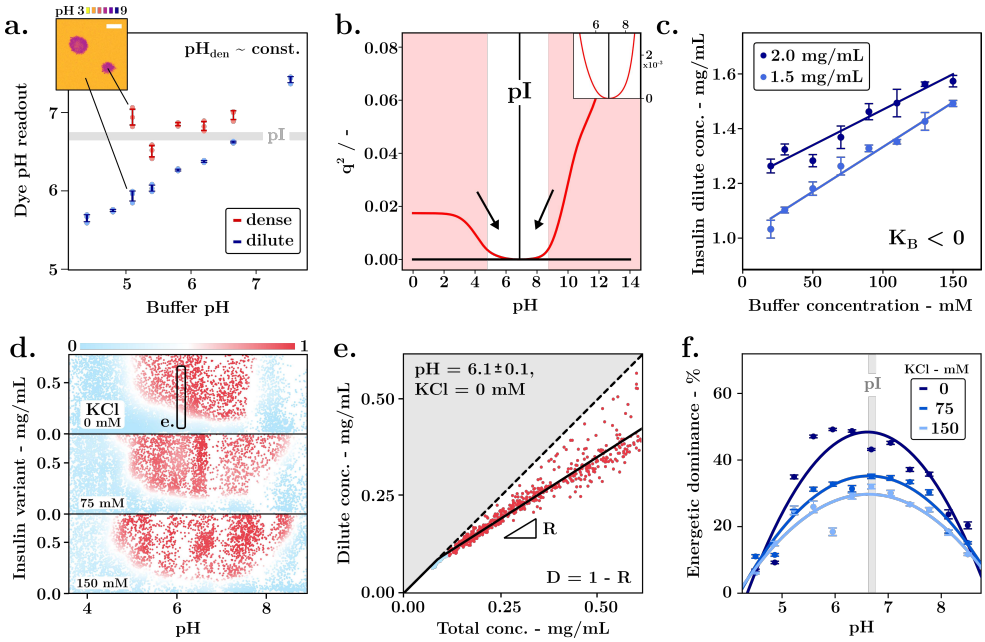

Figure 2: Dense phase pH gradients minimise intermolecular electrostatic repulsion. **(a)** Dense and dilute phase pH readouts of insulin variant samples at varying pH conditions using the ratiometric pH sensing dye SNARF-4F. Inset: representative image colour-coded for the pH readout. Scale bar: 10 µm. **(b)** Change in the square of the sequence charge density (q) used as a proxy for the repulsion free energy with varying pH for insulin variant. **(c)** Insulin dilute phase concentration change with varying H/S buffer concentration at fixed pH = 6.4 at a total insulin concentration of 2 and 1.5 mg/mL. **(d)** Effect of the addition of KCl on the pH responsive phase boundary of the insulin variant (n_{75mM} = 12623, n_{150mM} = 9969) with the data points being shaded for their phase separation probability (P_{PS} = 1: red, P_{PS} = 0.5: white, P_{PS} = 0: blue). **(e)** Change in c_{dil} of insulin variant with changing total concentration at a fixed pH and salt concentration (see marking box in d.). R - dilute phase response gradient. **(f)** Quantification of the protein dominance with varying pH and salt concentrations for the insulin variant.

To quantify the effect of the counter ions on the energetics we next applied a previously published dominance analysis approach[19,29,33]. Specifically, this allows us to map the relative free energy gain (D_i) of individual components to the overall driving force of phase separation. To do so, we sectioned the pH phase scan data into slices of fixed pH conditions and plotted the change in the peptide c_{dil} with increasing peptide c_{tot} (Fig. 2e). Below the critical concentration (c_{sat}) this will give a linear correlation with a response gradient (R) of one as $c_{dil} = c_{tot}$, while above the c_{sat} this gradient will decrease as $c_{dil} < c_{tot}$. A response gradient of R = 0, for example indicates that the excess peptide will be completely converted to the dense phase ($c_{dil} = c_{sat}$ = const.), giving a pure one component system (D = 1 – R = 1). At a pH of 6.1 and in the absence of salt, for example, we observe a D_p = 0.46 for insulin glargine, showing that even under salt-free conditions other factors contribute to the free energy gain associated with phase separation. We also find that D_p is maximised at pHs around the peptide pI (Fig. 2f), highlighting that other factors become much less relevant to the driving force of phase separation at low peptide net charge. Upon addition of KCl this trend is conserved but we observe a general decrease in the peptide dominance with increasing the KCl concentrations. This suggests that other components, especially ions, are starting to increasingly contribute to driving phase separation and corroborating our previous observation of phase separation at more extreme pH's due to modulation of the electrostatic repulsion.

PGL3 condensation creates acidic pH microenvironments

We reason that the necessity to reduce electrostatic repulsion via pH gradients in the dense phase will also translate to other protein phase transitions. To test this, we investigated the condensation of PGL3, a protein critical to the germ cell development of *Caenorhabditis elegans* by scaffolding the formation of P-granules[35]. Specifically, we induced PGL3 phase separation by decreasing the ionic strength down to 5 mM KCl at fixed pH (50 mM TRIS, pH = 7.3) in presence of SNARF-4F to read out the condensate pH. Indeed, a spectral shift of SNARF-4F was observed, from pH = 7.44 ± 0.11 in the dilute phase to more acidic pH = 6.40 ± 0.05 in the dense (Fig. 3a, b). This was further confirmed when evaluating large numbers of condensates as we obtain a clear gaussian distribution at acidic pH (Fig. 3c, see SI Fig. 11 for condensate mask). The resulting pH gradient brings the condensed phase pH significantly closer to the isoelectric point of the protein (pI_{PGL3} = 5.1), indicating that the decreasing intermolecular PGL3 repulsion is critical to condensation. Accordingly, the q^2 profile shows a minimum at acidic pH indicating that PGL3

repulsion will be decreased by the observed dense phase pH gradient (see Fig. 3d). Interestingly, the q^2 profile is highly asymmetrical for PGL3 as the increase in q^2 is much smaller at pH > pI, indicating that the intermolecular repulsion is expected to not increase as much at slightly higher pHs. Together with electrostatic shielding of the ions present this explains why the effective dense phase pH at around pH 6 to 6.5 still suffices to favouring phase separation.

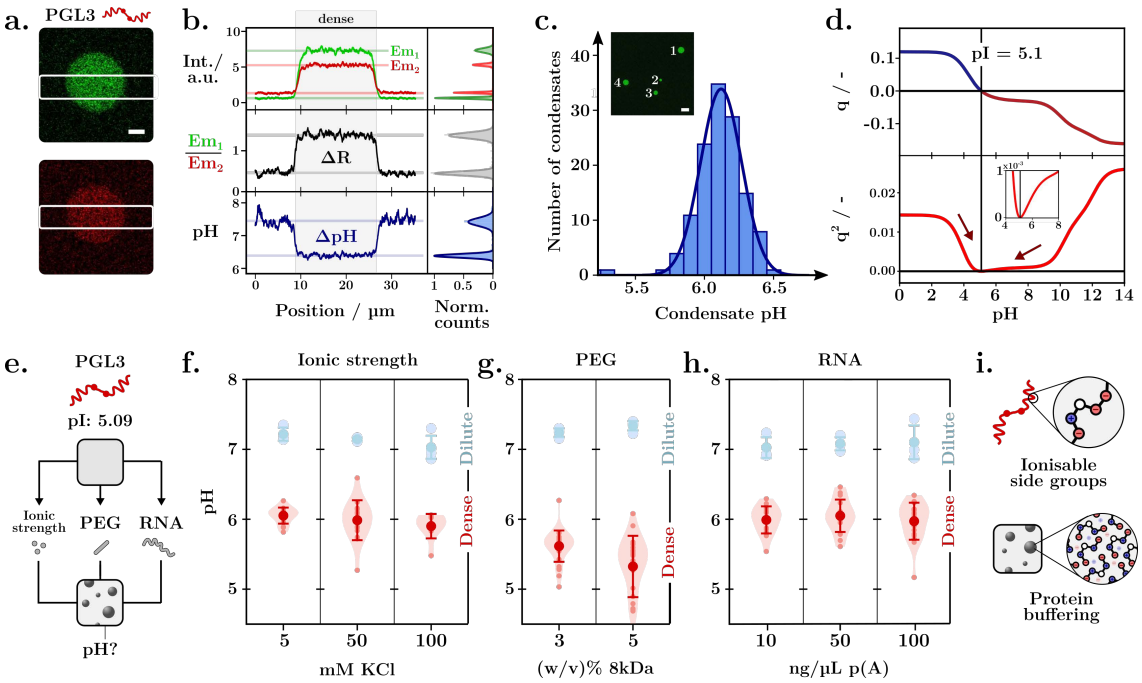

Figure 3: PGL3 condensates sustain acidic pH microenvironments buffered towards the protein pI. (a) Confocal imaging of PGL3 condensates at 2 µM protein concentration and 1 mM KCl, 50 mM TRIS buffer at pH = 7.3 using the ratiometric pH sensing dye SNARF-4F. Scale bar: 5 µm. **(b)** Spatially resolved ratiometric dye signal in dense and dilute phase for both emission channels, the emission intensity ratio and the corresponding pH, including readout histograms of all signal traces. **(c)** Condensate pH distribution for large numbers of PGL3 condensates formed at low ionic strength conditions from (a). Inset: Confocal image of multiple condensates. Scale bar: 20 µm. **(d)** Net charge density and net charge density square of PGL3 protein with varying pH. **(e)** Schematic of the experimental workflow to probe the effect of various environmental conditions on the obtained PGL3 condensate pH. **(f-h)** PGL3 dense and dilute phase pH at varying KCl, PEG, and RNA concentrations, respectively. Fixed total protein concentration and buffer condition (2 µM protein, pH=7.3, 50 mM TRIS buffer, 150mM KCl in g and h). **(i)** Schematic representation of the mechanism of action underlying the protein dense phase buffering driving pH gradients.

We next subjected PGL3 to phase separation under varying KCl, PEG and RNA concentrations (Fig. 3e), to understand the impact of different solute constitutions on the observed pH gradient. Changes in the KCl concentration between 5-100 mM at fixed protein and buffer conditions did not change pH_{den} significantly as all conditions display a dense phase pH < 6.5 (see Fig 3f). Hence, for PGL3, independent of the environmental ionic strength, the formation of a dense phase pH gradient is still necessary to promote protein-protein interactions. When triggering PGL3 phase separation by introducing polyethylene glycol (PEG) we also find an acidic dense phase pH closer to the protein pI (Fig. 3g), showing that the formation of a PGL3 dense phase requires sustaining a pH gradient even at higher ionic strengths and when using a different phase separation trigger. At 5% PEG, for example, $pH_{den} = 5.3 \pm 0.4$ was observed comparable to the low ionic strength condition.

Heteromolecular interactions are a critical contributor to protein phase separation and PGL3 in particular is known to undergo RNA dependent co-phase separation[35]. To investigate condensate pH gradients under these conditions we triggered PGL3 phase separation by addition of polyadenylic acid RNA (p(A)) to find $pH_{den} \sim 6$ (Fig. 3h), indicating a pH gradient under heterotypic condensation. Crucially, nucleic acids are also characterised by a low pI because of their phosphate backbone groups. Therefore, heterotypic PGL3-RNA interactions will also become more favourable at lower pH because of a decrease in repulsion. Similarly, the co-partitioned RNA will also contribute to decreasing the dense phase pH via buffering of its phosphate backbones. Taken together, PGL3 condensates form acidic pH microenvironments under a wide range of phase separation conditions, highlighting the ability of PGL3 to minimise electrostatic repulsion via regulating the dense phase pH towards its pI through the protein's innate buffering capacity (see Fig. 3i scheme).

Effective repulsion and heterotypic interactions modulate condensate pH gradients

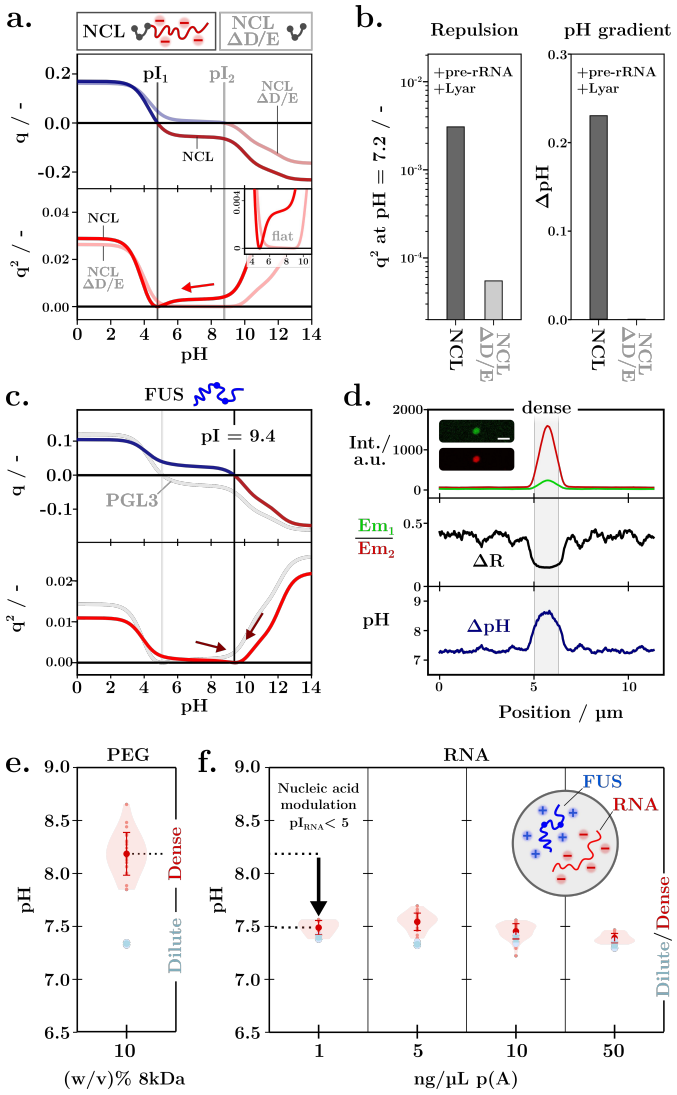

Figure 4: Condensate pH gradients are controlled by degree of intermolecular repulsion and heteromolecular interactions. (a) Comparison net charge density and net charge density square for the NCL protein and ΔD/E NCl deletion mutation (deletion of a sequence region rich in acidic amino acids – D/E tract). **(b)** Comparison of the net charge density square at pH = 7.2 for NCL with and without the D/E tract (left). Observed *in vitro* pH gradient from [26] for condensates formed with 15 µM NCL protein, 1 µM pre-rRNA and 0.2 µM Lyar protein. **(c)** Net charge density and net charge density square of FUS protein with varying pH. **(d)** Spatially resolved ratiometric dye signal in dense and dilute phase for both emission channels, the emission intensity ratio and the corresponding pH, for a FUS condensate formed at 10 (w/v)% PEG, 2.5 µM protein, 150 mM KCl and 50 mM TRIS buffered at pH=7.3. Scale bar: 2 µm **(e)** FUS dense and dilute phase pH at 10 (w/v)% PEG, 2.5 µM protein, 150 mM KCl and 50 mM TRIS buffered at pH=7.3 **(f)** FUS dense and dilute phase pH at varying p(A) RNA concentrations and 2 µM protein, 150 mM KCl and 50 mM TRIS buffered at pH=7.3.

Our observations suggest that the extent of the protein-protein repulsion will dictate the observed pH gradient. To investigate the correlation between the charge repulsion and condensate pH, we refer to previously published work on a NCL protein system associated with nucleolar phase separation[26] (see Fig. 4a). The negatively charged wildtype NCL protein (pI = 4.8) forms a slightly acidic dense phase when undergoing phase separation with RNA, in presence of low concentrations of another client protein (LYAR, 70-fold lower concentration). This is expected when evaluating the q^2 profile of NCL which highlights a clear minimum at low pH (Fig. 4b) resembling our observations on PGL3. In comparison, a deletion mutant of NCL where a region rich in negatively charged amino acids was removed (ΔD/E NCl, pI = 8.8), does not display a pH gradient ($pH_{den} = pH_{dil} = 7.2$). Although ΔD/E NCl displays a high pI, it effectively has zero net charge between pH 5-9 (Fig. 4a). This results in a q^2 term of close to zero over a large pH range (Fig. 4b), indicating that ΔD/E NCl will experience minimal repulsion in the dense phase even at environmental pH. This suggests that dense phase pH requires regulation only when the level of electrostatic repulsion at the environmental pH becomes too unfavourable (Fig. 4b).

Based on these observations, highly positively charged proteins could yield the formation of alkaline condensate pH gradients. To test this, we turned to the FUS protein with pI = 9.4 (see Fig. 4c) which has been associated with aberrant phase transitions in the emergence of amyotrophic lateral sclerosis (ALS)[36].

Indeed, when triggering FUS phase separation by addition of PEG we find an alkaline dense phase pH of $pH_{den} > 8$ using SNARF-4F readouts (see Fig. 4d, e). Similarly to PGL3, FUS sets up a pH gradient towards its pI under a range of other homotypic phase separation conditions (see SI Fig. 12), indicating that the formation of a pH gradient is necessary to favour intermolecular attraction. Accordingly, FUS phase separation is expected to compete with the environmental buffer strength as also shown for the peptide model system. When probing for FUS phase separation while only varying the concentration of TRIS buffer, we find that decreasing the TRIS concentration promotes phase separation (SI Fig. 13). The reduction of the buffer concentration allows for the protein to regulate the pH_{den} more easily to favour intermolecular interactions.

We next set out to study the effect of co-phase separation of the positively charged FUS protein with negatively charged RNA. Compared to the alkaline homotypic FUS condensates we find a decrease in the pH gradient when triggering FUS phase separation by addition of p(A) RNA to effectively $pH_{den} \sim pH_{dil}$ (see Fig. 4f). The formation of a pH gradient is no longer required because the heterotypic FUS-RNA interactions enable attraction even at the environmental pH. This decreases the requirement for both highly charged species to form a pH gradient yielding an effective pI of the mixture of macromolecules somewhere between the pI's of species individually. Interestingly, pH_{den} displays no significant change when varying the RNA concentration. This suggests that the dense phase composition and FUS/RNA stoichiometry may not vary with changing the RNA concentration as large excesses of either species would likely necessitate the formation of a pH gradient to balance the homotypic repulsion. The content of FUS and RNA is likely balanced towards comparable strengths therefore enabling interactions without having to pay the energetic penalty of sustaining a pH gradient.

We next sought to compare the scaffolding behaviour of the FUS/RNA to the dense phase pH modulation of a client species. Here, we look towards the previously discussed NCL system which undergoes phase separation in an RNA-dependent manner while the dense phase pH gradient was characterised at varying concentrations of the positively charged co-factor LYAR (pI = 8.9). At high concentrations comparable to that of NCL, LYAR decreases the effective pH yielding a dense phase pH comparable to the environmental pH[26], which displays a similar effect to the FUS/RNA system. However, at low concentrations, LYAR does not affect the observed dense phase pH. The pH gradient is equivalent to that observed for NCL/RNA phase separation in absence of LYAR. This indicates that the presence of clients with different physico-chemical properties to the condensate scaffolds makes fine tuning of pH_{den} possible.

Phase separation-prone proteins in general are highly charged

Having established that condensates commonly sustain pH gradients, we next sought to understand if this is a more widespread phenomenon that could expand the available diversity of intracellular electrochemical gradients. Given the importance of the protein pI to the formation of condensate pH gradients we set out to characterise the abundance and distribution of pI's in naturally occurring proteins. When plotting the pI distribution of the human proteome (n = 20324), we observe a distinct bimodal shape (see Fig. 5a), highly devoid in proteins with low net charge around physiological pH. Only 8.5% of all human proteins have an isoelectric point between 7-8, which is likely a necessity for guaranteeing cytosolic solubility. Crucially, the bimodal distribution is also conserved through expression, as shown by mass spectrometry (MS) quantification of proteins present in the U2OS cell lysate[37] (SI Fig. 14), suggesting that the available intracellular protein content still displays a similar distribution. In the context of phase separation, however, this indicates that condensates in general will be constituted of highly charged macromolecules driving the formation of dense phase pH gradients.

To trace back the origins of the bimodal protein pI distribution we turned to the physico-chemical properties of the amino acids themselves. When considering the pK_a's of the naturally charged amino acids it becomes apparent that none of them reach charge neutrality at physiological conditions (see SI Fig. 15). Hence, combinations of these building blocks behave similarly, indicating that proteins in general are biased towards high charge at physiological pH. Even independent species such as yeast (*C. cerevisiae*) or even extremophiles (*Acidithiobacillus thiooxidans, Alkalihalobacillus clausii*) as shown in SI Fig. 16, that face a necessity to adjust to much harsher environmental conditions display a lack of proteins with $7 < pI < 8$. To further investigate the bias of protein selection we generated large sets of randomised sequences to find that the bimodal protein pI distribution remains a characteristic feature (see Fig. 5a, SI Fig. 17) even when selecting arbitrary sequences not accounting for the natural amino acid abundance. This, however, yields a much narrower bimodal distribution indicating that the natural selection of proteins is biased towards increasing protein net charge as confirmed when applying human amino acid abundances (SI Fig. 17).

To characterise how this apparent protein charge bias translates to phase separation-prone proteins we use the phase separation predictor DeePhase[38], to quantify the phase separation propensity of all human proteins. We then plotted the distribution of protein pI for a subset of 3950 proteins with high homotypic phase separation propensity (Score between 0-1; high phase separation propensity threshold > 0.8; e.g.

FUS: 0.99). This similarly gives a distinct bimodal shape (see Fig. 5a) but a slightly smaller fraction of phase separation prone proteins with a pI between 7 and 8 (7.2 % for PS-prone proteins, 8.5 % full human proteome). Hence, despite phase separation-prone proteins not requiring increased physical stability compared to the full human proteome these proteins still appear to be highly charged in general. A disparity in the probability density between acidic and basic pI values can potentially be attributed to a bias of DeePhase for homotypic phase separating proteins as high pI proteins are more likely to engage in protein-RNA interactions. Together this highlights that the formation of condensate pH gradients is likely much more common, and that a huge variety of pH gradients could be accessible through phase separation.

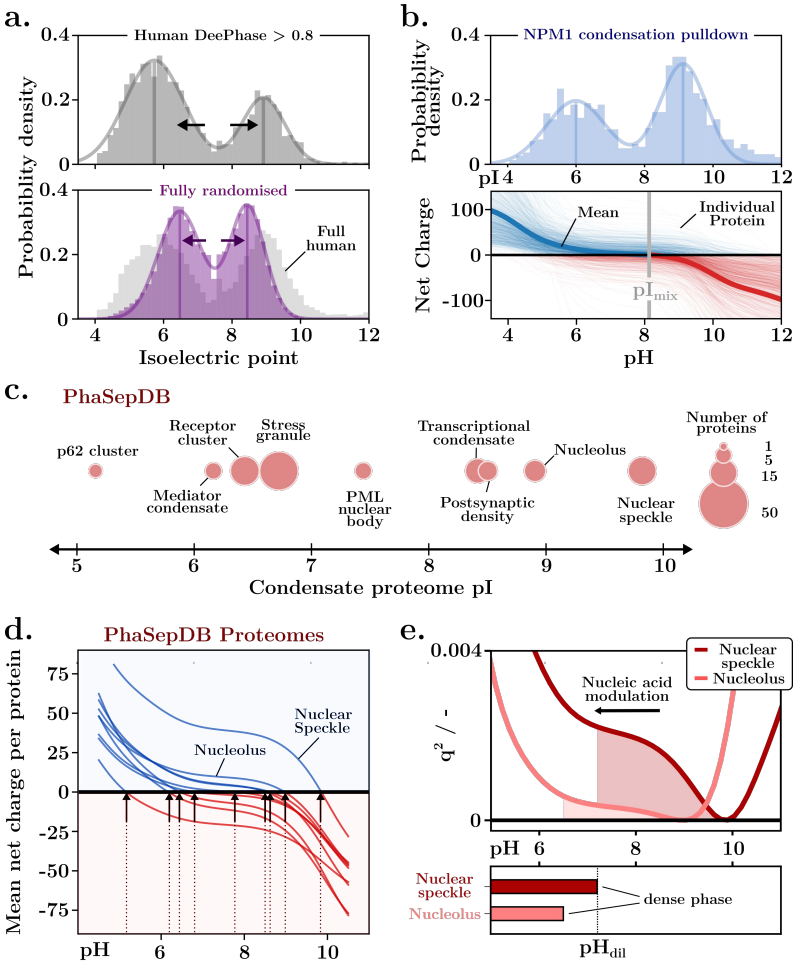

Figure 5: Charge bias of phase separation-prone proteins translates to intracellular condensate pH gradients. **(a)** Isoelectric point distribution of highly phase separation-prone proteins as quantified using DeePhase Scoring (DeePhase Score > 0.8)[38] (top) compared to fully randomised sequences and the human proteome (bot). **(b)** Isoelectric point distribution of quantitative proteomics data obtained from NPM1 condensation in cellular lysate from [39] (top). Net charge change with pH of the full set of individual proteins obtained from (b) compared to the mean. The point at which the mean net charge curve crosses the x-axis is referred to as the condensate mixture pI (bot). For correction for the relative abundance of proteins see SI Fig. 18. **(c)** Comparison of condensate proteome mixture pIs of complex condensate proteomes as obtained via evaluation of protein sequences from PhaSepDB[40]. **(d)** Visualisation of condensate proteome net charge change with pH for complex condensate proteomes obtained from PhaSepDB[40]. **(e)** Comparison of nucleolus and nuclear speckle net charge density square profiles and correlation to observed *in vivo* pH gradient from [26].

Condensate proteomes cover a large pI range

In more complex environments, condensates commonly recruit numerous client proteins, necessitating consideration of a more complex set of constituents beyond the scope of currently available *in vitro* reconstitution approaches. To investigate the electrochemical diversity of these condensate proteomes (i.e. the collection of sequences recruited into a condensates), we analysed published NPM1 pulldown proteomics data[39]. Here, condensation of NPM1 in cellular lysate was induced to mimic the composition of granular compartments inside the nucleolus. The detected sequences display a bias towards proteins with alkaline pI as 63 % of the individual proteins and 68% of total number of proteins detected display a pI > 7 (Fig. 5b, SI Fig. 18 for relative abundance). NPM1 as the primary scaffold, however, displays an acidic pI of ~ 5 suggesting that the combination of other recruited clients can determine the condensate proteome's electrochemical properties. Overall, the enrichment of positively charged proteins is expected given their propensity to interact with negatively charged nucleic acids in the nucleus.

To quantify the proteome's response to changes in pH, we then calculate the charge of each individual protein from the amino acid pKa's. Normalising the charge sum of all proteins with the total number of proteins provides us with the mean net charge per effective condensate protein (see Fig. 5b, SI Fig. 18 relative abundance correlated). The point at which the condensate proteome net charge is neutralised is of particular interest as it resembles the effective isoelectric point of the mixture. This yields $pI_m = 8.7$ when considering the relative abundance of the proteins present in the condensate, and pI = 8.2 when not accounting for each protein's relative dense phase abundance. This suggests that the approximation of using only the individual proteins without accounting for their relative abundance data may function well as a proxy for the quantitative abundance data in this case. Although this might not hold true in all systems, it is unsurprising that client partitioning specificity is heavily biased by the electrochemical properties of the proteins in the dense phase. Here, the formation of electrochemical gradients, for example favouring the interaction of positively charged proteins with RNA, is likely to drive further recruitment of sequences with similar physico-chemical properties.

To further quantify the electrochemical properties of complex condensates, we then apply this approach to nine condensate proteomes acquired from PhaSepDB (see SI for details) including that of nucleoli and nuclear speckles (see Fig. 5c, d). The obtained condensate proteome pI for the nucleolus at 8.9 correlates well with the values obtained for the simple NPM1 lysate pulldown data, suggesting consistency in the electrochemical property profile for these types of condensates independently of the underlying methodological approach or consideration of the relative abundances. In comparison, the stress granule proteome displays a slightly acidic mixture pI of 6.8, indicating that the electrochemical properties of condensates *in vivo* can display substantial variation. This is particularly interesting given that stress granules are also characterised by RNA binding and recruitment. The electrochemical differences between condensates become even more noticeable when considering the full set of these nine systems as their mixture pI spans a range of almost five pH units, indicating that condensates present a large range of electrochemical properties.

We subsequently wanted to understand how the observed condensate proteome pI translates to condensate pH *in vivo*. We turned to published data where the pH_{den} of both nucleoli and nuclear speckles was determined via expressing the ratiometric protein-based pH sensing dye pHluorin2 [30]. Our condensate proteome analysis suggests that both these systems are characterised by a high positively charge content with mixture pIs of 8.9 and 9.9, respectively. Both condensation systems are known for their propensity to recruit nucleic acids from the nucleus, therefore, limiting the necessity for generating an alkaline pH gradient as also shown for our simplified FUS system. When evaluating the q^2 profiles of both condensation systems, we find that the nucleolus proteome displays a much less rapid increase in q^2 compared to the nuclear speckle (Fig. 5e). Hence, the nucleolus is expected to be much more susceptible to the observed acidic pH shift introduced by nucleic acids. Indeed, intracellularly, nuclear speckles were shown not to sustain a pH gradient across dense and dilute phase, whereas nucleoli create slightly acidic microenvironments (Fig. 5e). Our analysis also suggests that recruitment of negatively charged nucleic acids will also be buffered much less effectively by the positively charged proteins in the nucleolus compared to the nuclear speckle, which explains the formation of an acidic pH gradient. As such, *in silico* evaluation of condensate proteomes can also inform us of the observed electrochemical condensate properties *in vivo*.

Condensates as a source of dynamic electrochemical buffering

We next sought to characterise the potential effector range and functional diversity of electrochemical gradients condensates could offer. To do so, we used the Protein Condensate Atlas, which comprises predictions for condensate systems generated through an unsupervised clustering of a protein-protein interaction network[41]. Filtering for clusters that contain a high proportion of predicted phase-separating proteins (> 50%) gives a set of 85 condensate proteomes (Fig. 6a, SI Fig. 19, 20). We then subjected these proteomes to quantification of the pH-dependent net charge changes and the condensate cluster pI (Fig. 6b), to find mixture pI's ranging from around 5 – 10 (Fig. 6c). This suggests that condensates could potentially allow for the formation of a large diversity of dynamic pH microenvironments suited to the specific function and property of the individual system. The distribution of condensate pI values highlights a double gaussian shape devoid of clusters at physiological pH, closely correlated to the pI distribution of the phase separation-prone proteins. A disparity in peak heights between acidic and basic pH values can potentially be attributed to the clustering of protein-protein interaction networks favouring prediction of condensates that are less dependent on protein-RNA interactions.

Given the importance of the change in repulsion with pH to the observed behaviour, we next determined the q^2 profile for all of the 85 condensate clusters. To enable comparability, we centred these profiles at the mixture pI of the individual cluster (see Fig. 6d), to find that the q^2 terms can vary up to five orders of magnitude at ~ 1.5 units below the respective mixture pI. This suggests that condensates will not only be able to set up a range of pH gradients, but also display distinct response behaviour to changes in pH. To

further characterise the potential diversity of condensate buffering responsiveness, we calculated the net charge change with pH at the condensate mixture pI: $-(\Delta C/\Delta pH)_{pI}$ (Fig. 6e). This encapsulates the change in the average net charge caused by environmental perturbations, which varied by a factor of 20 across the 85 condensate clusters. We observed a clear correlation between the condensate cluster pI and $-(\Delta C/\Delta pH)|_{pI}$, as the gradient gets shallower for condensate pIs closer to neutral pH (Fig. 6e). Condensate clusters with a bias towards, for example, high pI will have an enrichment of AAs that change their protonation state at high pH, thereby increasing $-(\Delta C/\Delta pH)|_{pI}$. Despite this underlying bias, condensate clusters with pI ~ 6, for example, still display a diversity in $-(\Delta C/\Delta pH)|_{pI}$ of around half an order of magnitude. Hence, even individual systems experiencing charge neutrality at the same pH, can display largely different responses to slight changes in the environmental conditions. Taken together, this analysis suggests that condensate systems not only might be capable of establishing distinct electrochemical environments but also create unique responsiveness to environmental changes (Fig. 6f).

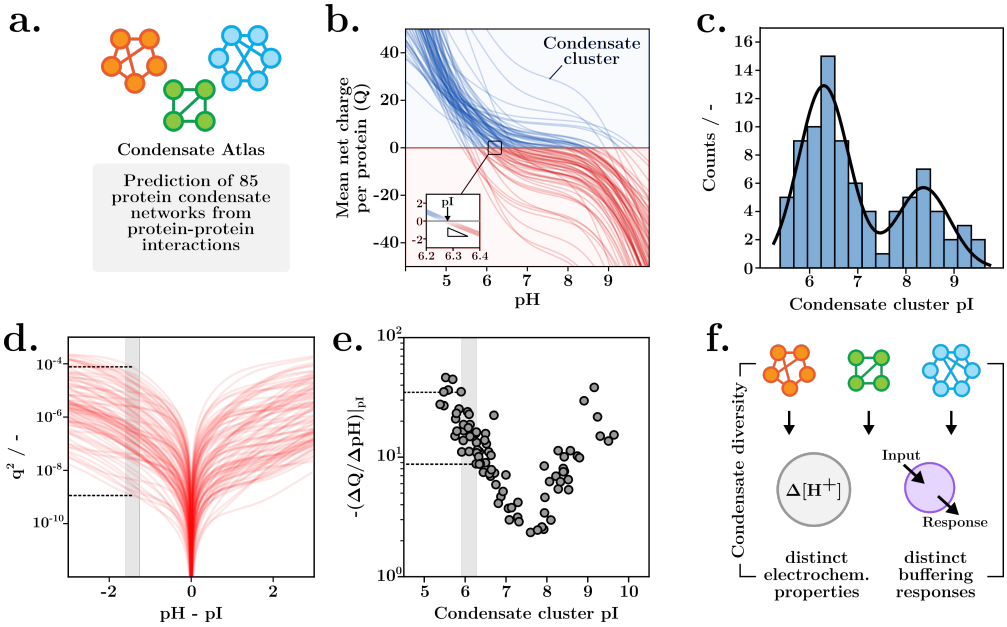

Figure 6: Predicted condensate proteomes present a massive diversity in electrochemical properties. **(a)** Schematic representation of the Protein Condensate Atlas from [41] used to evaluate the theoretical range of electrochemical properties presented by complex condensates. **(b)** Condensate proteome net charge change with pH of all 85 condensate clusters. The inset highlights the condensate cluster pI as well as the charge/pH response gradient at the pI $(\Delta C/\Delta pH)|_{pI}$. **(c)** Histogram of condensate cluster pIs obtained from the Protein Condensate Atlas. **(d)** Comparison of the net charge density square for all 85 condensate clusters. Grey box: highlighting the range of q^2 values of different condensate cluster at similar pH – pI. **(e)** Change in the charge responsiveness to pH pertubations $-(\Delta C/\Delta pH)|_{pI}$ against condensate cluster pI. Grey box: highlighting the range of $-(\Delta C/\Delta pH)|_{pI}$ values of different condensate cluster with similar pI. **(f)** Schematic representation highlighting the diversity of potential pH gradients and pH buffering responses display by complex condensates.

Discussion and Conclusion

In this work, we show that biomolecular phase separation commonly creates pH gradients between the dense and dilute phase without active processes in order to minimise intermolecular electrostatic repulsion.

Experimentally, we extract information on the pH responsive phase behaviour of biomolecules via the development of a microfluidics-based platform capable of continuously scanning pH phase space while providing information on partitioning and energetics. Our approach yields thousands of data points per minute and so promises to be a valuable technical addition to a variety of applications. This could include formulation optimisation as testing for variations in pH conditions, while critical to ensuring safety and efficacy, is currently a highly laborious process[42,43].

Investigations on a peptide model system reveal that polypeptide chains can regulate the dense phase pH towards their pI independent of the environmental pH condition. Evaluation of a proxy for the repulsive free energy highlights that sustaining these pH phase gradients is necessary to minimise electrostatic repulsion (Fig. 7a). Considering increased protein concentrations in the order of 100-1000-fold in the dense phase, it is unsurprising that the amino acid side groups can exert their inherent acidity/alkalinity to establish the dense phase pH. Further investigations also show that regulation of the dense phase pH is aided by exclusion of the small molecule buffer from the dense phase. This makes the formation of dense phase pH gradients an emergent property, where the formation of a distinct phase is necessary to enable differential solute partitioning. This is of particular interest when consider protein folding or aggregation, which similarly require charge neutralisation for assembly. Here electrostatic repulsion is commonly overcome by pK_a shifts of residues placed in the protein core or the assembly interface[44–49].

We subsequently show that the formation of dense phase pH gradients extends to the phase separation of FUS and PGL3. Using ratiometric dye readouts, we show that both these proteins generate condensates that buffer pH towards the respective pI, setting up alkaline and acidic gradients of over a unit in pH. Further investigations highlight that the formation of pH gradients is conditional to the electrochemical properties of the protein. Specifically, sustaining pH gradients only becomes necessary under sufficient levels of electrostatic repulsion at the environmental pH. This is particularly interesting considering the various dynamic tools nature has available to modify the effective physico-chemical properties of proteins such as post translational modifications[50–52]. Here, phosphorylation modifications, for example, could function as tools for finetuning dense phase electrochemical gradients by adding negative charges, reducing the pI of the protein and therefore the pH of the dense phase.

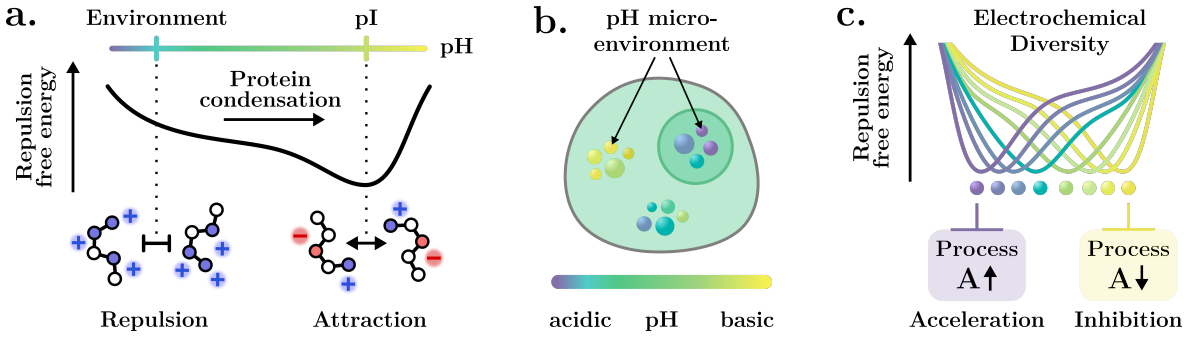

Figure 7: Condensates as pH buffering microenvironments. (a) Intermolecular electrostatic repulsion in collective interactions can be minimised by shifting the dense phase pH towards the biomolecular isoelectric point. **(b)** Biomolecular condensates are dynamically fine-tuned pH microenvironments without necessitating membrane enclosure or the presence of active processes. **(c)** The diversity of electrochemical properties of complex condensates suggests potential for segregation and selective acceleration or inhibition of a range of functional processes, as highlighted visually by the modulation of a generic process A.

Investigations of multicomponent protein/nucleic acid systems highlight that heterotypic interactions modulate condensates pH gradients. FUS/RNA co-condensation, for example, shows that the condensate pH is connected to the overall biomolecular component mixture, as addition of the negatively charged RNA drastically reduces the pH of the high pI FUS condensates. The attraction between the high pI protein and low pI RNA will be maximised at pH conditions where both macromolecules are present at opposing charges. Therefore, in the case of the negatively charged PGL3 protein, we still observe an acidic pH gradient upon phase separation with RNA, as both macromolecules will be less charged at lower pH, hence, reducing repulsion and enhancing the heterotypic interaction. This suggests that condensate pH gradients can effectively be controlled by composition considering the physico-chemical properties of the individual components. Hence, controlling the level of RNA recruitment or even the specific nucleic acid type or structure are expected to also regulate condensate pH gradients.

Using a bioinformatics approach, we capture the sequence and physico-chemical complexity of *in vivo* condensation systems, which recruit a large diversity of clients. This analysis reveals that the electrochemical properties of condensates might span 5 units in their mixture isoelectric point. We also show that our *in silico* electrochemical reconstitution approach is informative of intracellular condensate pH gradients sustained by nucleoli and nuclear speckles. Lastly, we explore the effector range and diversity of condensate pH buffering, by evaluating the electrochemical properties of 85 predicted human condensate networks. This analysis highlights that condensates might not only be able to dynamically set up a large diversity of pH microenvironments but also present distinct responses to environmental changes (Fig. 7b).

Effects of pH on the rate of biologically essential reactions, either by regulating enzyme activity or direct catalytic effects of H^+ and OH^-, are well documented. Membrane bound organelles commonly display distinct pH conditions aimed at optimising processes relevant to their function[7–12]. While biomolecular condensation yields drastic increases in viscosity, the molecular scale dynamics of condensed systems still remain rapid enough to support efficient reactions[53]. As such, condensates have been repeatedly shown to affect chemical reaction rates[54–57] and recent evidence further highlights that their local electrochemical properties such as differences in pH can even modulate redox reactions[25]. Our work further reveals that condensation presents a generic mechanism for the formation of pH gradients, which can be finetuned through composition, and indicates that intracellular complex condensates present a massive diversity of electrochemical properties. Together, this highlights biomolecular condensation as a means for the efficient segregation and optimisation of a range of functional processes through changes in the emergent chemical microenvironment.

References

1. Mitchell, P. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* **191**, 144–148 (1961).
2. Dixon, M. The effect of pH on the affinities of enzymes for substrates and inhibitors. *Biochem J* **55**, 161–170 (1953).
3. Chiti, F. & Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **75**, 333–366 (2006).
4. Yang, A.-S. & Honig, B. On the pH Dependence of Protein Stability. *Journal of Molecular Biology* **231**, 459–474 (1993).
5. Casey, J. R., Grinstein, S. & Orlowski, J. Sensors and regulators of intracellular pH. *Nature reviews Molecular cell biology* **11**, 50–61 (2010).
6. Madshus, I. H. Regulation of intracellular pH in eukaryotic cells. *Biochemical journal* **250**, 1 (1988).
7. Santo-Domingo, J. & Demaurex, N. The renaissance of mitochondrial pH. *Journal of General Physiology* **139**, 415–423 (2012).
8. Akhmedov, D. *et al.* Mitochondrial matrix pH controls oxidative phosphorylation and metabolism-secretion coupling in INS-1E clonal beta cells. *FASEB J* **24**, 4613–4626 (2010).
9. Paroutis, P., Touret, N. & Grinstein, S. The pH of the Secretory Pathway: Measurement, Determinants, and Regulation. *Physiology* **19**, 207–215 (2004).
10. de Duve, C. & Wattiaux, R. Functions of Lysosomes. *Annual Review of Physiology* vol. 28 435–492 (1966).
11. Mindell, J. A. Lysosomal Acidification Mechanisms*. *Annual Review of Physiology* vol. 74 69–86 (2012).
12. Chanat, E. & Huttner, W. B. Milieu-induced, selective aggregation of regulated secretory proteins in the trans-Golgi network. *Journal of Cell Biology* **115**, 1505–1519 (1991).
13. Poznanski, J., Szczesny, P., Ruszczyńska, K., Zielenkiewicz, P. & Paczek, L. Proteins contribute insignificantly to the intrinsic buffering capacity of yeast cytoplasm. *Biochemical and biophysical research communications* **430**, 741–744 (2013).
14. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology* **18**, 285–298 (2017).
15. Boeynaems, S. *et al.* Protein Phase Separation: A New Phase in Cell Biology. *Trends in Cell Biology* **28**, 420–435 (2018).
16. Alberti, S. & Dormann, D. Liquid-Liquid Phase Separation in Disease. *Annu Rev Genet* **53**, 171–194 (2019).
17. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid-liquid phase separation in biology. *Annual review of cell and developmental biology* **30**, 39–58 (2014).
18. Alberti, S., Gladfelter, A. & Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419–434 (2019).
19. Hannes Ausserwöger *et al.* Quantifying collective interactions in biomolecular phase separation. *bioRxiv* 2023.05.31.543137 (2023) doi:10.1101/2023.05.31.543137.
20. Welsh, T. J. *et al.* Surface Electrostatics Govern the Emulsion Stability of Biomolecular Condensates. *Nano Lett* **22**, 612–621 (2022).
21. Martin, R. M. *et al.* Principles of protein targeting to the nucleolus. *Nucleus* **6**, 314–325 (2015).
22. Mitrea, D. M. *et al.* Methods for Physical Characterization of Phase-Separated Bodies and Membrane-less Organelles. *Journal of Molecular Biology* **430**, 4773–4805 (2018).
23. Cakmak, F. P., Choi, S., Meyer, M. O., Bevilacqua, P. C. & Keating, C. D. Prebiotically-relevant low polyion multivalency can improve functionality of membraneless compartments. *Nature Communications* **11**, 5949 (2020).
24. Choi, S., Knoerdel, A. R., Sing, C. E. & Keating, C. D. Effect of Polypeptide Complex Coacervate Microenvironment on Protonation of a Guest Molecule. *J. Phys. Chem. B* **127**, 5978–5991 (2023).
25. Dai, Y. *et al.* Interface of biomolecular condensates modulates redox reactions. *Chem* **9**, 1594–1609 (2023).
26. King, M. R. *et al.* Macromolecular condensation organizes nucleolar sub-phases to set up a pH gradient. *Cell* **187**, 1889-1906.e24 (2024).
27. Dingfelder, F., Henriksen, A., Wahlund, P.-O., Arosio, P. & Lorenzen, N. Measuring Self-Association of Antibody Lead Candidates with Dynamic Light Scattering. *Methods Mol Biol* **2313**, 241–258 (2022).
28. Arter, W. E. *et al.* Biomolecular condensate phase diagrams with a combinatorial microdroplet platform. *Nature Communications* **13**, 7845 (2022).
29. Daoyuan Qian, Hannes Ausserwoger, Tomas Sneideris, Rohit Pappu, & Tuomas P. J. Knowles. Dominance metric in multi-component binary phase equilibria. *bioRxiv* 2023.06.12.544666 (2023) doi:10.1101/2023.06.12.544666.
30. Qian, D. *et al.* Tie-Line Analysis Reveals Interactions Driving Heteromolecular Condensate Formation. *Phys. Rev. X* **12**, 041038 (2022).
31. Adame-Arana, O., Weber, C. A., Zaburdaev, V., Prost, J. & Jülicher, F. Liquid phase separation controlled by pH. *Biophysical Journal* **119**, 1590–1605 (2020).
32. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *Journal of molecular biology* **157**, 105–132 (1982).
33. Daoyuan Qian *et al.* Linking modulation of bio-molecular phase behaviour with collective interactions. *bioRxiv* 2023.11.02.565376 (2023) doi:10.1101/2023.11.02.565376.
34. Feynman, R. P. (Richard P., 1918-1988. *The Feynman Lectures on Physics.* (Reading, Mass. : Addison-Wesley Pub. Co., c1963-1965., 1965).
35. Saha, S. *et al.* Polar Positioning of Phase-Separated Liquid Compartments in Cells Regulated by an mRNA Competition Mechanism. *Cell* **166**, 1572-1584.e16 (2016).
36. Patel, A. *et al.* A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066–1077 (2015).
37. Beck, M. *et al.* The quantitative proteome of a human cell line. *Molecular Systems Biology* **7**, 549 (2011).
38. Saar, K. L. *et al.* Learning the molecular grammar of protein condensates from sequence determinants and embeddings. *Proceedings of the National Academy of Sciences* vol. 118 e2019053118 (2021).
39. Freibaum, B. D., Messing, J., Yang, P., Kim, H. J. & Taylor, J. P. High-fidelity reconstitution of stress granules and nucleoli in mammalian cellular lysate. *Journal of Cell Biology* **220**, e202009079 (2021).
40. Hou, C. *et al.* PhaSepDB in 2022: annotating phase separation-related proteins with droplet states, co-phase separation partners and other experimental information. *Nucleic Acids Research* **51**, D460–D465 (2023).
41. Saar, K. L. *et al.* Protein Condensate Atlas from predictive models of heteromolecular condensate composition. *accepted Nature Communications* (2024).
42. Shire, S. J., Shahrokh, Z. & Liu, J. Challenges in the development of high protein concentration formulations. *Journal of Pharmaceutical Sciences* **93**, 1390–1402 (2004).
43. Frokjaer, S. & Otzen, D. E. Protein drug stability: a formulation challenge. *Nature Reviews Drug Discovery* **4**, 298–306 (2005).

44. Tan, Y.-J., Oliveberg, M., Davis, B. & Fersht, A. R. Perturbed pKA-values in the Denatured States of Proteins. *Journal of Molecular Biology* **254**, 980–992 (1995).
45. Isom, D. G., Castañeda, C. A., Cannon, B. R. & García-Moreno E., B. Large shifts in pKa values of lysine residues buried inside a protein. *Proceedings of the National Academy of Sciences* **108**, 5260–5265 (2011).
46. Kesvatera, T., Jönsson, B., Thulin, E. & Linse, S. Measurement and modelling of sequence-specific pKavalues of lysine residues in calbindin D9k. *Journal of molecular biology* **259**, 828–839 (1996).
47. Shammas, S. L. *et al.* Perturbation of the stability of amyloid fibrils through alteration of electrostatic interactions. *Biophysical journal* **100**, 2783–2791 (2011).
48. Bartik, K., Redfield, C. & Dobson, C. M. Measurement of the individual pKa values of acidic residues of hen and turkey lysozymes by two-dimensional 1H NMR. *Biophysical journal* **66**, 1180–1184 (1994).
49. Buell, A. K. *et al.* Electrostatic effects in filamentous protein aggregation. *Biophysical journal* **104**, 1116–1126 (2013).
50. Deribe, Y. L., Pawson, T. & Dikic, I. Post-translational modifications in signal integration. *Nature Structural & Molecular Biology* **17**, 666–672 (2010).
51. Walsh, C. *Posttranslational Modification of Proteins: Expanding Nature's Inventory.* (Roberts and Company Publishers, 2006).
52. Walsh, C. T., Garneau-Tsodikova, S. & Gatto Jr, G. J. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angewandte Chemie International Edition* **44**, 7342–7372 (2005).
53. Galvanetto, N. *et al.* Extreme dynamics in a biomolecular condensate. *Nature* **619**, 876–883 (2023).
54. Küffner, A. M. *et al.* Acceleration of an Enzymatic Reaction in Liquid Phase Separated Compartments Based on Intrinsically Disordered Protein Domains. *ChemSystemsChem* **2**, e2000001 (2020).
55. Gil-Garcia, M. *et al.* Local environment in biomolecular condensates modulates enzymatic activity across length scales. *Nature Communications* **15**, 3322 (2024).
56. Küffner, A. M. *et al.* Sequestration within biomolecular condensates inhibits Aβ-42 amyloid formation. *Chem. Sci.* **12**, 4373–4382 (2021).
57. Tibble, R. W., Depaix, A., Kowalska, J., Jemielity, J. & Gross, J. D. Biomolecular condensates amplify mRNA decapping by biasing enzyme conformation. *Nature Chemical Biology* **17**, 615–623 (2021).

Conflicts of interest

T.P.J.K. is the C.T.O. of Transition Bio Inc., G.I. and N.L.Z. are employees of Novo Nordisk and S.A and A.A.H. are members of the scientific advisory board of Dewpoint Therapeutics Inc. The work reported here was not influenced by these affiliations.

Acknowledgements

The research leading to these results has received funding from Global Research Technologies, Novo Nordisk A/S (H.A., T.P.J.K.). T.S. is supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant MicroREvolution (agreement no. 101023060) and the Frances and Augustus Newman Foundation. The authors thank Prof. Rohit Pappu for valuable discussions and feedback.

Author contributions

H.A., R.S. and T.P.J.K. designed and conceptualised the study. H.A., R.S., T.S., C.M.F., E.d.C., A.Z.B, and M.O. performed investigations. T.S., K.L.S., T.M.F., S.W., J.M.I.-A., G.K., G.I., S.A., A.A.H., N.L.Z. and T.P.J.K. provided materials and methods. H.A., R.S. and D.Q. analysed the data. H.A. and T.P.J.K. wrote the original draft of the paper. All authors discussed the results, reviewed, edited and contributed to the final manuscript.

Correspondence

Correspondence should be addressed to Tuomas P. J. Knowles (tpjk2@cam.ac.uk).

Data availability

The raw data and analysis code underlying this study will be made available upon request.