Article

The sequence–structure–function relationship of intrinsic ERα disorder

Zhanwen Du[1,6], Han Wang[2,6], Shuqi Luo[1,6], Zixi Yun[2], Chen Wu[2], Wangfei Yang[3], Matthias Buck[4], Wenwei Zheng[3], Alexandar L. Hansen[5], Hung-Ying Kao[2] & Sichun Yang[1✉]

The oestrogen receptor (ER or ERα), a nuclear hormone receptor that drives most breast cancer[1], is commonly activated by phosphorylation at serine 118 within its intrinsically disordered N-terminal transactivation domain[2,3]. Although this modification enables oestrogen-independent ER function, its mechanism has remained unclear despite ongoing clinical trials of kinase inhibitors targeting this region[4–6]. By integration of small-angle X-ray scattering and nuclear magnetic resonance spectroscopy with functional studies, we show that serine 118 phosphorylation triggers an unexpected expansion of the disordered domain and disrupts specific hydrophobic clustering between two aromatic-rich regions. Mutations mimicking this disruption rescue ER transcriptional activity, target-gene expression and cell growth impaired by a phosphorylation-deficient S118A mutation. These findings, driven by hydrophobic interactions, extend beyond electrostatic models and provide mechanistic insights into intrinsically disordered proteins[7], with implications for other nuclear receptors[8]. This fundamental sequence–structure–function relationship advances our understanding of intrinsic ER disorder, crucial for developing targeted breast cancer therapeutics.

Breast cancer, the leading cause of cancer-related deaths among women globally, is predominantly oestrogen receptor-positive (ER+)[1,9]. Serine 118 phosphorylation (pS118) by MAPK[2] or CDK7 (ref. 3) enables ligand-independent ER activation. Although CDK7 inhibitors targeting this phosphorylation have shown promise in clinical trials[6], the mechanistic basis of phosphorylation-driven activation remains unclear. We discover that Ser118 phosphorylation drives conformational changes through the disruption of hydrophobic interactions between aromatic-rich clusters, showing how hydrophobic mutations can restore ER-mediated transcription.

The ER functions through two distinct domains: an intrinsically disordered N-terminal transactivation domain (NTD) and a structurally folded C-terminal ligand-binding domain (LBD)[10]. Although antioestrogens target the LBD[11,12], persistent pSer118 maintains ER activity despite treatment[4,13]. This resistance mechanism has prompted CDK7 inhibitor development[5]. However, unlike direct androgen receptor NTD inhibitors[14,15], current CDK7 inhibitors affect the ER-NTD indirectly.

Current structural techniques like cryoelectron microscopy[16] and computational tools such as AlphaFold[17] cannot effectively characterize these disordered domains, and limited biophysical data[18–21] have constrained our understanding of phosphorylation-triggered activation. Our findings show that pSer118 regulates ER activity by disruption of specific hydrophobic interactions rather than electrostatic effects, separating aromatic-rich clusters within the ER-NTD. This mechanism provides insights into phosphorylation-regulated disorder in transcription factor function.

Residual structure by amino acid clusters

To characterize the biophysical properties of the disordered ER-NTD (184 residues; Fig. 1a and Extended Data Fig. 1a), we integrated size-exclusion chromatography-coupled, small-angle X-ray scattering (SEC-SAXS) and nuclear magnetic resonance (NMR) spectroscopy. SEC-SAXS measurements (Extended Data Fig. 1b,c) yielded precise dimensional parameters of ER-NTD, improving on our previous flow-cell set-up[20]. Analysis showed a Flory exponent (v) of 0.51 and a radius of gyration (R_g) of 33.9 ± 0.2 Å (Extended Data Fig. 1d,e), determined through empirical molecular form factor fitting[22]. The Kratky plot demonstrates ER-NTD's structural disorder relative to folded C-terminal domains (Fig. 1b and Extended Data Fig. 1d), whereas pairwise distance distribution derived from SEC-SAXS data indicates a more compact conformation than that predicted by flexible-meccano calculations[23] (Fig. 1b), suggesting the presence of residual structures and tertiary interactions within the disordered ER-NTD.

Our NMR analysis at 400 μM protein concentration achieved near-complete (over 98%) resonance assignments (Fig. 1c and Extended Data Fig. 2a), advancing previous studies that lacked assignments[18–20]. The ^{1}H-^{15}N heteronuclear single-quantum coherence (HSQC) spectrum shows narrow proton chemical shift dispersion, characteristic of intrinsic disorder (Fig. 1c). Analysis of C_α and C_β chemical shifts shows minimal persistent secondary structures, except for two short alanine segments (A64–A68 and A86–A88) showing transient helical features with propensity of approximately

[1]Case Comprehensive Cancer Center and Department of Nutrition, Case Western Reserve University School of Medicine, Cleveland, OH, USA. [2]Department of Biochemistry and Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, OH, USA. [3]College of Integrative Sciences and Arts, Arizona State University, Mesa, AZ, USA. [4]Department of Physiology and Biophysics, Case Western Reserve University School of Medicine, Cleveland, OH, USA. [5]Campus Chemical Instrument Center, The Ohio State University, Columbus, OH, USA. [6]These authors contributed equally: Zhanwen Du, Han Wang, Shuqi Luo. ✉e-mail: sichun.yang@case.edu
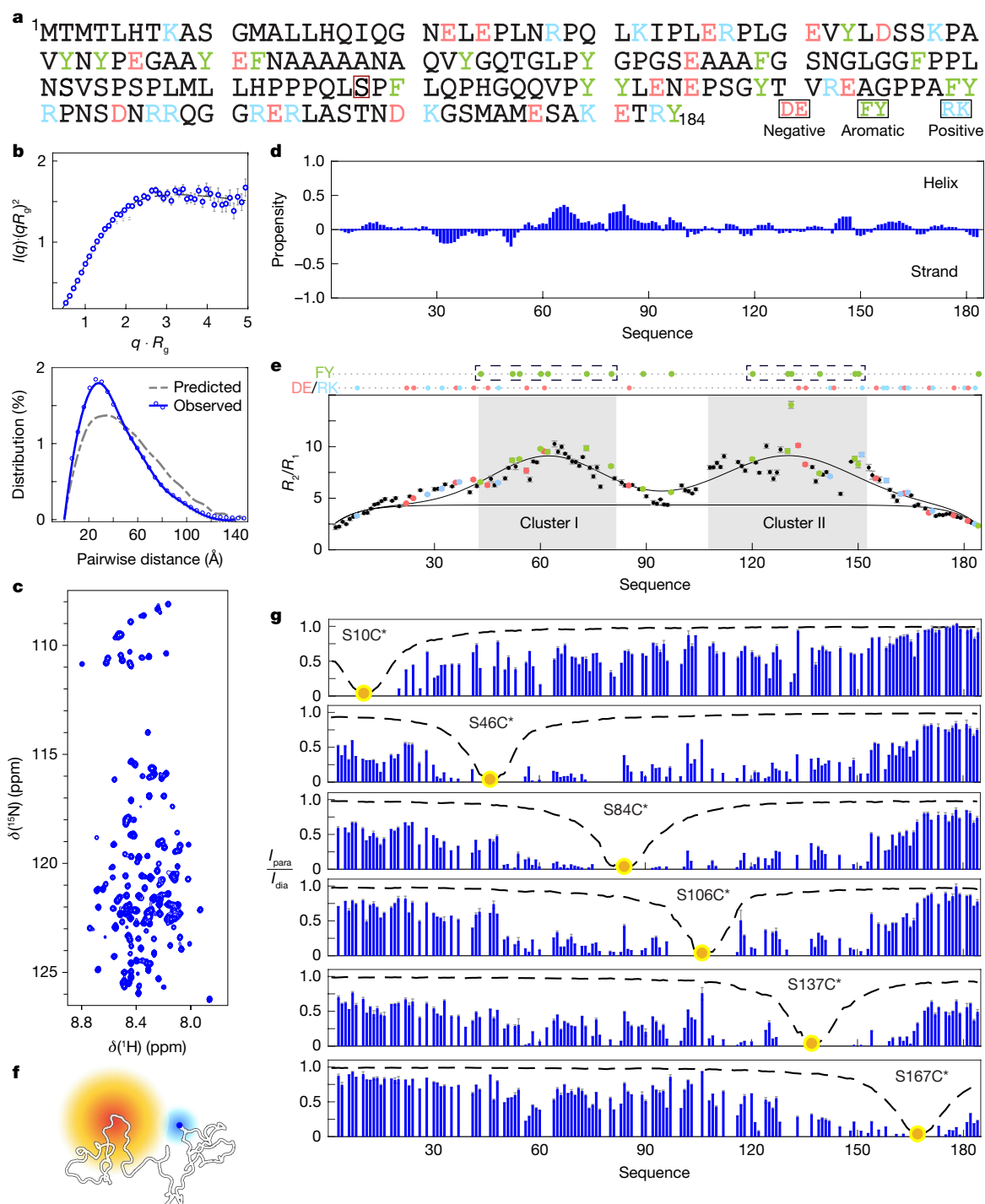
Fig. 1 | Hydrophobic clustering in ER-NTD. a, ER-NTD sequence highlighting charged (negative, light red; positive, light blue) and aromatic residues (light green), with Ser118 in red box. **b**, Top, Kratky plot of SEC-SAXS data (4 °C). Error bars, $I(q)$ propagated uncertainties. Bottom, pairwise distance distribution (solid line, SAXS-EOM[52]; circles, GNOM[53]; dashed line, flexible-meccano[23]). **c**, ^{1}H-^{15}N HSQC spectrum (850 MHz) (see Extended Data Fig. 2a for assignments). **d**, Secondary structure propensity from ^{1}H, ^{15}N and ^{13}C (C$^\alpha$, C$^\beta$ and C′) chemical shifts using disordered protein algorithm[54]. Low scores indicate minimal structure. **e**, ^{15}N amide R_2/R_1 ratios showing two clusters—cluster I (residues 43–81) and cluster II (108–152)—identified by two-cluster Gaussian fitting (Supplementary Methods). Shaded regions, higher R_2/R_1 ratios; top, aromatic residues. Error bar, standard deviation propagated from R_1 and R_2 uncertainties. **f**, Schematic of paramagnetic effect between nitroxide-spin (yellow) and NMR-active nucleus (blue). **g**, PRE profiles following MTSL spin labelling at six cysteine mutants. Peak intensity ratio (I_{para}/I_{dia}) indicates spin label–amide proton distances. Yellow circles, spin label position; dashed lines, predicted random coil effect[23]. $I(q)$, normalized scattering intensity; q, scattering vector amplitude.

30% (Fig. 1d). These well-resolved NMR spectra provide unprecedented residue-specific detail of ER-NTD conformational features. Measurements of longitudinal (R_1) and transverse (R_2) relaxation rates (Extended Data Fig. 3a) identified two regions with restricted backbone mobility[24,25]: cluster I (residues 43–81) and cluster II (residues 108–152), characterized by elevated R_2/R_1 ratios (Fig. 1e). These clusters correspond to residual structures and explain the compact conformation observed in SEC-SAXS data.
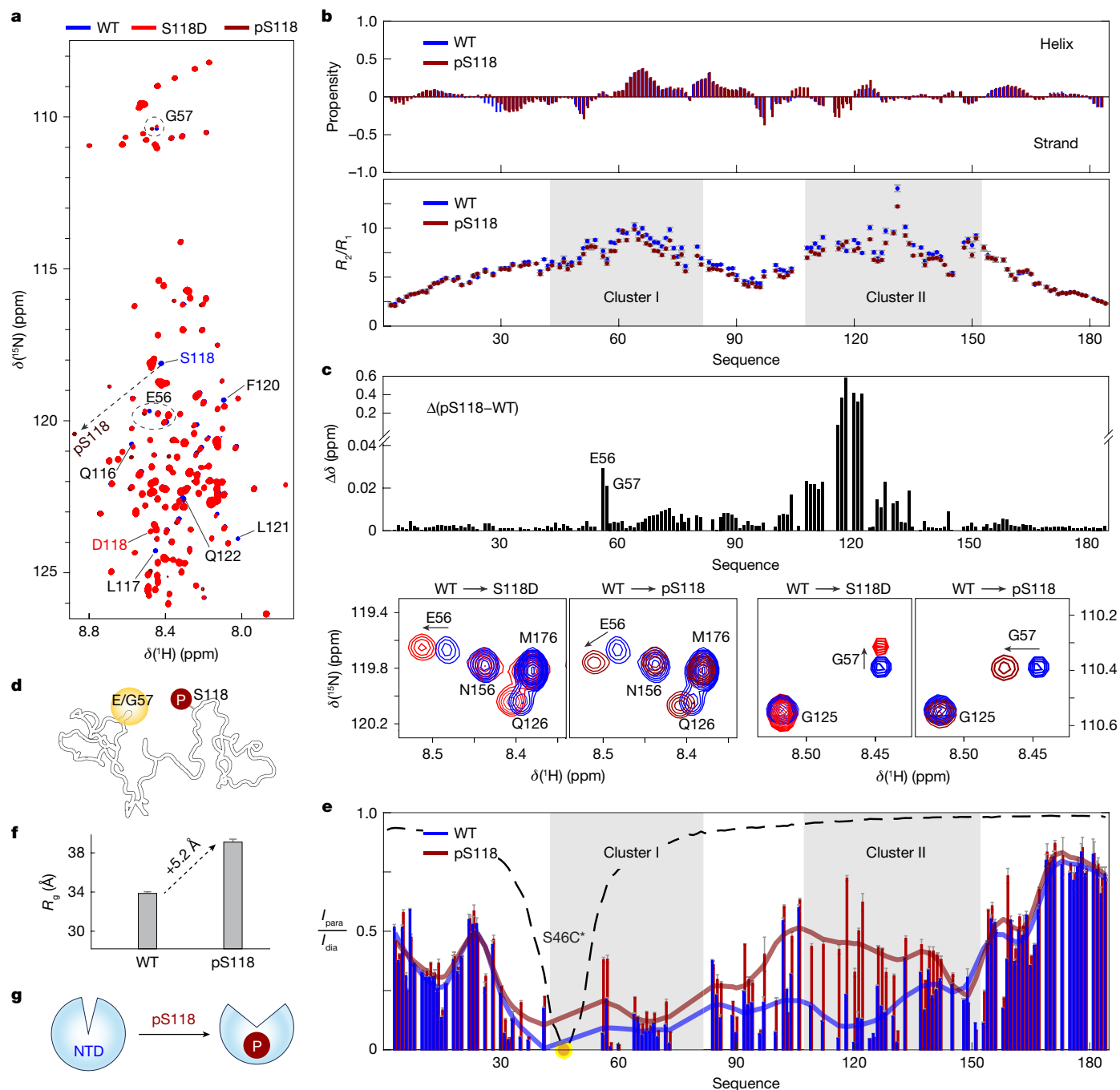
Fig. 2 | pSer118 induces conformational changes. a, ^{1}H-^{15}N HSQC spectra overlay: WT (blue), S118D (red), and pS118 (dark red), with notable chemical shift changes annotated. Circled peaks magnified in **c**. **b**, WT versus pS118 comparison: secondary structure propensity (top) and ^{15}N amide R_2/R_1 ratios (bottom). **c**, Top, E56 and G57 backbone amide chemical shift changes following pSer118. $\Delta\delta = \sqrt{(\Delta\delta H^2 + (\Delta\delta N \times 0.154)^2}$, where $\Delta\delta H$ and $\Delta\delta N$ are proton and nitrogen changes, respectively, relative to unbound protein. Bottom, WT and S118D/pS118 spectral overlap. ^{13}C chemical shift changes are given in Extended Data Fig. 6a,b. **d**, Long-range propagation of pSer118 effects to E56/G57. **e**, PRE profiles from S46C and S84C spin labelling showing increased separation between clusters I and II following phosphorylation (Extended Data Fig. 6c). Yellow circle denotes spin-labelling position, shaded lines are a visual aid. **f**, pSer118 increases R_g by 5.2 ± 0.4 Å (Extended Data Fig. 1e). **g**, Phosphorylation-induced ER-NTD expansion.

Hydrophobic aromatic clustering

To investigate amino acid interactions between clusters I and II, we conducted paramagnetic relaxation enhancement (PRE) measurements (Fig. 1f). PRE detects distances between a nitroxide spin label and nearby amino acids within approximately 12–25 Å (refs. 26,27), with closer residues showing substantial to complete resonance broadening. Exploiting the absence of native cysteine residues in ER-NTD, we positioned MTSL ((1-oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl)

methanesulfonothioate) spin labels at six cysteine-comparable serine sites distributed across the primary sequence (Fig. 1g). The resulting PRE profiles showed low-intensity ratios between paramagnetic and diamagnetic states for residues 40–150 within mid-sequence regions, with these intensity attenuations remaining consistent across protein concentrations (Extended Data Fig. 2b,c), indicating predominantly intramolecular interactions.

The observed attenuations spanned a wider region than predicted, providing evidence for long-range interactions within the disordered
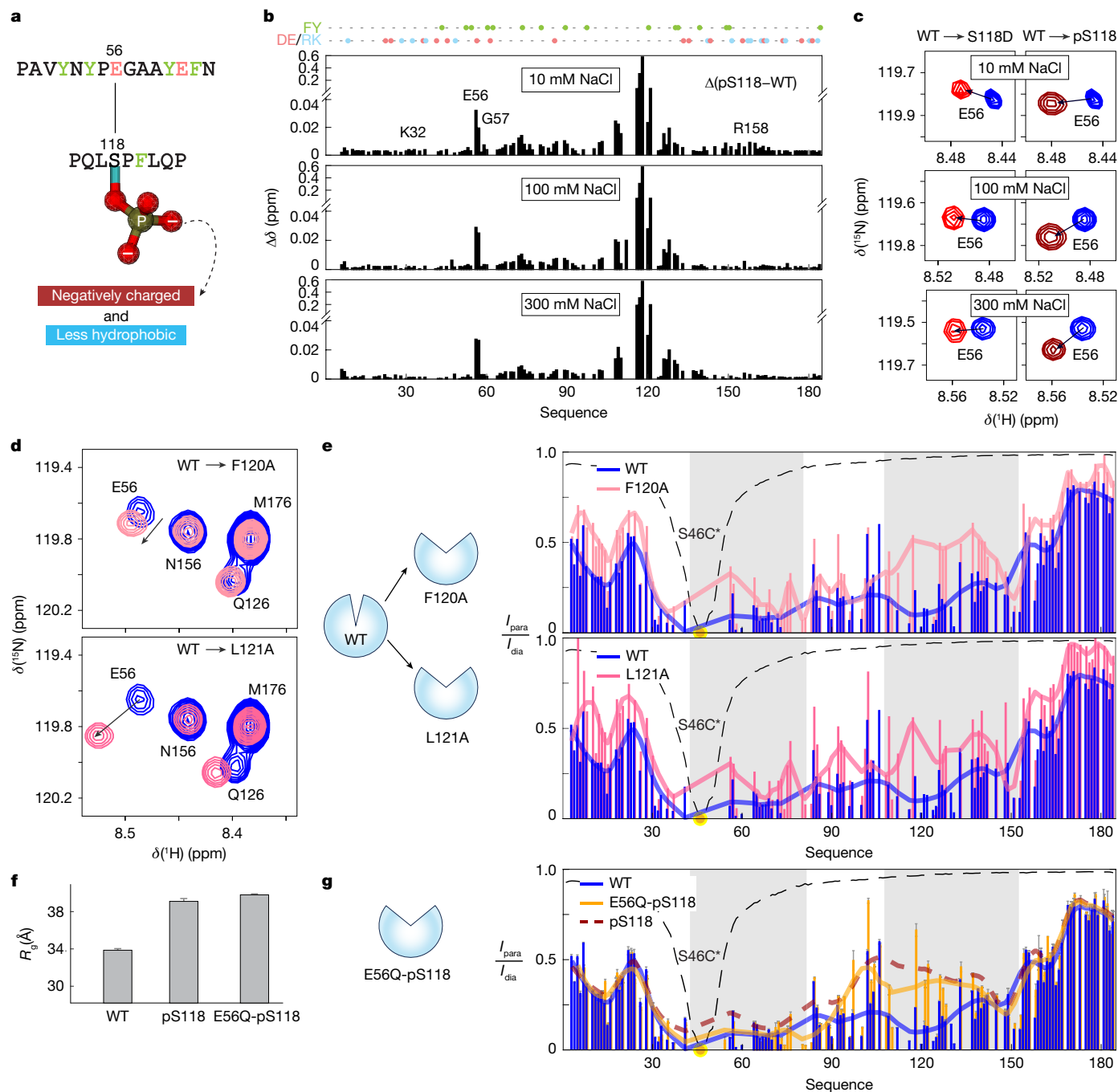
Fig. 3 | Hydrophobic disruption drives ER-NTD conformational changes. **a**, Structural context of pS118 and E56/G57 interactions. Ball-and-stick, phosphoryl group of pS118; light green, aromatic residues; red, charged residues. **b**, Chemical shift changes ($\Delta\delta$) following pSer118 at varying salt concentrations (Methods). **c**, E56 chemical shift changes in HSQC spectra for S118D (left) and pS118 (right) under three salt conditions. **d**, E56 chemical shift changes induced by mutations in F120A (top) and L121A (bottom). **e**, PRE measurements with S46C spin labelling showing F120A and L121A mutations mimicking phosphorylation effects. Yellow circles, spin label position; solid lines, visual aid (WT blue, mutants pink). HSQC, 850 MHz, 4 °C. **f**,**g**, Charge-neutralizing E56Q mutation showing minimal impact on R_g (**f**; Extended Data Fig. 1e) and cluster separation (**g**; S46C spin-labelling PRE) compared with pS118.

ER-NTD. Analysis showed that clusters I and II contain 13 of the 16 aromatic amino acids (Fig. 1e). PRE measurements demonstrated clear interactions between these clusters, suggesting that the residual structures arise from aromatic amino acid clustering, bridged by specific hydrophobic interactions between these aromatic-rich regions.

pSer118 drives cluster separation

We examined residue-specific effects of pS118 on ER-NTD conformation. ER-NTD proteins were phosphorylated in vitro using recombinant MAPK[28] (Extended Data Fig. 4a), and resonances for pS118-NTD were assigned (Fig. 2a and Extended Data Fig. 5a). Chemical shift changes were minimal and were localized near Ser118 (Extended Data Fig. 6a), with no significant alterations in local secondary structure propensity or cluster formation (Fig. 2b and Extended Data Fig. 5b) requiring exploration by future Carr−Purcell−Meiboom−Gill (CPMG) measurements. However, we observed substantial long-range effects: pS118 triggers conformational changes extending to distant Glu56 and Gly57 sites in cluster I, despite being located in cluster II (Fig. 2c,d). Intercluster distance changes assessed using PRE measurements with spin labelling
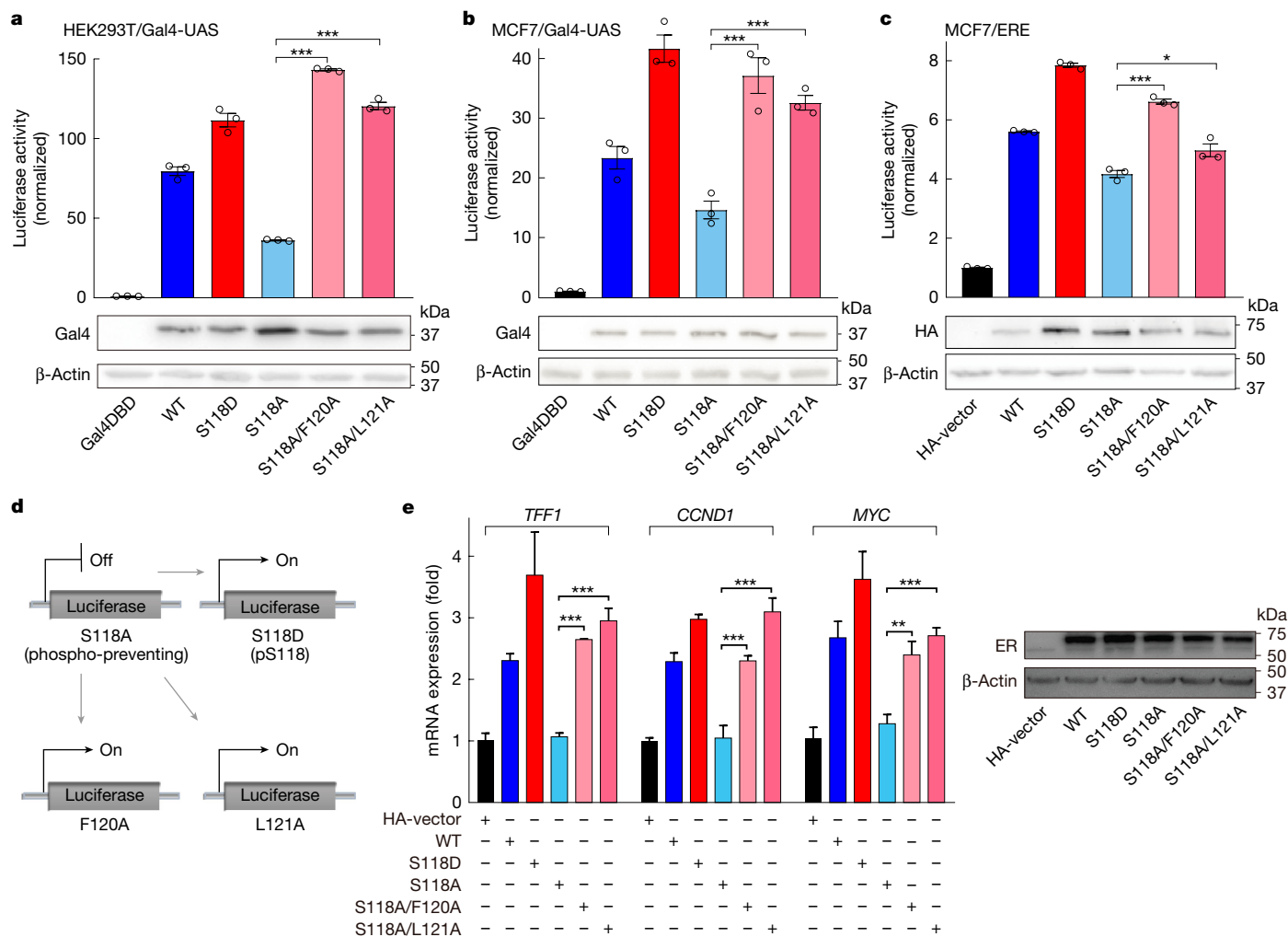
Fig. 4 | Hydrophobic mutations rescue S118A-induced transcriptional deficits. a,b, F120A and L121A restore S118A-impaired reporter activity in ER-NTD (HEK293T (**a**) and MCF7 cells (**b**)). Immunoblots confirmed protein expression. Mean ± s.e.m. from four technical repeats performed across three biological replicates. For gel source data, see Supplementary Fig. 4. **c**, F120A and L121A restore S118A-impaired reporter activity in full-length ER (MCF7 cells). Immunoblots confirmed protein expression. **d**, Schematic showing

F120A and L121A compensation for S118A deficiency. **e**, F120A and L121A restore S118A-reduced ER target gene expression (*TFF1*, *CCND1* and *MYC*) in MCF7 cells. Immunoblots confirmed protein expression. Mean ± s.e.m. from three technical repeats across three biological replicates. One-way analysis of variation with Dunnett's test, *$P < 0.05$, **$P < 0.01$, ***$P < 0.005$. mRNA, messenger RNA.

at S46C (edge of cluster I) and S84C (outside both clusters) (Fig. 2b) showed increased cluster separation following phosphorylation, through greater distances in PRE profiles at S46C (Fig. 2e) and S84C (Extended Data Fig. 6c). SEC-SAXS data confirmed this expansion, showing a significant increase (5.2 ± 0.4 Å) in R_g for pS118 (Fig. 2f). These findings demonstrate that phosphorylation promotes ER-NTD expansion through increased cluster seperation (Fig. 2g) while maintaining local cluster integrity (Fig. 2b).

The phosphomimetic S118D mutation elicits conformational changes comparable to pS118. Consistent with enhanced ER transcription activity following pSer118 by MAPK or CDK7 (refs. 2,3), S118D exhibits two-to-three-fold increased ER-NTD activity versus wild type (WT) in HEK293 and MCF7 cells (Extended Data Fig. 7a,b). This effect extends to full-length ER (Extended Data Fig. 7c), whereas phosphorylation-deficient S118A retains only 50–60% of WT activity. Multiple lines support S118D as a valid phosphomimetic: S118D proteins interact with Pin1 (Extended Data Fig. 4b), which selectively binds phosphorylated ER-NTD[19]. Resonance assignments of S118D (Fig. 2a and Extended Data Fig. 5b) show chemical shift changes at Glu56/Gly57 (Fig. 2c and Extended Data Fig. 6b) similar to pS118, without altering secondary structure propensity (Extended Data Fig. 5b).

S118D increases the separation between clusters I and II (Extended Data Fig. 6d), although less so than pS118 (Extended Data Fig. 6c). SEC-SAXS data confirm this expansion, showing that S118D causes a 3 Å increase in R_g over WT or S118A (Extended Data Fig. 1e), compared with a 5 Å increase for pS118 (Fig. 2g).

Hydrophobic disruption of cluster interactions

pSer118 affects structure through two potential mechanisms: electrostatic repulsion with negative residues (Glu56 and Glu61) and disruption of hydrophobic interactions between clusters I and II (Fig. 3a). Whereas terminal charged residues (Lys32 and Arg158) show salt-dependent chemical shift changes, cluster I residues (Glu56 and Gly57) remain salt-insensitive (Fig. 3b), suggesting that nonelectrostatic factors drive changes.

Three experiments established hydrophobic disruption as the primary mechanism. First, non-ionic detergent *n*-dodecyl β-D-maltoside (DDM) alters PRE profiles in cluster II regions (Extended Data Fig. 2d). Second, alanine substitutions near Ser118 (F120A and L121A) mimic phosphorylation effects, showing similar chemical shift changes and increased cluster separation (Fig. 3d,e). Third, alanine mutations in
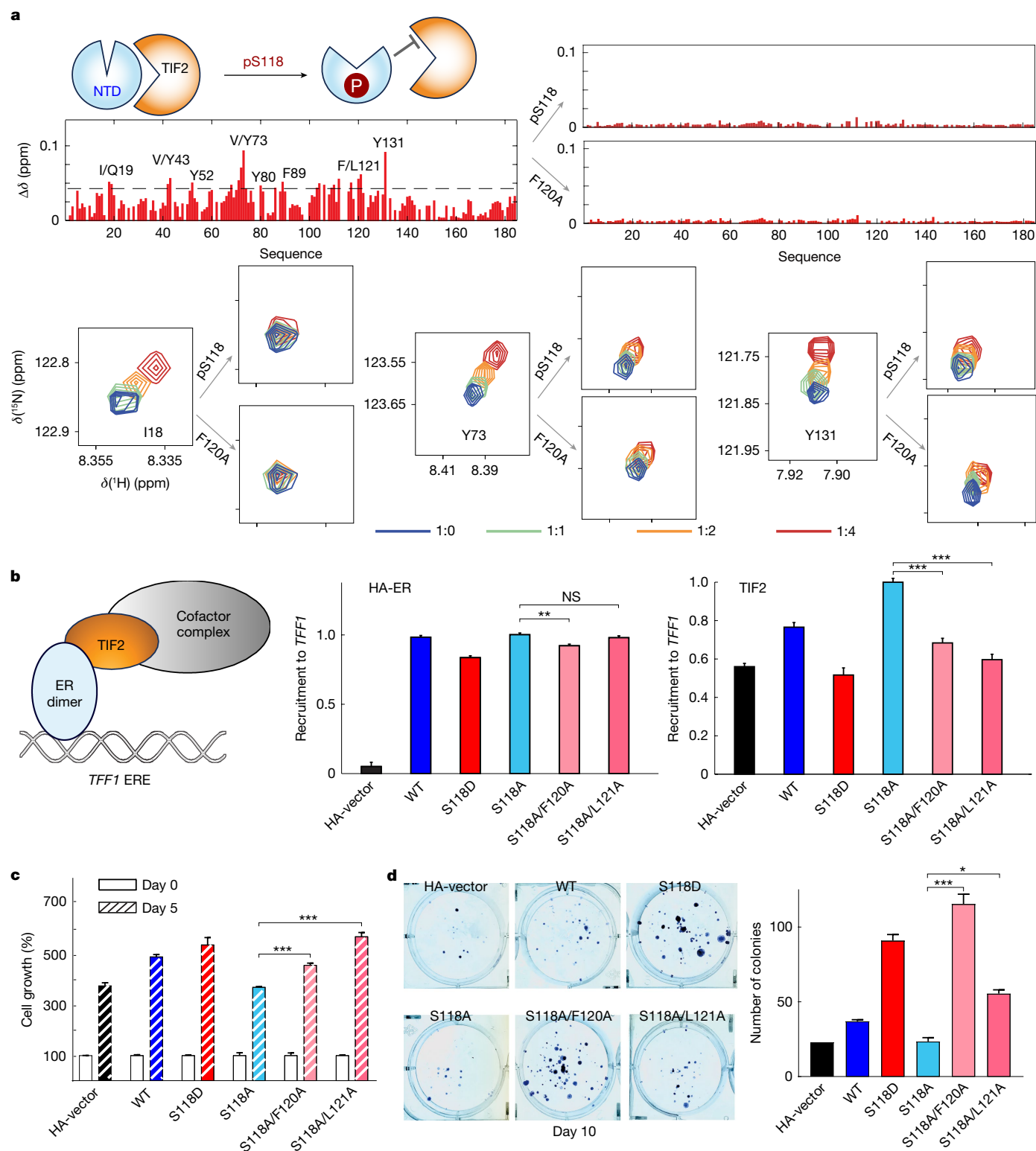
Fig. 5 | Hydrophobic mutations modulate cofactor interactions and cell growth. a, Top, F120A alters ER-NTD–TIF2-QRD interactions similarly to pSer118. Chemical shift changes ($\Delta\delta$) shown for ER-NTD variants (WT, pS118 and F120A) following TIF2-QRD binding (1/4 ratio). Annotated residues, above-average changes. Additional data are provided in Extended Data Fig. 9d and Supplementary Fig. 6. Bottom, HSQC spectra at various NTD/TIF2 ratios. **b**, Effects of F120A and L121A on ER-S118A and TIF2 recruitment to *TFF1* promoter (MCF7 cells). Left, ER–TIF2 interaction schematic. Middle, ER variant association unchanged versus WT. Right, F120A and L121A reverse S118A-induced TIF2

recruitment increase, mimicking S118D. *TFF1* ERE was quantified by CUT&RUN/ qPCR. Mean ± s.e.m. from four technical repeats. **c,d**, F120A and L121A restore S118A-impaired MCF7 cell growth (**c**; mean ± s.e.m. from three technical repeats across three biological replicates) and colony formation (**d**; mean ± s.e.m. from two technical repeats across two biological replicates; DMEM medium, crystal violet quantification at 450 nm). Additional data on 4-hydroxytamoxifen are provided in Extended Data Fig. 7d. One-way analysis of variation with Dunnett's test, *$P < 0.05$, **$P < 0.01$, ***$P < 0.005$. NS, not significant.
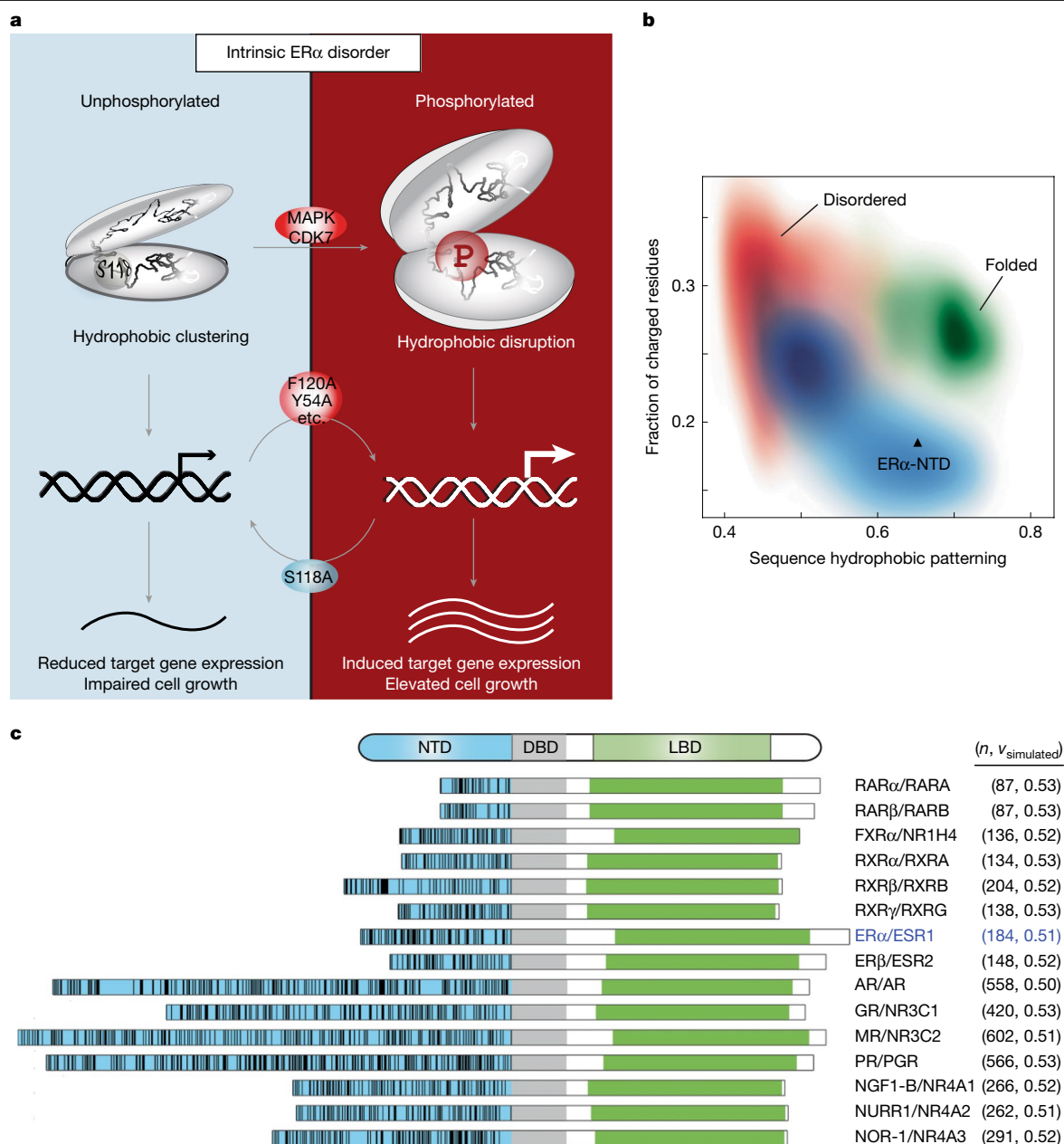
Fig. 6 | Hydrophobic patterning in nuclear receptor NTDs. a, pSer118 alters ER-NTD hydrophobic clustering, inducing long-range conformational changes. This clamshell-like motion modulates ER-mediated gene expression and cellular phenotypes. **b**, Nuclear receptor NTDs showing higher hydrophobic content than typical IDPs. Contour plot comparing fraction of charged residues versus sequence hydrophobic patterning[55] for 28,058 IDPs[7] (red), 2,360 folded proteins[56] (green) and 48 nuclear receptor NTDs (blue). One-third of nuclear receptors and 9.4% of IDPs share ERα-NTD-like characteristics (fraction of charged residues below 0.22, sequence hydrophobic patterning above 0.6). **c**, Left, nuclear receptors with similar NTD characteristics (blue) showing DNA-binding domain (DBD, grey), LBD (green) and hydrophobic residue positions (black lines: A, V, I, L, M, F, Y, W). Right, receptors with NTD residue count (n) and Flory exponent ($v_{simulated}$) from coarse-grained simulations (Extended Data Fig. 10; values provided in Supplementary Table 1); ER is highlighted in light blue.

cluster I show cumulative effects, with double mutants (Y52A/Y54A and Y60A/F62A) causing greater separation than single mutants (Y54A and Y60A) (Extended Data Fig. 8a–c).

The charge-neutralizing E56Q mutation confirms this mechanism—both E56Q-pS118 and pS118 show similar increases in the radius of gyration and cluster separation versus WT (Fig. 3f,g), establishing hydrophobic disruption as the primary driver of structural changes.

Hydrophobic rescue of ER function

To investigate whether reduced hydrophobicity mimics pS118 effects, we compared phosphorylation-deficient (S118A) and phosphomimetic

(S118D) mutations. S118D exhibited enhanced transcriptional activity versus S118A (Fig. 4a–d and Extended Data Fig. 7a–c), validating our phosphomimetic approach. Both S118A/F120A and S118A/L121A double mutants rescued S118A-impaired transcription in HEK293 and MCF7 cells, in both ER-NTD and full-length contexts (Fig. 4a–d). Additional mutations Y52A/Y54A and Y60A/F62A in the S118A background similarly enhanced transcriptional activity (Extended Data Fig. 8d–f), demonstrating hydrophobic disruption effects across both aromatic-rich clusters. Quantitative PCR (qPCR) with reverse transcription analyses of ER target genes (*TFF1*, also known as *pS2* (ref. 29), *CCND1* (ref. 30) and *MYC*[31]) showed that both S118A/F120A and S118A/L121A restored expression to WT and S118D levels (Fig. 4e).

We next investigated ER interactions with transcription intermediary factor TIF2, which binds ER-NTD through its Q-rich domain (QRD) and ER-LBD through its nuclear receptor interaction domain (Extended Data Fig. 9a,b)[32–34]. Chemical shift analyses showed that both pSer118 and F120A significantly reduced ER-NTD–TIF2-QRD interactions (Fig. 5a and Extended Data Fig. 9d). At the *TFF1* promoter, which contains estrogen-responsive elements (EREs), S118A enhanced TIF2 recruitment compared with WT ER[35,36]. Both S118D and hydrophobic mutations reversed this effect while maintaining a WT-like promoter association (Fig. 5b). These results demonstrate that ER-NTD hydrophobic clustering modulates cofactor interactions and transcriptional activity.

Phenotypic restoration of cell growth

To elucidate the biological significance of the hydrophobic-driven mechanism, we investigated how F120A and L121A mutations affect cell growth and colony formation in MCF7 cells. Whereas S118A failed to promote growth, S118D enhanced growth comparable to, or higher than, WT ER (Fig. 5c). F120A or L121A alongside S118A restored growth-promoting capacity to WT and S118D levels (Fig. 5c), with enhanced promotion in 4-hydroxytamoxifen-treated cells (Extended Data Fig. 7d). Colony formation assays showed fewer colonies with S118A than WT ER, whereas S118D enhanced colony formation. S118A/F120A and S118A/L121A double mutations restored colony formation to levels comparable to, or higher than, WT ER (Fig. 5d). Additional mutations Y52A/Y54A and Y60A/F62A (in S118A background) also rescued cell growth and colony formation impaired by S118A (Extended Data Fig. 8g,h), supporting hydrophobic disruption effects across both clusters.

These functional rescues, combined with structural evidence (Fig. 3e and Extended Data Fig. 8b,c), demonstrate a fundamental sequence–structure–function relationship of intrinsic ER disorder. Single-site phosphorylation on disordered ER-NTD alters hydrophobic clustering, leading to significant structural and functional consequences (Fig. 6a). Mutations disrupting hydrophobic interactions near Ser118, or across both clusters, mimic phosphorylation effects, rescuing the functional deficiency caused by the phosphorylation-deficient S118A mutation.

Discussion

Ser118 phosphorylation, frequently observed in therapy-resistant ER[+] breast cancer[4,13], enhances ER ligand-independent activity and represents a validated therapeutic target, as demonstrated by CDK7 inhibitors[5]. We demonstrate that this phosphorylation triggers specific long-range conformational changes in the disordered ER-NTD, shifting its ensemble towards expanded structures (Fig. 6a). Although some intramolecular contacts persist (Fig. 2b,e), this reorganization significantly impacts ER-mediated transcription and growth. This modulation between hydrophobic clusters shows a new regulatory mechanism among intrinsically disordered proteins/regions (IDPs)[35,37,38], with ER-NTD exemplifying rare regulation across multiple disordered regions, distinct from conventional phosphorylation effects in folded proteins[36,39].

Our findings extend beyond conventional electrostatic models[40–42], establishing hydrophobic interactions as the primary driver of phosphorylation-induced changes in ER-NTD. Systematic analysis shows that, whereas salt variations minimally affected the Glu56/Gly57 region (Fig. 3b), hydrophobic disruption through non-ionic DDM (Extended Data Fig. 2d) and targeted alanine substitutions (Fig. 3d,e and Extended Data Fig. 8) produced significant effects. This mechanism provides insights into ER phase-separation behaviour[43], although regulatory mechanisms vary across nuclear receptors[44].

Quantitative analysis of hydrophobic patterning shows its presence in 9.4% of analysed IDPs[7] (Fig. 6b) and in about 15 nuclear receptor NTDs[8] (Fig. 6c and Extended Data Fig. 10), despite size variations and low sequence homology. The ER-NTD maintains its disorder while engaging in specific protein interactions, particularly with TIF2 (refs. 45,46), which bridges non-interacting ER-NTD and ER-LBD[34,47] (Extended Data Fig. 9b,c). Ser118 phosphorylation precisely modulates these interactions through conformational changes, with effects propagating throughout the interactome[48,49]. Although targeting of disordered NTDs remains challenging, as evidenced by androgen receptor NTD inhibitor development[14,15], and current ER degraders[50,51] affect the NTD indirectly, our study establishes a framework for direct ER-NTD targeting through conformational modulation.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-08400-1.

1. Ali, S. & Coombes, R. C. Endocrine-responsive breast cancer and strategies for combating resistance. *Nat. Rev. Cancer* **2**, 101–112 (2002).
2. Kato, S. et al. Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase. *Science* **270**, 1491–1494 (1995).
3. Chen, D. et al. Activation of estrogen receptor alpha by S118 phosphorylation involves a ligand-dependent interaction with TFIIH and participation of CDK7. *Mol. Cell* **6**, 127–137 (2000).
4. Patel, H. et al. ICEC0942, an orally bioavailable selective inhibitor of CDK7 for cancer treatment. *Mol. Cancer Ther.* **17**, 1156–1166 (2018).
5. Sava, G. P., Fan, H., Coombes, R. C., Buluwela, L. & S, Ali, S. CDK7 inhibitors as anticancer drugs. *Cancer Metastasis Rev.* **39**, 805–823 (2020).
6. Coombes, R. C. et al. Dose escalation and expansion cohorts in patients with advanced breast cancer in a Phase I study of the CDK7-inhibitor samuraciclib. *Nat. Commun.* **14**, 4444 (2023).
7. Tesei, G. et al. Conformational ensembles of the human intrinsically disordered proteome. *Nature* **626**, 897–904 (2024).
8. Mangelsdorf, D. J. et al. The nuclear receptor superfamily: the second decade. *Cell* **83**, 835–839 (1995).
9. Giaquinto, A. N. et al. Breast cancer statistics, 2022. *CA Cancer J. Clin.* **72**, 524–541 (2022).
10. Tora, L. et al. The human estrogen receptor has two independent nonacidic transcriptional activation functions. *Cell* **59**, 477–487 (1989).
11. Guan, J. et al. Therapeutic ligands antagonize estrogen receptor function by impairing its mobility. *Cell* **178**, 949–963 (2019).
12. Shiau, A. K. et al. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **95**, 927–937 (1998).
13. Jeselsohn, R. et al. Allele-specific chromatin recruitment and therapeutic vulnerabilities of ESR1 activating mutations. *Cancer Cell* **33**, 173–186 (2018).
14. Andersen, R. J. et al. Regression of castrate-recurrent prostate cancer by a small-molecule inhibitor of the amino-terminus domain of the androgen receptor. *Cancer Cell* **17**, 535–546 (2010).
15. De Mol, E. et al. EPI-001, a compound active against castration-resistant prostate cancer, targets transactivation unit 5 of the androgen receptor. *ACS Chem. Biol.* **11**, 2499–2505 (2016).
16. Yi, P. et al. Structural and functional impacts of ER coactivator sequential recruitment. *Mol. Cell* **67**, 733–743 (2017).
17. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
18. Warnmark, A., Wikstrom, A., Wright, A. P. H., Gustafsson, J. A. & Hard, T. The N-terminal regions of estrogen receptor alpha and beta are unstructured in vitro and show different TBP binding properties. *J. Biol. Chem.* **276**, 45939–45944 (2001).
19. Rajbhandari, P. et al. Regulation of estrogen receptor alpha N-terminus conformation and function by peptidyl prolyl isomerase Pin1. *Mol. Cell. Biol.* **32**, 445–457 (2012).
20. Peng, Y. et al. A metastable contact and structural disorder in the estrogen receptor transactivation domain. *Structure* **27**, 229–240 (2019).
21. Zheng, W., Du, Z., Ko, S. B., Wickramasinghe, N. P. & Yang, S. Incorporation of D(2)O-induced fluorine chemical shift perturbations into ensemble-structure characterization of the ERalpha disordered region. *J. Phys. Chem. B* **126**, 9176–9186 (2022).
22. Riback, J. A. et al. Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water. *Science* **358**, 238–241 (2017).
23. Ozenne, V. et al. Flexible-meccano: a tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. *Bioinformatics* **28**, 1463–1470 (2012).
24. Klein-Seetharaman, J. et al. Long-range interactions within a nonnative protein. *Science* **295**, 1719–1722 (2002).
25. Martin, E. W. et al. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694–699 (2020).
26. Battiste, J. L. & Wagner, G. Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* **39**, 5355–5365 (2000).

27. Tang, C., Iwahara, J. & Clore, G. M. Visualization of transient encounter complexes in protein-protein association. *Nature* **444**, 383–386 (2006).

28. Likhite, V. S., Stossi, F., Kim, K., Katzenellenbogen, B. S. & Katzenellenbogen, J. A. Kinase-specific phosphorylation of the estrogen receptor changes receptor interactions with ligand, deoxyribonucleic acid, and coregulators associated with alterations in estrogen and tamoxifen activity. *Mol. Endocrinol.* **20**, 3120–3132 (2006).

29. Brown, A. M., Jeltsch, J. M., Roberts, M. & Chambon, P. Activation of pS2 gene transcription is a primary response to estrogen in the human breast cancer cell line MCF-7. *Proc. Natl Acad. Sci. USA* **81**, 6344–6348 (1984).

30. Eeckhoute, J., Carroll, J. S., Geistlinger, T. R., Torres-Arzayus, M. I. & Brown, M. A cell-type-specific transcriptional network required for estrogen regulation of cyclin D1 and cell cycle progression in breast cancer. *Genes Dev.* **20**, 2513–2526 (2006).

31. Shang, Y., Hu, X., DiRenzo, J., Lazar, M. A. & Brown, M. Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell* **103**, 843–852 (2000).

32. Voegel, J. J. et al. The coactivator TIF2 contains three nuclear receptor-binding motifs and mediates transactivation through CBP binding-dependent and -independent pathways. *EMBO J.* **17**, 507–519 (1998).

33. Schaufele, F. et al. Temporally distinct and ligand-specific recruitment of nuclear receptor-interacting peptides and cofactors to subnuclear domains containing the estrogen receptor. *Mol. Endocrinol.* **14**, 2024–2039 (2000).

34. Benecke, A., Chambon, P. & Gronemeyer, H. Synergy between estrogen receptor alpha activation functions AF1 and AF2 mediated by transcription intermediary factor TIF2. *EMBO Rep.* **1**, 151–157 (2000).

35. Bah, A. & Forman-Kay, J. D. Modulation of intrinsically disordered protein function by post-translational modifications. *J. Biol. Chem.* **291**, 6696–6705 (2016).

36. Pawson, T. & Scott, J. D. Protein phosphorylation in signaling–50 years and counting. *Trends Biochem. Sci.* **30**, 286–290 (2005).

37. Lum, J. K., Neuweiler, H. & Fersht, A. R. Long-range modulation of chain motions within the intrinsically disordered transactivation domain of tumor suppressor p53. *J. Am. Chem. Soc.* **134**, 1617–1622 (2012).

38. Kulkarni, P. et al. Phosphorylation-induced conformational dynamics in an intrinsically disordered protein and potential role in phenotypic heterogeneity. *Proc. Natl Acad. Sci. USA* **114**, E2644–E2653 (2017).

39. Motlagh, H. N., Wrabl, J. O., Li, J. & Hilser, V. J. The ensemble nature of allostery. *Nature* **508**, 331–339 (2014).

40. Uversky, V. N., Gillespie, J. R. & Fink, A. L. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* **41**, 415–427 (2000).

41. Das, R. K. & Pappu, R. V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl Acad. Sci. USA* **110**, 13392–13397 (2013).

42. Muller-Spath, S. et al. From the cover: charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc. Natl Acad. Sci. USA* **107**, 14609–14614 (2010).

43. Boija, A. et al. Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842–1855 (2018).

44. Ahmed, J., Meszaros, A., Lazar, T. & Tompa, P. DNA-binding domain as the minimal region driving RNA-dependent liquid-liquid phase separation of androgen receptor. *Protein Sci.* **30**, 1380–1392 (2021).

45. Hong, H., Kohli, K., Trivedi, A., Johnson, D. L. & Stallcup, M. R. GRIP1, a novel mouse protein that serves as a transcriptional coactivator in yeast for the hormone binding domains of steroid receptors. *Proc. Natl Acad. Sci. USA* **93**, 4948–4952 (1996).

46. Dutertre, M. & Smith, C. L. Ligand-independent interactions of p160/steroid receptor coactivators and CREB-binding protein (CBP) with estrogen receptor-alpha: regulation by phosphorylation sites in the A/B region depends on other receptor domains. *Mol. Endocrinol.* **17**, 1296–1314 (2003).

47. Norris, J. D., Fan, D., Stallcup, M. R. & McDonnell, D. P. Enhancement of estrogen receptor transcriptional activity by the coactivator GRIP-1 highlights the role of activation function 2 in determining estrogen receptor pharmacology. *J. Biol. Chem.* **273**, 6679–6688 (1998).

48. Kim, M. et al. A protein interaction landscape of breast cancer. *Science* **374**, eabf3066 (2021).

49. Rega, C. et al. Exploring the spatial landscape of the estrogen receptor proximal proteome with antibody-based proximity labeling. *Mol. Cell. Proteomics* **23**, 100702 (2024).

50. Wittmann, B. M., Sherk, A. & McDonnell, D. P. Definition of functionally important mechanistic differences among selective estrogen receptor down-regulators. *Cancer Res.* **67**, 9549–9560 (2007).

51. Gough, S. M. et al. Oral estrogen receptor PROTAC vepdegestrant (ARV-471) is highly efficacious as monotherapy and in combination with CDK4/6 or PI3K/mTOR pathway inhibitors in preclinical ER⁺ breast cancer models. *Clin. Cancer Res.* **30**, 3549–3563 (2024).

52. Bernado, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* **129**, 5656–5664 (2007).

53. Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503 (1992).

54. Tamiola, K., Acar, B. & Mulder, F. A. Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J. Am. Chem. Soc.* **132**, 18000–18003 (2010).

55. Zheng, W., Dignon, G., Brown, M., Kim, Y. C. & Mittal, J. Hydropathy patterning complements charge patterning to describe conformational preferences of disordered proteins. *J. Phys. Chem. Lett.* **11**, 3408–3415 (2020).

56. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).

Methods

Recombinant protein expression and purification

The NTD of human ER (ER-NTD, residues M1–Y184) was subcloned into the bacterial expression vector pMCSG7 and expressed as inclusion bodies in *Escherichia coli* BL21(DE3) cells using lysogeny broth medium. Protein expression was induced at optical density $OD_{600} = 0.8$ by the addition of 0.1 mM isopropyl-ß-D-thiogalactopyranoside (IPTG), followed by overnight incubation at 25 °C. Mutants were generated using the Phusion site-directed mutagenesis kit (Thermo Fisher, catalogue no. F541) and verified by DNA sequencing.

For isotopic ^{15}N and ^{13}C labelling, cells were cultured in M9 minimal medium supplemented with 1 g l[-1] [15]NH$_4$Cl and 3 g l[-1] ^{13}C-glucose (Cambridge Isotope Laboratories, catalogue nos. CLM-1396-50 and NLM-467-50, respectively). Harvested cells were lysed by high-pressure homogenization (Avestin EmulsiFlex-C3) in 50 mM Tris pH 8.0 and 500 mM NaCl buffer with protease inhibitors (Millipore Sigma, catalogue no. 11836153001). Inclusion bodies were solubilized in 50 mM sodium phosphate pH 7.4, 5 mM imidazole and 1 mM phenylmethanesulfonyl fluoride (PMSF) with 6 M guanidine hydrochloride, and clarified by centrifugation at 4 °C.

Untagged and His-tagged ER-NTD proteins were purified from solubilized inclusion bodies. Following the purification protocol[57], untagged proteins were dialysed into 20 mM Tris pH 7.5, 10 mM NaCl and 0.1 mM PMSF, and purified by anion exchange using a HiTrap Q column (GE Healthcare, catalogue no. 29-0513-25) and gel filtration with an Enrich SEC650 column (Bio-Rad, catalogue no. 7801650). His-tagged proteins were purified on cobalt resin (Thermo Fisher, catalogue no. 89964), eluted with imidazole and dialysed into 20 mM Tris pH 7.4, 10 mM NaCl and 0.1 mM PMSF. The His-tag was removed by overnight TEV protease digestion at 4 °C, and further purified through the same anion-exchange and size-exclusion steps. Purified ERα-NTD proteins were stored at either 4 or −80 °C.

In vitro protein phosphorylation

Purified ER-NTD proteins were phosphorylated in vitro by incubation with recombinant, activated MAP kinase 1 (MAPK1/ERK2) possessing an N-terminal glutathione *S*-transferase (GST) tag (Millipore Sigma, catalogue no. 14-550). Phosphorylation followed a previously published protocol[19], using 3 mg of ER-NTD proteins and 5 µg of activated MAPK1 in a kinase reaction buffer (25 mM Tris pH 7.5, 25 mM MgCl$_2$, 10 mM dithiothreitol and 100 µM ATP). Reactions proceeded for 4 h at 30 °C. Site-specific phosphorylation at S118 was verified by immunoblot using a phosphorylation-ER Ser118 antibody (Cell Signaling, catalogue no. 2511). Following the reaction, GST-MAPK1 enzyme was removed by incubation with glutathione-agarose resin (Macherey-Nagel, catalogue no. NC0389541).

SEC-SAXS

SEC-SAXS measurements were performed on protein samples at concentrations ranging from 2 to 4 mg ml[-1]. The running buffer, consisting of 20 mM sodium phosphate, 150 mM NaCl, 0.5 mM EDTA and 0.1 mM PMSF pH 7.4, was precooled to 4 °C and maintained on ice during data acquisition. Samples were injected into either an Enrich SEC650 column (Bio-Rad, catalogue no. 7801650) or a Superdex 75 Increase 5/150 column (Cytiva, catalogue no. 29148722), at flow rates ranging from 0.35 to 0.70 ml min[-1]. Scattering images were collected at either 0.5 or 2-s exposure per frame during column elution. Background subtraction was applied to the elution peak to obtain the final protein-scattering profile.

For WT, S118D and S118A samples, data were collected at the BioCAT-18ID beamline of the Advanced Photon Source (Argonne, IL, USA) using a Pilatus3-X-1M detector with X-rays at an energy of 13 keV. For pS118 and E56Q-pS118 samples, data collection was performed at the 16-ID (LiX) beamline of the National Synchrotron Light Source II (Upton, NY, USA) using a combination of Pilatus3 S 1 M and Pilatus3 900-K detectors with X-rays at an energy of 15 keV.

NMR spectroscopy

Isotope-labelled proteins were prepared in 20 mM sodium phosphate buffer pH 7.4, 100 mM NaCl, 0.5 mM EDTA, 0.1 mM PMSF and 5% (v/v) D$_2$O. NMR data were acquired using either a Bruker Avance III HD 700 or 850-MHz spectrometer equipped with a TXO- or TCI triple-resonance cryogenic probe, respectively, at a temperature of 4 °C, to reduce the line broadening of solvent-exposed residues due to their amide hydrogen exchange with the solvent.

For ^{15}N,[13]C ER-NTD samples at a concentration of 400 µM, a set of standard triple-resonance backbone assignment experiments[58] were recorded at 850 MHz, using sensitivity-enhanced gradient coherence selection[59,60], semiconstant time acquisition in the ^{15}N dimension[61] and non-uniform sampling following either Poisson-gap[62] or exponentially weighted sampling schedules. Briefly, the typical direct ^{1}H dimension contained 1,024 complex points for a 100-ms acquisition time, whereas maximum data size of the indirect dimensions ([15]N × [13]C), in terms of both complex points (acquisition times) and sparsity levels, was as follows: ^{15}N HSQC with 100 ms (380; 47.4%), HNCO with 100 × 50 ms (380 × 150; 0.53%), HNCA with 100 × 14 ms (380 × 90; 1.75%), HNCOCA with 100 × 14 ms (380 × 90; 0.88%), HNCACO with 100 × 50 ms (380 × 150; 1.05%), HNCACB with 100 × 14 ms (380 × 240; 1.31%) and CBCACONH with 100 × 6.7 ms (380 × 115; 1.37%). Additional three-dimensional NNH-NOESY spectra were acquired with 380 complex points for 100-ms acquisition times in indirect ^{15}N dimensions using a sparsity level of 0.83%. The typical experimental times for spectra were as follows: 22 min for HSQC, 10 h for HNCO, 37 h for HNCA, 28 h for HNCOCA, 38 h for HNCACO, 75 h for HNCACB, 38 h for CBCACONH and 84 h for NNH-NOESY.

In addition to proton-detected spectra, [13]CO-detected hCACON and hCANCO spectra[63] were recorded at 700 MHz for resonance assignments of most proline residues. Data were collected using virtual decoupling, and without non-uniform sampling, at acquisition times (complex points) of 103 ms (512) × 30 ms (60) × 5.6 ms (30) for [13]CO × [15]N × [13]Cα dimensions. All NMR spectra were recorded at 4 °C.

Data were acquired and processed using BRUKER Topspin (v.3.6) and NMRPipe (v.1.0)[64]. Spectra were analysed using NMRfam-SPARKY (v.1.470) with Sparky 3.190 (ref. 65).

Spectral referencing was performed directly for the ^{1}H dimension using the temperature-dependent water resonance frequency, whereas ^{13}C and ^{15}N were referenced indirectly. Combined chemical shift changes for ^{15}N HSQC were calculated as $\Delta\delta = \sqrt{(\Delta\delta H^2 + (\Delta\delta N \times 0.154)^2}$, where $\Delta\delta H$ and $\Delta\delta N$ represent the changes in proton and nitrogen chemical shift, respectively, compared with the free or WT protein spectrum.

Paramagnetic relaxation enhancement

Single-cysteine mutants were labelled with the paramagnetic nitroxide spin label MTSL (Toronto Research Chemicals, catalogue no. O875000), following the established protocol[66]. Protein samples (100–200 µM) were buffer exchanged into a labelling buffer (50 mM Tris-HCl pH 7.4 and 50 mM NaCl) supplemented with 2.5 mM freshly prepared dithiothreitol (DTT) to prevent cysteine oxidation. Immediately before labelling, DTT was removed by passing the samples through a Zeba spin-desalting column equilibrated with DTT-free labelling buffer. MTSL was dissolved in acetonitrile (0.1% v/v) to a minimum concentration of 2 mM, and over 20-fold molar excess was added to the protein solution. The reaction mixture was incubated overnight at 4 °C on a rotating device, protected from light to ensure efficient labelling. Excess MTSL was removed using a Zeba spin-desalting column, and samples were concentrated to the desired level using an Amicon Ultra-15 centrifugal filter (4 °C). Spin-labelling efficiency (over 95%) was confirmed by electrospray ionization mass spectrometry. Diamagnetic control samples were prepared by reducing the paramagnetic samples

with a tenfold molar excess of L-ascorbic acid (Millipore Sigma, catalogue no. A7631), with overnight incubation at 4 °C.

Two-dimensional ^{1}H-^{15}N HSQC spectra were acquired for each MTSL-label, single-cysteine mutant in both paramagnetic and diamagnetic states at 4 °C. Comparison of chemical shifts between labelled and unlabelled samples confirmed that MTSL labelling had not significantly perturbed the protein's structural ensemble. PRE effects were quantified by calculating the peak intensity ratio (I_{para}/I_{dia}) between paramagnetic and diamagnetic spectra for each resolved cross-peak. Errors in intensity ratios were propagated using root-mean-square deviation between observed and Gaussian-fit peak heights.

To investigate the role of hydrophobicity in long-range interactions, PRE measurements were conducted for ER-NTD with S46C spin labelling (30 μM) with and without DDM (120 μM; Anatrace, catalogue no. D310). ^{1}H-^{15}N HSQC spectra were recorded under identical conditions for DDM-containing and DDM-free samples.

The effect of paramagnetic relaxation on cross-peak intensity[26,67] is expressed as

$$\frac{I_{para}}{I_{dia}} = \frac{R_2\left(^1H^N\right) \times e^{-\Gamma_2 \cdot t_o}}{R_2\left(^1H^N\right) + \Gamma_2},$$

where t_o = 10.87 ms (combined INEPT time) and $R_2(^1H^N)$ is the intrinsic relaxation rate of the amide proton. Γ_2, paramagnetic rate enhancement, is related to the distance between the amino proton and spin label by

$$R = \left[\frac{K}{\Gamma_2}\left(4\tau_c + \frac{3\tau_c}{1+\omega_h^2\tau_c^2}\right)\right]^{\frac{1}{6}},$$

where $K = 1.2 \times 10^{-32}$ cm^6 s^{-2} for the MTSL spin label[26] and ω_h is the Larmor frequency of proton spin at 850 MHz. The apparent correlation time (τ_c) for electron-amide proton interactions is 5.9 ± 1.4 ns (mean ± s.d.).

^{15}N relaxation measurement

Protein samples were used at a concentration of 60 μM for R_1 and $R_{1\rho}$ (spin-lattice relaxation rate in the rotating frame) measurements, at 850 MHz and 4 °C, using the HSQC-based sequences of Lakomek et al.[68] but with modifications for non-deuterated proteins[69]. Experiments were performed with ^{1}H and ^{15}N sweep width (acquisition times) of 12 ppm (100 ms) and 21 ppm (88 ms), respectively, using eight signal-averaging scans and 2.5-s recovery delay. Delays for R_1 ($R_{1\rho}$) relaxation were 0, 280, 560 × 2 and 840 × 2 ms (0, 55, 110 × 2 and 165 × 2) with repeats for statistics. Magnetization was aligned with the effective field in $R_{1\rho}$ experiments using 2-ms adiabatic half passages[70] at a calibrated B$_1$ field of 1.08 kHz centred at 121 ppm. R_1 and $R_{1\rho}$ rates were obtained by fitting mono-exponential decays using in-house Python scripts. Transverse R_2 rates were derived from R_1 and $R_{1\rho}$ using the equation

$$R_{1\rho} = R_1\cos^2\theta + R_2\sin^2\theta,$$

where $\theta = \tan(\omega_1/\Omega)$ is the effective tilt angle of B1 field, ω_1 is B1 field strength and Ω is ^{15}N signal offset from spinlock carrier frequency[71]. Errors were propagated from signal amplitude variations.

In vitro GST pulldown

The GST-tagged proteins were cloned into pGEX-4T-1 vector and expressed in bacterial Rosetta (DE3) pLysS-competent cells. Expression was induced with 0.2 mM IPTG overnight at 20 °C. Purification was performed using a glutathione spin column (Thermo Fisher, catalogue no. 16105). Untagged NTD proteins (10 μg) were incubated with GST-tagged proteins (100 μg) and immobilized overnight at 4 °C on glutathione-Sepharose beads in 20 mM Tris-HCl pH 8.0, 150 mM NaCl,

10% glycerol, 1% Nonidet P-40 (Sigma-Aldrich, catalogue no. N-3516) and 2 mM EDTA. Subsequently, the resin was washed three times with the same binding buffer to remove unbound proteins. Pulled-down complexes were denatured in tricine sample buffer, separated by 4–20% SDS–polyacrylamide gel electrophoresis and visualized using GelCode blue staining (Thermo Fisher, catalogue no. 24594).

Transient transfection reporter assay

To assay Gal4-NTD reporter activity, HEK293 or MCF7 cells were cultured in DMEM supplemented with 10% fetal bovine serum (FBS). Cells were seeded at 2×10^6 per well on 6-cm plates and transfected with 1 μg of Gal4-NTD, 3 μg of 5× Gal4 BS-TATA-luciferase plasmid (Addgene, catalogue no. 46756) and 0.3 μg of pRL-Renilla-luciferase plasmid (Promega, catalogue no. E2231), using Lipofectamine 2000 (Invitrogen, catalogue no. 11668027). Twenty-four hours after transfection, cells were washed with 1× PBS and grown in DMEM with 10% FBS for a further 24 h, and lysates prepared using 1× Passive Lysis Buffer (Promega, catalogue no. E1941).

For full-length ER assays, MCF7 cells in phenol red-free DMEM with 10% charcoal-stripped FBS were transfected with 2.25 μg of HA-tagged, full-length ER expression vector, 6.75 μg of 3× ERE-TATA-luciferase plasmid (Addgene, catalogue no. 11354) and 0.67 μg of pRL-Renilla-luciferase plasmid. Cells were seeded at 7×10^6 per well on 10-cm plates. Twenty-four hours after transfection, cells were washed with 1× PBS, split and cultured in phenol red-free DMEM with 10% charcoal-stripped FBS. Cells were treated with either vehicle (0.1% ethanol), 100 nM E2 or 1 μM 4-hydroxytamoxifen (Cayman, catalogue no. 17308) before lysate preparation.

Firefly and Renilla luciferase activities were measured using the Dual-Luciferase Reporter Assay kit (Promega, catalogue no. E1910), and firefly activity was normalized to Renilla luciferase. Mean ± s.e.m. was calculated from four technical repeats with three biological replicates.

Nucleocytoplasmic fractionation and whole-cell lysate preparation

Cells were lysed with NP-40 buffer (10 mM Tris pH 7.4, 10 mM NaCl, 3 mM MgCl$_2$, 0.5% NP-40 and 5% glycerol) and centrifuged to obtain cytoplasmic fractions. Nuclear fractions were isolated by lysing the resulting pellets in RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS and 50 mM Tris pH 7.4). Both lysis buffers included protease (Millipore Sigma, catalogue no. 04693116001) and phosphatase inhibitors (Millipore Sigma, catalogue no. 4906845001). MCF7 and HEK293T cells were lysed in RIPA buffer with protease (Millipore Sigma, catalogue no. 04693116001) and phosphatase (Millipore Sigma, catalogue no. 4906845001) inhibitors to obtain whole-cell lysates.

Immunoblotting

Proteins were separated by 10% SDS–polyacrylamide gel electrophoresis and transferred to polyvinylidene difluoride membranes (Bio-Rad, catalogue no. 1620177). Membranes were immunoblotted overnight at 4 °C with primary antibodies against ER (abcam, catalogue no. ab16660), phospho-ER-Ser118 (Cell Signaling, catalogue no. 2511), β-actin (Cell Signaling, catalogue no. 4967), Lamin B1 (Santa Cruz, catalogue no. sc-374015), β-tubulin (Santa Cruz, catalogue no. sc-166729), Gal4 (Santa Cruz, catalogue no. sc-510), HA-tag (Santa Cruz, catalogue no. sc-7392), His-tag (Millipore Sigma, catalogue no. SAB4301134) and GST-tag (Millipore Sigma, catalogue no. SAB4301139). Secondary anti-IgG antibodies (Bio-Rad, catalogue no. 1706515 or 1706516) were applied at room temperature before imaging on a ChemiDoc MP system (Bio-Rad, catalogue no. 12003154).

RNA extraction and qPCR with reverse transcription

To determine the effects of WT and mutant ER on the expression of ER target genes, cells on 6-cm plates were transfected with 2 μg of either

HA-tagged ER expression vector or empty vector control (HA-vector) using Lipofectamine 3000 (Thermo Fisher, catalogue no. L3000001) and 400 µl of Opti-MEM (Gibco, catalogue no. 31985062). Twelve hours after transfection, culture medium was replaced with regular DMEM supplemented with 10% FBS, and cells were grown for an additional 24 h before harvest. Cells were collected by scraping in cold 1× PBS and pelleted. Cell pellets were collected by scraping in cold 1× PBS, centrifuged, resuspended and evenly split for RNA and protein analysis.

Total RNA was isolated using the RNeasy Mini Kit (Qiagen, catalogue no. 74104) and reverse transcribed using iScript supermix (Bio-Rad, catalogue no. 1708841). Quantitative PCR was performed on a Bio-Rad CFX Connect using iQ SYBR Green supermix (Bio-Rad, catalogue no. 1708880).

The following primers were used:

TFF1 forward 5′-ATACCATCGACGTCCCTCCA-3′, reverse 5′-AAGCGT GTCTGAGGTGTCCG-3′;

MYC forward 5′-GGTGCTCCATGAGGAGACA-3′, reverse 5′-CCTGC CTCTTTTCCACAGAA-3′;

CCND1 forward 5′- GGATGCTGGAGGTCTGCGA-3′, reverse 5′-AGAG GCCACGAACATGCAAG-3′ and *GAPDH* forward 5′-CCACAGTCCATGC CATCA-3′, reverse 5′-GGATGACCTTGCCCACAG-3′.

Calculations were based on mean ± s.e.m. from three technical repeats with three biological replicates.

Cofactor interactions and chemical shift perturbation

The Q-rich domain of transcriptional mediators/intermediary factor 2 (TIF2-QRD, residues G1180–N1288) was subcloned into pMCSG7 bacterial expression vector with an N-terminal 6xHis-tag, and expressed in bacterial BL21(DE3) cells as inclusion bodies using TB medium. TIF2-QRD was purified similarly to ER-NTD as described above, including tag cleavage.

For chemical shift perturbation measurements, ^{15}N-labelled ER-NTD and unlabelled TIF2-QRD were exchanged into 20 mM sodium phosphate buffer pH 7.4, containing 100 mM NaCl. TIF2-QRD was titrated into 30 µM ER-NTD at ER-NTD/TIF2-QRD molar ratios of 1:0, 1:1, 1:2 and 1:4. Two-dimensional ^{1}H-^{15}N HSQC spectra were recorded at each titration point. Combined chemical shift changes ($\Delta\delta$) were calculated as $\Delta\delta = \sqrt{(\Delta\delta H^2 + (\Delta\delta N \times 0.154)^2}$, where $\Delta\delta H$ and $\Delta\delta N$ are the changes in proton and nitrogen chemical shift, respectively, relative to the free unbound protein spectrum.

CUT&RUN assay followed by qPCR

CUT&RUN assays were performed according to the kit protocol (Cell Signaling, catalogue no. 86652). Briefly, MCF7 cells were transfected with WT or mutant HA-tagged ER expression plasmids overnight and cultured for 48 h in phenol red-free medium supplemented with 5% charcoal-stripped FBS. Cells were then treated with 10 nM E2 for 150 min. For each immunoprecipitation reaction, 1×10^5 cells were immobilized on Concanavalin A beads and incubated with either 2 µl of primary antibody anti-HA (Cell Signaling, catalogue no. 3724) or anti-TIF2 (Cell Signaling, catalogue no. 96687) for 2 h at 4 °C. Chromatin was fragmented by further incubation with pAG-MNase digestion. Fragmented DNA was purified and analysed by qPCR (normalized to input) using PowerUp SYBR Green Master Mix (Thermo Fisher, catalogue no. A25742) on a QuantStudio-3 real-time PCR system with *TFF1* promoter primers (forward 5′-CTAGACGGAATGGGCTTCAT-3′, reverse 5′-TTGGCCGTGACAACAGTG-3′). calculations are based on mean ± s.e.m. from four technical repeats.

Cell growth

MCF7 cells were transfected with vector alone, WT or mutant ER expression plasmids. After transfection, 2×10^3 MCF7 cells per well were seeded in 96-well plates. Following 12 h allowed for attachment, 10 µl of Cell Counting Kit-8 (CCK-8) reagent (Dojindo, catalogue no. CK04-05) was added and plates incubated for 2 h at 37 °C. Absorbance

at 450 nm was measured as baseline (day 0). Subsequently, media were changed to regular DMEM with 10% FBS, 0.1% ethanol (vehicle) or 1 µM 4-hydroxytamoxifen, and refreshed every 2 days. On day 5, CCK-8 reagent was added and absorbance at 450 nm determined. Calculations are based on mean ± s.e.m. from three technical repeats with three biological replicates.

Colony formation

At 24 h after transfection with 2 µg of the respective plasmids using Lipofectamine 3000 (Thermo Fisher, catalogue no. L3000001), MCF7 cells were trypsinized. Viable cells were counted with a haemocytometer, and 500 per well seeded into six-well plates. After 10 days, culture supernatant was removed and cells were washed with PBS, fixed with 4% paraformaldehyde for 15 min then washed again with PBS. Colonies were stained with 0.2% crystal violet solution (Millipore Sigma, catalogue no. C0775) for 15 min, followed by additional PBS washes to remove background staining. Stained colonies were quantified by digital image analysis using NIH ImageJ software, and defined as entities with more than 50 cells. Calculations are based on mean ± s.e.m. from two technical repeats with two biological replicates.

Cell line information

Cell lines MCF7 (ATCC HTB-22), T47D (ATCC HTB-133) and HEK293T (ATCC CRL-3216) were obtained directly from the American Type Culture Collection (ATCC). MCF7 and T47D are human breast adenocarcinoma lines, and HEK293T is a human embryonic kidney epithelial line. Cell lines were authenticated by ATCC and validated in-house by comparison of morphology and growth characteristics against supplier specifications. All lines were regularly tested for mycoplasma contamination using a mycoplasma PCR detection kit (MP Biomedicals, catalogue no. 3050301).

Reporting summary

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

Data availability

NMR assignments are available in the Biological Magnetic Resonance Data Bank under accession code nos. 51972 (WT), 51976 (S118D) and 51977 (pS118). SEC-SAXS data are deposited in the Small Angle Scattering Biological Databank under accession code nos. SASDSK2 (WT), SASDSF5 (S118D) and SASDVC5 (pS118). All additional data are present in the main text or Supplementary Information. Expression plasmids can be obtained from S.Y. under a material transfer agreement with Case Western Reserve University.

57. Palmer, I. & Wingfield, P. T. Preparation and extraction of insoluble (inclusion-body) proteins from *Escherichia coli*. *Curr. Protoc. Protein Sci.* **Chapter 6**, 6 3 1–6 3 20 (2012).
58. Sattler, M., Schleucher, J. & Griesinger, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 93–158 (1999).
59. Palmer, A. G. III, Cavanagh, J., Wright, P. E. & Rance, M. Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *J. Magn. Reson. (1969)* **93**, 151–170 (1991).
60. Kay, L., Keifer, P. & Saarinen, T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663–10665 (1992).
61. Grzesiek, S. & Bax, A. Amino acid type determination in the sequential assignment procedure of uniformly 13C/15 N-enriched proteins. *J. Biomol. NMR* **3**, 185–204 (1993).
62. Hyberts, S. G., Takeuchi, K. & Wagner, G. Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data. *J. Am. Chem. Soc.* **132**, 2145–2147 (2010).
63. Bermel, W., Bertini, I., Felli, I. C. & Pierattelli, R. Speeding up 13C direct detection biomolecular NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 15339–15345 (2009).
64. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
65. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).
66. Sjodt, M. & Clubb, R. T. Nitroxide labeling of proteins and the determination of paramagnetic relaxation derived distance restraints for NMR studies. *Bio Protoc.* **7**, e2207 (2017).

Article

67. Clore, G. M. Practical aspects of paramagnetic relaxation enhancement in biological macromolecules. *Methods Enzymol.* **564**, 485–497 (2015).
68. Lakomek, N. A., Ying, J. & Bax, A. Measurement of (1)(5)N relaxation rates in perdeuterated proteins by TROSY-based methods. *J. Biomol. NMR* **53**, 209–221 (2012).
69. Gairi, M. et al. An optimized method for (15)N R(1) relaxation rate measurements in non-deuterated proteins. *J. Biomol. NMR* **62**, 209–220 (2015).
70. Mulder, F. A., Skrynnikov, N. R., Hon, B., Dahlquist, F. W. & Kay, L. E. Measurement of slow (micros-ms) time scale dynamics in protein side chains by (15)N relaxation dispersion NMR spectroscopy: application to Asn and Gln residues in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc.* **123**, 967–975 (2001).
71. Palmer, A. G. III & Massi, F. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. *Chem. Rev.* **106**, 1700–1719 (2006).
72. Yang, S. Methods for SAXS-based structure determination of biomolecular complexes. *Adv. Mater.* **26**, 7902–7910 (2014).
73. Huang, W. et al. Multidomain architecture of estrogen receptor reveals interfacial cross-talk between its DNA-binding and ligand-binding domains. *Nat. Commun.* **9**, 3520 (2018).
74. Srinivasan, S. et al. Ligand-binding dynamics rewire cellular signaling via estrogen receptor-alpha. *Nat. Chem. Biol.* **9**, 326–332 (2013).
75. Weikum, E. R., Liu, X. & Ortlund, E. A. The nuclear receptor superfamily: a structural perspective. *Protein Sci.* **27**, 1876–1892 (2018).
76. Tesei, G., Schulze, T. K., Crehuet, R. & Lindorff-Larsen, K. Accurate model of liquid-liquid phase behavior of intrinsically disordered proteins from optimization of single-chain properties. *Proc. Natl Acad. Sci. USA* **118**, e2111696118 (2021).
77. Tesei, G. & Lindorff-Larsen, K. Improved predictions of phase behaviour of intrinsically disordered proteins by tuning the interaction range. *Open Res. Eur.* **2**, 94 (2023).
78. Zheng, W. et al. Inferring properties of disordered chains from FRET transfer efficiencies. *J. Chem. Phys.* **148**, 123329 (2018).

Extended Data Fig. 1 | SEC-SAXS analysis of ER-NTD disorder. (**a**) ER modular domains: NTD, DBD (C domain), hinge region (D domain), and LBD (E domain). (**b**) SEC-SAXS data acquisition schematic[72]: ER-NTD sample flows through a chromatography column coupled to SAXS measurements. (**c**) SEC-SAXS data of ER-NTD at 4 °C using pre-cooled buffer. $I(q)$: scattering intensity with error bars, normalized by zero-angle scattering; q: scattering vector amplitude. (**d**) Kratky plot comparing ER-NTD SEC-SAXS data (blue) to folded DBD-hinge-LBD fragment (CDE, gray; ref. 73). v: Flory exponent; R_g: radius of gyration. (**e**) R_g values for ER-NTD variants (WT, S118D, S118A, pS118, E56Q-pS118) from SEC-SAXS at 4 °C. Dashed line: empirical molecular form factor method fitting[22].
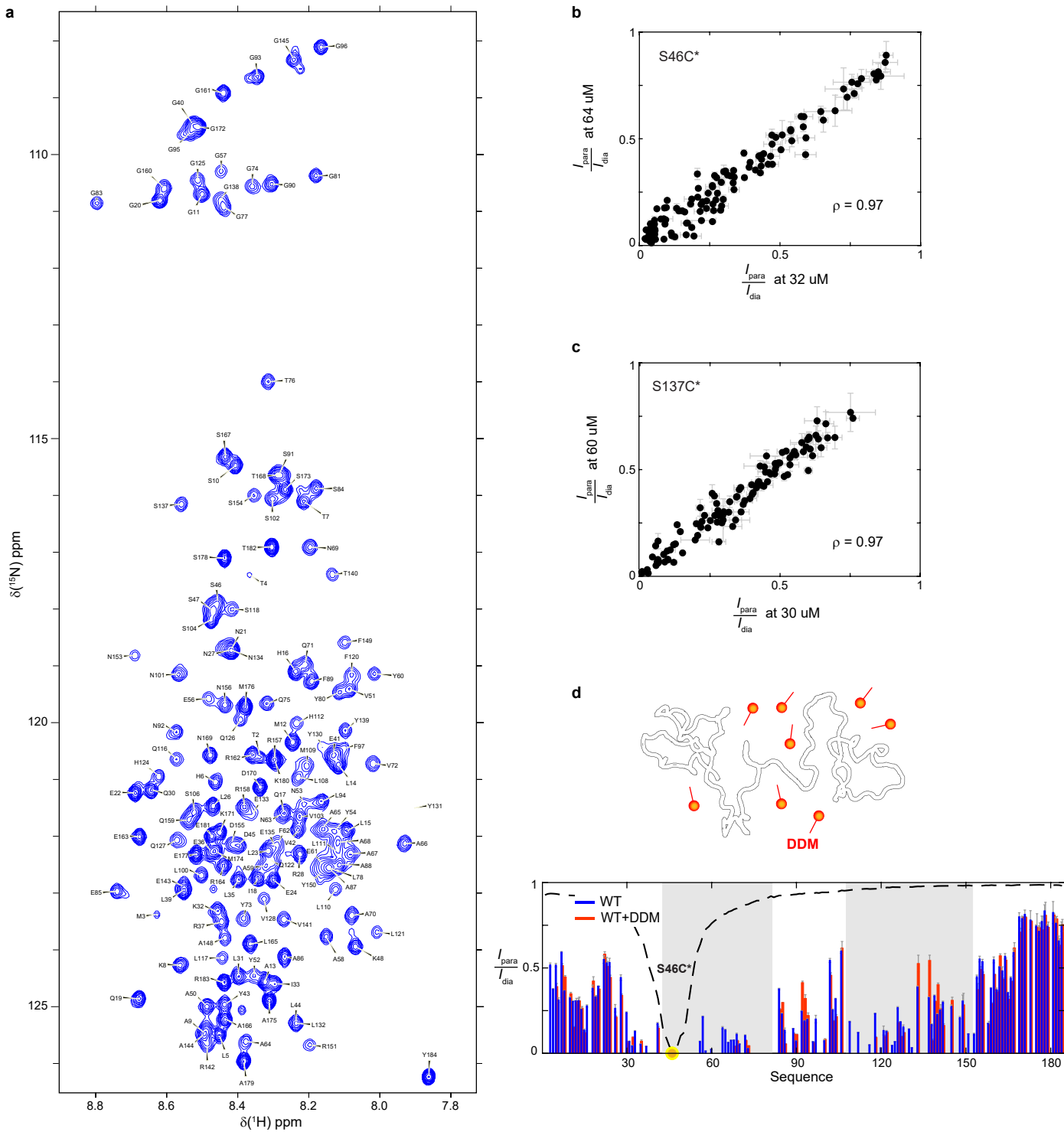
Extended Data Fig. 2 | ER-NTD resonance assignments and PRE analysis.
(**a**) ^{1}H-^{15}N HSQC spectrum (850 MHz) showing backbone amide NH chemical
shift assignments (BMRB: 51972). (**b,c**) PRE intensity ratios (I_{para}/I_{dia}) comparing
at two protein concentrations: S46C (32 μM vs 64 μM) and S137C (30 μM vs
60 μM) sites. High Pearson correlation coefficients (ρ) indicate consistent
ratios across concentrations, demonstrating negligible intermolecular
interactions. (**d**) Non-ionic detergent DDM (120 μM) effects on long-range
interactions probed by S46C spin labeling. Protein concentration: 30 μM;
yellow circle: spin label position.
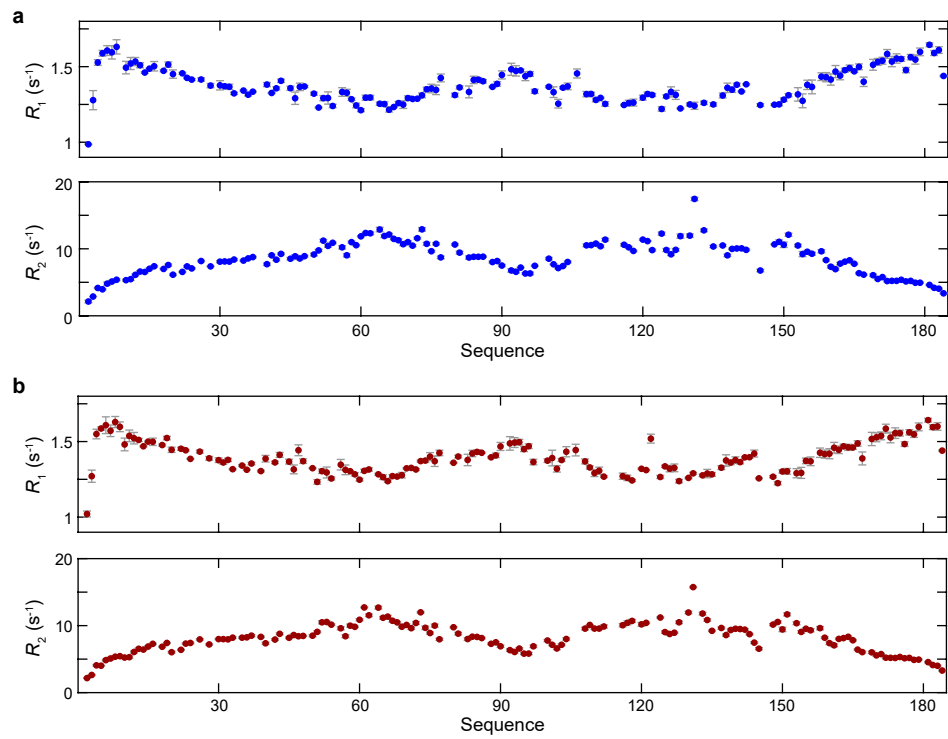
Extended Data Fig. 3 | Backbone dynamics analysis by ^{15}N relaxation. (**a**) Longitudinal (R_1) and transverse (R_2) relaxation rates of wildtype ER-NTD as a function of residue number (850 MHz). (**b**) ^{15}N relaxation profiles (R_1, R_2) of pS118-MTD (850 MHz).

Article
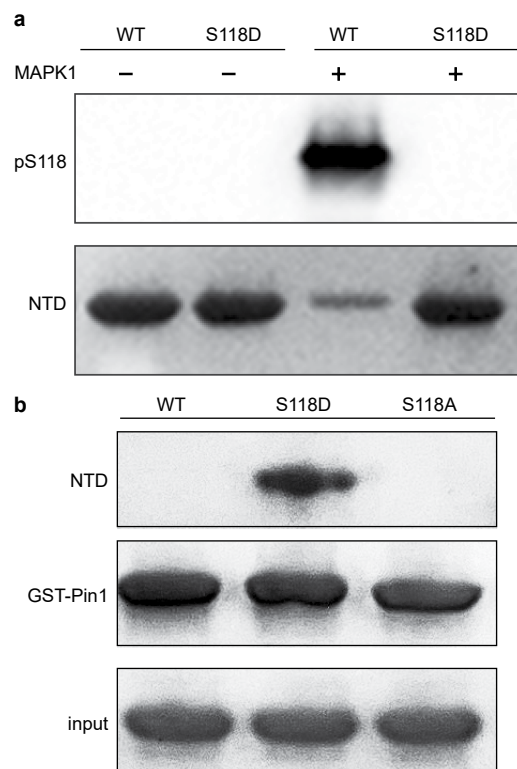
Extended Data Fig. 4 | Analysis of ER-NTD Ser118 phosphorylation and phosphomimetic mutation. (**a**) Western blot analysis of purified ER-NTD (WT or S118D) incubated with activated MAPK1/ERK2 kinase, using phospho-Ser118-specific antibody. Total proteins visualized by Ponceau S staining. For gel source data, see Supplementary Fig. 1. (**b**) GST-Pin1 pulldown assays with purified WT, S118A, or S118D proteins. Pin1, which selectively binds phosphorylated Ser118[19], shows robust binding to S118D but not to unphosphorylated or S118A proteins, validating S118D as a phosphomimetic mutation. For gel source data, see Supplementary Fig. 1.
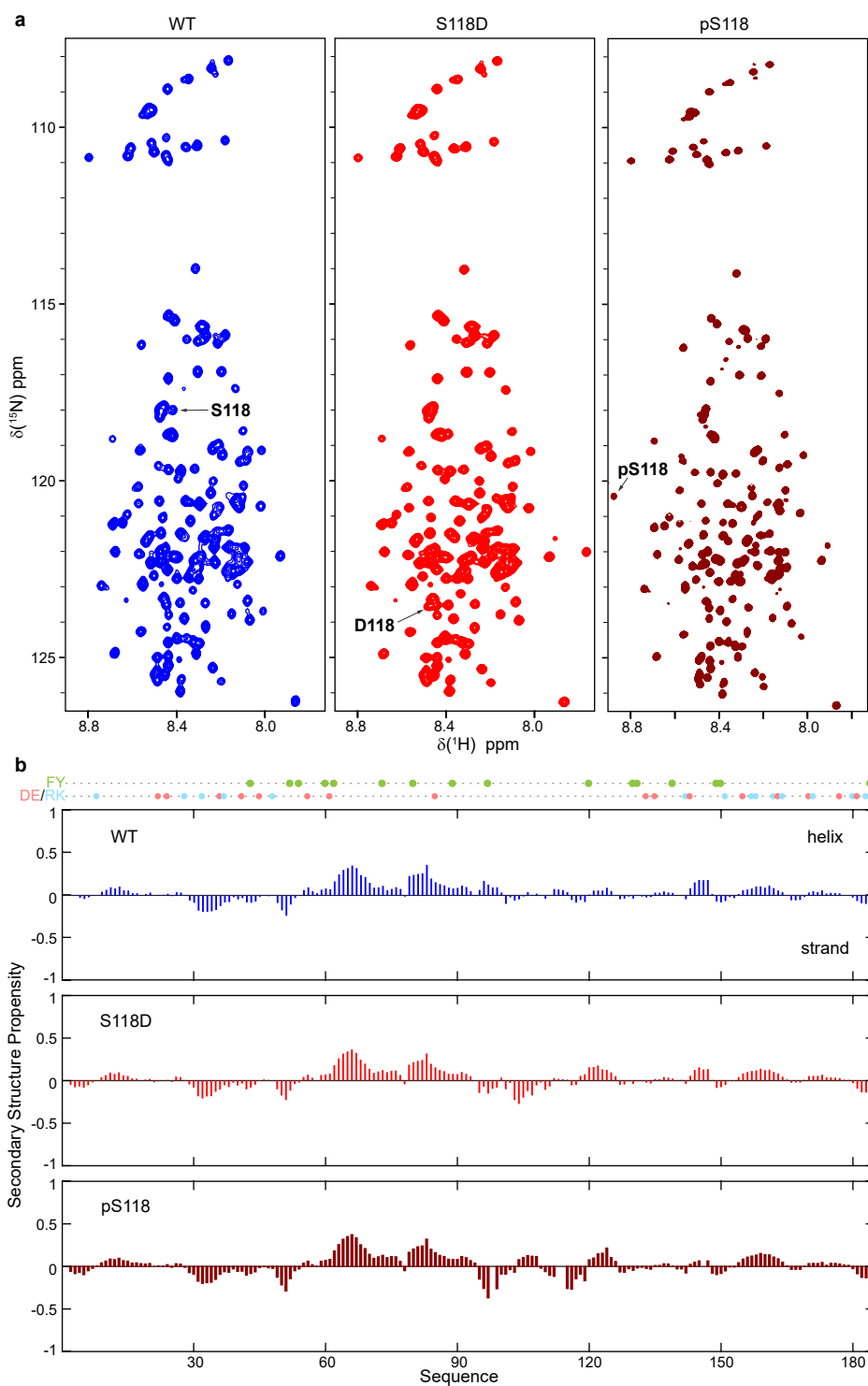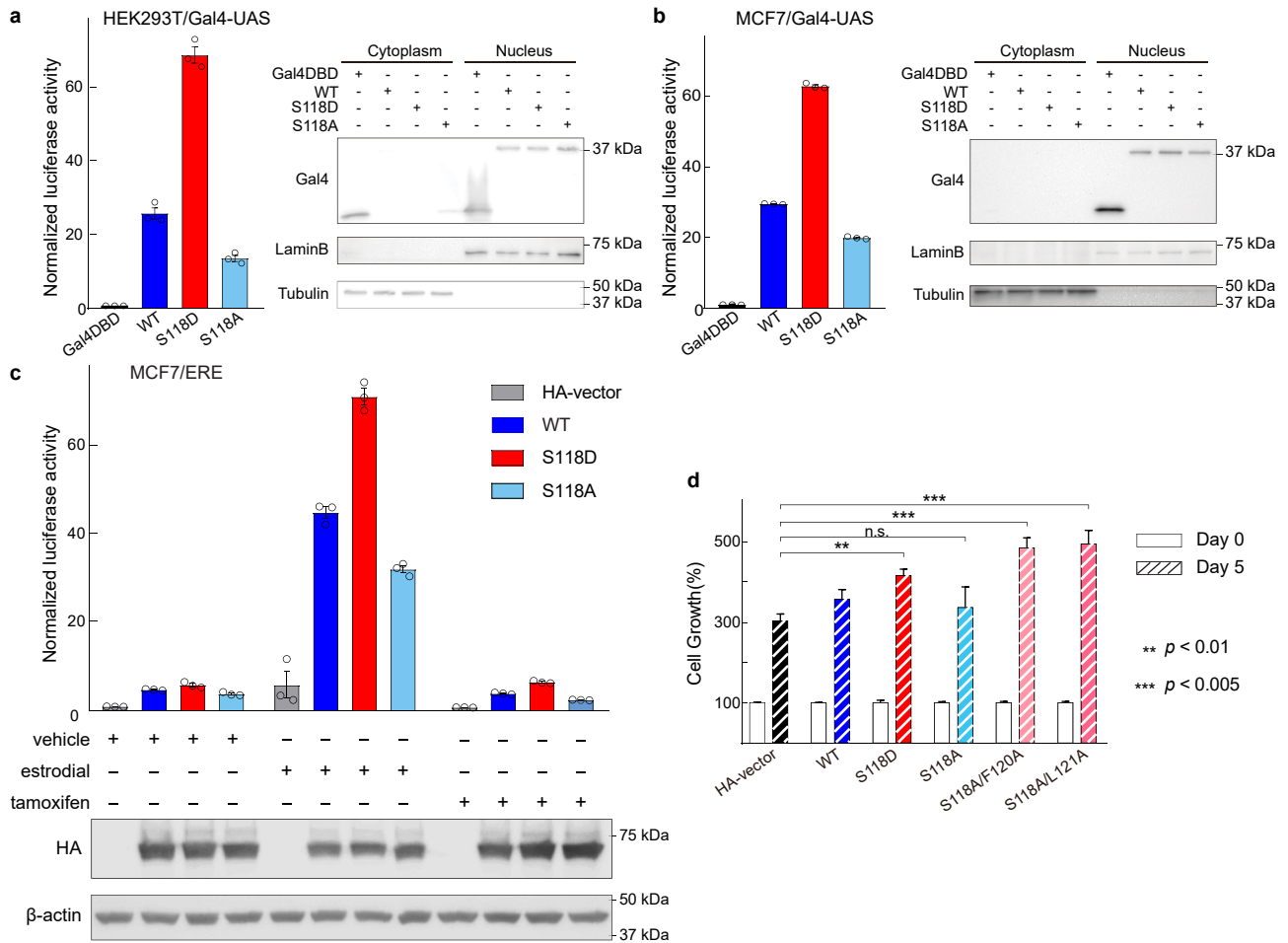
Extended Data Fig. 5 | NMR analysis of WT, S118D, and pS118 ER-NTD.
(**a**) ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled WT, S118D, and pS118 proteins. BMRB accession codes: 51972 (WT), 51976 (S118D), 51977 (pS118). (**b**) Secondary structure propensity (SSP) calculated from ^{1}H, ^{15}N, and ^{13}C (C$^\alpha$, C$^\beta$, and C′) chemical shifts[54]. Low SSP scores indicate absence of persistent secondary structure.
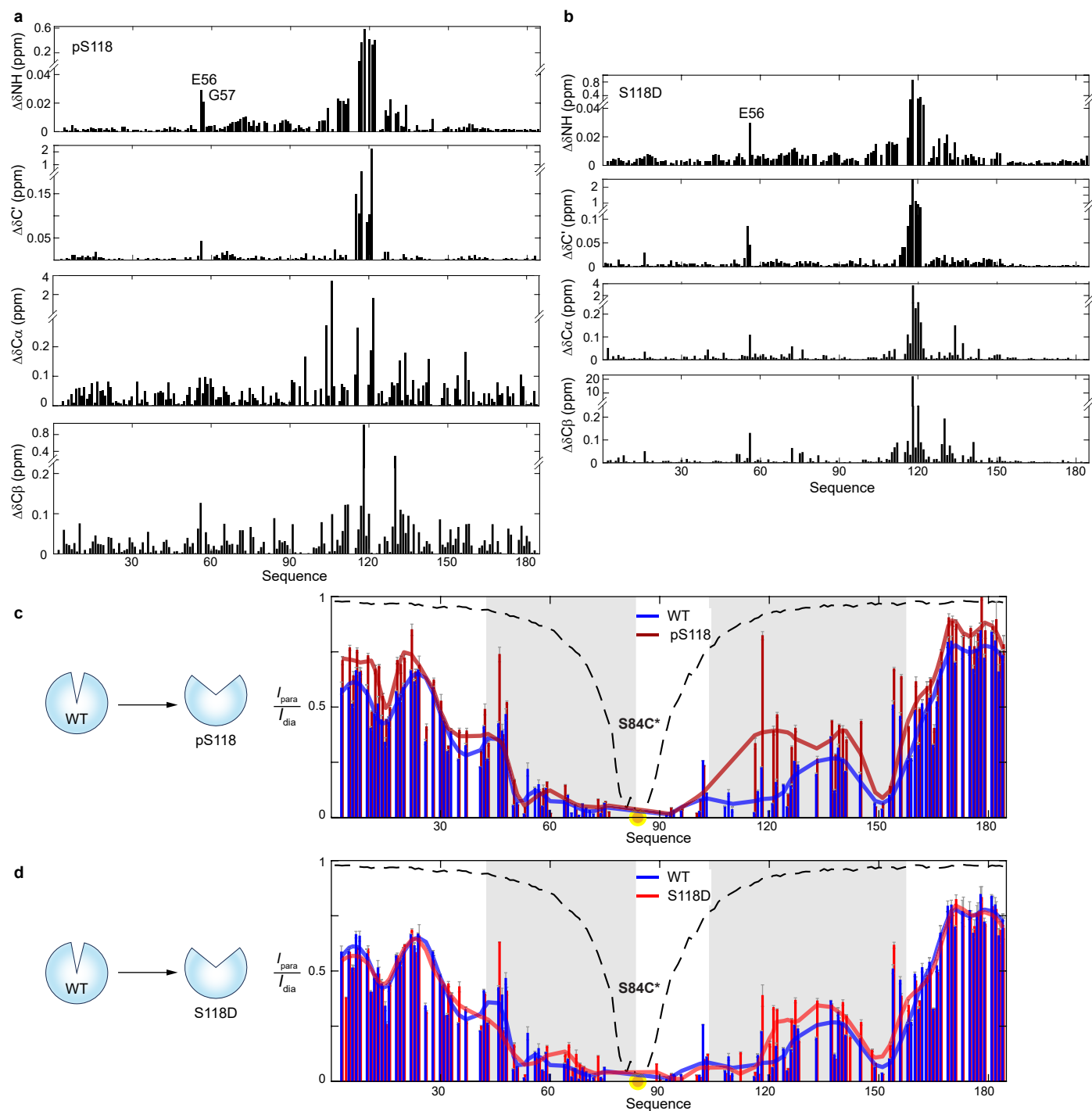
Extended Data Fig. 6 | S118 modification effects on ER-NTD long-range interactions. (a,b) Chemical shift changes induced by pS118 and S118D relative to WT, calculated using ^{1}H, ^{15}N, ^{13}C (C^{α}, C^{β}, and C') chemical shifts. $\Delta H/N = \sqrt{\Delta \delta H^2 + (\Delta \delta N \times 0.154)^2}$, where $\Delta \delta H$ and $\Delta \delta N$ represent backbone amide ^{1}H and ^{15}N chemical shift changes. **(c,d)** Long-range interaction changes probed by S84C spin labeling: PRE profiles of pS118 (dark red) and S118D (red) compared to WT (blue). Protein concentration: 30 μM; yellow circle: spin label position.

Extended Data Fig. 7 | Functional analysis of S118D and hydrophobic mutations. (**a, b**) Transcription activity of Gal4-NTD variants (WT, S118D, or S118A) in HEK293T cells and MCF7 cells with western blot validation. Mean ± SEM from four technical repeats performed across three biological replicates. For gel source data, see Supplementary Fig. 2. (**c**) Full-length ER variant (WT, S118D, S118A) transcriptional activity in MCF7 cells treated with vehicle, 100 nM E2, or 1 μM 4OH-tamoxifen (12 h). Western blot confirms protein expression. For gel source data, see Supplementary Fig. 2. (**d**) Growth rescue by F120A and L121A mutations in S118A MCF7 cells treated with 1 μM 4OH-tamoxifen (5 days). Cell numbers quantified by CCK-8 assay (450 nm absorbance). Mean ± SEM from four technical repeats performed across three biological replicates. One-way ANOVA with Dunnett's test: $*p < 0.05$; $**p < 0.01$; $***p < 0.005$.
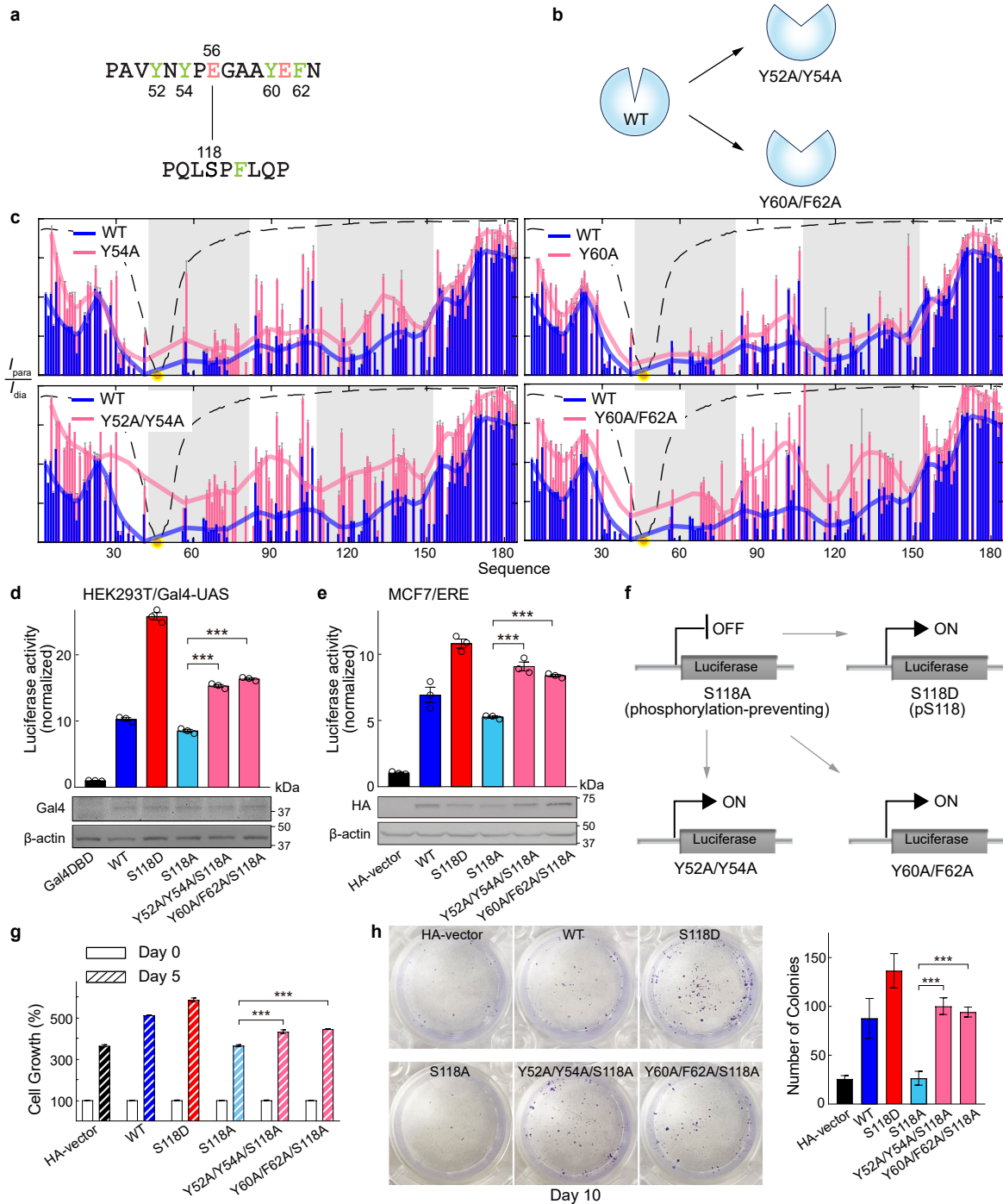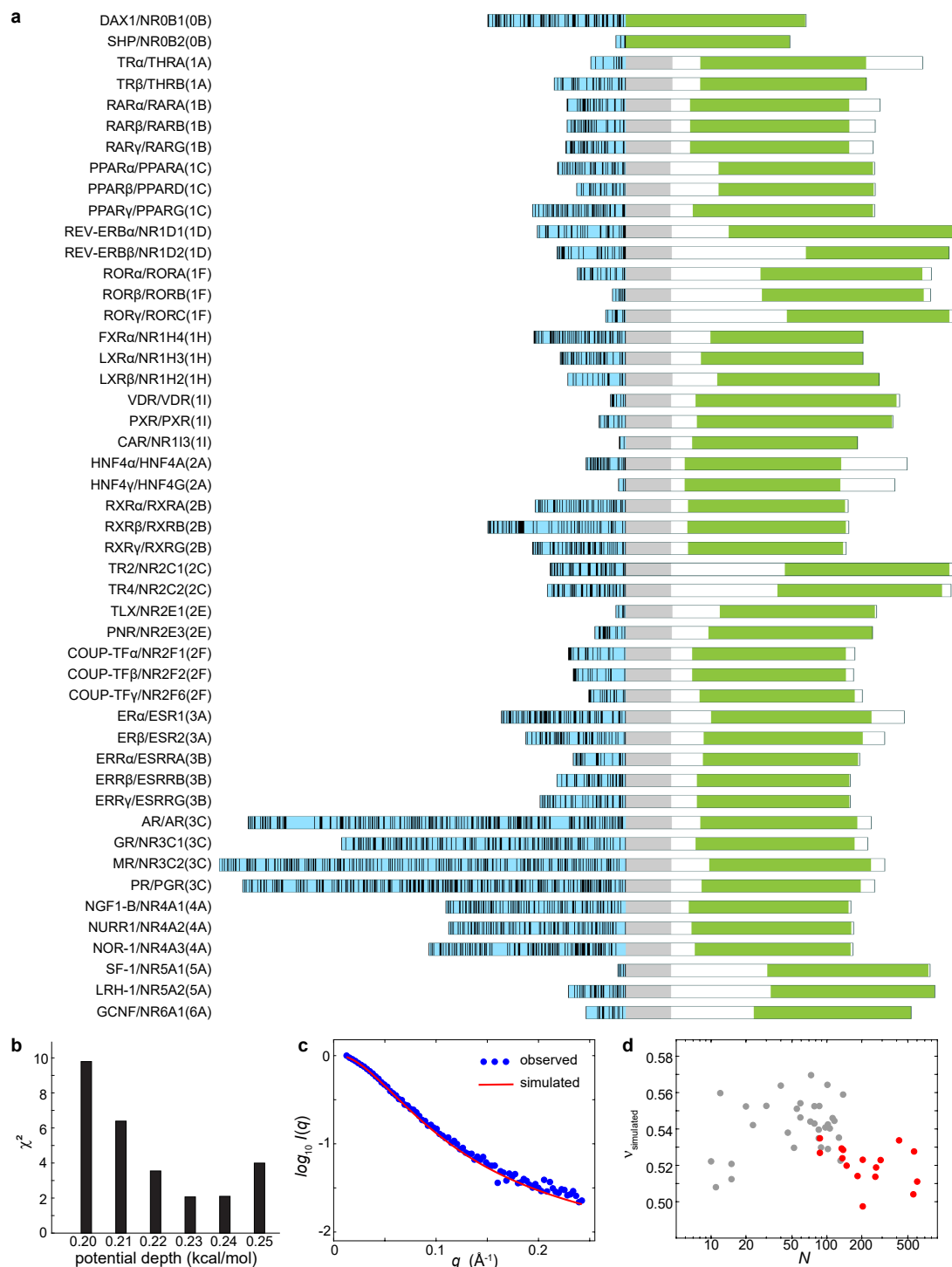
Extended Data Fig. 8 | Hydrophobic aromatic residue mutations modulate ER-NTD function. (**a**) Structural proximity of hydrophobic aromatic residues (Y52, Y54, Y60, F62; green) to E56 (red) and serine 118. (**b,c**) PRE analysis (S46C spin labeling) shows Y52A/Y54A and Y60A/F62A double mutations induce greater conformational changes than single mutations in non-phosphorylated NTD. Yellow circle: spin label position; solid lines: visual aid. HSQC: 850 MHz, 4 °C. (**d,e**) Y52A/Y54A and Y60A/F62A rescue S118A-impaired reporter activity in ER-NTD (HEK293T cells) and full-length ER (MCF7 cells). Western blots confirm protein expression. Mean ± SEM from four technical repeats performed across three biological replicates. For gel source data, see Supplementary Fig. 3. (**f**) Schematic: Y52A/Y54A or Y60A/F62A rescue transcription activity of phosphorylation-deficient S118A mutant. (**g,h**) Y52A/Y54A and Y60A/F62A restore S118A-impaired cell growth (Mean ± SEM from three technical repeats across three biological replicates) and colony formation in MCF7 cells (Mean ± SEM from two technical repeats across two biological replicates; DMEM media, crystal violet quantification at 450 nm). One-way ANOVA with Dunnett's test: *$p < 0.05$; **$p < 0.01$; ***$p < 0.005$.
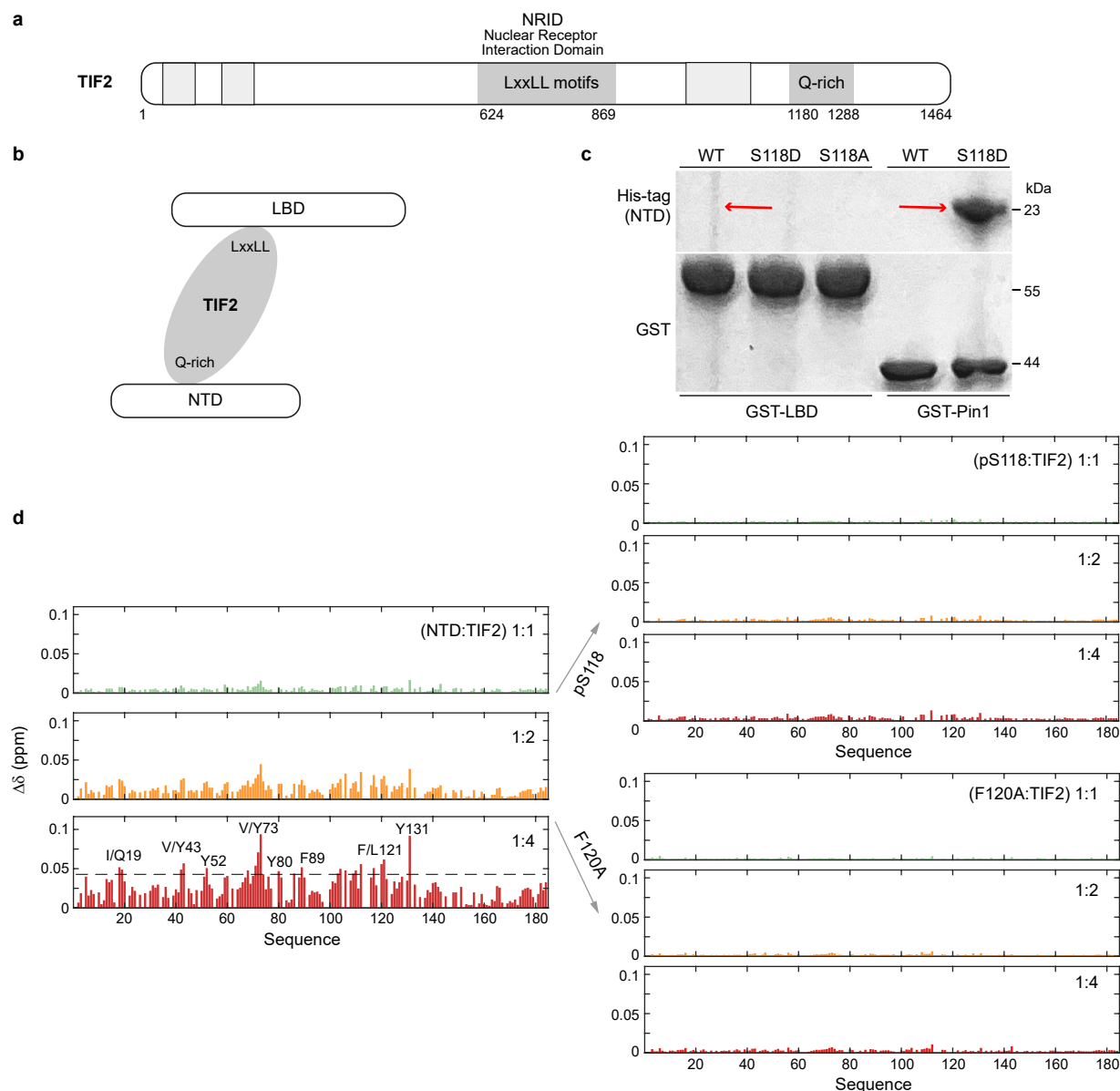
Extended Data Fig. 9 | Analysis of TIF2 interactions and NTD chemical shifts. (a) TIF2 domains showing NRID (via LxxLL motifs) and Q-rich domain (QRD)[32,33]. **(b)** Model of TIF2 binding to ER-NTD (via QRD) and ER-LBD (via NRID)[34]. **(c)** GST-LBD pulldowns with His-tagged NTD proteins (WT, S118D, S118A), analyzed by western blot with GST-Pin1 controls. For gel source data, see Supplementary Fig. 5. ER-LBD purified per Nettles protocol[74]. **(d)** Chemical shift changes in ER-NTD variants (WT, F120A, pS118) upon TIF2-QRD titration (1:0 to 1:4 molar ratios). $\Delta\delta = \sqrt{(\Delta\delta H^2 + (\Delta\delta N \times 0.154)^2}$, where $\Delta\delta H$ and $\Delta\delta N$ are proton and nitrogen changes relative to unbound protein. Full spectra in Supplementary Fig. 6.

Extended Data Fig. 10 | Analysis of nuclear receptor NTD diversity.
(**a**) Domain organizations of 48 human nuclear receptors[75] showing NTD (blue), DBD (gray), and LBD (green), with marked hydrophobic residue positions in NTDs. (**b**) CALVADOS2[76,77] model optimization for ER-NTD using SEC-SAXS data. χ^2 (root-mean-square deviation between simulated and experimental SAXS intensities) vs. ε (pairwise interaction potential depth). Optimal $\varepsilon = 0.23$ kcal/mol at 310 K, 150 mM ionic strength, pH 7.4, 1 µs total (Supplementary Methods).

(**c**) Comparison of simulated and experimental SAXS intensities for ER-NTD (CALVADOS2, $\varepsilon = 0.23$ kcal/mol). (**d**) Analysis of NTDs (excluding two DBD-lacking sequences): residue count (N) vs. simulated Flory exponent ($\nu_{\text{simulated}}$) from adjusted CALVADOS2 model ($\varepsilon = 0.23$ kcal/mol) using R_g-N power-law relation[78]. Red: NTDs with FCR < 0.22 and SHP > 0.6 (ER-NTD-like); gray: others. FCR and SHP values in Supplementary Table 1.

Corresponding author(s): Sichun Yang

Last updated by author(s): Oct 31, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	NMR data: BRUKER Topspin 3.6 (acquisition), NMRPipe 11.0 (processing) SEC-SAXS data: Pilatus3-X-1M, Pilatus3 S 1M, and Pilatus3 900K detectors Western blot imaging: ChemiDoc MP system Real-time PCR: QuantStudio-3 and CFX Connect Real-Time PCR Detection System Reporter assay: Varioskan LUX Multimode Microplate Reader
Data analysis	NMR data were analyzed using NMRfam-SPARKY version 1.470 in conjunction with Sparky 3.190; SAXS data were processed and analyzed using the ATSAS software suite; Statistical analyses of in-cell data were performed using GraphPad Prism version 8.0.1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

All manuscripts must include a <u>data availability statement</u>. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our <u>policy</u>

NMR assignments are available in the Biomolecular NMR database (BMRB) under accession codes 51972 (wild-type), 51976 (S118D), and 51977 (pS118).

SEC-SAXS data for the wild-type and mutants are available in the Small Angle Scattering Biological Databank (SASBDB) under accession codes SASDSK2 (wild-type), SASDSF5 (S118D), SASDSE5 (S118A), and SASDVC5 (pS118).

The sequence analysis databases used in this study, including (a) 48 nuclear receptor N-terminal domains (NTDs), (b) 2,360 representative folded proteins, and (c) 28,058 intrinsically disordered proteins/regions, have been made publicly available in a GitHub repository at https://github.com/wzhenglab/ERalphaNTD-2024.

Research involving human participants, their data, or biological material

Policy information about studies with <u>human participants or human data</u>. See also policy information about <u>sex, gender (identity/presentation), and sexual orientation</u> and <u>race, ethnicity and racism</u>.

Reporting on sex and gender	NA
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	NA
Recruitment	NA
Ethics oversight	NA

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For transient transfection reporter assays, a sample size of approximately 2,000,000 cells per well was used; Cell proliferation studies were conducted by seeding 2,000 cells per well in 96-well plates; Colony formation assays were performed by seeding 500 cells per well into 6-well plates.
Data exclusions	No data were excluded from the analyses.
Replication	Transient transfection reporter assays were performed with three biological replicates, each consisting of four technical repeats; RNA extraction and quantitative RT-PCR experiments were conducted with three biological replicates, each comprising three technical repeats; CUT&RUN assays, followed by qPCR, were carried out with four technical repeats; Colony formation assays were performed with two biological replicates, each consisting of two technical repeats.
Randomization	For cell-based experiments, samples were randomly allocated into different experimental groups. This was achieved by maintaining cell cultures under identical conditions, passaging cells simultaneously, and harvesting and pooling them to create a single uniform cell suspension. The suspension was then thoroughly mixed to ensure an even distribution of cells across the experimental groups.
Blinding	The investigators were blinded to group allocation during data collection and analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	ER (Abcam, #ab16660), phospho-ER-Ser118 (Cell Signaling, #2511), β-actin (Cell Signaling, #4967), Lamin B1 (Santa Cruz, #sc-374015), β-tubulin (Santa Cruz, #sc-166729), Gal4 (Santa Cruz, #sc-510), HA-tag (Santa Cruz, #sc-7392), His-tag (Millipore Sigma, #SAB4301134), GST-tag (Millipore Sigma, #SAB4301139), anti-IgG antibodies (Bio-Rad, #1706515 or #1706516), HA-Tag Rabbit mAb (Cell Signaling, #3724), TIF2 (Cell Signaling, #96687)
Validation	All antibodies used in this study were validated by their respective manufacturers for specificity, sensitivity, and reproducibility. For Western blot experiments, antibodies were used following the manufacturers' recommended conditions, such as dilution ratios, blocking buffers, and incubation times, as detailed in their product datasheets. To further validate the specificity of the anti-ER antibody (Abcam, Cat# ab16660), its performance was tested using an ER knockdown cell line.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	Human breast adenocarcinoma cell lines MCF7 (ATCC HTB-22) and T47D (ATCC HTB-133), as well as human embryonic kidney epithelial cell line HEK293T (ATCC CRL-3216), were obtained directly from the American Type Culture Collection (ATCC).
Authentication	All cell lines were authenticated by the supplier (ATCC). Upon receipt, cell line identity was further validated by assessing morphology and growth characteristics, which were compared to the expected features provided by the supplier.
Mycoplasma contamination	All cell lines were routinely tested for mycoplasma contamination using the Myco-Sniff-Valid™ Mycoplasma PCR Detection Kit (MP Biomedicals, Cat# 3050301).
Commonly misidentified lines (See ICLAC register)	The cell lines used in this study (MCF7, T47D, and HEK293T) are not listed in the International Cell Line Authentication Committee (ICLAC) database of commonly misidentified cell lines.

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA