Cardiac stress leads to regulation of Filamin C dimerisation via an ancient phosphorylation-modulated interaction with HSPB7

Zihao Wang[1,2], Guodong Cao[1,2,†], Miranda P. Collier[1,2,†], Xingyu Qiu[1,2,†], Sophie Broadway-Stringer[3], Dominik Šaman[1,2], Jediael Z.Y. Ng[4], Navoneel Sen[1,2], Amar J. Azad[3], Charlotte Hooper[3], Johannes Zimmermann[5], Michael McDonough[6], Jürgen Brem[6], Patrick Rabe[6], Haigang Song[1,2], T. Reid Alderson[1,2], Christopher J. Schofield[6], Jani R. Bolla[7], Kristina Djinovic-Carugo[8,9], Dieter O. Fürst[10], Bettina Warscheid[5], Matteo T. Degiacomi[11], Timothy M. Allison[12], Georg K.A. Hochberg[4,13,14], Carol V. Robinson[1,2], Katja Gehmlich[3,15,*], Justin L.P. Benesch[1,2,*]

[1] *Department of Chemistry, Dorothy Crowfoot Hodgkin Building, University of Oxford, Oxford, UK*

[2] *Kavli Institute for Nanoscience Discovery, University of Oxford, Oxford, UK*

[3] *Institute of Cardiovascular Sciences, University of Birmingham, Edgbaston, Birmingham, UK*

[4] *Evolutionary Biochemistry Group, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch Straße 10, 35043 Marburg*

[5] *Biochemistry II, Theodor Boveri-Institute, Biocenter, Chemistry and Pharmacy, University of Würzburg, 97074 Würzburg, Germany*

[6] *Department of Chemistry, Chemistry Research Laboratory, Oxford, UK*

[7] *Department of Biology, University of Oxford, Oxford, UK*

[8] *European Molecular Biology Laboratory, 38000 Grenoble, France.*

[9] *Department of Structural and Computational Biology, Max Perutz Labs, University of Vienna, 1030 Vienna, Austria.*

[10] *Institute for Cell Biology, University of Bonn, Bonn, Germany*

[12] *Department of Physics, Durham University, Durham, UK*

[13] *Biomolecular Interaction Centre and School of Physical and Chemical Sciences, University of Canterbury, Christchurch, NZ*

[13] *Department of Chemistry, Philipps-University Marburg; Hans-Meerwein-Str. 4, 35043 Marburg, Germany*

[14] *Center for Synthetic Microbiology, Philipps-University Marburg, Karl-von-Frisch Straße 4, 35043 Marburg*

[15] *Division of Cardiovascular Medicine, Radcliffe Department of Medicine and British Heart Foundation Centre of Research Excellence Oxford, University of Oxford, Oxford, UK*

[†] These authors contributed equally.

[*] Correspondence : k.gehmlich@bham.ac.uk and justin.benesch@chem.ox.ac.uk

Abstract

The biomechanical properties and responses of tissues underpin a variety of physiological functions and pathologies. In striated muscle, the actin-binding protein filamin C (FLNC) is a key protein whose variants causative for a wide range of cardiomyopathies and musculoskeletal pathologies. Seemingly a multi-functional protein that interacts with a variety of partners, how FLNC is regulated at the molecular level is not well understood. Here we have investigated its interaction with HSPB7, a cardiac-specific molecular chaperone whose absence is embryonically lethal. We found that FLNC and HSPB7 interact in cardiac tissue under biomechanical stress, forming a strong hetero-dimer whose structure we have solved by means of X-ray crystallography. Our quantitative analyses show that the hetero-dimer out-competes the FLNC homo-dimer interface, potentially acting to abrogate the ability of the protein to cross-link the actin cytoskeleton, and to enhance its diffusive mobility. We show that phosphorylation of FLNC at threonine 2677, located at the dimer interface and associated with cardiac stress, acts to favour the homo-dimer. Conversely, phosphorylation at tyrosine 2683, also at the dimer interface, has the opposite effect and shifts the equilibrium towards the hetero-dimer. Evolutionary analysis and ancestral sequence reconstruction reveals this interaction and its mechanisms of regulation to date around the time primitive hearts evolved in chordates. Our work rationalises on the molecular level how FLNC might switch between stabilising functions in the cell, and reveals how HSPB7 acts as a specific molecular chaperone that regulates FLNC.

Introduction

The beating of the heart applies continual cycles of biomechanical force to the contractile machinery and cytoskeletal support proteins that comprise cardiac muscle tissue. The molecular transitions that these proteins undergo, during both physiological and stress conditions, encode a variety of sensing and response behaviors, so-called mechano-signalling. The breakdown of, or aberrant, mechano-signalling pathways can lead to serious cardiac dysfunctions, including cardiomyopathies and heart failure[1].

A key protein within the cardiac cytoskeleton that is heavily involved in mechano-signalling is filamin C (FLNC; also known as γ-FLN, ABP-L, or FLN2). FLNC is found primarily in striated muscle cells, while the other two filamin paralogs, FLNB and FLNA, are expressed in all other cell types[2]. FLNC has important cellular roles in signalling pathways, cell migration, differentiation and the assembly of the myofibrillar Z-discs[3,4], and contributes to rapid stabilisation and repair of myofibrillar damage[5]. Genetic variants in *FLNC* are associated with cardiomyopathy[6] and musculoskeletal myopathies[7], and its aberrant expression is related to cancer cell invasiveness[8,9].

FLNC's diverse roles apparently stem from it being both an actin-binding protein, and a platform for binding a plethora of proteins ranging from membrane proteins to cytoskeletal network components, allowing it to integrate cellular architecture and signalling pathways[3]. The actin-binding domain of FLNC is located at the N-terminal end, and is followed by 24 immunoglobulin-like (Ig) domains (d1-d24). While d1-15 is thought to form a relatively elongated structure at rest, d16-24 appears to comprise a an extendable mechanosensitive region that can bind to a variety of interaction partners[3]. The C-terminal domain, d24, is the site of homo-dimerisation of FLNC[10]. The dimer formation of

FLNC is necessary to cross-link actin, and genetic variants in d24 that hinder dimerisation are associated with myofibrillar myopathy[11–13]. Numerous sites of phosphorylation have been identified on FLNC, in particular in d20, the hinge 2 region between d23 and d24, and d24. Those sites whose molecular consequences have been characterised appear to regulate its stability and interaction with different binding partners under stress conditions[14–16]. This nuanced link between phosphorylation events and the mechanical responses of FLNC adds a layer of regulatable adaptability to its role in cardiac mechano-signalling.

FLNC has been shown to interact with HSPB7, also known as cardiovascular heat shock protein (cvHSP), via d24[17–19]. HSPB7 is a member of the small Heat-Shock Protein (sHSP) family, and is expressed specifically in heart and skeletal muscle[19,20]. The expression level of HSPB7 increases with age, and it is located in cardiomyocytes, either diffusely or at the Z-disc and the intercalated discs[18,21,22]. HSPB7 appears to play a crucial role in regulating the length of thin filaments during heart development and preventing the aggregation of monomeric actin[22]. Knockout, or missense and truncating variants, of HSPB7 leads to the formation of large aggregates of FLNC and a significant increase in actin filament bundles, which results in embryonically lethal cardiac defects in mice[17,18,22,23].

While less studied than most human sHSPs, HSPB7 appears to have features uncharacteristic of this family of molecular chaperones. Though it shares the common architecture of sHSPs (an α-crystallin domain (ACD) flanked by N- and C-terminal regions[24]), HSPB7 is unusual as it has been observed to be oligomerisation incompetent[25,26]. HSPB7 also does not appear to display the canonical sHSP function of broadly inhibiting heat-induced protein aggregation[26,27]. These observations suggest that HSPB7 is not a promiscuous oligomeric chaperone, as in the archetypal sHSP, but instead may be

specialised to a particular cellular role.

Here we have tested this hypothesis by examining the association between HSPB7 and FLNC from the tissue to atomic levels. We validated the interaction between the two proteins in cardiac tissue, and observed up-regulation and co-localisation of the two proteins in mouse models of biomechanical stress. Using a combination of structural approaches, we localised the binding between the two proteins to arise from a strong interaction between FLNCd24 and the HSPB7 ACD, and solved the structure of the resulting hetero-dimer. We found that the equilibrium between FLNC homo-dimer and FLNC-HSPB7 hetero-dimer is modulated by phosphorylation of FLNC, and identified key amino-acid contacts that mediate this competitive binding. Our study reveals an ancient and uniquely specific sHSP target interaction which appears to be a key regulatory mechanism for FLNC in its role of maintaining cytoskeletal integrity.

Results

Biomechanical stress leads to up-regulation and interaction of FLNC and HSPB7 in mouse hearts

To test the interaction between FLNC and HSPB7 *in vivo*, we used three mouse models of biomechanical stress and assayed the location and abundance of each protein. The first model is a muscle LIM protein (MLP) knock-out (KO) mouse. MLP is found exclusively in striated muscle, and its absence results in a distinctive tissue morphology that mirrors dilated cardiomyopathy and heart failure in humans[16]. Immunoprecipitation (IP) experiments using an anti-FLNC antibody followed by western blotting, resulted in the detection of both FLNC and HSPB7 in the wild-type (WT) and MLP KO cardiac tissue (**Fig. 1A**). This demonstrates that these proteins interact in the mouse heart.

Noticeably more HSPB7 was pulled down from the MLP KO tissue. To determine whether this was due to increased expression under cardiac impairment, we evaluated levels of HSPB7 and FLNC in WT and MLP KO hearts. Western blots revealed significantly higher levels of both proteins in the MLP KO mice (**Fig. 1B**, left). We performed similar blots on tissue from mice that had been subject to biomechanical stress either due to transverse aortic constriction (TAC) surgery (which leads to pressure overload and induces heart failure) or chronic isoprenaline/epinephrine (IsoPE) treatment (which increases blood pressure and heart rate)[16]. In each case, we observed higher levels of FLNC and HSPB7 compared to the respective controls (**Fig. 1B**).

Next, we immuno-fluorescently stained frozen ventricular sections from the three mouse models for both HSPB7 and FLNC. In each we found FLNC and HSPB7 to co-localise to varying degrees, and to do so primarily at the Z-discs and intercalated discs (**Fig. 1C**). Combined these results demonstrate that under biomechanical stress, FLNC and HSPB7 are up-regulated and interact with each other at

sites of traction in cardiac tissue.

HSPB7 exists as a monomer with limited anti-aggregation activity encoded by the N-terminal region

To investigate the nature of the interaction between FLNC and HSPB7, we turned to *in vitro* experiments on the purified proteins. We first expressed and purified the full-length HSPB7 (isoform 2, the longest of the three[19]) as well as two truncated constructs: one missing the N-terminal region ($HSPB7_{\Delta N}$, residues 78-175 of isoform 2) and one additionally missing the C-terminal region, i.e. just the core ACD ($HSPB7_{ACD}$, residues 78-162 of isoform 2) (**Extended Data Fig 1A**). The N- and C-terminal regions are predicted to be disordered, and not to make substantive contacts with the ACD (**Extended Data Fig 1B**). We obtained native mass spectrometry (MS) data for each construct at ~10 μM, and found that all three forms of HSPB7 were primarily monomeric (**Extended Data Fig 1C**). This contrasts with equivalent experiments on other human sHSPs, which at micromolar concentrations are typically oligomeric in their full-length form, while their ACDs form dimers[16,28–30].

HSPB7 has been shown to be largely ineffective in preventing the heat-induced aggregation of proteins, in contrast to the canonical chaperone activity of other sHSPs[26]. However, one study showed some suppression of aggregation of citrate synthase[26], so we used this target protein to perform a comparative assay of our three constructs. We found that full-length HSPB7 was able to suppress aggregation, albeit at high relative ratios (>1:1) of chaperone to target. However, neither $HSPB7_{\Delta N}$ nor $HSPB7_{ACD}$ displayed any anti-aggregation activity, even at ratios of 4:1 (**Extended Data Fig. 1D**). This demonstrates that the N-terminal region encodes any canonical chaperone activity that HSPB7 has, in line with previous data[25,31].

HSPB7 does not oligomerise due to the failure of its C-terminal region to bind the ACD

To identify the source of this unusual lack of oligomerisation by HSPB7, we examined the regions of its sequence that mediate assembly in others sHSPs. Typically, sHSPs oligomerise via an interaction mediated by an "IXI motif", where X refers to a variable amino acid, and the I refers typically (but not exclusively) to Ile[32]; in human sHSPs the motif is generally manifested as IPI/V. The IXI motif in the C-terminal region of one monomer facilitates assembly by binding into the ACD β4-β8 groove of a neighbouring monomer[29,32]. HSPB7 contains the sequence IKI at its extreme C-terminus. To test whether this can bind to the ACD, we incubated a peptide mimicking the last eight residues of the C-terminus (residues 168-175, TFRTEIKI) with $HSPB7_{ACD}$ and performed native MS experiments. We observed no significant interaction between the two (**Extended Data Fig. 2**). Control experiments under the same conditions for the ACDs of HSPB1 and HSPB5 (human sHSPs that oligomerise[32]) showed binding of their cognate peptides with ≤100 μM K_D, consistent with previous reports [28,29].

To test whether this absence of binding was due to the IXI motif or the $HSPB7_{ACD}$ β4-β8 groove, we performed experiments in which we incubated either $HSPB1_{ACD}$ or $HSPB5_{ACD}$ with the HSPB7 peptide, and $HSPB7_{ACD}$ with either the HSPB1 or HSPB5 peptide. We found that neither of the ACDs were able to bind the HSPB7 peptide to a significant extent, whereas HSPB7 was able to bind both HSPB1 and HSPB5 peptides (**Extended Data Fig. 2**). This indicates that the failure of HSPB7 to oligomerise is due to its C-terminal region being different to that of other human sHSPs, either because it does not contain the IXI motif in the IPI/V form found in the other human sHSPs, or due to the IXI motif being at the extreme end of the sequence and hence lacking flanking residues that provide additional stabilisation of the peptide-ACD interaction[29,33].

The HSPB7 structure reveals its failure to dimerise is due to a destabilised pairing of β6+7 strands

To understand why none of the HSPB7 constructs dimerise to an appreciable extent, we turned to X-ray crystallography. We undertook crystallisation trials on $HSPB7_{ACD}$ and $HSPB7_{\Delta N}$, but did not obtain any protein crystals. We speculated that this might be due to a solvent-accessible cysteine (C131) leading to non-specific intermolecular interactions during crystallisation, and hence generated the C131S mutant of both constructs. Of these, $HSPB7_{ACD}^{C131S}$ returned good crystals, which diffracted to 2.2 Å and could be solved by molecular replacement in the $P4_12_12$ space group (**Extended Table. 1**). The fold of the protein is Ig-like, with six β-strands packed into a β-sandwich arrangement, very similar to that observed for other sHSP ACDs[34,35] (**Fig. 2**). Each asymmetric unit contains three monomer chains; two of the three monomers are well resolved, while the third displays disorder in the amino acids that comprise the loop linking the β5 and β6+7 strands (residues 112-118, LAADGTV), a phenomenon that has also been observed in other sHSPs upon monomerisation[36].

The monomers within the asymmetric unit contact each other through crystallographic interfaces between neighbouring anti-parallel β4 and β6+7 strands that form long chains of monomers in the crystal. The lack of a head-to-head dimer in the crystal form is different to most sHSP ACD structures that have been determined, but is consistent with our native MS data that showed monomers to be the dominant form in solution. A rationale for this might be that the β6+7 strand is shorter in HSPB7 compared to the canonical human sHSPs and is missing charged amino acids at positions 118 and 129 that form inter-monomer salt bridges in HSPB1 and HSPB5 (**Fig. 2C**). These differences likely act to reduce the strength of a putative anti-parallel interaction of two β6+7 strands (as per the canonical metazoan sHSP dimer), with the result that HSPB7 exists primarily as a monomer.

HSPB7 binds to FLNC d24 strongly by out-competing homo-dimerisation.

Having developed an understanding of the HSPB7 structure, we proceeded by expressing and purifying domain 24 of FLNC ($FLNC_{d24}$), which forms the dimerisation interface of FLNC and has been identified as the interaction site for HSPB7[12,17]. Native MS measurements of the $FLNC_{d24}$ construct revealed it to be dimeric, in line with the literature[10]. We then incubated it at a 1:1 ratio with each of HSPB7, $HSPB7_{\Delta N}$ and $HSPB7_{ACD}$ in turn, and observed the appearance of new charge state series that could be assigned unambiguously and exclusively to hetero-dimers of $FLNC_{d24}$ and the HSPB7 construct (**Fig. 3A**, **Extended Data Table 2)**. These hetero-dimers were formed by $HSPB7_{ACD}$ to the same abundance as in the other constructs, revealing the interaction to be mediated by the ACD of the sHSP.

We assayed the strength of this interaction quantitatively for $HSPB7_{\Delta N}$ and $HSPB7_{ACD}$ by performing experiments at a range of $FLNC_{d24}$ concentrations; the instability of full length HSPB7 (some unfolding is observed in the native MS spectra, and we found the protein could readily precipitate in vitro) precluded reliable equivalent experiments for it. This allowed us to obtain dissociation constants (K_D) for the $FLNC_{d24}$:$HSPB7_{\Delta N}$ hetero-dimer as 6.7 ± 1.4 nM and $FLNC_{d24}$:$HSPB7_{ACD}$ hetero-dimer as 3.9 ± 0.9 nM, significantly tighter binding than we measured for the $FLNC_{d24}$ homo-dimer (75.9 ± 8.4 nM), by >6 $kJmol^{-1}$ (see later, **Fig. 4**). Control experiments using the ACD of HSPB1 and HSPB5 displayed negligible binding to $FLNC_{d24}$ (**Extended Data Fig. 3**). These results show that FLNC and HSPB7 form a strong, specific hetero-dimer through the interaction of $FLNC_{d24}$ with the ACD of HSPB7.

The structure of the HSPB7:FLNC hetero-dimer reveals the molecular mechanism of competition.

Given the nature of the hetero-dimer between $HSPB7_{ACD}$ and $FLNC_{d24}$ as composed of two well-structured units interacting with nanomolar affinity, we considered them a strong candidate for co-crystallisation. We were successful in obtaining crystals which diffracted to 2.9 Å, and were able to solve the structure in the $P3_221$ space group (**Extended Data Table. 1**). The asymmetric unit comprised a single hetero-dimer, with the interface centred around a parallel pairing of $FLNC_{d24}$ strand C with $HSPB7_{ACD}^{C131S}$ β4, and the donation of sidechains from each strand into an adjacent hydrophobic groove on each monomer (β4-β8 groove on $HSPB7_{ACD}$, and C-D groove on $FLCN_{d24}$) (**Fig. 3B**). Note that this is a different site and mode of association to the dimerisation of sHSP ACDs, which occurs via antiparallel pairing of β6+7 strands. Specifically, the β4-β8 groove accepts M2667 and M2669 from FLNC, while the groove between strands C and D accepts I102. The interface is thus mainly comprised of hydrophobic contacts made between the monomers' sidechains (**Fig. 3C**), with additional stability conferred by polar contacts between the chains (**Extended Data Table. 3**).

To validate the hetero-dimer interface, we performed hydrogen-deuterium exchange (HDX)-MS experiments, whereby the solvent accessibility of back-bone amides in proteins is probed in solution[37]. We obtained excellent sequence coverage for both $HSPB7_{ACD}$ (97%) and $FLNC_{d24}$ (98%), allowing us to compare the uptake of deuterium across the sequence of both proteins when they were labelled in complex versus alone. We found that the rate of deuterium uptake of several peptide products in each of $HSPB7_{ACD}$ and $FLNC_{d24}$ was slowed when they were co-incubated, consistent with them forming a hetero-dimer whose interface protects interacting regions from exchange (**Fig. 3D-F**). For $FLNC_{d24}$, the most protected region at all time-points is residues 2658-2682, which spans from the middle of

strand B to the start of strand D (**Fig. 3D**). For HSPB7$_{ACD}$, the most protected region encompasses residues 101-117, incorporating both β4 and β5 (**Fig. 3D**). These results are in close accord with the dimer interface we determined in the crystal structure as being centred on the pairing of strands C and β4.

Bearing in mind that HDX-MS operates at a resolution defined by the enzymatic digestion (and hence unaffected residues can appear protected simply by being within the same peptide product as residues that actually are), mapping the protection from deuteration measured onto our structure emphasises how protection is highest near the crystallographic interface (**Fig. 3F**). This also shows that the C131S mutation that we introduced to aid crystallisation, but was absent in the construct used in HDX-MS, did not alter the location of the dimer interface. The protection measured between HSPB7 β3 and β4 (residues 90-100) can be rationalised by the salt-bridge formed between D100 and FLNC H2686. The protection at the C-terminal end of HSPB7$_{ACD}$, from the β8-9 loop to the C-terminus (residues 147-162) appears to arise from an interaction between the loop and the Gly-Ser residue pair at the extreme N-terminus of our FLNC$_{d24}$ construct that are residual from its production as a His-tagged protein, and are likely artefactual. Together, our crystallography and HDX-MS data provides a validated structure that rationalises the tight binding of FLNC d24 to the ACD of HSPB7.

Phosphorylation of FLNC domain 24 regulates its homo- and hetero-dimerisation

Numerous phosphorylation sites have been discovered on FLNC, many of which become modified upon cellular stress of various types. Examining these in light of the overlapping homo- and hetero-dimerisation interfaces of FLNC$_{d24}$ revealed two sites for further interrogation, pT2677 and pY2683. The former has been reported in studies examining the up-regulation of phosphorylation in infarct

hearts[38], and Covid-19 infection[39] (a disease in which there is significant cardiovascular involvement[40], including stress cardiomyopathy[41]). pY2683 has been reported in studies examining exercise-induced signalling[42,43], or phosphorylation in a variety of tumor tissues including bone cancer, osteosarcoma and glioma[44]. Based on manual analysis of tandem-MS spectra from the deposited datasets, we were able to validate the presence of both phosphorylation sites (**Extended Data Fig. 4)**.

To mimic the effect of phosphorylation at each of these sites we expressed and purified the $FLNC_{d24}$ T2677D and Y2683E mutants for *in vitro* experiments. Native MS data for each of the mutants and the WT $FLNC_{d24}$ at the same protein concentration showed differing amounts of monomer and dimer in each (**Fig. 4A**, **Extended Data Table 1**). Compared to the WT, $FLNC_{d24}^{T2677D}$ showed an increase in abundance of homo-dimer, while $FLNC_{d24}^{Y2683E}$ showed a decrease. From titration experiments we obtained the corresponding K_Ds as 75.9 ± 8.4 nM (WT), 8.5 ± 1.0 nM (T2677D) and 4047 ± 557.0 nM (Y2683E) (**Fig. 4B**). This suggests that phosphorylation of each site differentially regulates the homo-dimerisation of FLNC, with T2677 stabilising the dimer (by 5.4 $kJmol^{-1}$) and Y2683 weakening it (by 9.9 $kJmol^{-1}$) (**Fig. 4C**).

We next performed similar experiments for the hetero-dimer. We found that the interaction with $HSPB7_{ACD}$ is even stronger for $FLNC_{d24}^{Y2683E}$ (2.2 ± 0.6 nM) than the WT (3.9 ± 0.9 nM), a stabilisation of 1.5 $kJmol^{-1}$ (**Fig. 4C**). However, the opposite is true for $FLNC_{d24}^{T2677D}$ (14.1 ± 0.5 nM), which is destabilised relative to the WT by 3.2 $kJmol^{-1}$. Indeed, together with the stabilisation conferred by the phosphor-mimic to the homo-dimer, this means that $FLNC_{d24}^{T2677D}$ is actually favours homo- versus hetero-dimerisation (**Fig. 4C**).

To validate these findings, we obtained comparative HDX-MS data for the phospho-mimics and WT

$FLNC_{d24}$ in complex with $HSPB7_{ACD}$. Focusing on $HSPB7_{ACD}$ only (because $FLNC_{d24}$ has a different sequence in the three constructs and hence generates different, and thus not directly comparable, peptides) we found that the peptide diagnostic of the dimer interface (encompassing residues 101-117 and incorporating both β4 and β5, **Fig. 3D-E**) showed decreased protection from deuteration for $FLNC_{d24}^{T2677D}$ versus WT, and increased protection of $FLNC_{d24}^{Y2683E}$ (**Extended Data Fig. 5**). This is consistent with the order of interface strengths that we observed in the native MS experiments, and underscores how phosphorylation at these two residues can regulate the interaction with HSPB7.

Molecular dynamics analysis reveals the residue contacts altered by phosphorylation

To isolate the amino acid contacts that underpin these phosphorylation-induced changes in FLNC homo- and hetero-dimerisation we turned to molecular dynamics (MD) simulations. We first generated a T2677- or Y2683-phosphorylated $FLNC_{d24}$ homo-dimer structure[45] ($FLNC_{d24}^{pT2677}$, $FLNC_{d24}^{pY2683}$), by adding a doubly charged phosphate group to the side chains of either T2677 or Y2683 in the WT $FLNC_{d24}$ structure (**Fig. 5A**). We also did the same for the hetero-dimers, generating $FLNC_{d24}^{pT2677}$:$HSPB7_{ACD}$ and $FLNC_{d24}^{pY2683}$:$HSPB7_{ACD}$ from our $FLNC_{d24}$:$HSPB7_{ACD}$ structure.

We ran 1-μs MD simulations under the same conditions for all six structures. All remained stable over the trajectories, allowing us to interrogate them to understand how they differ in the inter-monomer interactions due to the introduction of the charged phosphate groups. Examining the homo-dimers first (**Fig. 5B**, **Extended Data Fig. 6-7**), and comparing each mutant individually with the wild-type, we found that pT2677 forms a salt-bridge with R2690 on the opposite monomer. This likely strengthens the dimer interface in the phosphorylated form. In contrast, while the WT Y2683 is able to form an inter-molecular hydrogen bond with E2681 across the dimer interface, pY2683 does not. Instead the

negative charges on the phosphate group reside within a predominantly negatively charged environment, and are coulombically repelled by the surrounding residues. These effects combine to destabilise the homo-dimer interface of $FLNC_{d24}^{pY2683}$ relative to the WT.

Examining the hetero-dimers next (**Fig. 5C**, **Extended Data Fig. 6-7**), we noted that phosphorylation of T2677 twists the loop in which it resides which leads to the loss of a hydrogen bond between the FLNC G2674 backbone carbonyl and the HSPB7 N107 sidechain (**Fig. 5C**, upper). We also see strengthened interaction between HSPB7 T143 (on β8) and FLNC D2712 (in the loop between strands F and G) in pT2677. The increase in occupancy of this hydrogen bond (formed 10% of the time in the WT hetero-dimer, 38% in pT2677 form) is however insufficient to offset the destabilisation of the hetero-dimer caused by the loss of the G2674-N107 bond. In the pY2683 hetero-dimer, the phosphorylated pY2683 binds with $HSPB7_{ACD}$ R113 via a salt bridge which is not formed in the WT (**Fig. 5C**, lower). We also find that E99 from HSPB7 forms bonds with H2686, N2689 and R2690 in the phosphorylated form, further stabilising the dimer interface. These results therefore provide residue-specific detail for the phosphorylation-mediated regulation of stability of FLNC homo- and hetero-dimerisation (**Fig. 4**).

FLNC:HSPB7 hetero-dimerisation is an ancient regulated interaction

Given the apparently strong, specific and regulatable interaction between $FLNC_{d24}$ and HSPB7, we were interested to examine the evolutionary conservation of this complex. Therefore, we constructed phylogenetic trees of HSPB7, from which we could trace it into organisms as distantly related to humans as the tunicate *Styela* (a jawless chordate which posses a heart), but not in the cephalochordate *Branchiostoma* (**Extended Data Fig. 8A**). Interestingly, when examining the set of 113 HSPB7

sequences, the residues that we identified as key in acting to prevent dimerisation of human HSPB7 (A118 and H129) are highly conserved across all HSPB7s, while in the other sHSPs they are highly conserved as Asp and Arg, respectively (**Extended Data Fig. 8B**). This strongly suggests that HSPB7 is predominantly monomeric in all organisms in which it is found.

To assess the interaction between FLNC and HSPB7 we designed an *in silico* competitive-binding assay (**Extended Data Fig. 8C-D**). In our approach we supplied AlphaFold two chains of both $FLNC_{d24}$ and HSPB7, such that it had the opportunity to either assemble the proteins into hetero- or homo-dimers. For the human pair, this approach identified hetero-dimerisation as favoured, and returned a model that matched our crystal structure closely. A negative control experiment, using HSPB5 (which our experiments show does not bind to $FLNC_{d24}$, **Extended Data Fig. 3**), led to homo-dimers of it and FLNC instead.

The alignment of these results with our experiments gave us confidence in our computational approach, and we thus inferred the dimerisation mode for various $FLNC_{d24}$:HSPB7 pairs along this phylogenetic tree. We found that hetero-dimerisation was preferred in extant organisms as distantly related to humans as *C. milli* (elephant shark; in the class chondrichthyes). Moreover, the *C. milli* hetero-dimer model is extremely similar to the human hetero-dimer (**Extended Data Fig. 8D**). To estimate when this hetero-dimer first emerged, we used ancestral sequence reconstruction to resurrect the HSPB7 sequences in the (long since extinct) ancestors of humans and *C. milli*, and humans and the tunicate *Styela*. AlphaFold also returns hetero-dimer models of these ancestral HSPB7s (with either human or *C. milli* $FLNC_{d24}$) of equivalent structure to the extant forms (**Extended Data Fig. 8D**). These results suggest that HSPB7's interaction with FLNC likely evolved along the branch leading to the last

common ancestor of tunicates and humans. This animal was a chordate that likely had a primitive tubular heart[46], in line with HSPB7's present day function in humans.

Having elucidated that the interaction between HSPB7 and FLNC is >400 million years old, we asked whether there is evidence for the regulatory fine-tuning we uncovered for T2677 and Y2683 to be equally ancient. We noted that both sites also exist in the sequence of $FLNC_{d24}$ from *C. milli* and *Styela*, suggesting this to be the case. We also examined the sequence of $FLNC_{d24}$ in archetypal organisms across the relevant diversity of organisms, including reptiles, amphibians, fish and birds. In all cases the two sites were conserved, suggesting that they are available for phosphorylation, assuming the presence of an appropriate kinase[47]. This evolutionary analysis therefore implies that not only is the interaction between HSPB7 and FLNC long-established, but so too is its regulation by post-translational modification during stress-signalling.

Discussion

FLNC dimerisation and hence diffusional mobility is regulated by HSPB7

In this study we have found that FLNC and HSPB7 interact in murine cardiac tissue, with both proteins up-regulated under biomechanical stress *in vivo*. This is consistent with previous studies that have found that FLNC and HSPB7 can co-localise and interact in cells[17,18]. The interaction between FLNC and HSPB7 has previously been shown to occur at d24 of FLNC[17]; our results confirm this and map the binding of $FLNC_{d24}$ to the ACD of an HSPB7 monomer. Close examination of the $FLNC_{d24}$:$HSPB7_{ACD}^{C131S}$ hetero-dimer structure with the published structure of the $FLNC_{d24}$ homo-dimer (PDB ID:1V05)[10] shows that the interface surfaces overlap very closely, and thereby provides a structural rationale for the competitive binding between the two proteins (**Fig. 6**). Notably, our experimental data show that the hetero-dimer has a low nM K_D, significantly stronger than the FLNC homo-dimer. Taken together, these results can be viewed as HSPB7 effectively acting to shift the equilibrium between FLNC dimer and monomer towards dissociation, with the dissociated FLNC monomer then becoming tightly capped by HSPB7 (**Fig. 6C**).

The full-length FLNC:HSPB7 hetero-dimer (26 domains, ~310 kDa) is much smaller than the FLNC homo-dimer (50 domains, ~582 kDa). This difference in size leads to a larger diffusion coefficient for the free hetero-dimer, with the result that FLNC is more mobile in solution in the presence of HSPB7. A consequence of up-regulation of HSPB7 during biomechanical stress might therefore be to increase the mobility of FLNC, such that it translocates more rapidly to force-induced microlesions in the sarcomeric structure, where it can act to stabilise the cytoskeleton in advance of permanent repair by the production of new cytoskeletal proteins[5].

Phosphorylation of FLNC near the interface modulates dissociation in concert with HSPB7

Phosphorylation has also been identified as regulatory mechanism that influences the mobility of FLNC[14,15,48]. Notably, the phosphorylation of serine and threonine residues on FLNC, for instance by protein kinase C, can have influence along the length of the protein, rigidifying it and reducing its mobility[48]. This includes hindering the proteolytic cleavage between d23 and d24[14,48]; and inhibiting the binding with FILIP1 required to initiate degradation of FLNC[15]. Here we investigated two sites of phosphorylation of FLNC associated with non-basal conditions: T2677 and Y2683. We chose them due to their proximity to the dimer interface within d24, reasoning that they might lead to a shift in the FLNC monomer-dimer equilibrium. Our results show phosphorylation at T2677 strengthening the homo-dimer, and phosphorylation at Y2683 weakening it. Remarkably, we observed the opposite effect on the FLNC:HSPB7 hetero-dimer. The combined effect of these two changes on the coupled equilibria is that, in the presence of HSPB7, stabilisation of the FLNC homo-dimer is increased in the case of pT2677, whereas its dissociation is heightened for pY2683 (**Fig. 6C**). The enhancement by HSPB7 of phosphorylation-based regulation adds strength to the hypothesis of interplay between molecular chaperones and kinase-mediated signalling pathways[49].

Phosphorylation at T2677 was reported (and validated in our manual analysis of the data) in cardiovascular pathologies. Our insight predicts that this phosphorylation will act to stabilise the FLNC homo-dimer, reducing its mobility in the cell, and acting to solidify the cytoskeletal linkages that FLNC mediates. pY2683, instead, was identified in exercise-induced and cancer phospho-proteomics screens. In general, phosphorylation on tyrosine is common in cancer, due to the activation of tyrosine kinases which leads to modulation of a variety of cellular activities[50,51]. Our data suggests that the weakened

FLNC homo-dimer interface that we observed upon Y2683 phosphorylation may increase the mobility of FLNC, and might facilitate its participation in cancer cell locomotion, proliferation and cytoskeletal rearrangement.

These insights provide a mechanistic rationale for how FLNC might switch between its multiple apparent functions, with the monomer representing a mobile state that is inherently incapable of crosslinking actin filamins within the cytoskeleton but can translocate to sites where it is required, and the dimer being less mobile, but able to provide cytoskeletal stabilisation through actin-crosslinking. This equilibrium therefore represents an opportunity to modulate the function of FLNC, with possible implications for both muscle disease and cancer.

HSPB7 evolved its unique interaction with FLNC through loss of self-interaction

Our elucidation of the specific association made by HSPB7 with FLNC is unusual in the context of our understanding of sHSPs. This degree of specificity and binding strength to a target protein has not been observed for a member of the sHSP family before, and it is also notable that the target is natively folded rather than partially unfolded or *en route* to aggregation[26,32,52]. Access to FLNC is facilitated by the ACD not binding other regions of the HSPB7 sequence, either inter- or intra-molecularly. As a result HSPB7 is monomeric at low μM concentrations *in vitro*, comparable to the abundance at which it is found in cells that express it[53]. In line with this, even upon over-expression in HEK293 cells, HSPB7 was found to exist as low molecular weight species[25].

Our phylogenetic analyses suggest this interaction is >400 million years old. The most likely evolutionary trajectory for HSPB7 is the duplication of an ancestral sHSP followed by degeneration

of its oligomerisation interfaces. This degenerative loss then may have unmasked a latent ability to bind FLNC, which is prevented in other sHSPs through their self-interactions, and consolidated by the presence of an additional hydrophobic residue (I102 in humans) that engages in the hetero-dimer interface. Given the promiscuous binding characteristics and broad tissue distribution of canonical sHSPs (e.g. HSPB1, HSPB5 in vertebrates), it stands to reason that their duplication provides good opportunity for sub- and neo-functionalisation into chaperones tailored to the specific regulatory needs of a particular target protein. We anticipate therefore to find other such examples within the sHSP and wider molecular chaperone families.

Methods

Protein expression and purification

Full-length human HSPB7 (UniProt ref: Q9UBY9-2, Isoform 2, A67 to AAHPTA, 175 residues) encoded in pET23a was transformed into *E. coli* BL21(DE3) pLysS cells (Agilent) and expressed in LB media (with 0.5% glucose, 50 mg/mL ampicillin, and 37 mg/mL chloramphenicol) for 16 hours at 22 °C following cell culture growth to OD600 between 0.6-0.8 and induction with 0.5 mM IPTG. Cells were lysed using French Press Microfluidizer in 20 mM Tris, 1 mM EDTA, 5 mM β-mercaptoethanol (BME) at pH 7.4 with EDTA-free protease inhibitor cocktail. The lysate was centrifuged at 20,000 g for 20 min at 4 °C, which pelleted the desired protein into inclusion bodies. The pellet was washed with the lysis buffer supplemented with 1% triton X-100 twice, and then once with 1 M urea, to remove membrane-associated contaminants. The protein was then solubilised in 6 M urea and 10 mM sodium phosphate at 4 °C, and refolded by removing urea stepwise in dialysis buffers containing 3M, 2M, and 1M urea, with 100 mM NaCl, 20 mM Tris, 5 mM BME (pH 7.5) buffering the process.

$HSPB7_{ACD}$ (residues 78-162 of isoform 2 UniProt ref: Q9UBY9-2, both WT and C131S forms), $HSPB7_{\Delta N}$ (residues 78-175 of isoform 2, WT and C131S) and $FLNC_{d24}$ (residues 2630-2724, WT and phosphomimic mutants T2677D and Y2683E) encoded in pET28a(+)with an N-terminal His_6 tag were transformed into *E. coli* BL21(DE3) cells (Agilent). Cells were used to inoculate a 10 mL LB culture with kanamycin (30 μg/mL) at 37 °C overnight and transferred to a 1 L large culture with the same antibiotic. The culture was shaken at 37°C until the OD600 reached 0.6-0.8, followed by induction with IPTG to a final concentration of 0.5 mM. After shaking overnight again, at 22 °C, the cells were harvested and lysed using Microfluidizer with the addition of EDTA-free protease inhibitor cocktail

and buffer containing 300 mM NaCl, 50 mM Tris and 20 mM imidazole (pH 8.0). The lysate was then spun at 20,000 g for 20 min at 4°C and loaded to a HisTrap HP column (Merck). The protein was eluted with the buffer with 300 mM NaCl, 50 mM Tris and 500 mM imidazole (pH 8.0) and dialysed to the loading buffer with TEV protease at room temperature overnight. Cleavage of the His_6 tag leaves a residual Gly-Ser overhang on the N-terminus of the constructs. The sample was again loaded to the HisTrap HP column to remove the free His_6 tag and uncleaved proteins, and exchanged to PBS buffer with 5 mM Tris (2-carboxyethyl) phosphine (TCEP) for $FLNC_{d24}$ constructs or buffer containing 50 mM Tris (pH 8.0) for HSPB7 constructs. HSPB7 constructs were loaded to HiTrap Q anion exchange column (Merck) and eluted with a gradient of 150-200 mM NaCl in the elution buffer. The pure protein was then exchanged into PBS buffer with 5 mM TCEP for storage.

Crystallisation and structure determination

Freshly prepared $HSPB7_{ACD}^{C131S}$ protein that had never been frozen was buffer-exchanged to 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, pH 8.0), concentrated to 10 mg/mL and used for crystallisation. Crystallisation broad screens were prepared using an Art Robbins Phenix liquid handler and MRC SWISSCI 96-3 well sitting-drop vapor diffusion plates (Hampton Research, US). Drops containing 100:200 nL, 100:100 nL, and 200:100 nL (protein to well) were set with an 80µL precipitant well volume. The crystal for structural determination of $HSPB7_{ACD}$ C131S was discovered and harvested after setting up for 10 days from a well containing 0.2 M magnesium formate dihydrate and 20% w/v polyethylene glycol (PEG) 3350 (JCSG-plus A5, Molecular Dimensions, protein: precipitant ratio 2:1). Crystals were cryoprotected using 10% glycerol in well solution, mounted on a nylon loop and then cryo-cooled by rapid plunging into liquid N_2. Diffraction data were

collected and auto processed at Diamond Light Source beamline I04. The diffraction data were integrated, merged and scaled automatically using Xia2 processing pipelines at Diamond Light Source[54]. The Phenix software package[55] was used for data analysis as follows. Molecular Replacement in PHASER using a search model generated with the Robetta comparative modeling server[56] which was based on the HSPB6 (PDBID: 4JUS) structure was used to obtain initial phases. Refinement was carried out using iterative cycles of fitting in COOT[57] and refinement sequentially in Phenix-refine, PDB-REDO[58] and CCP4 REFMAC[59]. The final structure has R and R_{free} values at 0.253 and 0.280, respectively.

For crystallisation of the $HSPB7_{ACD}^{C131S}$:$FLNC_{d24}$ hetero-dimer, both proteins (purified on the same day) were mixed at a ratio of 4:1 based on monomer and injected onto a gel-filtration column (Superdex 200, 16/600, Merck) equilibrated in 20 mM HEPES buffer (pH 8.0). The eluate corresponding to the hetero-dimer was concentrated to 10 mg/mL and used for crystallisation. Crystal screens were set as described above. Crystals appeared after one week in a condition containing 0.1 mM HEPES and 20% w/v PEG 8000, pH 7.0 (The Protein Complex Suite F1, Molecular Dimensions, protein: precipitant ratio 2:1). The crystal was harvested and flash-cooled without adding cryo-protectant. Diffraction data were collected at Diamond Light Source beamline I03, and auto processed by Xia2 using the processing pipelines at Diamond Light Source. The $FLNC_{d24}$ chain in the hetero-dimer was phased using PHASER with $FLNC_{d24}$ (PDBID: 1V05) as the initial search model and was subsequently fed to CCP4 PHASER[59] with $HSPB7_{ACD}^{C131S}$ chain A to solve the hetero-dimer structure. Iterative cycles of fitting and refinement were carried out using COOT and Phenix-refine to R and R_{free} values of 0.224 and 0.291, respectively. Data collection and refinement statistics are shown in **Extended Data Table 1**.

Quantitative native MS

All proteins except for the full-length HSPB7 were buffer-exchanged into 200 mM ammonium acetate (pH 6.9) using biospin (Bio-Rad); full-length HSPB7 was buffer-exchanged into 1 M ammonium acetate (pH 6.9). C-terminal peptides (Biomatik) were dissolved in 200 mM ammonium acetate (pH 6.9). Native MS was performed using nano-electrospray infusion in positive ion mode, using gold-coated capillaries that were prepared in house, on a Synapt G1 (Micromass/Waters) modified for the transmission of intact non-covalent protein complexes. Experiments were performed according to standard native MS procedures[60], with key instrumental parameters as follows: capillary, 1.5 kV; sampling cone, 20-35 V; extraction cone, 3.0-3.3 V; backing pressure, 3.8-4.1 mbar; trap gas (argon) flow, 1.5-3.0 mL/min; trap, 6-8 V; transfer, 6-12 V. For the peptide titration experiments, acetonitrile was added to the source housing to help preserve non-covalent interaction[61].

Theoretical and measured masses are shown in **Extended Data Table 2**. Three technical repeats were collected for each ratio in the titration experiments, and peak intensities were extracted manually. We fitted the titration curves using either GraphPad Prism 8.0.1 or home-built Python code. The fitting of the ACD-peptide titrations used a Hill slope model for a 1:1 binding, with B_{max} and h values constrained to 1.0, while the competitive homo- and hetero-dimerisation was considered as a system of coupled equilibria (**Extended Data Methods**).

Aggregation inhibition assay

Citrate synthase (CS) from the porcine heart (Sigma-Aldrich Co Ltd.) and HSPB7 constructs were buffer-exchanged into the reaction buffer with 40 mM HEPES and 1mM DTT, pH 7.5. The final

concentration of CS was maintained at 1 μM, while the concentrations of HSPB7 constructs were varied from 0.25-4 μM. The samples were transferred to the 96-microwell plates, and aggregation of CS was induced at 45 °C and monitored through the chance in apparent absorbance at 340 nm due to light scattering for at least two hours in a Fluostar Optima plate reader (BMG Lab Technologies). Three technical repeats were conducted for each ratio.

Molecular Dynamics simulations

The crystal structures of wild-type $FLNC_{d24}$ homo-dimer (PDB ID: 1V05) and $FLNC_{d24}$:$HSPB7_{ACD}^{C131S}$ hetero-dimer were used as the initial structures for MD simulations. The structures of phosphorylated residues were modelled using Visual Molecular Dynamics (VMD). Each complex was immersed in a TIP3P periodic cubic water box extending 10 Å beyond the protein boundaries in every direction, neutralised with 0.15 M NaCl. Atomic interactions were described according to the Amber FF14SBforce field[62], combined with the phosAA19SB force field to describe phosphorylated residues.

MD simulations of all complexes were run using the NAMD2.14b2 engine[63]. All bonds were restrained using the SHAKE algorithm, enabling a 2 fs timestep, and long-range electrostatic interactions were treated with Particle Mesh Ewald with a 12 Å cutoff. The systems were first minimised via conjugate gradient, with 2000 steps for the wild type dimers and 10000 steps for modified dimers, then simulated in the nPT ensemble (300 K, 1 atm) for 0.5 ns with alpha carbons fixed by a 10 kcal mol^{-1} constraint. Constant temperature and pressure were maintained by Langevin dynamics, with 1 ps^{-1} as the damping constant, 200 fs as the piston period, and 50 fs as the piston decay. The pressure of each system was then equilibrated by removing the above constraints and running a 1 ns simulation in the nVT ensemble.

Finally, for each system a 1 μs production run in the nPT ensemble was carried out, with frames collected every 1 ns.

Simulations were analysed with custom Python scripts using the MDAnalysis, numpy, and matplotlib packages[64,65]. For the hydrogen-bonding contact analyses, the cutoff values used for identifying hydrogen bonds were set as 1.2 Å for the donor-hydrogen distance, 3.0 Å for the hydrogen-acceptor distance and 150° as the donor-hydrogen-acceptor angle. The hydrogen-bond occupancy of any given pair of residues was calculated as a percentage, taking into account the possibility that residue pairs could form more than one hydrogen bond.

To identify the contact surface within $FLNC_{d24}$ homo- and hetero-dimers, we calculated both the average shortest distance and minimum shortest distance from each residue in $FLNC_{d24}$ to the partner monomer during each frame of the MD trajectories. The shortest distance is defined as being between any heavy atom on the $FLNC_{d24}$ residue in question, and any heavy atom on the partner monomer. To display this on a surface rendering of $FLNC_{d24}$, we coloured each residue with the saturation negatively linearly correlated to the value of the average shortest distance from the partner monomer.

HDX-MS

$HSPB7_{ACD}$, $FLNC_{d24}$ and the hetero-dimer were all prepared in 40 mM HEPES at pH 7.5. To form the hetero-dimer, $HSPB7_{ACD}$ and $FLNC_{d24}$ were mixed at 1:1 (25 μM : 25 μM, for $HSPB7_{ACD}$ measurement) or 4:1 (50 μM :12.5 μM, for $FLNC_{d24}$ measurement) molar ratios. To carry out the HDX-MS experiments, 5 μL sample volumes were each exposed to labelling in 55 μL of D_2O buffer (40mM HEPES, pD 7.5) for a time course of 30, 60, 500, 1000 and 5000 s, followed by mixing with one

equivalent of pre-chilled quench solution (60 µL, 40 mM HEPES, 5 mM TCEP, pD 2.0). 80 µL of each of the quenched samples were then injected in turn into a nanoACQUITY UPLC System (Waters) and digested online using an Enzymate™ BEH Pepsin Column (2.1 x 30 mm, Waters) at room temperature. The digested peptides were trapped by a BEH C18 trap column (1.7 µm, 2.1 x 5 mm, Waters) for 3 minutes for desalting. A BEH C18 analytical column (1.7 µm, 1 x 100 mm, Waters) was then used to separate the peptides using a linear gradient from 3% to 35% of acetonitrile with 0.1 % formic acid at a flow rate of 40 µL/min. To minimise back-exchange, the temperature was kept below 4 °C after digestion. A wash buffer of 1.5 M guanidinium chloride pH 2.8, 4 % acetonitrile, 0.8% formic acid was used to clean the pepsin column between each run. MS data were obtained on a Synapt G2-Si (Micromass/Waters) operating in positive ion mode, and fragmentation spectra were obtained in MS^E mode. Peptides from un-deuterated samples were analysed and identified using ProteinLynx Global Server 2.5.1 (Waters). DynamX 3.0 (Waters) and Deuteros 2.0[66] were used to analyse and visualise the HDX-MS data. The peptide-level significance test was used as the model, with the alpha value set at 0.01 (i.e. 99% confidence).

Phylogenetic analyses and ancestral sequence reconstruction

To infer the phylogenetic tree of HSPB7 and homologs, amino acid sequences were gathered by online BLASTP on 19 September 2023 using human (NP_055239.1) and elephant shark (XP_007905160.1) HSPB7 amino acid sequences as initial queries. Gathered HSPB7 sequences were further used to interrogate genomes from representative taxa in each major clade, resulting in an expanded set of 471 HSPB amino acid sequences spanning Bilateria and Cnidaria. To distinguish between homologs, all HSPB ACD domains were first aligned using MUSCLE (v.3.8.31) before tree inference by FastTree

(v.2.1.11). Sequences clustering together in each HSPB clade were manually curated to remove lineage-specific indels before profile-alignment with MAFFT (v7.505).

The maximum likelihood (ML) phylogeny of all HSPB ACD domains was inferred using RaxML-NG (v.1.1) and the Jones-Taylor-Thornton model, as determined by automatic best-fit evolutionary model selection in IQ-TREE (v.2.2.0.3), with gamma-distributed among site rate variation and fixed base frequencies. Lastly, to obtain the final tree, hagfish and *Styela* HSPB7 sequences were constrained to the base of vertebrates due to incongruence of their positions in the ML tree with known species relationships.

Ancestral sequences and posterior probabilities (PP) of ancestral states were reconstructed at internal nodes leading to the HSPB7 lineage in our constrained tree as implemented in IQ-TREE, using the same evolutionary model as per tree inference. Gap assignment of ancestral sequences was determined using Fitch parsimony with PastML (v.1.9.34) given our final tree. Ancestral sequences contain states with the highest PP at all sites selected. For alternative ancestors, states with the second best PP if PP > 0.20 at each site were selected, with all other sites containing ML states.

The extant FLNC domain 24 sequences were gathered using the online BLASTP tool on 4 January 2024. The sequence for human FLNC d24 (NP_055239.1) was used as a query sequence and the search was restricted to the organisms selected as representative of their taxa. From the sequences found, the hit with highest percentage identity was chosen as the FLNC d24 homolog of that organism. The resulting sequence IDs are: elephant shark (XP_042191161.1), *Styela* (XP_039265544.1), zebrafish (AAH90688.1), *Xenopus* (AAH99062.1), green anole (XP_016864456.1), chicken (NP_989904.1).

In silico competitive binding assay

All structure predictions were generated used ColabFold[67], a fast and reliable software that combines the fast homology search of MMseqs2 with AlphaFold2[68]. We ran AlphaFold locally on a Dell Precision 5820 workstation installed with an Nvidia RTX4000 GPU card. We used 10 recycles through the network to improve the prediction and the *--amber* flag to relax the predicted structure using the AMBER force field. The structures generated were ranked using pTM (estimate of the Template Modelling score). To generate multiple structures we used different seeds, and the best ranked structure was used for visualisation and further investigation.

Because AlphaFold inevitably produces a structure of a protein complex when two (or more) chains are inputted, distinguishing false from true positives is a major challenge. Hence, we designed a competition assay where we provided two chains of each of HSPB7 and FLNC as input to AlphaFold. Therefore homo-dimerisation (giving $HSPB7_2$ and $FLNC_2$) and hetero-dimerisation (giving 2 copies of FLNC:HSPB7) outcomes are both possible. The resulting models from the competition were inspected and assigned as homo- or hetero-dimers.

Western blotting, immunofluorescence and co-IP

Western blotting, immuno-fluorescence on cryosections from mouse cardiac tissue and co-immunoprecipitations were performed as described previously[16]. For the detection of HSPB7 we used the rabbit antibody 15700-1-AP (Proteintech).

Acknowledgments

This work was supported by the Medical Research Council [MR/V009540/1 to ZW, SBS, KG and JLPB]; British Heart Foundation (BHF) [FS/12/40/29712 to KG]; the Oxford BHF Centre of Research Excellence [RE/13/1/30181 to KG]; the Wellcome Trust [201543/B/16/Z to KG]; the Leverhulme trust [grant no. RPG-2021-246 to DS, JZYN, GKAH, JLPB] the German Research Foundation (DFG) [FOR 2743, project P09 (WA1598/6) to BW] and [FOR 2743 - FU339/11-2 and FU339/13-1 to DOF]. The Institute of Cardiovascular Sciences, University of Birmingham, has received an Accelerator Award by the British Heart Foundation [AA/18/2/34218]. JRB acknowledges support from the Royal Society University Research Fellowship. We also thank the staff of I03 and I04 at Diamond Light Source for their support.

Author contributions

Conceptualisation: ZW, KD-C, DOF, BW, GKAH, KG, JLPB. Investigation: ZW, GC, MPC, XQ, SB-S, JZYN, NS, AJA, CH, JB. Formal analysis: ZW, GC, XQ, DS, JZYN, MM, PR, HS, JRB, MTD, TMA. Writing – original draft: ZW, KG, JLPB. Writing – review & editing: all authors. Supervision: CJS, BW, MTD, GKAH, KG, JLPB.

Competing interests

The authors have no competing interests

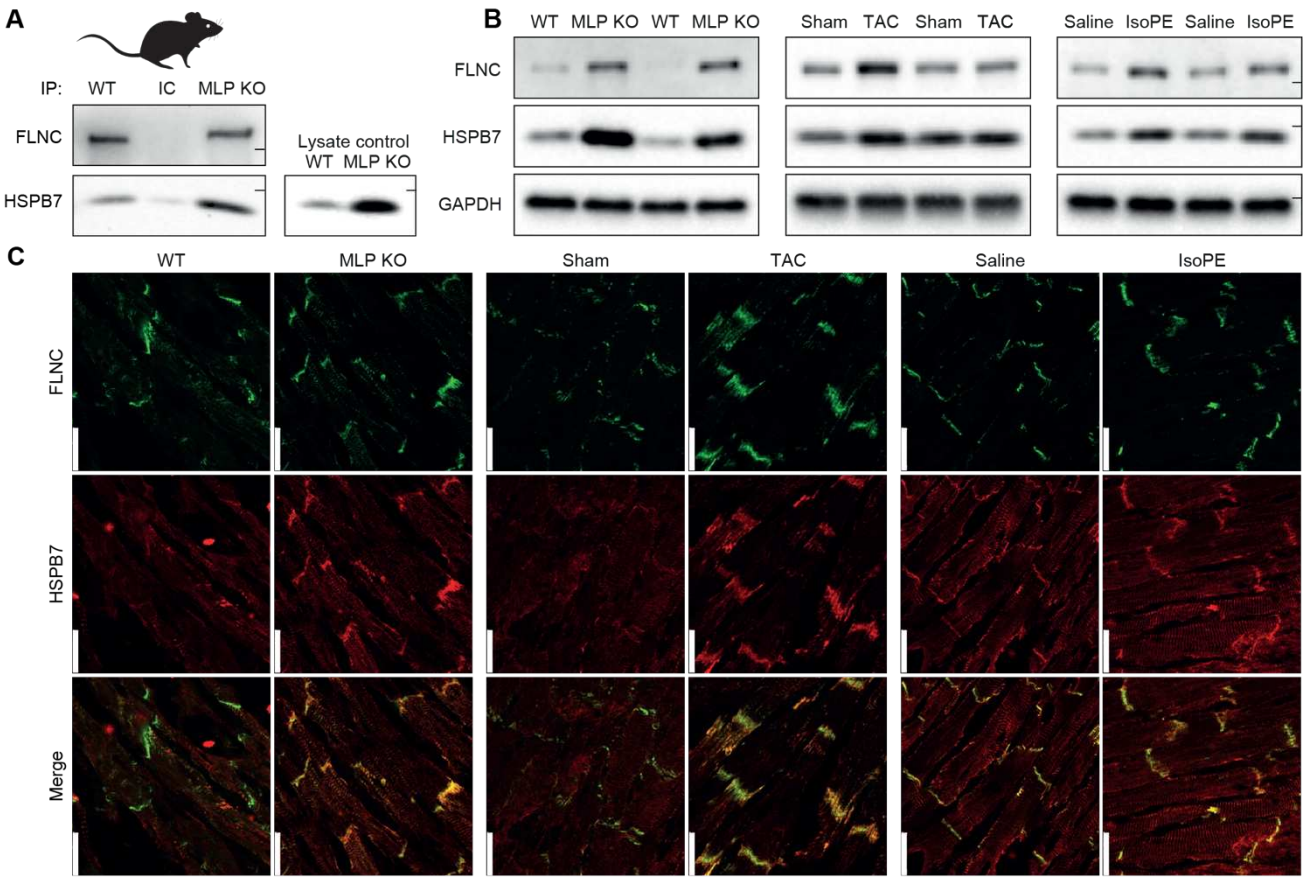

Fig. 1 | HSPB7 and FLNC are up-regulated and interact in biomechanical stress models in mouse.

A Co-immunoprecipitation of HSPB7 and FLNC from MLP KO mouse ventricular tissue. The co-precipitation of the two proteins when precipitating FLNC suggests they form (at least part of) a complex. An isotype control, where a non-specific antibody (rather than for FLNC) raised in the same species was used as a negative control and showed no bands, whereas a positive control using an aliquot of lysate showed clear bands for HSPB7. Molecular weight bands are marked as ticks, and are 250 kDa (FLNC) and 20 kDa (HSPB7). **B** Western blots for FLNC and HSPB7 in MLP KO, TAC and ISO/PE mouse ventricular tissue. GAPDH was used as the loading control. Sham and saline are the respective specific negative controls for TAC and ISO/PE. In all stress treatments, HSPB7 and FLNC are observed at higher levels. Molecular weight bands are marked as ticks, and are 250 kDa (FLNC),

20 kDa (HSPB7), and 37 kDa (GAPDH). **C** Immuno-fluorescence of HSPB7 and FLNC from frozen sections of ventricular tissues from TAC, ISO/PE and MLP KO mouse models. Sections were stained for FLNC (Green) and HSPB7 (Red). In all stress treatments, the two proteins are up-regulated and co-localise. Scale bars are 25 μm.

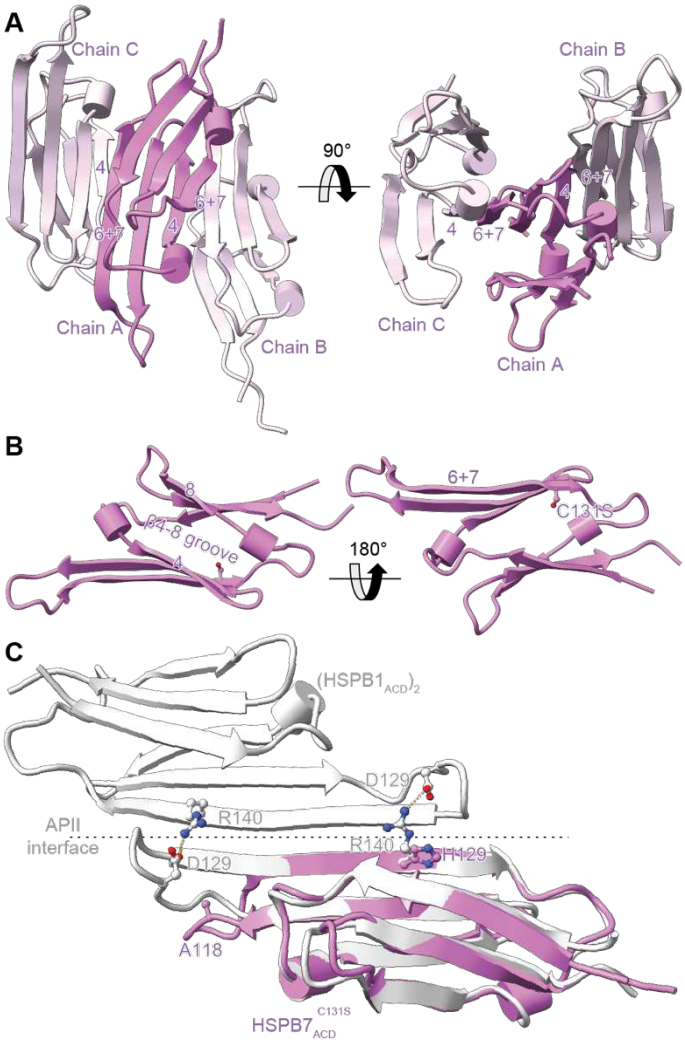

Fig. 2 | Crystal structure of the HSPB7 monomeric ACD. A The asymmetric unit of $HSPB7_{ACD}^{C131S}$contains three chains (A-C), packed through the neighbouring β4 and β6+7 strands. **B** The Ig-fold is very similar to that of other sHSP ACDs: it contains six β strands, with a groove between β4 and β8 and an extended "β6+7" strand. **C** Overlay of the $HSPB7_{ACD}^{C131S}$ (purple) structure with that of the closely related HSPB1 (white) reveals how the absence in HSPB7 of specific salt bridges that are conserved in the other HSPBs (**Extended Data Fig. 4B**) have weakened the dimer interface such that monomers are the dominant species in solution.

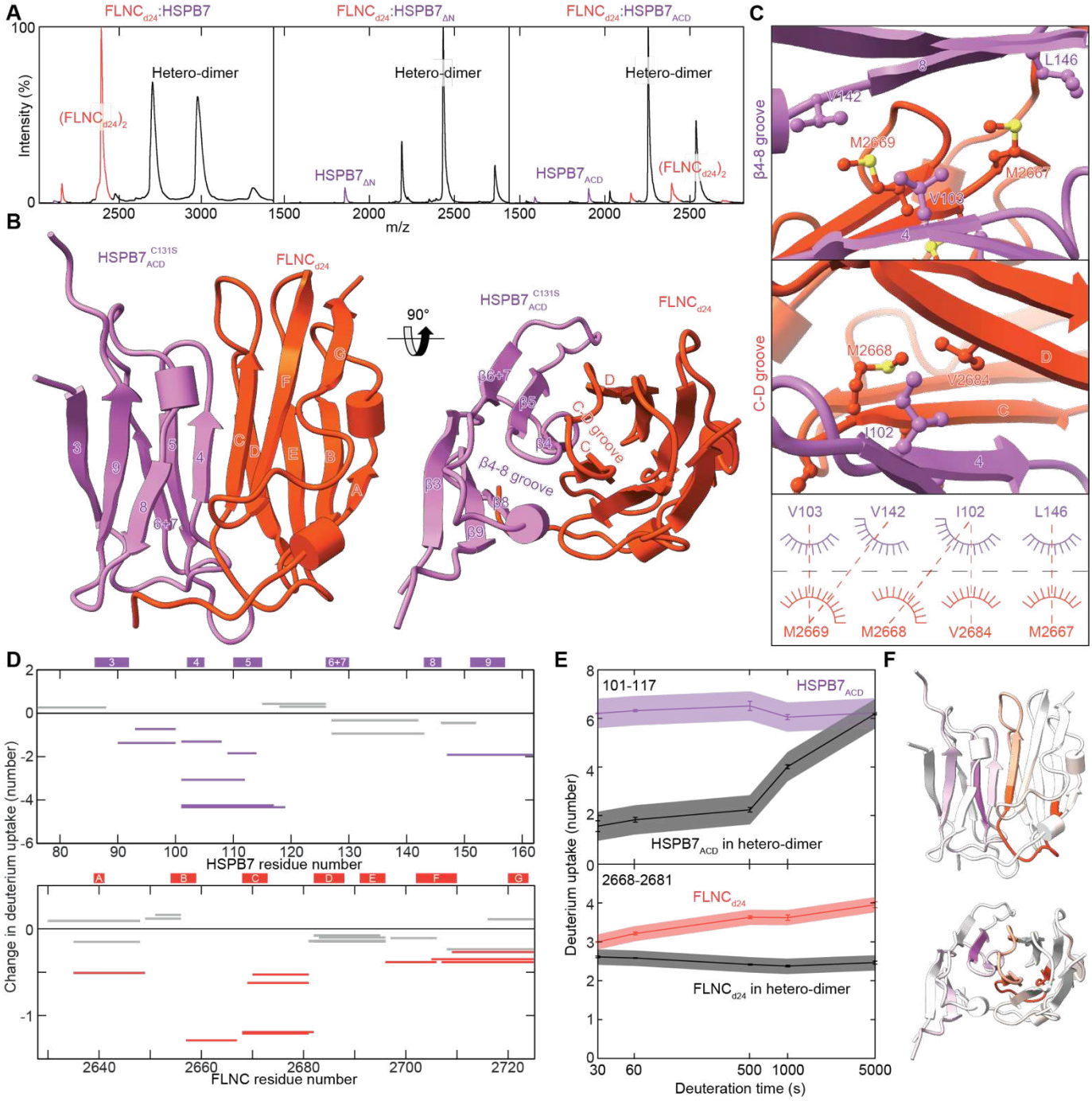

Fig. 3 | Hetero-dimerisation of HSPB7 and FLNC. A Native MS experiments show that $FLNC_{d24}$ (red) forms hetero-dimers (black) with HSPB7, $HSPB7_{\Delta N}$ and $HSPB7_{ACD}$ (purple). No complexes between the $FLNC_{d24}$ homo-dimer and HSPB7 are observed, revealing how homo- and hetero-dimerisation are directly competitive. Proteins were mixed at a monomer ratio of 1:1, but the instability of full length HSPB7 means it was somewhat depleted in solution such that $FLNC_{d24}$ was effectively

in excess (and $FLNC_{d24}$ is visible in the spectrum). **B** The $FLNC_{d24}$:$HSPB7_{ACD}^{C131S}$ hetero-dimer structure reveals an interface that involves the HSPB7 β4, β8 strands and $FLNC_{d24}$ strands C and D, centred on the parallel pairing of strands C and β4. **C** The hydrophobic β4-8 groove on HSPB7 accepts M2667 and M2669 from FLNC (top), while the groove between strands C and D on $FLNC_{d24}$ accepts I102 from HSPB7 (middle). This leads to a network of hydrophobic interactions that stabilise the hetero-dimer (bottom). **D** HDX-MS analysis of $HSPB7_{ACD}$ and $FLNC_{d24}$ in the hetero-dimer compared to in isolation. Woods plots showing the difference in deuterium uptake of $HSPB7_{ACD}$ (top) and $FLNC_{d24}$ (bottom) in the two states, at a labelling time of 500 s. The y-axis is calculated as the uptake for the isolated proteins minus that in the hetero-dimer; negative values denote protection from exchange in the hetero-dimer. Error bars refer to the standard deviation of three repeats at each time-point, and the shaded bands 99% confidence. **E** Plot of deuterium uptake versus exposure time of the most protected peptide in $HSPB7_{ACD}$ (upper) and $FLNC_{d24}$ (lower) in the hetero-dimer compared to this same peptide in the isolated protein. Because HSPB7 is monomeric, in the absence of FLNC this peptide is solvent exposed and therefore has consistently high uptake values. **F** Mapping the uptake difference at 500 s (as a fraction of the theoretical maximum) onto the hetero-dimer protein structure shows that protection is centered on the dimer interface that we found in our crystal structure.

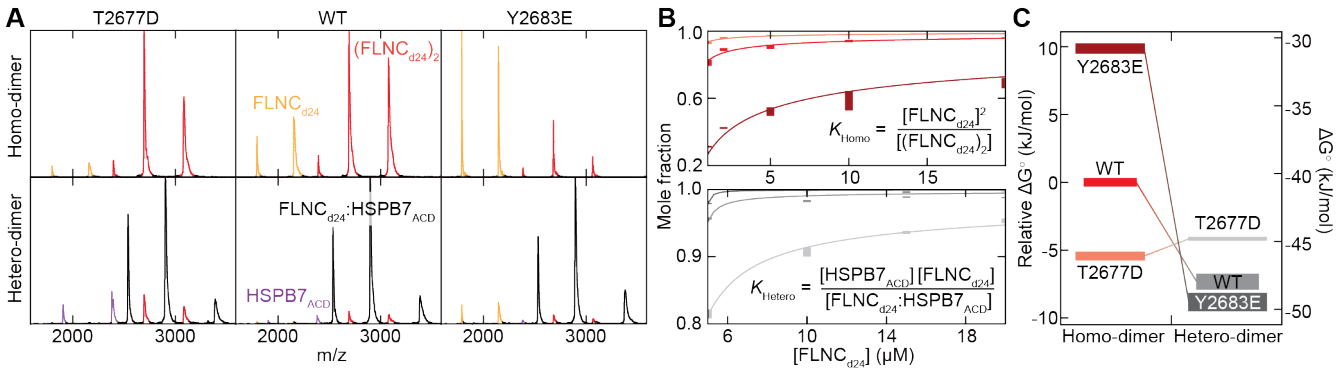

Fig. 4 | Mimicking phosphorylation at T2677 and Y2683 has opposite effects in modulating FLNC$_{d24}$ homo- and hetero-dimerisation. A Native mass spectra at 1 μM of FLNC$_{d24}$ WT (middle) and the phospho-mimics T2677D (left) and Y2683E (right) show differences in the relative abundances of monomer (yellow) and dimer (red) (upper row). T2677D displays a stabilised homo-dimer, relative to WT; Y2683 a destabilised dimer. Native MS spectra at 5 μM of each of HSPB7$_{ACD}$ and either FLNC$_{d24}$ WT or one of the phospho-mimics (lower row). Hetero-dimers are the dominant species in each of the spectra, but a notable abundance of FLNC (red, and yellow) and HSPB7$_{ACD}$ (purple) can be observed in the case of the T2677D mutant, and to a lesser extent for WT and Y2683E. **B** Titration curves for homo- (upper) and hetero-dimerisation (lower). WT: red and medium grey; T2677D: salmon and light grey; Y2683D: burgundy and dark grey. Error bars represent ± 1 standard deviation of the mean. **C** Free energy diagram showing the differences in stability for homo- and hetero-dimers in each of the three forms of FLNC. The vertical thickness of the bars corresponds to the error in our estimates. Lower free energies correspond to higher stability.

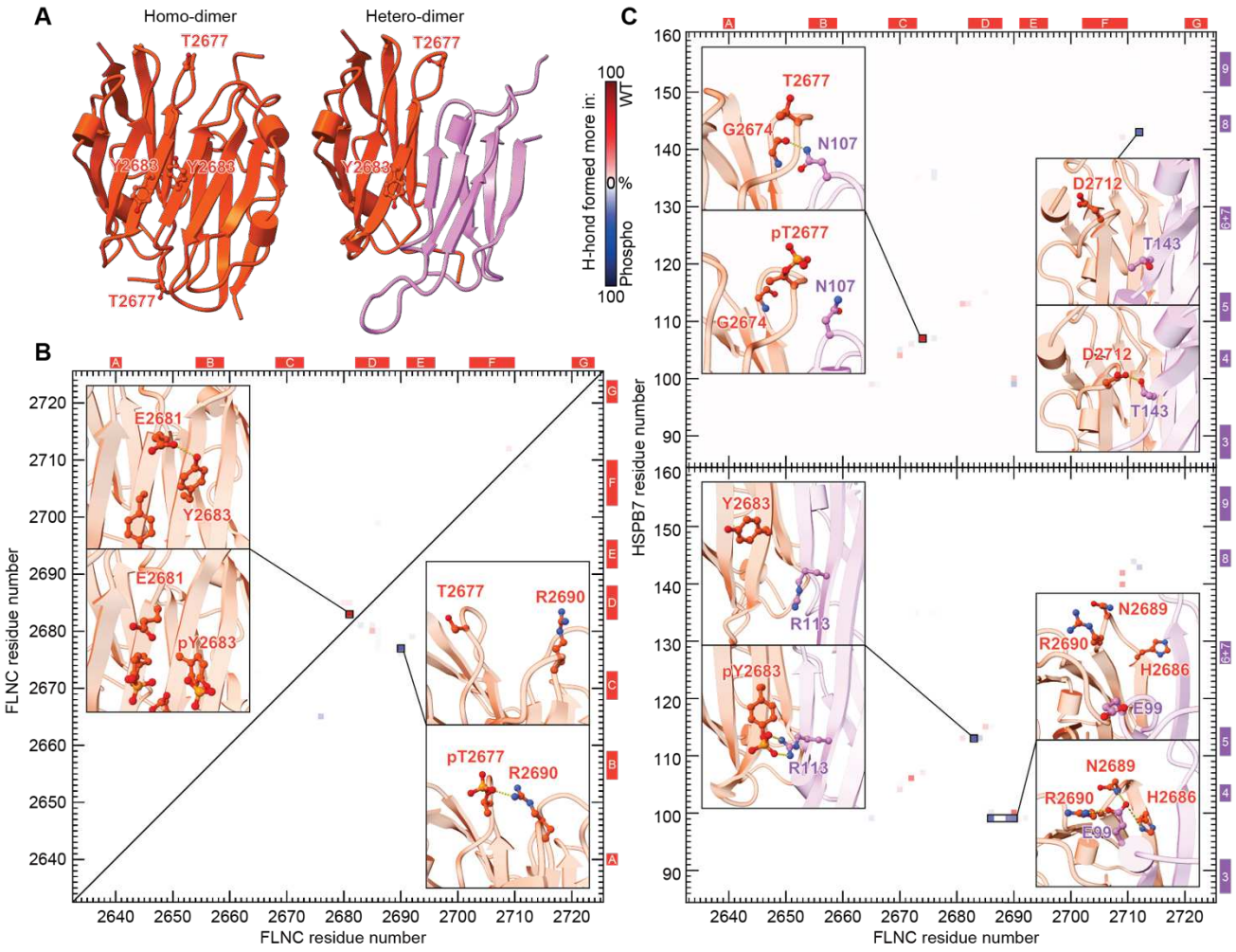

Fig.5 | MD simulations reveal how phosphorylation modulates the strength of the FLNC$_{d24}$ homo- and hetero-dimer interfaces. A The location of the T2677 and Y2683 phosphorylation sites on the FLNC$_{d24}$ homo-dimer (PDB 1V05) and FLNC$_{d24}$:HSPB7$_{ACD}^{C131S}$ hetero-dimer structures is near the interface. FLNC: orange; HSPB7: purple. **B** Difference map of hydrogen bond occupancy between WT and either pT2677 (lower right triangle) or pY2683 FLNC$_{d24}$ (upper left) homo-dimers. Each pixel corresponds to a particular inter-monomer contact between residue pairs. Blue: hydrogen bond formed more in the phosphorylated form; red: hydrogen bond formed more in the WT. A highly occupied hydrogen bond formed across the WT homo-dimer interface (between E2681 and Y2683) is lost upon phosphorylation of Y2683. And a new interaction with R2690 is formed upon phosphorylation of T2677 (**insets**). **C** Difference maps of hydrogen bond occupancy between WT and pT2677 (upper) or

pY2683 $FLNC_{d24}$:$HSPB7_{ACD}^{C131S}$ hetero-dimers (lower). Colouring as in **B**. The most significant differences are highlighted with representative frames from the trajectories showing the contacts made (**insets**). For instance, a highly occupied hydrogen bond formed across the WT hetero-dimer interface (between G2674 and N107) is lost upon phosphorylation of T2677; and a new interaction is formed by pY2683 with R113 that is absent in the WT. Differential interactions are also observed between residues that do not directly involved the phosphorylated site.

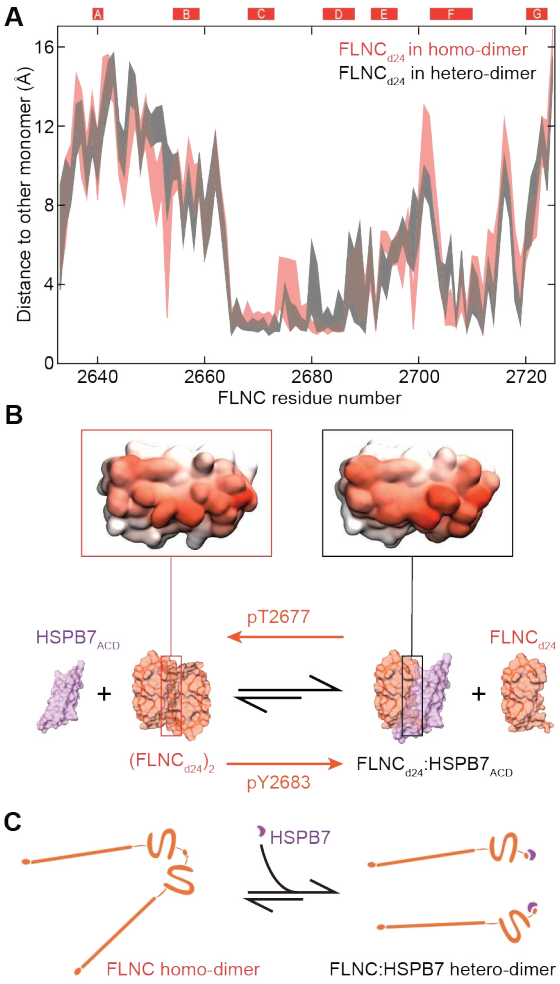

Fig. 6 | Overlapping interfaces for homo- and hetero-dimerisation lead to an equilibrium that can be regulated by phosphorylation to adjust FLNC mobility. A Plot of the distances between subunits in the homo- (orange) and hetero-dimer (black), as a function of sequence position in our MD simulations. The bands are defined by the closest (lower value) and average (upper value) distance over the length of the simulation. The two bands overlap well, demonstrating the similarity of the interfaces in the two dimers. **B** Colouring residues on the surface in proportion to the residual closest distances of a FLNC monomer reinforces the similarity in binding site (**insets**). This allows us a simple view of the equilibria involved, where $FLNC_{d24}$ homo- and hetero-dimerisation are competitive. The

equilibrium favours the hetero-dimer for WT FLNC, and can be shifted in either direction by phosphorylation. **C** The shifting of the equilibrium effectively adjusts the state of FLNC, either stabilising it in the actin-crosslinking-competent homo-dimer form, or effectively releasing it from this role by generating the much smaller hetero-dimer, which is likely much more mobile in the cell.

.

References

1. Tarone, G. & Brancaccio, M. Keep your heart in shape: Molecular chaperone networks for treating heart disease. *Cardiovasc Res* **102**, 346–361 (2014).

2. Zhou, A. X., Hartwig, J. H. & Akyürek, L. M. Filamins in cell signaling, transcription and organ development. *Trends Cell Biol* **20**, 113–123 (2010).

3. Razinia, Z., Mäkelä, T., Ylänne, J. & Calderwood, D. A. Filamins in Mechanosensing and Signaling. *Annu Rev Biophys* **41**, 227–246 (2012).

4. Van Der Ven, P. F. M. *et al.* Characterization of muscle filamin isoforms suggests a possible role of γ-Filamin/ABP-L in sarcomeric Z-disc formation. *Cell Motil Cytoskeleton* **45**, 149–162 (2000).

5. Leber, Y. *et al.* Filamin C is a highly dynamic protein associated with fast repair of myofibrillar microdamage. *Hum Mol Genet* **25**, 2776–2788 (2016).

6. Agarwal, R. *et al.* Filamin C Cardiomyopathy Variants Cause Protein and Lysosome Accumulation. *Circ Res* **129**, 751–766 (2021).

7. Shatunov, A. *et al.* In-frame deletion in the seventh immunoglobulin-like repeat of filamin C in a family with myofibrillar myopathy. *Eur J Hum Genet* **17**, 656–663 (2009).

8. Qiao, J. *et al.* Filamin C, a dysregulated protein in cancer revealed by label-free quantitative proteomic analyses of human gastric cancer cells. *Oncotarget* **6**, 1171–1189 (2015).

9. Pal, J. *et al.* Systematic analysis of migration factors by MigExpress identifies essential cell migration control genes in non-small cell lung cancer. *Mol Oncol* **15**, 1797–1817 (2021).

10. Pudas, R., Kiema, T. R., Butler, P. J. G., Stewart, M. & Ylänne, J. Structural basis for vertebrate filamin dimerization. *Structure* **13**, 111–119 (2005).

11. Ruparelia, A. A., Oorschot, V., Ramm, G. & Bryson-Richardson, R. J. FLNC myofibrillar myopathy results from impaired autophagy and protein insufficiency. *Hum Mol Genet* **25**, 2131–2142 (2016).

12. Schuld, J. *et al.* Homozygous expression of the myofibrillar myopathy-associated p.W2710X filamin C variant reveals major pathomechanisms of sarcomeric lesion formation. *Acta Neuropathol Commun* **8**, 1–19 (2020).

13. Vorgerd, M. *et al.* A mutation in the dimerization domain of filamin c causes a novel type of autosomal dominant myofibrillar myopathy. *Am J Hum Genet* **77**, 297–304 (2005).

14. Reimann, L. *et al.* Myofibrillar Z-discs are a protein phosphorylation hot spot with protein kinase C (PKCα) modulating protein dynamics. *Mol Cell Proteomics* **16**, 346–367 (2017).

15. Reimann, L. *et al.* Phosphoproteomics identifies dual-site phosphorylation in an extended basophilic motif regulating FILIP1-mediated degradation of filamin-C. *Commun Biol* **3**, 1–19 (2020).

16. Collier, M. P. *et al.* HspB1 phosphorylation regulates its intramolecular dynamics and mechanosensitive molecular chaperone interaction with filamin C. *Sci Adv* **5**, (2019).

17. Juo, L. Y. *et al.* HSPB7 interacts with dimerized FLNC and its absence results in progressive m. *J Cell Sci* **129**, 1661–1670 (2016).

18. Liao, W. C., Juo, L. Y., Shih, Y. L., Chen, Y. H. & Yan, Y. T. HSPB7 prevents cardiac conduction system defect through maintaining intercalated disc integrity. *PLoS Genet* **13**, e1006984 (2017).

19. Krief, S. *et al.* Identification and Characterization of cvHsp. *J Biol Chem* **274**, (1999).

20. Golenhofen, N., Perng, M. Der, Quinlan, R. A. & Drenckhahn, D. Comparison of the small heat shock proteins αb-crystallin, MKBP, HSP25, HSP20, and cvHSP in heart and skeletal muscle.

Histochem Cell Biol **122**, 415–425 (2004).

21. Doran, P., Gannon, J., O'Connell, K. & Ohlendieck, K. Aging skeletal muscle shows a drastic increase in the small heat shock proteins αB-crystallin/HspB5 and cvHsp/HspB7. *Eur J Cell Biol* **86**, 629–640 (2007).

22. Wu, T. *et al.* HSPB7 is indispensable for heart development by modulating actin filament assembly. *Proc Natl Acad Sci U S A* **114**, 11956–11961 (2017).

23. Mercer, E. J., Lin, Y. F., Cohen-Gould, L. & Evans, T. Hspb7 is a cardioprotective chaperone facilitating sarcomeric proteostasis. *Dev Biol* **435**, 41–55 (2018).

24. Kappé, G. *et al.* The human genome encodes 10 α-crystallin-related small heat shock proteins: HspB1-10. *Cell Stress Chaperones* **8**, 53–61 (2003).

25. Vos, M. J. *et al.* HSPB7 is the most potent polyQ aggregation suppressor within the HSPB family of molecular chaperones. *Hum Mol Genet* **19**, 4677–4693 (2010).

26. Mymrikov, E. V., Daake, M., Richter, B., Haslbeck, M. & Buchner, J. The chaperone activity and substrate spectrum of human small heat shock proteins. *J Biol Chem* **292**, 672–684 (2017).

27. Vos, M. J., Kanon, B. & Kampinga, H. H. HSPB7 is a SC35 speckle resident small heat shock protein. *Biochim Biophys Acta Mol Cell Res* **1793**, 1343–1353 (2009).

28. Hochberg, G. K. A. *et al.* The structured core domain of αB-crystallin can prevent amyloid fibrillation and associated toxicity. *Proc Natl Acad Sci U S A* **111**, (2014).

29. Hilton, G. R. *et al.* C-terminal interactions mediate the quaternary dynamics of αB-crystallin. *Philos Trans R Soc B* **368**, 1–13 (2013).

30. Baldwin, A. J. *et al.* Probing dynamic conformations of the high-molecular-weight αb-crystallin heat shock protein ensemble by NMR spectroscopy. *J Am Chem Soc* **134**, 15343–15350 (2012).

31. Muranova, L. K., Shatov, V. M., Bukach, O. V. & Gusev, N. B. Cardio-Vascular Heat Shock Protein (cvHsp, HspB7), an Unusual Representative of Small Heat Shock Protein Family. *Biochemistry (Moscow)* **86**, S1–S11 (2021).

32. Collier, M. P. & Benesch, J. L. P. Small heat-shock proteins and their role in mechanical stress. *Cell Stress Chaperones* **25**, 601–613 (2020).

33. Freilich, R. *et al.* Competing protein-protein interactions regulate binding of Hsp27 to its client protein tau. *Nat Commun* **9**, (2018).

34. Baranova, E. V. *et al.* Three-dimensional structure of α-crystallin domain dimers of human small heat shock proteins HSPB1 and HSPB6. *J Mol Biol* **411**, 110–122 (2011).

35. Laganowsky, A. *et al.* Crystal structures of truncated alphaA and alphaB crystallins reveal structural mechanisms of polydispersity important for eye lens function. *Protein Sci* **19**, 1031–1043 (2010).

36. Alderson, T. R., Ying, J., Bax, A., Benesch, J. L. P. & Baldwin, A. J. Conditional Disorder in Small Heat-shock Proteins. *J Mol Biol* **432**, 3033–3049 (2020).

37. Masson, G. R. *et al.* Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat Methods* **16**, 595–602 (2019).

38. Reitz, C. J. *et al.* Proteomics and phosphoproteomics of failing human left ventricle identifies dilated cardiomyopathy-associated phosphorylation of CTNNA3. *Proc Natl Acad Sci U S A* **120**, (2023).

39. Bouhaddou, M. *et al.* The Global Phosphorylation Landscape of SARS-CoV-2 Infection. *Cell* **182**, 685-712.e19 (2020).

40. Bonow, R. O., Fonarow, G. C., O'Gara, P. T. & Yancy, C. W. Association of Coronavirus Disease 2019 (COVID-19) with Myocardial Injury and Mortality. *JAMA Cardiol* **5**, 751–753 (2020).

41. Shah, R. M., Shah, M., Shah, S., Li, A. & Jauhar, S. Takotsubo Syndrome and COVID-19: Associations and Implications. *Curr Probl Cardiol* **46**, (2021).

42. Nelson, M. E. *et al.* Phosphoproteomics reveals conserved exercise-stimulated signaling and AMPK regulation of store-operated calcium entry. *EMBO J* **38**, (2019).

43. Hoffman, N. J. *et al.* Global Phosphoproteomic Analysis of Human Skeletal Muscle Reveals a Network of Exercise-Regulated Kinases and AMPK Substrates. *Cell Metab* **22**, 922–935 (2015).

44. Bai, Y. *et al.* Phosphoproteomics identifies driver tyrosine kinases in sarcoma cell lines and tumors. *Cancer Res* **72**, 2501–2511 (2012).

45. Ae, N. H., Horn, A. H. C., Ae, H. L. & Sticht, H. AMBER force-field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine. *J Mol Model* **12**, 281–289 (2006).

46. Ferrández-Roldán, A. *et al.* Cardiopharyngeal deconstruction and ancestral tunicate sessility. *Nature* **599**, 431–435 (2021).

47. Manning, G., Plowman, G. D., Hunter, T. & Sudarsanam, S. Evolution of protein kinase signaling from yeast to man. *Trends Biochem Sci* **27**, 514–520 (2002).

48. Raynaud, F., Jond-Necand, C., Marcilhac, A., Fürst, D. O. & Benyamin, Y. Calpain 1-gamma filamin interaction in muscle cells: a possible in situ regulation by PKC-alpha. *Int J Biochem Cell Biol* **38**, (2006).

49. Gabai, V. L., Sherman, M. Y. & Sherman, M. Y. highlighted topics Molecular Biology of Thermoregulation Invited Review: Interplay between molecular chaperones and signaling

pathways in survival of heat shock. *J Appl Physiol* **92**, 1743–1748 (2002).

50. Shor, A. C., Agresta, S. V., D'Amato, G. Z. & Sondak, V. K. Therapeutic potential of directed tyrosine kinase inhibitor therapy in sarcomas. *Cancer Control* **15**, 47–54 (2008).

51. Du, Z. & Lovly, C. M. Mechanisms of receptor tyrosine kinase activation in cancer. *Mol Cancer* **17**, 1–13 (2018).

52. Haslbeck, M., Weinkauf, S. & Buchner, J. Small heat shock proteins: Simplicity meets complexity. *J Biol Chem* **294**, 2121–2132 (2019).

53. Kim, M. S. *et al.* A draft map of the human proteome. *Nature* **509**, 575–581 (2014).

54. Winter, G. Xia2: An expert system for macromolecular crystallography data reduction. *J Appl Crystallogr* **43**, 186–190 (2010).

55. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861–877 (2019).

56. Song, Y. *et al.* High-resolution comparative modeling with RosettaCM. *Structure* **21**, 1735–1742 (2013).

57. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132 (2004).

58. Joosten, R. P., Joosten, K., Murshudov, G. N. & Perrakis, A. PDB-REDO: Constructive validation, more than just looking for errors. *Acta Crystallogr D Biol Crystallogr* **68**, 484–496 (2012).

59. Agirre, J. *et al.* The CCP4 suite: integrative software for macromolecular crystallography. *Acta Crystallogr D Struct Biol* **79**, 449–461 (2023).

60. Kondrat, F. D. L., Struwe, W. B. & Benesch, J. L. P. Native mass spectrometry: Towards high-

throughput structural proteomics. In *Structural Proteomics: High-Throughput Methods: Second Edition* (2014). doi:10.1007/978-1-4939-2230-7_18.

61. Hopper, J. T. S., Sokratous, K. & Oldham, N. J. Charge state and adduct reduction in electrospray ionization-mass spectrometry using solvent vapor exposure. *Anal Biochem* **421**, 788–790 (2012).

62. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **11**, 3696–3713 (2015).

63. Phillips, J. C. *et al.* Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys* **153**, (2020).

64. Gowers, R. J. *et al.* MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proc. 15th Python in Science Conf* (2016).

65. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J Comput Chem* **32**, 2319–2327 (2011).

66. Lau, A. M., Claesen, J., Hansen, K. & Politis, A. Deuteros 2.0: Peptide-level significance testing of data from hydrogen deuterium exchange mass spectrometry. *Bioinformatics* **37**, 270–272 (2021).

67. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat Methods* **19**, 679–682 (2022).

68. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).