

OPEN

Solution structure of the autophagy-related protein LC3C reveals a polyproline II motif on a mobile tether with phosphorylation site

Carsten Krichel^{1,2}, Christina Möckel^{1,2}, Oliver Schillinger^{1,3}, Pitter F. Huesgen⁴, Heinrich Sticht⁵, Birgit Strodel^{1,3}, Oliver H. Weiergräber¹, Dieter Willbold^{1,2} & Philipp Neudecker^{1,2}

(Macro-)autophagy is a compartmental degradation pathway conserved from yeast to mammals. The yeast protein Atg8 mediates membrane tethering/hemifusion and cargo recruitment and is essential for autophagy. The human MAP1LC3/GABARAP family proteins show high sequence identity with Atg8, but MAP1LC3C is distinguished by a conspicuous amino-terminal extension with unknown functional significance. We have determined the high-resolution three-dimensional structure and measured the backbone dynamics of MAP1LC3C by NMR spectroscopy. From Ser18 to Ala120, MAP1LC3C forms an α -helix followed by the ubiquitin-like tertiary fold with two hydrophobic binding pockets used by MAP1LC3/GABARAP proteins to recognize targets presenting LC3-interacting regions (LIRs). Unlike other MAP1LC3/GABARAP proteins, the amino-terminal region of MAP1LC3C does not form a stable helix α_1 but a “sticky arm” consisting of a polyproline II motif on a flexible linker. Ser18 at the interface between this linker and the structural core can be phosphorylated *in vitro* by protein kinase A, which causes additional conformational heterogeneity as monitored by NMR spectroscopy and molecular dynamics simulations, including changes in the LIR-binding interface. Based on these results we propose that the amino-terminal polyproline II motif mediates specific interactions with the microtubule cytoskeleton and that Ser18 phosphorylation modulates the interplay of MAP1LC3C with its various target proteins.

Macroautophagy - hereafter termed autophagy - is an intracellular lysosomal degradation pathway conserved in eukaryotes¹. Upon initiation of autophagy a newly emerging double membrane structure, the phagophore, engulfs cytoplasmic targets and closes to form a double membrane vesicle, the autophagosome, which then fuses with lysosomal organelles for cargo degradation. The protein Atg8 (autophagy-related 8) is essential for autophagosome genesis in the yeast *Saccharomyces cerevisiae*². The carboxy-terminal residue of Atg8, Arg117 (Fig. 1), is removed by the *S. cerevisiae* cysteine protease Atg4 to expose a conserved glycine, Gly116, at the C-terminus. Subsequently, this exposed glycine can enzymatically be conjugated covalently to the membrane lipid phosphatidylethanolamine and thereby tethered to the growing phagophore³.

In humans, the Atg8 gene has diversified into at least eight orthologs, which can be divided into two subfamilies: the GABARAP (γ -aminobutyric acid type A receptor associated protein)-like and MAP1LC3 (microtubule associated proteins 1 light chain 3)-like proteins^{4,5}. The functional role of this diversification is still poorly

¹ICS-6 (Strukturbiochemie) and JuStruct, Forschungszentrum Jülich, 52425, Jülich, Germany. ²Institut für Physikalische Biologie and BMFZ, Heinrich-Heine-Universität Düsseldorf, 40225, Düsseldorf, Germany. ³Institut für Theoretische Chemie und Computerchemie, Heinrich-Heine-Universität Düsseldorf, 40225, Düsseldorf, Germany. ⁴ZEA-3 (Analytik), Forschungszentrum Jülich, 52425, Jülich, Germany. ⁵Institut für Biochemie, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91054, Erlangen, Germany. Correspondence and requests for materials should be addressed to D.W. (email: d.willbold@fz-juelich.de) or P.N. (email: p.neudecker@fz-juelich.de)

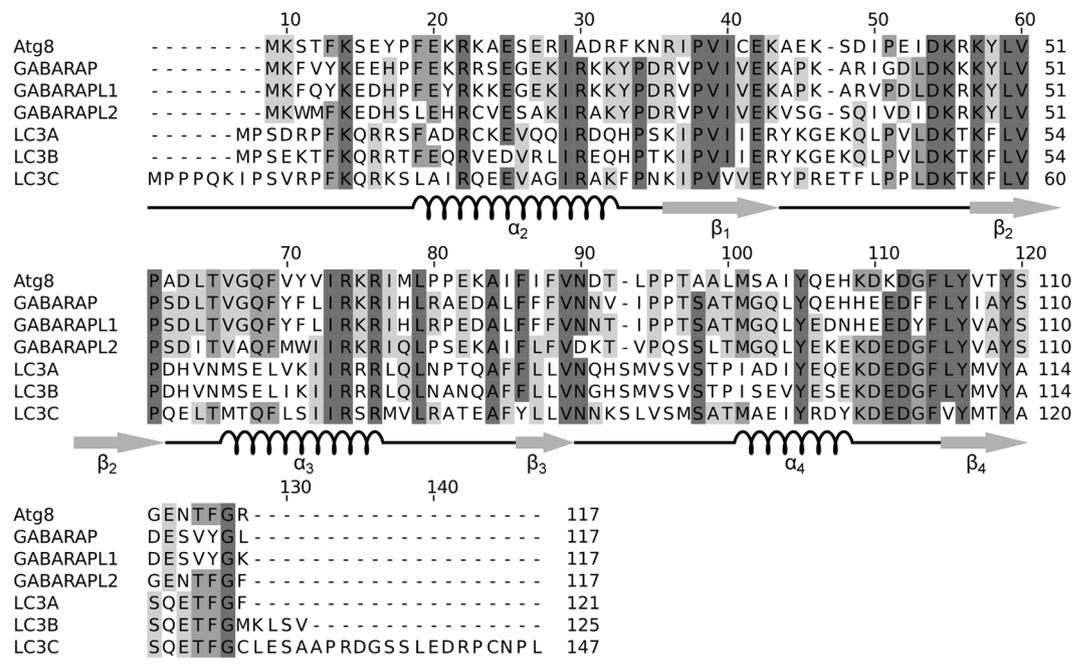


Figure 1. Sequence alignment of *S. cerevisiae* Atg8 (Uniprot P38182) and the canonical human homologues GABARAP (O95166), GABARAPL1/GEC1 (Q9H0R8), GABARAPL2/GATE-16 (P60520), and LC3A (Q9H492), LC3B (Q9GZQ8), and LC3C (Q9BXW4). Highly conserved residues are highlighted in shades of gray; darker shading indicates higher conservation. All members of the Atg8 family can be cleaved by the cysteine protease family ATG4 after the conserved Gly126, which can subsequently be conjugated to a phospholipid moiety for membrane anchoring. The alignment was performed using Jalview version 2⁹⁸. The regular secondary structure elements of LC3C are indicated below the alignment.

understood. Like Atg8, these proteins are initially expressed as longer precursor proteins and then cleaved by the *H. sapiens* ATG4 family of cysteine proteases to expose the conserved glycine (Gly126 in the case of MAP1LC3C; Fig. 1) at the C-terminus. The high-resolution three-dimensional structure of several of these proteins has been determined experimentally by X-ray crystallography and/or NMR spectroscopy (reviewed by Weiergräber *et al.*⁵). In particular, solution structures have been reported for yeast Atg8^{6–8}, for the human GABARAP-like proteins GABARAP⁹ and GABARAPL1¹⁰, and for the MAP1LC3-like proteins MAP1LC3A¹¹, MAP1LC3A in complex with SQSTM1¹², and MAP1LC3B in complex with optineurin¹³. Recently, crystal structures for MAP1LC3C in the free form¹⁴ (PDB 3WAM) as well as in complex with Atg13¹⁴ (PDB 3WAP) and in complex with NDP52¹⁵ (PDB 3VWV) were published. All these Atg8 homologs share a ubiquitin-like core structural motif consisting of a four-stranded β -sheet flanked by two α -helices, which is extended by an amino-terminal α -helical domain. The human MAP1LC3 (or simply LC3) proteins differ most significantly in the amino-terminal region (NTR) preceding this α -helical domain, which is considerably longer in LC3C as compared to LC3A and LC3B (Fig. 1). While for the latter two proteins a role of the NTR in membrane tethering^{2,16} or in recognition of mitochondrial phospholipids during mitophagy has been reported¹⁷, the functional implications of the longer NTR found in LC3C are still poorly understood.

Autophagy research has historically focused on unspecific (bulk) degradation of cytosolic targets. Recently, however, mammalian autophagy receptor proteins have been described that mediate selective autophagic pathways^{18,19}. Recruitment of selective autophagic degradation targets to LC3-like proteins is mediated by the target's conserved LC3-interacting region (LIR) of four amino acids with a canonical sequence motif (W/Y/F-x-x-L/I/V)^{12,18,20}. In contrast to other LC3 proteins, LIR binding to LC3C in selective autophagy can also be mediated by a non-canonical LIR motif (CLIR)¹⁵. Upon binding to LC3 the two large hydrophobic side chains flanking the LIR of the target insert into two hydrophobic pockets on the surface of LC3 and the backbone of the LIR forms an additional, parallel β -strand to extend the core LC3 β -sheet separating the two hydrophobic pockets¹².

Autophagy is regulated by a variety of mechanisms, including the modulation of LC3 protein-protein interactions via post-translational modifications. On the one hand, the affinity of binding to LC3 can be enhanced by modification of the target LIR^{19,21}. On the other hand, post-translational modifications of the LC3 proteins themselves have been described. Intriguingly, the post-translational modification sites are predominantly localized within the NTR of LC3. Cherra *et al.*²² reported the down-regulation of autophagy upon phosphorylation of LC3A at Ser12 by protein kinase A (PKA), while Jiang *et al.*²³ studied the influence of phosphorylation of LC3B at Thr6 and Thr26 by protein kinase C (PKC) on autophagosome formation. Both studies concluded that an increased cellular protein kinase activity could attenuate autophagy. Structural studies, however, were beyond the scope of these works. LC3C shares the PKA phosphorylation site mentioned above (Ser18 in this case) with LC3A and LC3B²², and, unlike other LC3-proteins, contains a second PKC phosphorylation site at Ser9 as predicted

Statistics of the NMR ensemble	
Conformational restraints	
NOE distance restraints	
Total	1162
Intraresidual (i = j)	202
Sequential (i - j = 1)	387
Medium range (1 < i - j < 5)	266
Long range (i - j ≥ 5)	307
Ambiguous	57
Hydrogen bond distance restraints	102
Dihedral angle restraints	214
Residual restraint violations	
Average number of distance restraint violations per structure	
>0.1 Å	7.8
>0.3 Å	0.2 (max. 0.34 Å)
Average number of dihedral angle restraint violations	
>1°	10.2
>3°	0.4 (max. 3.77°)
Atomic RMSDs from the average structure^a [Å]	
Backbone atoms	0.72 ± 0.10
Heavy atoms	1.30 ± 0.12
MolProbity Ramachandran statistics [%]	
Favored regions	91.6
Allowed regions	99.2
Model content	
Total no. of residues	126
Ordered residue range	18–120
BMRB accession number	26603
PDB ID code	2NCN

Table 1. Statistics of the NMR ensemble (n = 10). ^aOrdered residue range (18–120).

by computational algorithms²⁴. Furthermore, Huang *et al.*²⁵ reported the influence of de-acetylation of a lysine residue in nuclear-localised LC3, suggesting further modulating options for LC3 activity in its interplay with the autophagic machinery. Other diverse post-translational modifications of autophagy-related proteins have been reviewed recently²⁶.

In order to provide a structural basis for understanding the functional role of the NTR of LC3C on the atomic level we have determined the three-dimensional structure of human MAP1LC3C directly in solution and investigated the backbone dynamics and the influence of *in vitro* phosphorylation by PKA using NMR spectroscopy.

Results and Discussion

Description of the structure in solution. As reported previously²⁷, inspection of the NMR spectra and analysis of the ¹⁵N relaxation and relaxation dispersion experiments (see below) reveal that LC3C exhibits extensive conformational heterogeneity in several regions of the protein, which results in significant conformational averaging and/or exchange line broadening of the NMR spectra (Supplementary Fig. S3). Accordingly, manual inspection of the experimental restraints was necessary to exclude artificially restrictive restraints derived from conformationally averaged spectral parameters such as NOEs (Supplementary Fig. S4) or ³J_{HNH α} couplings. An ensemble of ten solution structures was calculated from 1162 NOE distance restraints, 214 dihedral angle restraints, and 102 hydrogen bond distance restraints. Of note, all three α -helices feature a slowly exchanging backbone amide hydrogen donor involved in N-capping hydrogen bonds to side chain oxygen atoms (α_2 : Ile21 HN – Ser18 O γ ; α_3 : Gln68 HN – Thr65 O γ 1; α_4 : Glu103 HN – Thr100 O γ 1). Experimental restraints and statistics of the structural ensemble are summarized in Table 1.

The overall structure of LC3C features a globular, ubiquitin-like fold consisting of four β -strands, three α -helices, and polymorphic terminal regions (see Fig. 2). The core of LC3C is formed by a twisted, mixed parallel/antiparallel β -sheet consisting of two inner β -strands (β_1 , β_4) pairing in a parallel orientation, each paired by an antiparallel β -strand (β_2 and β_3 , respectively) on the outside. The three α -helices α_2 , α_3 , and α_4 (this numbering has been chosen for consistency with other members of the Atg8 family that have an additional α -helix at the N-terminus) are flanking the inner β -sheet on both faces, with helix α_3 in an orientation that is approximately perpendicular to that of α_2 and α_4 . It is this arrangement of the regular secondary structure elements that forms the two characteristic hydrophobic surface pockets, hp1 between α_2 and β_2 on one face of the β -sheet and hp2 between α_3 and β_2 on the other face (Fig. 2), which allow Atg8 family proteins to bind canonical or non-canonical LC3-interacting regions (LIRs) via β -strand pairing of the LIR backbone to the edge strand β_2 .

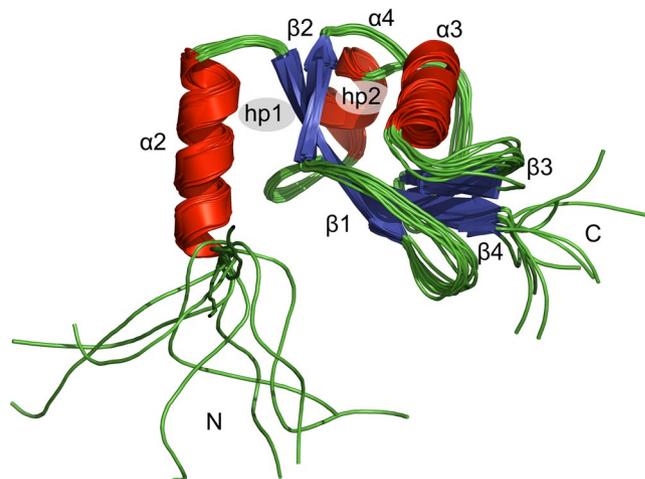


Figure 2. Solution structure of LC3C shown as a superposition of the backbone traces of the 10 accepted structures with a schematic representation of the secondary structure elements as identified with DSSP⁹⁹. The tertiary structure of LC3C consists of three α -helices (α_2 : Leu19 to Lys32, α_3 : Met66 to Arg76, α_4 : Met101 to Tyr108; red) and a central β -sheet of four β -strands (β_1 : Lys36 to Arg43, β_2 : Lys57 to Pro61, β_3 : Tyr86 to Val89, β_4 : Val115 to Ala120; blue). In the ubiquitin-like core of the protein the 10 accepted structures are in excellent agreement, whereas the amino (N) and carboxy (C) termini as well as some of the loops are highly flexible in solution. The location of the two hydrophobic pockets is indicated by hp1 and hp2. The overlay was performed using PyMOL (The PyMOL Molecular Graphics System, Version 1.7.2.1, Schrödinger, LLC).

while accommodating the two bulky hydrophobic side-chains of the LIR in hp1 and hp2¹². The side-chain of Lys55, which is located just N-terminal of strand β_2 and has been reported to restrict access to hp2 in order to contribute to LIR recognition¹⁴, is partially solvent-exposed in solution and sufficiently flexible across the ensemble of structures to allow the formation of a complex with a specifically recognized ligand (not shown).

The tertiary structure of LC3C is stabilized not just by hydrogen bonds but also by a variety of hydrophobic interactions involving aliphatic and aromatic side-chains. Tertiary contacts with aromatic side-chains give rise to several readily identifiable ¹H chemical shift outliers due to ring current effects. Most notably, the backbone amide proton of Arg46 displays a remarkable upfield shift (5.50 ppm), which is caused by the vicinity to the aromatic side-chain of Tyr44 and has been reported previously²⁷. Unfortunately, this chemical shift is too close to the residual ¹H₂O signal (4.82 ppm at 20.0 °C) not to be affected by band-selective water flip-back pulses²⁸ and the amide resonance of Arg46 is therefore visible only with artificially reduced intensity in amide-detected NMR experiments (Supplementary Fig. S3). Additional upfield shifts of ¹H chemical shifts were also identified for one of the two methyl groups of Val40 (−0.36 ppm; Supplementary Fig. S4), which is positioned in between the aromatic side chains of Phe114 and Tyr116, for the ¹H δ 1 methyl group of Ile73 (−0.15 ppm), which is oriented towards Phe69, as well as for one of the two methyl groups of Val26 (0.18 ppm; Supplementary Fig. S4), which faces the aromatic ring of Phe114, and for one of the two protons of the ¹H δ methylene group of Lys109 (0.68 ppm), which is close to Tyr105. Strong downfield shifts of ¹H chemical shifts were found for one of the two protons of the ¹H β methylene group of Asp110 (3.79 ppm), which is in the vicinity of the aromatic side-chain of Tyr116, for Thr118 ¹H α (6.28 ppm), which is adjacent to the aromatic side-chain of Tyr119, and for Tyr119 ¹H α (6.15 ppm) itself.

Conformational dynamics. While the structural core from Ser18 to Ala120 is well defined in solution with an average atomic RMSD from the average structure of 0.72 Å for the backbone and 1.30 Å for all heavy atoms (Tab. 1), several regions of the protein are constrained by only very few long-range NOE restraints and therefore show markedly increased atomic RMSDs (Supplementary Fig. S5), namely the N-terminus (Met1 to Lys17), the C-terminus (Ser121 to Gly126), and – to a lesser extent – the loops between strands β_1 and β_2 (Tyr44 to Thr56), between helix α_3 and strand β_3 (Met77 to Phe85), and between strand β_3 and helix α_4 (Asn90 to Ala99). Inspection of the ¹⁵N relaxation data (Fig. 3) reveals that these regions are indeed highly flexible as indicated by several {¹H}¹⁵N NOE values below 0.65²⁹, ¹⁵N transverse relaxation rates R_2 significantly slower than the average over residues Ser18 to Ala120 of 25.0/s \pm 6.7/s, and ¹⁵N longitudinal relaxation rates R_1 significantly faster than the average over residues Ser18 to Ala120 of 0.87/s \pm 0.07/s (Fig. 3). Accordingly, “model-free” analysis³⁰ of the ¹⁵N relaxation data reveals large-amplitude backbone motions on the sub-nanosecond time-scale that are reflected in low generalized order parameters S^2 at both termini and in some loop regions (Fig. 3D). By contrast, the regular secondary structure elements are quite rigid on the sub-nanosecond time-scale, with S^2 values consistently exceeding 80%. The ratios between ¹⁵N transverse and longitudinal relaxation rates of the rigid structural core of LC3C at 800 MHz are best fit by an oblate axially symmetric rotational diffusion tensor with eigenvalues of 1.39×10^7 /s and 1.30×10^7 /s at 20.0 °C, corresponding to overall rotational autocorrelation times³¹ of $\tau_A = 12.0$ ns, $\tau_B = 12.1$ ns, and $\tau_C = 12.6$ ns. In the approximation of fully isotropic rotational diffusion this reduces to a single autocorrelation time of $\tau_{iso} = 12.3$ ns. Although prediction of the hydrodynamics of LC3C on the basis of its tertiary structure is not straightforward because of the flexibility of the long terminal regions, these

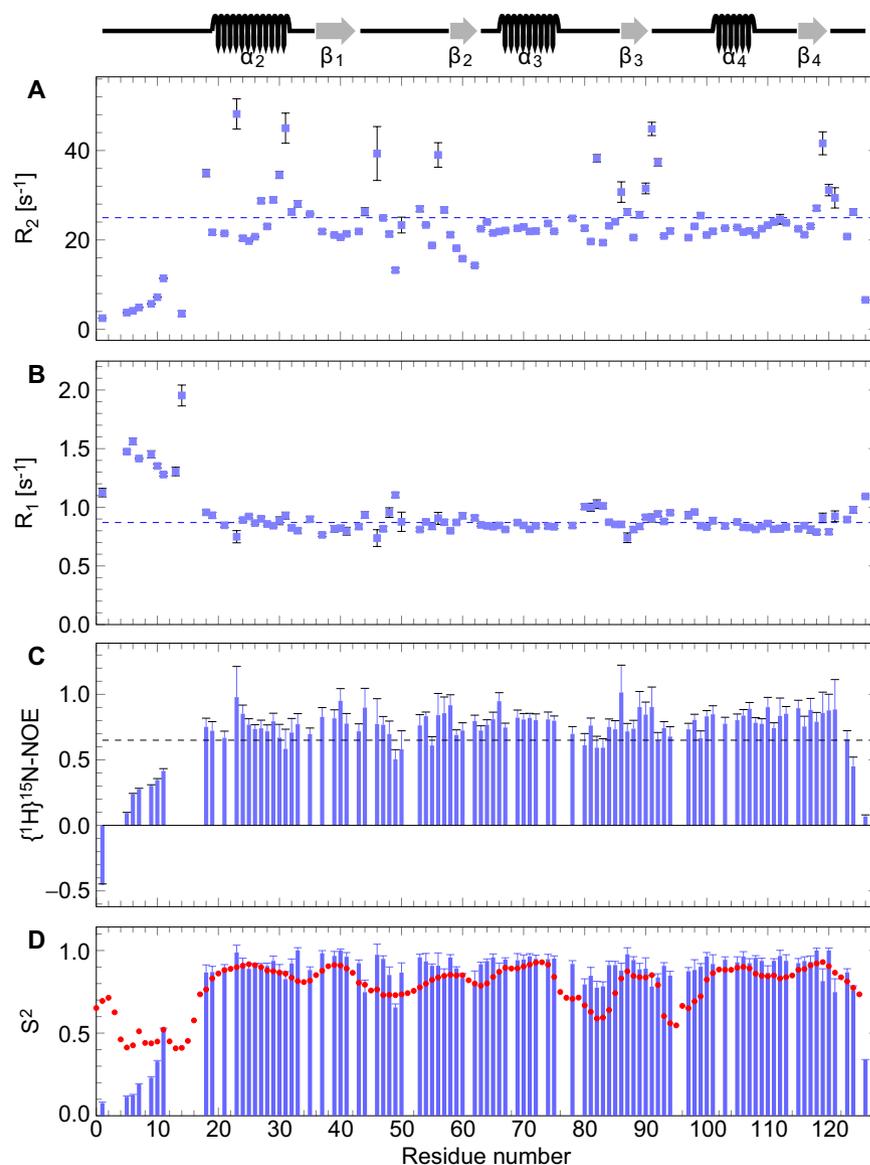


Figure 3. ^{15}N transverse (A) and longitudinal (B) relaxation rates and $\{^1\text{H}\}^{15}\text{N}$ heteronuclear NOE values (C) measured at 800 MHz and 20.0 °C, and generalized order parameters S^2 of the sub-nanosecond backbone amide motion (D). Dashed horizontal lines in (A,B) represent the average values over the ordered region (residues 18–120). $\{^1\text{H}\}^{15}\text{N}$ values below 0.65 (dashed horizontal line in (C)) indicate increased internal mobility²⁹. Experimental order parameters S^2 obtained from ^{15}N relaxation analysis (blue) are compared with order parameters predicted from the Random Coil Index (RCI)³³, S_{RCI}^2 (red circles), calculated with the default parameters as implemented in TALOS-N^{80,81}. The regular secondary structure elements are indicated above the graph.

autocorrelation times are slightly longer than but still consistent with the hydrodynamic properties expected for a predominantly monomeric protein the size of LC3C in the relatively viscous buffer used here. GABARAP, which is slightly smaller than LC3C (Fig. 1) and whose NTR is less mobile, shows an overall rotational autocorrelation time at 20 °C of $9.2 \text{ ns} \pm 0.4 \text{ ns}$ at low concentration and about 11 ns under conditions typically used for NMR spectroscopy³².

In addition to the generalized order parameters S^2 from ^{15}N relaxation, which are a measure of the motional restriction of the amide bond vector orientation with respect to the molecular frame on the sub-nanosecond time-scale on a scale of 0 (unrestricted motion) to 1 (rigid), we also calculated the Random Coil Index (RCI)³³ order parameters, S_{RCI}^2 , which are a measure of how different the backbone chemical shifts are from those of a disordered random coil on a scale of 0 (typical for a random coil) to 1 (typical for a well-ordered backbone conformation). In general, the generalized order parameters from ^{15}N relaxation, S^2 , are in good agreement with the RCI order parameters, S_{RCI}^2 , with one very conspicuous exception (Fig. 3D): Although the orientation of the NTR relative to the structural core of the molecule is largely disordered ($S^2 < 50\%$), the local backbone conformation of the first four residues is not consistent with a random coil ($S_{\text{RCI}}^2 > 60\%$). Instead, a local backbone overlay of these residues reveals that the three consecutive proline residues Pro2 to Pro4 adopt a canonical polyproline II (PPII)

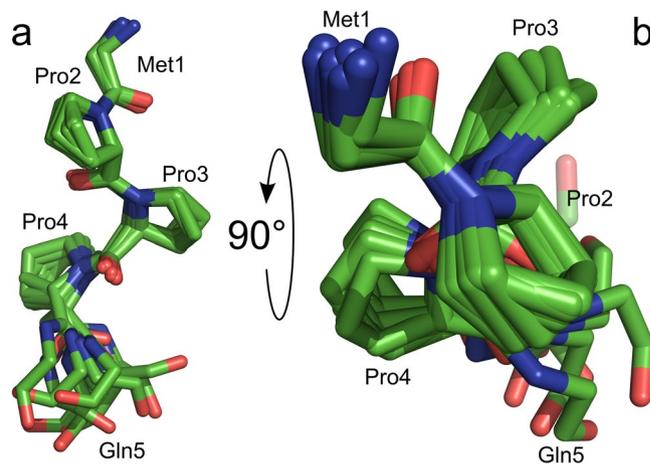


Figure 4. The amino-terminal polyproline II motif of LC3C shown as a superposition of the 10 accepted structures. Only the backbone atoms of the two adjacent residues, Met1 and Gln5, are shown. Carbon, nitrogen, and oxygen atoms are color-coded green, blue, and red, respectively. (A) Side view, (B) view from Met1 along the helix axis showing the typical cloverleaf-like arrangement of the proline rings due to the three-fold symmetry of the left-handed PPII helix. Pairwise alignment of backbone atoms from Pro2 to Pro4 was performed using PyMOL (The PyMOL Molecular Graphics System, Version 1.7.2.1, Schrödinger, LLC).

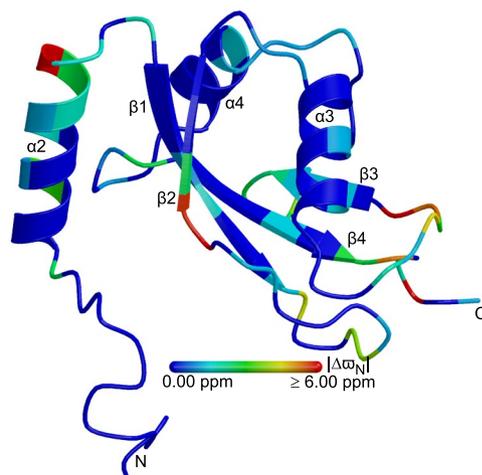


Figure 5. Schematic representation of the secondary structure of the lowest-energy solution structure of LC3C (similar view as in Fig. 2) colored according to the backbone amide ^{15}N chemical shift change between the major (population $98.37\% \pm 0.06\%$) and the minor (population $1.63\% \pm 0.06\%$) state, $|\Delta\omega_{\text{N}}|$, from 0.00 ppm (blue) to 6.00 ppm (red), or blue if $\Delta\omega_{\text{N}}$ could not be determined due to missing resonance assignments, resonance overlap, or if the exchange contributions to transverse relaxation were of insufficient magnitude. Values of $|\Delta\omega_{\text{N}}| > 1.5$ ppm (green to red) are localized to specific regions, indicating that the rest of the tertiary structure is not affected by the conformational exchange process. The figure was drawn with MolScript 2.1.2¹⁰⁰ and rendered with Raster3D 3.0¹⁰¹.

helical geometry (Fig. 4) with its typical backbone torsion angles $\Phi \approx -75^\circ$ and $\Psi \approx +120^\circ$ (Supplementary Fig. S6). From a structural point of view, the conformation of the unique NTR of LC3C (Fig. 1) in solution can therefore be described as an amino-terminal PPII motif attached to the ubiquitin-like core of the protein via a highly flexible 13-residue tether.

Whereas R_1 is solely sensitive to fast dynamics in the ns to ps regime, R_2 is also influenced by slower conformational exchange processes on the μs to ms time-scale³⁴. Inspection of the ^{15}N R_2 rates indeed shows many residues with ^{15}N transverse relaxation rates that are highly elevated by conformational exchange contributions (Fig. 3A). More detailed analysis using the “model-free” formalism³⁰ reveals that a large number of backbone amide groups covering virtually the entire structural core of the protein are affected by one or more conformational exchange processes on the μs to ms time-scale (data not shown), most prominently helix α_2 (especially its C-terminal side), the loop connecting strands β_1 and β_2 , the loops before and after strand β_3 , and the C-terminal region (Fig. 3A). These “hotspots” are corroborated by ^{15}N CPMG relaxation dispersion experiments

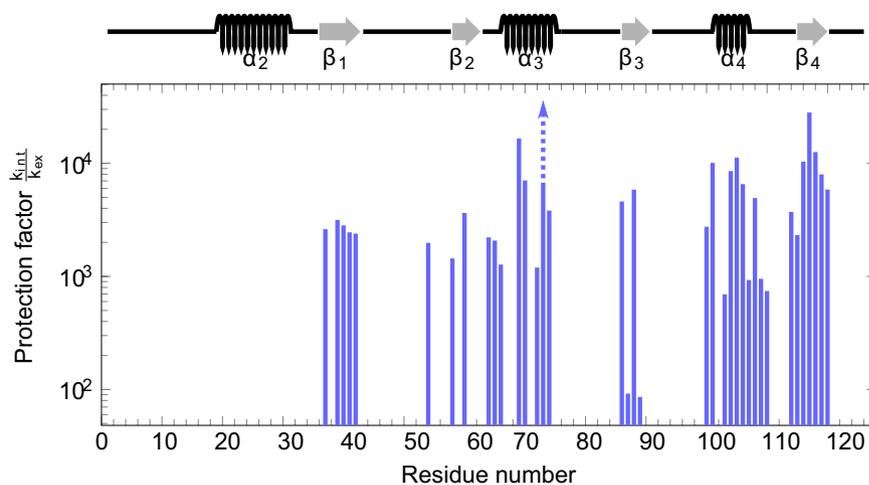


Figure 6. Logarithmic plot of the protection factors, $PF = k_{int}/k_{ex}$, calculated from the apparent backbone amide proton/deuteron (H/D) exchange rates, k_{ex} , as measured in a series of [^1H - ^{15}N] HSQC experiments at 20.0 °C (Supplementary Figs S9 and S10). Amide protons with protection factors below about 1×10^2 exchange in less than half an hour and hence too rapidly to be observed with sufficient resonance intensity for quantification. By contrast, amide protons involved in hydrogen bonding in the regular secondary structure elements, which are indicated above the graph, are highly protected from solvent exchange ($PF \approx 5 \times 10^2$ to 3×10^4). The backbone amide proton of Ile73 exchanges so slowly (Supplementary Fig. S9) that only an upper bound for k_{ex} and hence a lower bound for PF (indicated by the arrow) could be determined.

(Fig. 5). 57 backbone amide groups exhibit measurable and significant ^{15}N relaxation dispersions, which can be fit by a simultaneous two-state model with an exchange rate of $847/\text{s} \pm 46/\text{s}$ and a minor state population of $1.63\% \pm 0.06\%$ at 20.0 °C (Supplementary Fig. S7), suggesting that all these amide groups are involved in the same conformational exchange process, even though their location is non-contiguous in the tertiary structure (Fig. 5). With a root mean square deviation (RMSD) of 4.42 ppm, the ^{15}N chemical shift changes deviate markedly from those expected for a random coil³⁵ and are therefore inconsistent with a global or local unfolding process (Supplementary Fig. S8). Identification of the nature of this extensive conformational exchange process (for example, transient self-association or concerted/allosteric motions) has to await systematic quantification by relaxation dispersion spectroscopy for different nuclei at a series of different temperatures, which is beyond the scope of the present work.

Conformational exchange processes can also affect the backbone amide H/D exchange rates, k_{ex} . Amide protons involved in stable hydrogen bonds are protected from exchange with the solvent and therefore show very large protection factors, $PF = k_{int}/k_{ex}$, which measure how many times slower the protons exchange with the solvent in NMR buffer than under denaturing (random coil) conditions. However, if such a hydrogen-bonded (“closed”) conformation C is in equilibrium with an alternate conformation O that is not hydrogen-bonded (“open”) then solvent exchange will readily occur via the fraction p_o of the protein that populates the alternate conformation. In the case of conformational exchange on the micro- to millisecond time-scale at pH 6.0 and 20.0 °C it can be shown³⁶ that the resulting protection factors should be of the order of the inverse of the fractional population of the “open” state, $PF \approx 1/p_o$. As expected, the distribution of the highly protected backbone amide groups primarily reflects the hydrogen bonds stabilizing the regular secondary structure elements of LC3C (Fig. 6). By contrast, none of the amide groups in the flexible terminal regions is sufficiently protected from solvent exchange to allow experimental detection. Note that most of the amide protons in the loops identified above as “hotspots” of the conformational exchange on the millisecond time-scale also appear to exchange readily with the solvent (Fig. 6), indicating that the hydrogen bonding network of the regular secondary structure elements is stable and preserved by the conformational exchange processes; indeed, the high protection factors $PF \approx 2 \times 10^3$ to 3×10^4 found in all regular secondary structure elements except helix α_2 are at variance with those expected for an “open” state with a fractional population $p_o \gg 0.1\%$ large enough to give rise to significant exchange line broadening. In particular, this also rules out global unfolding of LC3C as the source of the observed exchange line broadening.

Comparison with X-ray structures. Overall, the NMR solution structure is in excellent agreement with the X-ray structures published for free LC3C(8–125)¹⁴ (PDB 3WAM), LC3C in complex with NDP52¹⁵ (PDB 3VWV), and LC3C(8–125) in complex with the LIR of Atg13¹⁴ (PDB 3WAP) (Fig. 7), with RMSDs of 0.65 Å, 0.89 Å, and 1.09 Å, respectively, for the backbone atoms of the structural core from Ser18 to Ala120. Minor differences are seen for the loops between strands β_1 and β_2 (Tyr44 to Thr56) and between helix α_3 and strand β_3 (Met77 to Phe85), which are involved in crystal contacts in the X-ray structures¹⁴ and undergo backbone motions on the pico- to nanosecond time-scale in solution (see above, Fig. 3). The largest differences, however, are observed for the NTR from the N-terminus to Lys17 (Fig. 7). In the crystal structure of LC3C in complex with NDP52¹⁵ (PDB 3VWV) residues Met1 to Pro12 could not be observed in the electron density map, consistent

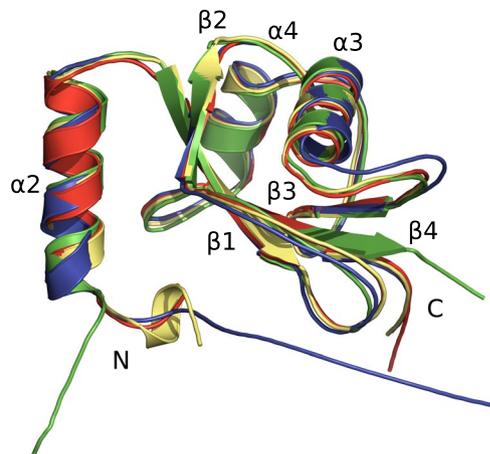


Figure 7. Comparison of a schematic representation of the secondary structure elements of the average solution structure of LC3C (green) with the crystal structures of LC3C in the free form (PDB 3WAM¹⁴, yellow; backbone RMSD 0.65 Å for residues Ser18 to Ala120) as well as in complex with NDP52 (PDB 3VWV¹⁵, red; 0.89 Å) and in complex with Atg13 (PDB 3WAP¹⁴, blue; 1.09 Å). The overlay and the calculation of backbone RMSDs were performed using PyMOL (The PyMOL Molecular Graphics System, Version 1.7.2.1, Schrödinger, LLC).

with our observation that this region exhibits no long-range order relative to the structural core of the protein (see above). The crystal structures of LC3C(8–125) in the free form and in complex with Atg13¹⁴ (PDB 3WAM and 3WAP, respectively) lack the additional six residues - including the polyproline motif - at the N-terminus of LC3C compared to LC3A and LC3B entirely (Fig. 1). In the absence of Atg13 no electron density could be found for residues before Arg11 but an additional helix α_1 is formed from Pro12 to Lys17, which is also found in many other proteins of the Atg8 family¹⁴. Whether this difference in conformation is a consequence of the different crystallization conditions, the different crystal packing, the different resolution of the electron density maps, or the presence of the LIR of Atg13 remains unclear. In solution, no medium-range NOE pattern characteristic of such an additional helix α_1 could be identified (Supplementary Fig. S5), and the low RCI order parameters S_{RCI}^2 are not indicative of any stable secondary structure in this region (Fig. 3D). However, some of the resonances in this region could not be identified in the NMR spectra²⁷, most likely as a result of exchange line broadening due to conformational heterogeneity, so it is possible that this region transiently forms a metastable α -helical conformation. In addition to the N-terminal region, there are also conformational differences in the C-terminal region: Whereas the C-terminal residues Ser121 to Gly126 fold back towards the face of the central β -sheet in the crystal structures (Fig. 7), this region is poorly defined in solution (Fig. 2) by the sparse medium- and long-range NOEs (Supplementary Fig. S5), which reflects high mobility on the ps to ns time-scale (see above) and line broadening due to conformational exchange on the μ s to ms time-scale (Figs 3A and 5). The mobility of the C-terminus in solution ensures that Gly126 (which is notably absent in the construct LC3C(8–125) used for two of the crystal structures¹⁴) is readily accessible for lipid conjugation by the ATG7/ATG3 system.

Phosphorylation. In contrast to the GABARAP-like protein subfamily, the LC3 proteins share an amino-terminal, conserved PKA phosphorylation site²². In LC3C, the target residue is Ser18, whose hydroxyl group provides the N-capping hydrogen bond acceptor³⁷ for helix α_2 (see above). We have used a series of [¹H-¹⁵N] HSQC spectra to follow chemical shift changes of the LC3C backbone amide groups during incubation with PKA *in vitro*. Subsequently, after >72 h, *in vitro* PKA-treated LC3C was proteolytically cleaved and analyzed by MALDI-LTQ-Orbitrap MS/MS. Glu-C hydrolysis of LC3C resulted in an amino-terminal fragment (GSMPPPQKIPSVRPFKQRKSLAIRQEE) for which a change in mass equivalent to the addition of a single phosphate group could be identified, and this fragment also gave rise to peaks with mass changes consistent with neutral losses of H₃PO₄ (mass change -97.97 Da) and HPO₃ (mass change -79.96 Da) that are characteristic for phosphorylated peptides.

The signal-to-noise ratio of the [¹H-¹⁵N] HSQC spectra gradually decreased to approximately a quarter of its original value over the incubation period of 63 h, although the solution remained clear and no new set of peaks with random coil chemical shifts that would indicate denaturation of the protein sample was observed to emerge. Initial chemical shift changes of several backbone amide resonances were already observed within the first 3.5 h after addition of PKA (Supplementary Fig. S11), but no further changes were detected after 10.5 h. Four qualitatively different patterns of chemical shift changes were found for different sets of residues: (i) Many of the backbone amide resonances show no significant chemical shift changes at all, including the intense resonances of the mobile amino-terminal residues from Met1 to Val10 (Fig. 8A), Met97 (Fig. 8D), and the carboxy-terminal residues Thr124 and Gly126, indicating that the conformation of these regions is unaffected by phosphorylation. (ii) By contrast, the amide groups of residues Arg11 to Gln15 exhibit two separate resonances with similar intensities in the presence of PKA, one at the position of the resonance in the unphosphorylated form plus an additional one with slightly different chemical shifts that appears to be present only with extremely low intensity close to the detection limit in the unphosphorylated form, as seen in Fig. 8B for Phe13. Accordingly, the effect

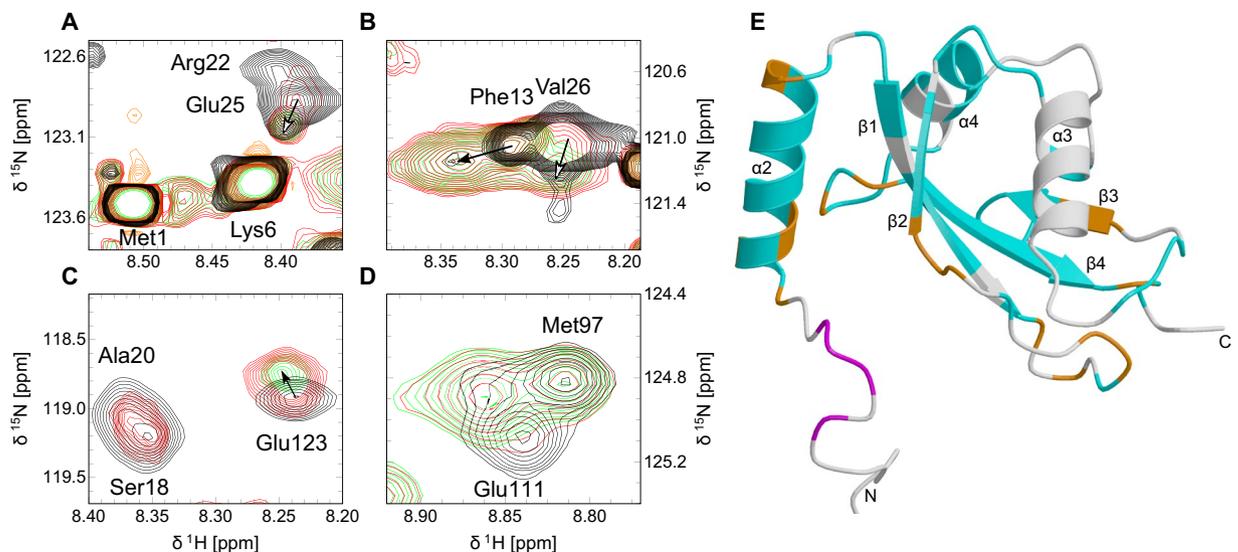


Figure 8. Overlay of sections of the $[^1\text{H}-^{15}\text{N}]$ HSQC spectra (Supplementary Fig. S11) of $[\text{U}-^{13}\text{C},^{15}\text{N}]$ LC3C before (black) and 3.5 h (red) and 7.0 h (green) after incubation with PKA to monitor the progress of *in vitro* phosphorylation. (A) Intense resonance signals of Met1 and Lys6 did not shift, while the shift of the resonance of Glu25 could be traced (open arrow) and the resonance of Arg22 disappears. (B) The populations of the split signal of Phe13 in the unphosphorylated state (black) shifted towards the minor peak position (filled arrow), while the resonance of Val26 could be traced to a new peak position (open arrow). (C) The intensities of the resonances of Ser18 and Ala20 decreased within 3.5 h and could not be traced to newly emerging peaks in the vicinity, in contrast to Glu123 (filled arrow) and (D) Glu111. The resonance of Met97 was not influenced by PKA reaction progress. (E) Split backbone amide resonances (magenta), traceable backbone amide resonances (cyan), and disappearing backbone amide resonances (orange) after incubation with PKA color-coded onto a schematic representation of the secondary structure of the lowest-energy solution structure of LC3C (same view as in Fig. 5). The figure was drawn with MolScript 2.1.2¹⁰⁰ and rendered with Raster3D 3.0¹⁰¹.

of phosphorylation of Ser18 on this region appears to be an increase in the population of an alternate conformation, which exchanges with the conformation dominating in the unphosphorylated form on a time-scale of about 10 ms or slower, or not at all. It is possible that this alternate conformation is essentially the additional helix α_1 that this region forms in the crystal structure of LC3C(8–125) (see above), although the relatively small chemical shift changes are also compatible with a subtler change in conformation. Note that two separate sets of NMR resonances resulting from slow conformational exchange are also observed for many residues located in or near helix α_1 of GABARAP, which does not share the PKA phosphorylation site (Fig. 1), in the absence of PKA⁹. (iii) Another large set of backbone amide resonances, such as those of Glu25 (Fig. 8A), Val26 (Fig. 8B), Asp110, Glu111 (Fig. 8D), Gly113, and Glu123 (Fig. 8C) display a gradual but complete transition over the first 10.5 h of incubation with PKA from the chemical shifts of the unphosphorylated form to a newly emerging resonance with chemical shifts close enough to the original resonance to be readily traced. (iv) Finally, another set of backbone amide resonances including Ser18 (Fig. 8C), Ala20 (Fig. 8C) and Arg22 (Fig. 8A) in the immediate vicinity of the phosphorylation site, as well as Ala31 and Lys32 at the C-terminal end of helix α_2 , Tyr44, Arg46, Phe49 and Leu50 in the loop connecting strands β_1 and β_2 , Lys55 to Lys57 in the LIR binding region, Tyr86, Leu87, Asp112, Phe114, and Ser121 disappear within 7.0 h after addition of PKA without any new resonance with similar chemical shifts emerging. We cannot rule out that amide resonances for these residues in the phosphorylated protein appear elsewhere in the spectrum because the quality of the NMR spectra of phosphorylated LC3C under these buffer conditions is insufficient for multidimensional assignment experiments. Alternatively, the disappearance of these resonances could be caused by unfavorable relaxation properties due to sample aggregation or chemical exchange on the μs to ms time-scale. The location of these resonances in the tertiary structure (orange in Fig. 8E) is intriguingly similar to the “hotspots” of conformational exchange identified by CPMG relaxation dispersion spectroscopy (Fig. 5), suggesting that the conformational effects of phosphorylation by PKA might be coupled to the conformational dynamics of unphosphorylated LC3C. Either way, the observed chemical shift changes indicate that phosphorylated LC3C is even more conformationally heterogeneous than unphosphorylated LC3C.

As described above, the side-chain of Ser18 provides an N-capping hydrogen bond acceptor (Fig. 9), so phosphorylation of Ser18 could have an influence on the stability of this hydrogen bond and, in turn, on the stability of helix α_2 in general, maybe even trigger a helix-to-coil transition³⁷. Moreover, helix α_2 is held in place by several favorable tertiary contacts, most notably electrostatic interactions of the side-chain of Arg22 with the negatively charged side-chains of Asp110, Glu111, and Asp112 (Fig. 9) in the loop between helix α_4 and strand β_4 , a rigid loop that shows little sign of backbone mobility on the ps to ns time-scale nor of line broadening from conformational exchange on the μs to ms time-scale (Figs 2, 3, 5). The chemical shift changes of the backbone amide resonance from Asp110 to Gly113 upon addition of PKA (see above) suggest that the negative charge of the phosphate group modification modulates these interactions.

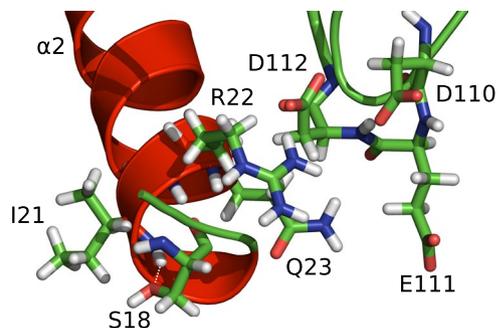


Figure 9. Electrostatic interactions of the acidic cluster formed by the side-chains of Asp110, Glu111, and Asp112 with helix α_2 residues Arg22 and Gln23 in the crystal structure of LC3C(8–125) (PDB 3WAM¹⁴; hydrogen atoms were added with PyMOL). The backbone amide proton of Ile21 forms an N-capping motif with the side-chain hydroxyl oxygen of phosphorylation target Ser18 (dashed line). The figure was created with PyMOL (The PyMOL Molecular Graphics System, Version 1.7.2.1, Schrödinger, LLC).

To test the plausibility of these hypotheses we performed comparative molecular dynamics (MD) simulations of LC3C in the unmodified (Ser18) and phosphorylated (pSer18) form. In the context of a random-coil peptide, the pK_a of phosphoserine is very close to 6.0, the pH of the NMR buffer used here, and the exchange between the protonated (monoanionic) and deprotonated (dianionic) states is fast on the chemical shift time-scale³⁸. Although both the pK_a and the exchange rate of the protonation equilibrium can be modulated by tertiary interactions we have to assume that both states are populated to a considerable degree in our NMR experiments³⁹; in fact, exchange between these two protonation states might well be the cause underlying some of the line broadening or peak splitting in the NMR spectra of phosphorylated LC3C described above. Therefore, we calculated three MD trajectories of 100 ns each, one with the unmodified serine (Ser18), one with the protonated phosphoserine (pSer18⁻), and one with the deprotonated phosphoserine (pSer18²⁻). The hydrogen bond between the amide group of Ile21 and the side-chain of Ser18 is indeed formed a considerable fraction of the time along the MD trajectory (Supplementary Fig. S12) but virtually absent from the MD trajectories for pSer18⁻ (Supplementary Fig. S13) and pSer18²⁻ (Supplementary Fig. S14), which corroborates our hypothesis that this hydrogen bond is destabilized by phosphorylation by PKA. However, helix α_2 remains stable for the entire duration of the MD trajectories regardless of the phosphorylation state of Ser18 (Supplementary Figs S12–S14) and no helix-to-coil transition is observed, at least not on the time-scale of 100 ns covered by the MD simulations. Instead of forming a hydrogen bond with the amide group of Ile21 the single negative charge of the monoanionic phosphate group of pSer18⁻ favors a salt bridge with the positively charged side-chain of Lys17 (Supplementary Fig. S13); intriguingly, the additional helix α_1 is populated a significant fraction of the MD trajectory calculated with pSer18⁻ (Supplementary Fig. S13), in support of our suspicion that phosphorylation of Ser18 might stabilize a helical conformation in this region (see above). Of course, these MD simulations do not rule out the possibility of any additional conformational changes on a time-scale slower than about 100 ns that might be triggered by phosphorylation. For example, the disappearance of the amide resonances of Ala31 and Lys32 at the C-terminal end of helix α_2 could readily be explained if this region acts as a hinge for repositioning of helix α_2 . In light of our experimental observation that the chemical shifts of several resonances in the LIR binding region of LC3C are strongly affected by phosphorylation (see above) a modulating effect of PKA on both, affinity as well as specificity of LC3C for its target LIRs appears highly likely.

Conclusions

High-resolution structure determination by NMR spectroscopy confirms that the well-ordered core (residues 18 to 120) of the autophagy-related protein MAP1LC3C adopts virtually the same tertiary structure in solution as in X-ray crystallography, although extensive exchange line broadening reveals the existence of alternate, low-populated conformations. By contrast, both termini are highly mobile. The mobility of the C-terminus guarantees full access of the lipid conjugation machinery to the C-terminal residue for membrane anchoring of LC3C. The N-terminus forms a polyproline II helix (Pro2, Pro3, Pro4) tethered to the ubiquitin-like core via a flexible linker, an arrangement that has been described as a “sticky arm” in the literature⁴⁰. Short proline-rich regions are a highly diverse protein-protein interaction motif found in a large number of proteins, often recognized by dedicated protein-protein interaction domains such as WW domains, SH3 domains, or profilins⁴⁰. In this context, ligand specificity is usually conferred by the spacing of the proline residues as well as the biophysical properties of the side-chains flanking the prolines, and sometimes further modulated by phosphorylation. While the N-terminal sequence motif of LC3C, MPPPQK (Fig. 1), deviates from the stereotypical WW and SH3 domain ligand motifs as compiled by Kay *et al.*⁴⁰, interfacing to such a protein-protein interaction domain cannot be ruled out as the functional role of the N-terminal PPII motif of LC3C. Intriguingly, though, proline-rich regions have also been identified to mediate interactions with microtubules, more specifically with the β -tubulin subunit⁴¹. Flexible proline-rich regions are involved in the interaction of the intrinsically disordered microtubule-associated protein Tau with microtubules, and phosphorylation has recently been shown to change the conformation and microtubule interaction mode of the second proline-rich region of Tau^{42,43}. The full name MAP1LC3 (microtubule associated proteins 1 light chain 3) reflects the fact that this subfamily was originally identified as proteins

associated with microtubules⁴⁴, and the interplay between autophagy and microtubules is well-documented⁴⁵, albeit still poorly understood. In the case of GABARAP the interaction with the negatively charged microtubular surface was mapped to the positively charged helix α_2 with 7 basic residues⁴⁶. Because this motif appears to be absent from LC3C, where only 3 of these 7 basic residues are conserved in helix α_2 from Ser18 to Phe33 (Fig. 1), it is plausible to hypothesize that the divergence of the amino-terminal region across the human GABARAP/ MAP1LC3 family (Fig. 1) provides for differential interaction with the cytoskeleton. The PKA phosphorylation site at Ser18 at the N-cap of helix α_2 of LC3C is well-positioned to further modulate such an interaction. The chemical shift changes upon phosphorylation do not indicate a direct influence on the PPII motif itself but are consistent with repositioning of helix α_2 and possibly also stabilization of a short additional helix α_1 in the NTR. Phosphorylation of Ser18 also has a large effect on the chemical shifts in the LIR binding region of LC3C and can therefore be expected to modulate affinity and specificity of the interaction of LC3C with its target LIRs or CLIRs.

Methods

Cloning, expression and purification. LC3C was cloned, expressed and purified as an amino-terminal GST-LC3C fusion protein as described in detail previously²⁷. Thrombin cleavage and final purification by cation exchange and size exclusion chromatography yielded highly pure samples of the cytosolic LC3C protein of 126 amino acids with a two-residue amino-terminal cloning artifact (Gly-1 and Ser0) (Supplementary Figs S1 and S2), which were concentrated to between 370 μ M and 700 μ M for NMR spectroscopy.

NMR spectroscopy. NMR samples of 370 μ M to 700 μ M [U -¹⁵N] or [U -¹³C,¹⁵N] LC3C were prepared in 20 mM PIPES, 150 mM NaCl, 0.1 mM EDTA and supplemented with 2% (v/v) glycerol-*d*₈ (Euriso-top, Gif-sur-Yvette, France) and 10% (v/v) ²D₂O at pH 6.0 (NMR buffer). The effect of the protein concentration and of the presence or absence of 2% (v/v) glycerol-*d*₈ on the position and qualitative line shape features of the NMR resonances is negligible (Supplementary Fig. S3). NMR experiments were recorded at 20.0 °C on Varian INOVA or Bruker AVANCE III spectrometers operating at ¹H frequencies of 600 MHz and 800 MHz and equipped with cryogenically cooled triple or quadruple resonance probes with pulse-field gradient capabilities. Sequence-specific ¹H, ¹⁵N, and ¹³C backbone and side-chain NMR resonance assignments and ³J_{H_NH α scalar coupling constants were reported previously (BMRB 26603)²⁷. ¹H-¹H distance information was derived from the following nuclear Overhauser effect spectroscopy (NOESY) experiments: 2D [¹H-¹H] NOESY⁴⁷ (120 ms mixing time) in 99% ²D₂O, 3D [¹H-¹⁵N] NOESY-HSQC⁴⁸ (120 ms mixing time) in 90% ¹H₂O/10% ²D₂O, 3D [¹H-¹³C] NOESY-HSQC⁴⁹ (120 ms mixing time) in 99% ²D₂O, 3D [¹H-¹⁵N] [¹H-¹⁵N] HSQC-NOESY-HSQC^{50,51} (150 ms mixing time) in 90% ¹H₂O/10% ²D₂O, and 3D [¹H-¹³C] [¹H-¹⁵N] HSQC-NOESY-HSQC⁵² (150 ms mixing time) in 90% ¹H₂O/10% ²D₂O. The ¹H₂O resonance was suppressed by excitation sculpting⁵³ in the 2D [¹H-¹H] homonuclear experiments and by gradient coherence selection in the heteronuclear experiments, quadrature detection in the indirect dimensions was achieved by States-TPPI⁵⁴ or the echo/antiecho method^{55,56}. All NMR spectra were processed using NMRPipe and NMRDraw⁵⁷ and analysed with NMRViewJ⁵⁸ and CcpNmr Analysis⁵⁹.}

To measure amide proton/deuteron (H/D) exchange a sample of 570 μ M [U -¹⁵N] LC3C in NMR buffer was freeze-dried and a series of seven consecutive [¹H-¹⁵N] HSQC experiments was recorded 0.3, 1.6, 2.8, 5.4, 7.9, 12.9, and 17.9 h after reconstitution of the lyophilized sample in ²D₂O. Signal intensities in the resulting [¹H-¹⁵N] HSQC spectra were quantified by three-way decomposition using MUNIN^{60,61} and fitted by mono-exponential decay functions using the Levenberg-Marquardt algorithm as implemented in MATLAB R2015b (The MathWorks, Inc., Natick, Mass., USA) to extract amide H/D exchange rates for semi-quantitative analysis⁶². Experimental H/D exchange rates, k_{ex} , were converted into protection factors $PF = k_{int}/k_{ex}$, where k_{int} are the intrinsic (unprotected) H/D exchange rates for these experimental conditions as predicted from the amino acid sequence⁶³ using the SPHERE server (<http://landing.foxchase.org/research/labs/roder/sphere>).

Protein backbone amide group dynamics on the pico- to nanosecond time-scale was probed by ¹⁵N spin relaxation experiments^{34,64} recorded on a sample of 510 μ M [U -¹⁵N] LC3C in NMR buffer at 800 MHz and 20.0 °C. ¹⁵N longitudinal relaxation rates, R_1 , were obtained from ¹⁵N inversion recovery experiments^{64,65} with 11 different inversion recovery times (3 of them collected in duplicate) between 80 ms and 1200 ms and a recycle delay of 2.5 s. ¹⁵N rotating frame relaxation rates, $R_{1\rho}$, were obtained from ¹⁵N spin lock experiments⁶⁶ with 10 different spin lock times (3 of them collected in duplicate) between 10 ms and 100 ms with a spin lock field strength of 2.02 kHz and a recycle delay of 3.0 s. Amide resonance intensities were quantified by three-way decomposition and fit by mono-exponential decay functions using MUNIN^{60,61} to extract the ¹⁵N relaxation rates R_1 and $R_{1\rho}$. ¹⁵N transverse relaxation rates, R_2 , were calculated from R_1 and $R_{1\rho}$ ⁶⁷. ¹H¹⁵N heteronuclear NOE values were calculated as the amide resonance intensity ratios in a pair of interleaved spectra recorded with and without proton saturation by applying a train of 120° pulses at a field strength of 11.6 kHz for the final 6.0 s of the recycle delay of 15.0 s⁶⁴. Uncertainties of the ¹H¹⁵N NOE values were estimated from the spectral noise background. The overall rotational diffusion tensor and “model-free” parameters³⁰ describing internal motion of the protein backbone such as the generalized order parameters for sub-nanosecond internal motion, S^2 , were determined by fitting the experimental ¹⁵N R_1 and R_2 rates and ¹H¹⁵N NOE values using Tensor 2.0⁶⁸ with the default parameters based on the lowest-energy structure of LC3C. Amide groups with ¹H¹⁵N NOE values below 0.65 and/or with R_2/R_1 ratios deviating by more than 10% from the mean value were considered to possess significantly increased internal mobility and excluded from the calculation of the rotational diffusion tensor²⁹. The presence of millisecond time-scale exchange processes was probed by ¹⁵N single-quantum Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion experiments^{69,70} recorded on this sample at 600 MHz as well as 800 MHz and 20.0 °C. In each ¹⁵N CPMG experiment 17 (14) different CPMG frequencies $\nu_{CPMG} = 1/(2\delta)$, where δ is the time between consecutive refocusing pulses, ranging from 41.7 Hz (50.0 Hz) to 2000.0 Hz (1000.0 Hz) were sampled during a constant-time relaxation interval of $T_{CPMG} = 48$ ms (40 ms) at 600 MHz (800 MHz). Amide resonance intensities $I(\nu_{CPMG})$ were quantified by three-way decomposition using MUNIN^{60,61} and converted into effective transverse relaxation

rates, $R_{2,\text{eff}}(\nu_{\text{CPMG}}) = -\ln(I(\nu_{\text{CPMG}})/I_0)/T_{\text{CPMG}}$, where I_0 is the corresponding resonance intensity in a reference spectrum recorded without the constant-time relaxation interval. Error estimates $\Delta R_{2,\text{eff}}$ for $R_{2,\text{eff}}$ were obtained from duplicate measurements at 3 different ν_{CPMG} values as described previously⁷¹, assuming a minimum relative error $\Delta R_{2,\text{eff}}/R_{2,\text{eff}}$ of 2.0% to account for offset effects and other systematic experimental imperfections. Global exchange parameters (exchange rates, equilibrium populations) and residue specific values (¹⁵N chemical shift differences, $\Delta\omega_{\text{N}}$, intrinsic relaxation rates, $R_{2,0}$) were extracted by a non-linear least-squares fitting procedure whereby experimental dispersion profiles, $R_{2,\text{eff}}(\nu_{\text{CPMG}}) = R_{2,0} + R_{\text{ex}}(\nu_{\text{CPMG}})$, were fit by those calculated from the evolution of magnetization during the CPMG interval by solving the Bloch-McConnell equations numerically for a two-site exchange model as described previously^{71,72}. Errors of the fitted parameters were calculated from the covariance matrix⁷³, an approach that has been shown to provide reasonable error estimates⁷⁴.

Structural restraints and structure calculation. Based on the almost complete assignment of the ¹H, ¹⁵N, and ¹³C resonances of LC3C published previously²⁷, the NOE cross peaks in the [¹H-¹⁵N] and [¹H-¹³C] NOESY-HSQC spectra were quantified with CcpNmr Analysis 2.4.1⁵⁹ and automatically assigned and converted into NOE distance restraints using ARIA 2.3^{75,76} in an iterative procedure. All resulting NOE assignments were inspected manually. 94 ³J_{HNH α} coupling constants were obtained from a quantitative HNHA experiment with a coherence transfer time of 12.3 ms and relaxation correction factor of 1.1^{77,78} as described previously²⁷. 32 of these ³J_{HNH α} in the well-ordered core of the protein were sufficiently different from the value of 7.0 Hz that is indicative of rotameric averaging⁷⁹ and had sufficiently small experimental uncertainties to be converted into Φ backbone torsion angle restraints using the Karplus relation⁷⁷ as implemented in CcpNmr Analysis with a tolerance of $\pm 30^\circ$. Additional Φ and Ψ backbone dihedral angle restraints were derived from the secondary chemical shifts (¹HN, ¹H α , ¹⁵N, ¹³CO, ¹³C α , ¹³C β) by TALOS-N^{80,81}. Amide hydrogen bond donors were identified as slowly exchanging protons in the H/D exchange experiment and the corresponding hydrogen bond acceptors were found by proximity in later stages of the structure calculation process and then incorporated as additional distance restraints. For each of the 51 hydrogen bonds the distance between the amide proton and the acceptor was restrained to between 1.5 Å and 2.3 Å and the distance between the amide nitrogen and the acceptor to between 2.5 Å and 3.5 Å. These experimental restraints served as input for the calculation of 100 structures using restrained molecular dynamics simulations with ARIA-optimized CNS 1.21⁸² using the CNS protocol parameters listed in Supplementary Table S1. The 10 structures showing the lowest energy values were further refined in an explicit water shell using the CSDX/OPLS hybrid force field as implemented in ARIA/CNS⁸³ and selected for further characterization. Structural models were visualized by PyMOL (The PyMOL Molecular Graphics System, Version 1.7.2.1, Schrödinger, LLC) and analyzed using ARIA/CNS, the NIH version 1.2.1⁸⁴ of X-PLOR 3.851⁸⁵, PROCHECK-NMR⁸⁶, MolProbity⁸⁷, and CING⁸⁸.

In-vitro phosphorylation. LC3C samples were phosphorylated *in vitro* using the catalytic subunit of murine PKA (NEB, Frankfurt a. M., Germany). A 650 μM sample of [^{U-¹³C, ¹⁵N}]-labeled LC3C in NMR buffer was supplemented with 4 mM MgCl₂ and subsequently 1 mM ATP (PKA buffer). This change in buffer conditions was monitored by [¹H-¹⁵N] HSQC spectra to rule out any major spectral differences that might indicate a substantial modulation of the structure or dynamics of LC3C. Finally, 1 μl PKA (circa 2500 U) was added and the reaction was monitored by a series of 17 consecutive [¹H-¹⁵N] HSQC experiments at 800 MHz and 20.0 °C over a period of 63 h. Additionally, a [¹H-¹³C] ct-HSQC experiment was recorded at 800 MHz and 20.0 °C after six days of incubation. Furthermore, the phosphorylation reaction was analyzed by proteolytic hydrolysis with trypsin and Glu-C of the [^{U-¹³C, ¹⁵N}]-labeled reaction product, followed by MALDI-LTQ-Orbitrap MS/MS analysis.

Molecular dynamics simulations. The lowest-energy solution structure after water refinement was used as a starting structure for molecular dynamics (MD) simulations. Three independent phosphorylation states of LC3C were simulated: unmodified serine 18 (Ser18), serine 18 mutated to a protonated phosphoserine (pSer18⁻), and serine 18 mutated to a deprotonated phosphoserine (pSer18²⁻). The Amber force field ff99SB-ILDN⁸⁹ was used in conjunction with the TIP3P water model⁹⁰. Parameters for bonded and van der Waals interactions of the phosphorylated serine residue were taken from the GAFF force field⁹¹ generated with acpype.py⁹² and the partial charges from Homeyer *et al.*³⁹. All MD simulations were performed using GROMACS 5.1⁹³. The protein was centered in a dodecahedral box with a minimum solute-to-wall distance of 1 nm. Sodium chloride ions were added to neutralize the systems and mimic a salt concentration of approximately 100 mM. The systems were first energy-minimized, equilibrated in the NPT ensemble (i. e., with a constant number of molecules, pressure, and temperature) for 1 ns and then simulated for 100 ns each in the NPT ensemble at 298 K (Nosé-Hoover thermostat⁹⁴) and 1.0 bar (Parrinello-Rahman barostat⁹⁵). Van der Waals and short-range electrostatic interactions were cut off at 12 Å and long-range electrostatics were treated with the particle mesh Ewald method⁹⁶. Bond lengths and bond angles of hydrogen atoms were constrained to their equilibrium values with the LINCS algorithm⁹⁷. The equations of motion were integrated with a velocity Verlet integrator and a time step of 2 fs. Atom positions were saved every 10 ps.

Data Availability

The atomic coordinates and experimental restraints have been deposited with the Protein Data Bank (access code: 2NCN; DOI: <http://dx.doi.org/10.2210/pdb2ncn/pdb>).

References

- Nakatogawa, H., Suzuki, K., Kamada, Y. & Ohsumi, Y. Dynamics and diversity in autophagy mechanisms: lessons from yeast. *Nat. Rev. Mol. Cell Biol.* **10**, 458–467 (2009).
- Nakatogawa, H., Ichimura, Y. & Ohsumi, Y. Atg8, a Ubiquitin-like Protein Required for Autophagosome Formation, Mediates Membrane Tethering and Hemifusion. *Cell* **130**, 165–178 (2007).
- Ichimura, Y. *et al.* A ubiquitin-like system mediates protein lipidation. *Nature* **408**, 488–492 (2000).

4. Shpilka, T., Weidberg, H., Pietrokovski, S. & Elazar, Z. Atg8: an autophagy-related ubiquitin-like protein family. *Genome Biol.* **12**, 226 (2011).
5. Weiergräber, O. H., Mohrlüder, J. & Willbold, D. Atg8 Family Proteins — Autophagy and Beyond in Autophagy - A Double-Edged Sword - Cell Survival or Death? (ed. Bailly, Y.) 13–45, doi:10.5772/55647 (InTech, 2013).
6. Schwarten, M., Stoldt, M., Mohrlüder, J. & Willbold, D. Solution structure of Atg8 reveals conformational polymorphism of the N-terminal domain. *Biochem. Biophys. Res. Commun.* **395**, 426–431 (2010).
7. Kumeta, H. *et al.* The NMR structure of the autophagy-related protein Atg8. *J. Biomol. NMR* **47**, 237–241 (2010).
8. Noda, N. N. *et al.* Structural Basis of Atg8 Activation by a Homodimeric E1, Atg7. *Mol. Cell* **44**, 462–475 (2011).
9. Stangler, T., Mayr, L. M. & Willbold, D. Solution Structure of Human GABAA Receptor-associated Protein GABARAP: implications for biological function and its regulation. *J. Biol. Chem.* **277**, 13363–13366 (2002).
10. Rozenknop, A. *et al.* Characterization of the Interaction of GABARAPL-1 with the LIR Motif of NBR1. *J. Mol. Biol.* **410**, 477–487 (2011).
11. Kouno, T. *et al.* Solution Structure of Microtubule-associated Protein Light Chain 3 and Identification of Its Functional Subdomains. *J. Biol. Chem.* **280**, 24610–24617 (2005).
12. Noda, N. N. *et al.* Structural basis of target recognition by Atg8/LC3 during selective autophagy. *Genes Cells* **13**, 1211–1218 (2008).
13. Rogov, V. V. *et al.* Structural basis for phosphorylation-triggered autophagic clearance of Salmonella. *Biochem. J.* **454**, 459–466 (2013).
14. Suzuki, H. *et al.* Structural Basis of the Autophagy-Related LC3/Atg13 LIR Complex: Recognition and Interaction Mechanism. *Structure* **22**, 47–58 (2014).
15. von Muhlinen, N. *et al.* LC3C, Bound Selectively by a Noncanonical LIR Motif in NDP52, Is Required for Antibacterial Autophagy. *Mol. Cell* **48**, 329–342 (2012).
16. Weidberg, H. *et al.* LC3 and GATE-16 N Termini Mediate Membrane Fusion Processes Required for Autophagosome Biogenesis. *Dev. Cell* **20**, 444–454 (2011).
17. Chu, C. T. *et al.* Cardiolipin externalization to the outer mitochondrial membrane acts as an elimination signal for mitophagy in neuronal cells. *Nat. Cell Biol.* **15**, 1197–1205 (2013).
18. Pankiv, S. *et al.* p62/SQSTM1 Binds Directly to Atg8/LC3 to Facilitate Degradation of Ubiquitinated Protein Aggregates by Autophagy. *J. Biol. Chem.* **282**, 24131–24145 (2007).
19. Wild, P. *et al.* Phosphorylation of the Autophagy Receptor Optineurin Restricts Salmonella Growth. *Science* **333**, 228–233 (2011).
20. Ichimura, Y. *et al.* Structural Basis for Sorting Mechanism of p62 in Selective Autophagy. *J. Biol. Chem.* **283**, 22847–22857 (2008).
21. Birgisdottir, A. B., Lamark, T. & Johansen, T. The LIR motif - crucial for selective autophagy. *J. Cell. Sci.* **126**, 3237–3247 (2013).
22. Cherra, S. J. *et al.* Regulation of the autophagy protein LC3 by phosphorylation. *J. Cell Biol.* **190**, 533–539 (2010).
23. Jiang, H., Cheng, D., Liu, W., Peng, J. & Feng, J. Protein kinase C inhibits autophagy and phosphorylates LC3. *Biochem. Biophys. Res. Commun.* **395**, 471–476 (2010).
24. Blom, N., Gammeltoft, S. & Brunak, S. Sequence and Structure-based Prediction of Eukaryotic Protein Phosphorylation Sites. *J. Mol. Biol.* **294**, 1351–1362 (1999).
25. Huang, R. *et al.* Deacetylation of Nuclear LC3 Drives Autophagy Initiation under Starvation. *Mol. Cell* **57**, 456–466 (2015).
26. Xie, Y. *et al.* Posttranslational modification of autophagy-related proteins in macroautophagy. *Autophagy* **11**, 28–45 (2015).
27. Krichel, C. *et al.* Sequence-specific 1H, 15N, and 13C resonance assignments of the autophagy-related protein LC3C. *Biomol. NMR Assign.* **10**, 41–43 (2016).
28. Grzesiek, S. & Bax, A. The Importance of Not Saturating H₂O in Protein NMR. Application to Sensitivity Enhancement and NOE Measurements. *J. Am. Chem. Soc.* **115**, 12593–12594 (1993).
29. Pawley, N. H., Wang, C., Koide, S. & Nicholson, L. K. An improved method for distinguishing between anisotropic tumbling and chemical exchange in analysis of 15N relaxation parameters. *J. Biomol. NMR* **20**, 149–165 (2001).
30. Lipari, G. & Szabo, A. Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 2. Analysis of Experimental Results. *J. Am. Chem. Soc.* **104**, 4559–4570 (1982).
31. Woessner, D. E. Nuclear Spin Relaxation in Ellipsoids Undergoing Rotational Brownian Motion. *J. Chem. Phys.* **37**, 647–654 (1962).
32. Möckel, C. *et al.* Integrated NMR, Fluorescence, and Molecular Dynamics Benchmark Study of Protein Mechanics and Hydrodynamics. *J. Phys. Chem. B* **123**, 1453–1480 (2019).
33. Berjanskii, M. V. & Wishart, D. S. A Simple Method To Predict Protein Flexibility Using Secondary Chemical Shifts. *J. Am. Chem. Soc.* **127**, 14970–14971 (2005).
34. Kay, L. E., Torchia, D. A. & Bax, A. Backbone Dynamics of Proteins As Studied by 15N Inverse Detected Heteronuclear NMR Spectroscopy: Application to Staphylococcal Nuclease. *Biochemistry* **28**, 8972–8979 (1989).
35. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S. & Sykes, B. D. 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR* **5**, 67–81 (1995).
36. Neudecker, P. *et al.* Structure of an Intermediate State in Protein Folding and Aggregation. *Science* **336**, 362–366 (2012).
37. Doig, A. J. & Baldwin, R. L. N- and C-capping preferences for all 20 amino acids in α -helical peptides. *Protein Sci.* **4**, 1325–1336 (1995).
38. Bienkiewicz, E. A. & Lumb, K. J. Random-coil chemical shifts of phosphorylated amino acids. *J. Biomol. NMR* **15**, 203–206 (1999).
39. Homeyer, N., Horn, A. H. C., Lanig, H. & 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).
40. Kay, B. K., Williamson, M. P. & Sudol, M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **14**, 231–241 (2000).
41. Gendreau, S., Schirmer, J. & Schmalzing, G. Identification of a tubulin binding motif on the P2X₂ receptor. *J. Chromatogr. B* **786**, 311–318 (2003).
42. Kadavath, H. *et al.* Tau stabilizes microtubules by binding at the interface between tubulin heterodimers. *Proc. Natl. Acad. Sci. (USA)* **112**, 7501–7506 (2015).
43. Schwalbe, M. *et al.* Structural Impact of Tau Phosphorylation at Threonine 231. *Structure* **23**, 1448–1458 (2015).
44. Mann, S. S. & Hammarback, J. A. Molecular Characterization of Light Chain 3. *J. Biol. Chem.* **269**, 11492–11497 (1994).
45. Mackeh, R., Perdiz, D., Lorin, S., Codogno, P. & Poüs, C. Autophagy and microtubules – new story, old players. *J. Cell. Sci.* **126**, 1071–1080 (2013).
46. Coyle, J. E., Qamar, S., Rajashankar, K. R. & Nikolov, D. B. Structure of GABARAP in Two Conformations: Implications for GABAA Receptor Localization and Tubulin Binding. *Neuron* **33**, 63–74 (2002).
47. Kumar, A., Ernst, R. R. & Wüthrich, K. A Two-Dimensional Nuclear Overhauser Enhancement (2D NOE) Experiment for the Elucidation of Complete Proton-Proton Cross-Relaxation Networks in Biological Macromolecules. *Biochem. Biophys. Res. Commun.* **95**, 1–6 (1980).
48. Zhang, O., Kay, L. E., Olivier, J. P. & Forman-Kay, J. D. Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR* **4**, 845–858 (1994).
49. Ikura, M., Kay, L. E., Tschudin, R. & Bax, A. Three-Dimensional NOESY-HMQC Spectroscopy of a 13C-Labeled Protein. *J. Magn. Reson.* **86**, 204–209 (1990).

50. Frenkiel, T., Bauer, C., Carr, M. D., Birdsall, B. & Feeney, J. HMQC-NOESY-HMQC, a Three-Dimensional NMR Experiment Which Allows Detection of Nuclear Overhauser Effects between Protons with Overlapping Signals. *J. Magn. Reson.* **90**, 420–425 (1990).
51. Ikura, M., Bax, A., Clore, G. M. & Gronenborn, A. M. Detection of Nuclear Overhauser Effects between Degenerate Amide Proton Resonances by Heteronuclear Three-Dimensional Nuclear Magnetic Resonance Spectroscopy. *J. Am. Chem. Soc.* **112**, 9020–9022 (1990).
52. Diercks, T., Coles, M. & Kessler, H. An efficient strategy for assignment of cross-peaks in 3D heteronuclear NOESY experiments. *J. Biomol. NMR* **15**, 177–180 (1999).
53. Hwang, T.-L. & Shaka, A. J. Water Suppression That Works. Excitation Sculpting Using Arbitrary Waveforms and Pulsed Field Gradients. *J. Magn. Reson. Series A* **112**, 275–279 (1995).
54. Marion, D., Ikura, M., Tschudin, R. & Bax, A. Rapid Recording of 2D NMR Spectra without Phase Cycling. Application to the Study of Hydrogen Exchange in Proteins. *J. Magn. Reson.* **85**, 393–399 (1989).
55. Kay, L. E., Keifer, P. & Saarinen, T. Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J. Am. Chem. Soc.* **114**, 10663–10665 (1992).
56. Schleucher, J., Sattler, M. & Griesinger, C. Coherence Selection by Gradients without Signal Attenuation: Application to the Three-Dimensional HNCO Experiment. *Angew. Chem. Int. Ed. Engl.* **32**, 1489–1491 (1993).
57. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
58. Johnson, B. A. & Blevins, R. A. NMRView: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**, 603–614 (1994).
59. Vranken, W. F. *et al.* The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins* **59**, 687–696 (2005).
60. Orekhov, V. Y., Ibraghimov, I. V. & Billeter, M. MUNIN: A new approach to multi-dimensional NMR spectra interpretation. *J. Biomol. NMR* **20**, 49–60 (2001).
61. Korzhnev, D. M., Ibraghimov, I. V., Billeter, M. & Orekhov, V. Y. MUNIN: Application of three-way decomposition to the analysis of heteronuclear NMR relaxation data. *J. Biomol. NMR* **21**, 263–268 (2001).
62. Andrec, M., Hill, R. B. & Prestegard, J. H. Amide exchange rates in *Escherichia coli* acyl carrier protein: Correlation with protein structure and dynamics. *Protein Sci.* **4**, 983–993 (1995).
63. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins* **17**, 75–86 (1993).
64. Farrow, N. A. *et al.* Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology 2 Domain Studied by 15N NMR Relaxation. *Biochemistry* **33**, 5984–6003 (1994).
65. Lakomek, N.-A., Ying, J. & Bax, A. Measurement of 15N relaxation rates in perdeuterated proteins by TROSY-based methods. *J. Biomol. NMR* **53**, 209–221 (2012).
66. Korzhnev, D. M., Skrynnikov, N. R., Millet, O., Torchia, D. A. & Kay, L. E. An NMR Experiment for the Accurate Measurement of Heteronuclear Spin-Lock Relaxation Rates. *J. Am. Chem. Soc.* **124**, 10743–10753 (2002).
67. Tjandra, N., Wingfield, P., Stahl, S. & Bax, A. Anisotropic rotational diffusion of perdeuterated HIV protease from 15N NMR relaxation measurements at two magnetic fields. *J. Biomol. NMR* **8**, 273–284 (1996).
68. Dosset, P., Hus, J.-C., Blackledge, M. & Marion, D. Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR* **16**, 23–28 (2000).
69. Loria, J. P., Rance, M. & Palmer, A. G. A Relaxation-Compensated Carr-Purcell-Meiboom-Gill Sequence for Characterizing Chemical Exchange by NMR Spectroscopy. *J. Am. Chem. Soc.* **121**, 2331–2332 (1999).
70. Tollinger, M., Skrynnikov, N. R., Mulder, F. A. A., Forman-Kay, J. D. & Kay, L. E. Slow Dynamics in Folded and Unfolded States of an SH3 Domain. *J. Am. Chem. Soc.* **123**, 11341–11352 (2001).
71. Neudecker, P. *et al.* Identification of a Collapsed Intermediate with Non-native Long-range Interactions on the Folding Pathway of a Pair of Fyn SH3 Domain Mutants by NMR Relaxation Dispersion Spectroscopy. *J. Mol. Biol.* **363**, 958–976 (2006).
72. Korzhnev, D. M. *et al.* Low populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* **430**, 586–590 (2004).
73. Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. *Numerical Recipes in C* (2nd ed., Cambridge University Press, 1992).
74. Neudecker, P., Korzhnev, D. M. & Kay, L. E. Assessment of the effects of increased relaxation dispersion data on the extraction of 3-site exchange parameters characterizing the unfolding of an SH3 domain. *J. Biomol. NMR* **34**, 129–135 (2006).
75. Habeck, M., Rieping, W., Linge, J. P. & Nilges, M. NOE Assignment With ARIA 2.0: The Nuts and Bolts in *Methods in Molecular Biology Vol. 278: Protein NMR Techniques* (ed. Downing, A. K.) 379–402 (Humana Press, 2004).
76. Rieping, W. *et al.* ARIA2: Automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23**, 381–382 (2007).
77. Vuister, G. W. & Bax, A. Quantitative J Correlation: A New Approach for Measuring Homonuclear Three-Bond J(HNH α) Coupling Constants in 15N-Enriched Proteins. *J. Am. Chem. Soc.* **115**, 7772–7777 (1993).
78. Zhang, W., Smithgall, T. E. & Gmeiner, W. H. Three-dimensional structure of the Hck SH2 domain in solution. *J. Biomol. NMR* **10**, 263–272 (1997).
79. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., Rance, M. & Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice* (2nd ed., Elsevier Academic Press, 2007).
80. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223 (2009).
81. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
82. Brünger, A. T. *et al.* Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **54**, 905–921 (1998).
83. Linge, J. P. & Nilges, M. Influence of non-bonded parameters on the quality of NMR structures: A new force field for NMR structure calculation. *J. Biomol. NMR* **13**, 51–59 (1999).
84. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73 (2003).
85. Brünger, A. T. *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR* (Yale University Press, 1992).
86. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
87. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **66**, 12–21 (2010).
88. Doreleijers, J. F. *et al.* CING: an integrated residue-based structure validation program suite. *J. Biomol. NMR* **54**, 267–283 (2012).
89. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958 (2010).

90. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
91. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
92. Sousa da Silva, A. W. & Vranken, W. F. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Research Notes* **5**, 367 (2012).
93. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
94. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
95. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
96. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
97. Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **4**, 116–122 (2008).
98. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
99. Kabsch, W. & Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **22**, 2577–2673 (1983).
100. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
101. Merrit, E. A. & Murphy, M. E. P. Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869–873 (1994).

Acknowledgements

This work has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Projektnummer 267205415 - SFB 1208 (projects A07 (B.S.), B02 (D.W.) and B03 (P.N.)) and INST 208/704-1 FUGG to purchase the hybrid computer cluster used in this study. The authors acknowledge access to the Jülich-Düsseldorf Biomolecular NMR Center.

Author Contributions

C.K., O.H.W., D.W. and P.N. designed the research. C.K. produced the NMR samples. C.K., C.M. and P.N. performed and analyzed the NMR spectroscopy. P.F.H. performed and analyzed the mass spectrometry. O.S., H.S. and B.S. performed molecular dynamics simulations and computational analysis. C.K., O.H.W. and P.N. were primarily responsible for writing the manuscript. All authors reviewed and edited the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-48155-8>.

Competing Interests: The authors declare no competing interests.

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



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

© The Author(s) 2019