DYNLL2 dynein light chain binds to an extended linear motif of myosin 5a tail that has structural plasticity

Andrea Bodor[1], László Radnai[2*], Csaba Hetényi[3], Péter Rapali[2], András Láng[1], Katalin E. Kövér[4],*

*András Perczel[1,5], Weixiao Y. Wahlgren[6], Gergely Katona[6], and László Nyitray[2**]*

[1]Laboratory of Structural Chemistry and Biology, Institute of Chemistry; [2]Department of Biochemistry, Eötvös Loránd University, 1117 Budapest, Pázmány Péter sétány 1/A, Hungary; [3]Molecular Biophysics Research Group, Hungarian Academy of Sciences, PázmányPéter sétány 1/C, Budapest, Hungary;[4]Department of Inorganic and Analytical Chemistry, University of Debrecen, Debrecen, Hungary; [5]MTA-ELTE Protein Modelling Research Group, Eötvös Loránd University, Budapest, Hungary; [6]Department of Chemistry and Molecular Biology, University of Gothenburg, Sweden.

* These authors made equal contribution to the present work

** Corresponding author phone/fax: 36-13812171/36-13812172, e-mail: nyitray@elte.hu

ABBREVIATIONS:

CT-HSQC: Constant Time Heteronuclear Single Quantum Coherence,

DSS: 4,4-dimethyl-4-silapentane-1-sulfonic acid,

DYNLL1/2: LC8 vertebrate LC8 dynein light chain isoforms,

HSQC: Heteronuclear Single Quantum Coherence,

IDP: intrinsically disordered protein,

myo5: myosin 5a,

MES: 2-(*N*-morpholino)ethanesulfonic acid,

NMR: Nuclear Magnetic Resonance,

NOE: Nuclear Overhauser Effect,

R_1: longitudinal relaxation rate constant,

R_2:transverse relaxation rate constant,

SCS: Secondary Chemical Shift,

STD: Saturation Transfer Difference,

T1: longitudinal relaxation time,

T2: transverse relaxation time.

ABSTRACT: LC8 dynein light chains (DYNLL) are conserved homodimeric eukaryotic hub proteins that participate in diverse cellular processes. Among the binding partners of DYNLL2, myosin 5a (myo5a) is a motor protein involved in cargo transport. Here we provide a profound characterization of the DYNLL2 binding motif of myo5a in free and DYNLL2 bound form by using NMR spectroscopy, X-ray crystallography and molecular dynamics simulations. In the free form the DYNLL2 binding region, located in an intrinsically disordered domain of the myo5a tail, has a nascent helical character. The motif becomes structured and folds into a β-strand upon binding to DYNLL2. Despite all differences of the myo5a sequence from the consensus binding motif, it accommodates into the same DYNLL2 binding groove as all other partners do. Interestingly, while the core motif shows similar interaction pattern in the binding groove as seen inother complexes, the flanking residues make several additional contacts, thereby lengthening the binding motif. The N-terminal extension folds back and partially blocks the free edge of the β-sheet formed by the binding motif itself. The C-terminal extension contacts the dimer interface and interacts with symmetry related residues of the second myo5a peptide. The involvement of flanking residues of the core binding site of myo5a could modify the quaternary structure of the full-length myo5a and affect its biological functions. The presented structural data widen our understanding of the diverse partner recognition of DYNLL proteins and provide an example of a Janus-faced linear motif.

The LC8 dynein light chains were originally described as the smallest subunit of the microtubule-associated cytoplasmic dynein motor complex.They show a highly conserved amino acid sequence throughout evolution and they were later shown to bind to more than 70 other proteins involved in various biological processes, therefore they are now considered hub proteins (*1, 2*).Vertebrates contain two closely related LC8 paralogs, DYNLL1 and DYNLL2 with partially overlapping functions(*1*). Under physiological conditions LC8proteins arestable homodimers (*3*).Atomic resolution structureswere determined both for the *apo* and ligand-bound form using X-ray crystallographyand NMR spectroscopy (*4-6*). Five β-strands make up two core β-sheets; one sideof each sheet is flanked by two α-helices, and the otherside forms the dimer interface. The β3-strand of one subunit pairs to theβ2'-strandof the other subunitin an antiparallel mode; and the dimer is stabilized by interface contacts through hydrophobic interactions and side chain H-bonds. Two equivalent grooves are formed and they represent the target binding pockets of the dimer.

LC8 binding motifs are usually localized near coiled-coil or other dimerization domains in intrinsically disordered segmentsof theirinteracting proteins. Unstructured polypeptide regions are often involved in interactions that are mediated by sequential binding sitesknown as linear motifs (*7, 8*). The LC8 binding motif presents a relatively weak consensus sequence $[DS]_{-4}K_{-3}X_{-2}[TVI]_{-1}Q_0[TV]_1[DE]_2$ with a conserved glutamine in the arbitrary chosen coreposition 0 (*9*). A few binding motifs differ considerably from the above consensus sequence; among them is myosin 5a in which position 0 is substituted by a methionine.

Myosin 5a (myo5a) is an intracellular motor protein involved in short-range vesicle transport along actin filaments playing various roles in different cell types (*10, 11*). Its structureis divided into three main parts: (i)the N-terminal motor domain responsible for force generation and ATP hydrolysis; (ii)the long, helical neck domain stabilized by the binding of six light chains; (iii) the coiled-coil tail ending in two globular cargo binding domains(*12*) (**Figure 1**).Within the long tail domain three coiled-coil segments, interrupted by flexible non-coiled-coil loops, are responsible for dimerization of the two heavy chains. The primary transcript of the human MYO5A geneis processed by alternative splicing (*13*). Three exons

within the tail domain (B, D, and F) are expressed in a cell type-specific manner; the predominant melanocyte- and brain-specific isoforms contains exons D, F, and exon B, respectively (*13*). The exon pattern of myo5a tail likely determines which cargo binds to the motor, because these sequences could serve as binding sites of adaptor proteins: exon D and F are necessary for the binding of Rab10 and melanophilin, respectively(*14*), (*15*). In a previous study, we identified the DYNLL2 binding motif of myo5a in a short sequence (Ile1280-Ile1294) situated between the medial and distal coiled-coils of the tail (*16*). The three-residue-long alternatively expressed exon B within the binding motif was shown to be necessary for the interaction with DYNLL2. Binding of DYNLL2 stabilizes the flanking coiled-coil sequences (*16, 17*). Detailed kinetic and thermodynamic characterization of this interaction using mono- and dimeric myo5a fragments revealed that the dissociation constant is somewhat weaker than that of the canonical motifs(~1-5 µM for a monomeric peptide) andbivalency of the ligand leads to pronounced avidity (*18*).

LC8 proteinsare subunits not only of the dynein (*19*), but also of the myo5a motor complex (*20*), and they are known binding partners of several proteins transported by these motors (e.g. nNOS (*21, 22*), Pak1 (*23-25*), Bmf (*26*), Bim (*27*), Bassoon (*28*), GKAP (*29*), and viral proteins (*30, 31*)), it was hypothesized that DYNLL may provide direct physical linkage between the motor and the cargos (*26, 28, 29*). However, as the majority of the binding partners (including the motors) are dimeric, and therefore they are bivalent ligands of LC8 and the formation of highly stable dimer-to-dimer complexes is preferred (*18*). Therefore, the validity of the original cargo adaptor hypothesis is questioned.

Here we extend our previous studies (*16*) and extensively characterize the DYNLL2 binding motif of myo5a and its complex with DYNLL2. NMR experiments show that the 23-residue-long DYNLL2-binding region of myo5a is unstructured with a nascent α-helical element in the*apo* form whichundergoes a disorder-to-order transitionand folds into a β-sheet upon complex formation. Our results suggest that the interacting surface in the DYNLL2-myo5a peptide complex is more extended compared to other LC8-partner complexes. NMR results were supported by molecular dynamics simulations performed on the free peptide and on the complex. Working with a shorter (14residues) myo5a peptide, we solved the

crystal structure of the DYNLL2 complex at 1.85Å resolution. The structure explains the role of the M_0 instead of the canonical Q_0 residue in the consensus binding sequence.

MATERIALS AND METHODS

Cloning, protein expression, peptide synthesis and purification. Human DYNLL2 (UniProt accession number: Q96FJ2) and a fragment of human myo5a (gshmSQKEAIQPKDDKNTMTDSTILLE, UniProt: Q9Y4I1, residues: 1275-1297, referred hereafter as M_L) was cloned, expressed and purified as previously described (*16*). For NMR experiments isotope labeled proteins were expressed in minimal media containing $^{15}NH_4Cl$ and/or ^{13}C glucose. The synthetic peptide (IQPKDDKNTMTDST, N-terminally acetylated, C-terminally amidated; referred hereafter as M_S) corresponding to residues 1280-1293 of human myo5a was obtained from Peptide 2.0 Inc. (Chantilly, USA). The peptides were dissolved in 0.1% TFA/water and further purified by reverse-phase HPLC on a Jupiter 300 C18 (300 Å 10x250 mm) column (Phenomenex, Torrance, USA). Elution was performed using a gradient (0% to 100%; 1%/min) of 0.1% TFA/acetonitrile. Fractions containing the peptide were pooled and lyophilized. The mass of the lyophilized powder was determined by direct measurement, and it was used later for concentration calculations. The identity and purity of the peptides were confirmed by mass spectrometry.

NMR measurements. NMR spectra were collected on a Bruker Avance II 500 spectrometer operating at 500.13 MHz for ^1H and on an Avance III 700 spectrometer operating at 700.17 MHz for ^1H, both equipped with 5-mm triple-resonance probe-heads with z-axis gradients. Sample temperatures varied in the 278-303 K temperature range. ^1H chemical shifts were referenced to the internal DSS standard, while ^{13}C and ^{15}N chemical shifts were referenced indirectly via the gyromagnetic ratios. Both ^{15}N- and ^{13}C,^{15}N-labeled M_L samples had the following compositions: ~0.8 mM peptide, 20 mM phosphate buffer, 20 mM NaCl, 3 mM NaN_3, 3 mM TCEP at pH 7.00, and 10% D_2O.

Sequential connectivity, backbone and side-chain resonance assignments were determined from standard triple resonance HNCA, HN(CO)CA, HNCACB, (H)CC(CO)NH measurements using uniformly

^{13}C,^{15}N-labeled peptide. ^{15}N T_1 and T_2 relaxation times and the ^{1}H-^{15}N NOE were measured using sensitivity and gradient improved pulse sequences.For T_1 measurements a series of 9 experiments with relaxation delay times ranging from 0.011 s to 2.701 s were applied, while for T_2 measurements relaxation delay times varied between 0.0 to 0.3 s. Estimation of errors was accomplished by repetition of measurements with identical relaxation delay. The relaxation time is determined from the decay of the intensity of the corresponding ^{15}N-NH cross peaks taken from the ^{1}H-^{15}N HSQC spectra, fitted with a two-parameter exponential decay in Sparky(*32*). NOE on/off spectra were recorded in an interleaved manner with a recycle delay of 6-7 s. The {^{1}H}-^{15}N NOE is calculated as the ratio of peak intensities from the spectra acquired with and without proton saturation.

Chemical shift mapping was performed by ^{1}H-^{15}N HSQC measurements. Combined chemical shift changes were calculated for the non-overlapping residues based on the following equation:

$$\Delta_{ppm} = [(\Delta\delta_{HN})^2 + (0.17 \cdot \Delta\delta_N)^2]^{\frac{1}{2}}$$

where $\Delta\delta_{HN}$ and $\Delta\delta_N$ represent differences in the chemical shifts for amide protons and nitrogens both for the free M_L and the DYNLL2-M_L complex, while 0.17 is the normalizing factor (*16*).

^{13}C- and ^{15}N-detected saturation transfer difference (2D STD-HSQC) experiments were acquired on a sample containing 1.5 mM ^{13}C,^{15}N-labeled M_L and 40 μM unlabeled DYNLL2. Selective saturation was achieved by a train of 50 ms E-BURP1 (90°) shaped pulse at -400 Hz (in the aliphatic region of the folded protein) or at 5500 Hz (the tryptophan indole peak of the protein) for a total saturation time of 2.0 s. The spectrum was recorded with 400-600 transients, and STD was monitored either by a ^{15}N HSQC or a ^{13}C CT-HSQC pulse sequence.Data were processed with TopSpin and resonance assignment was done by the CARA software (*33*), while relaxation parameterswere evaluated by the Sparky program (*32*). Reduced spectral density mapping analysis was conducted on the basis of Lefevreet al (*34, 35*). Spectral densities at given frequencies were calculated from the measured T_1, T_2 and heteronuclear NOE values for each amide-N atom.

Protein crystallization, X-ray diffraction, and model building. DYNLL2 was dialyzed twice against TBS (20 mM TRIS, 150 mM NaCl, 3 mM NaN_3, 5 mM DTT, pH 7.6) and concentrated using Amicon Ultra-4 filter with Ultracel-3 membrane (Millipore Corporation, Billerica, USA). The M_S peptide was dissolved in the same buffer and subsequently complexed with DYNLL2. DYNLL2 and M_S concentrations in the final solution of the complex were 1.5 mM and 2.0 mM, respectively.

Crystals were grown using the hanging drop vapor diffusion method at 293 K with a reservoir solution of 1.8 M $(NH_4)_2SO_4$, 0.02 M $CoCl_2$, 0.1 M MES pH 6.5 and a 3.5 µl drop composed of 1 µl reservoir solution and 2.5 µl protein solution. Crystals appeared within 2-3 days and reached a maximum size of 0.5 mm x 0.5 mm x 0.2 mm. The crystals were soaked in reservoir solution plus 20% glycerol for 1 min and subsequently flash frozen in liquid nitrogen. Single-crystal X-ray data were collected at 100 K with an ADSC Q315R CCD detector at ID29 of the ESRF (λ=0.93 Å). The oscillation range per image was 0.5°. Data were indexed, integrated and scaled to a resolution of 1.85 Å using XDS and XSCALE (*36*). Completeness of the data was 99.72% after merging the low and high resolution datasets. Crystals belonged to space group $P6_122$ with unit cell dimensions of a=b=45.2 Å, c=204.0 Å. Data statistics are summarized in **Table S1**.

The structure was solved by molecular replacement using the program PHASER (*37*) of CCP4 6.1.2 Program Suite (*38*). The structure of the PDB entry 1CMI was used as search model. After rigid body and restrained refinement with refmac5 (*39*), automated model building was carried out with Arp/wArp (*40*). During the refinement, one cobalt (II) ion and 134 water molecules were modeled. (The cobalt (II) ion was coordinated by four nitrogen atoms and one oxygen atom of DYNLL2. These nitrogens belong to G_{-2}, S_{-1}, H_0 residues, which are cloning artifacts.) The model was then systematically improved using iterative cycles of manual rebuilding with the program Coot (*41*). The model building was completed with R_{cryst} = 19.3% and R_{free} = 23.3%.

Molecular dynamics calculations. The M_L peptide was used for the molecular dynamics (MD) calculations. To gain information on the conformational dynamics, the peptide molecules were calculated both in unbound and in protein-bound forms using the GROMACS MD software (*42*). The starting

geometry of DYNLL2homodimer and the core (Q_{-8}-S_3) region of the two identical peptides and water molecules residing at the molecular surfaces were extracted from the crystallographic structure determined in the present study. The missing, terminal regions of the peptide were modeled using the TINKER program package (*43*) and some text editing settingto all-β backbone conformation. In this way, the terminal parts of the bound peptide were directed towards the bulk avoiding biased initial interactions with the protein partner. The AMBER03 (*44*) force field was applied along with neutralizing counter-ions and numerous TIP3P explicit water molecules (*45*)filling 9 Å spacing between the protein parts and the simulation box edges. The systems were refined with steepest descent and BFGSmolecular mechanics energy minimizations. The refined systems were subjected to 800-ns-MD simulations. A time step of 2 fs, constant temperature of 300 K, and LINCS bond constraints were applied. Temperature coupling was carried out using the v-rescale scheme. Long-range electrostatics was calculated using the particle-mesh Ewald method, and a van der Waals cutoff of 9 Å was used. Neighbor lists wereupdated every 10 fs. The resulting trajectory files were analyzed with programs of the GROMACS package and the DSSP program (*46*). Structural figures and movies were prepared using the PyMol program (*47*).

RESULTS

Selection of an extended DYNLL2 binding region of the myo5a heavy chain. We analyzed the sequence of myo5a by coiled-coil (COILS) (*48*), disorder (IUPred) (*49, 50*), and disordered binding motif (ANCHOR) (*51, 52*) predicting algorithms. The tail region (V915-V1855) begins just after the last IQ motif of the neck. As it was expected, COILS predicted three main coiled-coil regions in the tail: the proximal (V915-V1105), the medial (D1151-A1234) and the distal (V1338-G1446 in isoform 1 containing exons B and D); I1313-G1419 in isoform 2 corresponding to the predominant brain-specific splice form; I1313-G1471 in isoform 3 corresponding to the predominant melanocyte-specific splice form) coiled-coils.

The medial coiled-coil is disrupted by a proline residue (P1194; **Figure 1**). Similarly, two adjacent prolines disrupt the distal coiled-coil (P1392, P1393 in isoform 1 and 3; P1365, P1366 in isoform 2).

Interestingly, these disruptions are almost exactly in the center of their respective coiled-coil regions (before and after the breaks there are ~40-40 and ~50-50 residues in the case of the medial and distal coiled-coils, respectively). Moreover, coding sequences of the alternatively spliced exons D (in isoform 1 and 3) and F (in isoform 3) cause only negligible changes in the length of these coiled-coil segments: exon D extends the C-terminal part of the non-coiled-coil region between the medial and distal coiled-coils, while exon F introduces a new proline (P1423 in isoform 3) into the sequence, giving chance to the formation of a new, ~30-residue-long non-coiled-coil region flanked by the two segments of the distal coiled-coil.

Exon B encodes only three residues (DDK, 1284-1286) within the non-coiled-coil segment between medial and distal coiled-coils. IUPred predicted both long and short disorder around this region. Interestingly, while ANCHOR predicts binding sites within exon D and F, there is no binding site predicted within this disordered region. Taking into account the result of these predictions we have chosen the long disordered region between S1275 to E1297 as an extended DYNLL2 binding motif (**Figure 1**) for further NMR experimental analysis (M_L peptide, numbered here as: S_{-14}-E_8).

NMR characterization of the extended DYNLL2 binding motif of myo5a.

Resonance assignment. The 27-residue-long M_L consists of mainly hydrophilic residues with a small hydrophobic inner part ($A_{-10}L_9$) and a hydrophobic C-terminus ($I_5L_6L_7$). Note that the N-terminal four residues are from the expression vector.Confirming the IUPred analysis no tendency for ordering is observed. Both ^1H and ^1H-^{15}N HSQC spectra have low signal dispersion, most NMR resonances fall close to the average random coil value, resulting in signal overlap for several residues, as seen on Figure 3A. ^{15}N edited NOESY spectra showed only sequential connectivities, which is clear evidence for lack of secondary structural elements under the given experimental conditions. Secondary chemical shifts (SCS) were calculated using temperature and sequence corrected random coil chemical shift values (*53, 54*). SCS data for $\Delta C\alpha$ (**Figure 2A**), $\Delta(C\alpha$-$C\beta)$ and $\Delta H\alpha$ (Supplementary material table), along the amino acid sequence have values around zero, an indicator for disordered structure. On the other hand, in the D_{-5}-T_4 segment an inherent helical preference is observed: $\Delta H\alpha$ values are slightly negative and the

corresponding $\Delta C\alpha$, $\Delta C\alpha$-$C\beta$ values are slightly positive. Compared to the values obtained for a real helix, the helical tendency of this region is around 10% at 288K. However, this observation is valid for the M_L segment, and can not be safely stated for the long intact protein. On the other hand an increase in temperature causes higher mobility, this

Relaxation Parameters and Reduced Spectral Density Mapping Analysis. ^{15}N relaxation rates (R_1 and R_2 see Supplementary material table) of the free M_L show an unsymmetrical bell-like profile with a plateau decreasing towards the C-terminus. The N-terminus is highly mobile since corresponding peaks are below detection limit. The R_2/R_1 ratio along the sequence falls in the 1.0-1.4 region (**Figure 2B**) indicating that the ^{15}N relaxation is mostly dominated by higher frequency motions. Heteronuclear NOE values are more negative towards both N- and C-termini, showing a variation between -0.90 and -0.03, in accordance with fast motions and high flexibility. Higher NOE enhancements are obtained for the D_{-4}-T_{-1} region and the highest value at I_9 indicating that slower motions characterize these residues(**Figure 2C**).

Our results suggest that the relaxation of the I_9-T_{-1} region is not influenced, or it is the least influenced by the segmental motions of the backbone atoms (ps-ns scale). On the basis of these relaxation parameters the reduced spectral density mapping analysis allows a more direct evaluation of the relative degree of motion at ω_N, $0.87\omega_H$ and zero frequencies ($\omega_0=\nu_0 2\pi$). Restricted motions towards the µs time scale are indicated by larger J(0) values, as observed for the K_{-6}-N_{-2} region with local maximum at N_{-2}, while all other residues fall in the vicinity of the single motion limit curve (**Figure 2D**), indicating higher backbone flexibility. The overall conclusion of the NMR relaxation data is that free M_L is rather flexible and has a quite unstructured backbone, however, a weak, but noticeable helical tendency is detected in the central region accommodating the core recognition motif D_{-4}-T_1. We cannot directly extend this conclusion to the full length myosin 5a heavy chain, however, these results suggest that the connecting region between the medial and distal coiled-coil domain is mostly unstructured with some propensity to fold into an α-helix around D_{-4}-T_1.

Complex formation of the extended myo5a peptide with DYNLL2

Chemical shift mapping.Successive addition of unlabeled DYNLL2 to the $^{13}C,^{15}N$-labeled myo5a fragment results in the formation of a 2:2 complex (consisting of 232 residues in total). A molecule of such size tumbles slower in solution than does the M_L partner, and as a consequence it will present broader peaks in the spectra. **Figure 3A** shows the overlay of ^1H-^{15}N HSQC spectra obtained at 288 K for the free and DYNLL2 bound M_L. As this complex formation is slow on the NMR time scale, two sets of peaks appear, the narrow set belongs to the free, while the broad ones to the complexed M_L. The M_L peaks decrease in intensity while those of the complex increase. Several peaks undergo very large chemical shift variations; therefore, assignment for the DYNLL2-M_L complex peaks had to be obtained separately from triple-resonance measurements. Regions K_{-6}-K_{-3} and M_0-T_1 show the largest shifts (**Figure 3B**). ^1H-^{13}C CT HSQC measurements (**Figure S1**) reveal the disappearance of Cα, Cβ and Cγ carbons in the Q_{-8}-T_1 region, and decreasing intensities are observed for the I_{-9} environment and the D_{-2}-L_{-7} region, while intensities of the G_{-18}-A_{-10} and E_8 residues remain unchanged. On the other hand, calculated SCS values for Cα(**Figure 2A**) and Cα-Cβ(Supplementary material table) are highly negative for the K_{-6}-T_1 region indicating a β-sheet formation upon complexation. This concludes, that the K_{-6}-T_1 region is the primary bindingsite. In parallel, the intrinsic helicity characteristic of the free peptide transforms into a pronounced secondary structural element, adopting a β-sheet structure when binding to DYNLL2.

Saturation transfer difference (STD) measurements. Ligand-detected STD experiments are one of the most powerful methods in determining binding sites on the ligand, with best performance for K_d values in the μM to mM range (*55*). Note that the K_d of M_L to DYNLL2 was measured to be ≈9 μM (*18*). As DYNLL2 is a well folded protein, the region of choice for selective irradiation corresponds to the strongly shifted methyl resonances situated between −1 and −2 ppm. While the M_L fragment has no tryptophan residue, DYNLL2 possesses one, enabling selective saturation at around the indole environment at 10 ppm. Performing both the conventional ^1H STD and the group-selective STD experimentsweak STD signals were obtained. However, due to severe signal overlap in the obtained 1D STD spectra no unequivocal assignment could be assessed for the bound residues. Using $^{13}C,^{15}N$-labeled M_L, the initial^1H STD experiment could be appended with 2D ^{15}N HSQC or 2D ^{13}C CT-HSQC

sequences. As shown on **Figure 4A** STD effects were detected for the Q_{-8}, L_6, E_8 residues and at a lower extent for the L_7 backbone NH. The ^{13}C STD-(CT)-HSQC was proved to be more informative (**Figure 4B**), highlighting several affected methyl groups: the overlapping $C\gamma2s$ for L_9/I_5; the overlapping $C\delta1$ and $C\delta2$ for L_6/L_7; the overlapping $C\gamma2s$ for $T_{-1}/T_1/T_4$; $C\beta$ resonances for A_{-10}; $C\varepsilon$ for M_{-15}; $C\beta$ environments of H_{-16} (M_{-15} and H_{-16} are derived from the expression vector) and K_{-6}, $C\gamma s$ for E_{-11} and E_8, and $C\alpha$ for L_6 and L_7. Thus, STD results highlight the M_L-DYNLL2 interaction at a residue specific manner. Especially the hydrophobic contacts were revealed covering the T_4-E_8 and E_{-11}-L_9 segments, and some additional residues (T_1, T_{-1}, K_{-6}). The binding seems tighter at the C-terminus of M_L, and the role of the side-chains in complex formation will be discussed below.

Crystal structure of DYNLL in complex with the binding motif of myosin 5a. The inherent flexibility of the N- and C-terminal ends of the M_L hindered crystallization of the DYNLL2-M_L complex. However, removal of several residues from both termini resulted in the shorter M_s peptide (**Figure 1**) that proved to be suitable for crystallization and the structure of the DYNLL2-M_S complex could be solved at 1.85 Å resolution.

The overall conformation of DYNLL2 in the complex highly resembles those found in other complex structures of this family (1CMI (*5*); 3DVP (*56*); 3E2B, 2P2T (*57*); 2PG1(*58*); 3FM7, 3GLW (*59*); 3P8M, 2XQQ (*9*)). One M_S peptide is bound to each binding groove of the DYNLL2 (**Figure 5A**). The walls of the grooves are defined mainly by side chains of aromatic residues (F62, Y65, H68, F73, Y75 and Y77). At the bottom of the grooves six hydrogen bonds (M_S-D_{-5}(CO) to DYNLL-H68(NH), M_S-K_{-3}(NH) to V66(CO), M_S-K_{-3}(CO) to V66(NH), M_S-T_{-1}(NH) to S64(CO), M_S-T_{-1}(CO) to S64(NH), M_S-T_1(NH) to F62(CO)) connect the backbone atoms of M_S with the edge of the outer β-strands of DYNLL2, thereby extending the β-sandwich core by a new antiparallel β-strand (**Figure 5B**). The core binding motif begins with three residues corresponding to exon B ($D_{-5}D_{-4}K_{-3}$). The side chains of these residues make extensive contacts with DYNLL2: the carboxyl group of D_{-5} forms a hydrogen bond with the hydroxyl group of T70, the carboxyl group of D_{-4} forms hydrogen bonds simultaneously with the hydroxyl groups of T67

and Y65 and the methylene groups of K_{-3} have hydrophobic contacts with the aromatic ring of F73, while the amino group at the end of this side chain forms a hydrogen bond with N^{ε} of H68 (**Figure 5C**).

We havepreviously shown that the myo5a splice variant lacking exon B does not bind to DYNLL2(*16*). If D_{-5}, D_{-4} and K_{-3} were missing from M_S, they would be replaced in the complex by Q_{-8}, P_{-7} and K_{-6}, respectively. As the exchange of K_{-3} to another Lys would not cause any difference, the other two positions must be responsible for the dramatic affinity reduction. As stated above, D_{-5} forms one hydrogen bond with T70. Gln has a somewhat longer and more flexible side chain compared to Asp. Because both D_{-5} and T70 are located at the surface of the complex, it seems to be very likely that Gln could also make a similar hydrogen bond at this position without making significant strains in the structure due to the longer side chain. However, the exchange of D_{-4} to a Pro would make it impossible to form two hydrogen bonds (with the hydroxyl groups of T67 and Y65) and probably would force the N-terminal end of the peptide (Q_{-8} in this case) to a conformation where the backbone hydrogen bond at the preceding position (between Q_{-8} and H68) could not form either. The loss of these hydrogen bonds and the possible shortening of the β-strand explain why exon B is necessary for binding of myo5a to DYNLL2.

In addition to the three residues corresponding to exon B, five more (N_{-2}, T_{-1}, M_0, T_1, D_2) residues of M_S make direct contacts with DYNLL2 (**Figure 5**). The side chain of N_{-2} is within hydrogen bonding distance to two backbone carbonyl groups of DYNLL (K36 and S64) and to the amino group at the side chain of K36. As it was expected (*16*), the central TQT motif of the canonical sequences is replaced with the $T_{-1}M_0T_1$ motif at equivalent position. The hydroxyl group of the side chain of T_{-1} plays a special role in the complex; forms a hydrogen bond both with the hydroxyl group of S64 and with a water molecule deeply wedged between the peptide and DYNLL2. This water molecule is also within hydrogen bonding distance to a hydroxyl group (Y75) and a carbonyl oxygen (F62) of DYNLL2 (**Figure 5D**).

The non-canonical M_0 residue adopts a conformation very similar to the conformation of the typical Q_0, and it is able to participate in similar interactions at the surface of DYNLL2(see Discussion). The antiparallel β-strand conformation of the peptide ends at T_1, because there is a hydrogen bond between

the backbone nitrogen atom of F62 and the hydroxyl group of the side chain of T_1, instead of the carbonyl oxygen of T_1(**Figure 5B and 5D**). At this point the peptide turns outwards from the binding groove. Interestingly, neither $M_0(CO)$ nor $D_2(NH)$ involved in the β-sheet formation, but make hydrogen bonds with the hydroxyl groups of Y75 and Y77, respectively. These hydrogen bonds help to hold the planes of the respective peptide bonds in a "rotated" position, resulting in the side chain of T_1 to turn towards the bottom of the binding groove to break the hydrogen bonding network of the antiparallel β-strand, and leading the peptide out from the binding groove. This conformation is further stabilized by a salt bridge between D_2 and K9 (**Figure 5D**).

The most remarkable feature of the complex is the N-terminus of M_S (Q_{-8}, P_{-7} and K_{-6}). Although the quality of the electron density map was poor in this region, this short stretch of residues could be identified in the structure in an unexpected position. The peptide backbone makes a β-turn-like structure (hydrogen bond between the carbonyl oxygen of P_{-7} and the NH of D_{-4}), and the N-terminus of M_S folds back towards the core binding motif, making several favorable interactions with residues corresponding to exon B. The most important interactions stabilizing this conformation are two additional hydrogen bonds, one between the carbonyl oxygen of K_{-6} and the NH of K_{-3}, and another between the carbonyl oxygen of D_{-4} and the side chain amide nitrogen of Q_{-8}(**Figure 5D**). Although the weak electron density and high B factors of the atomic model indicate that the positions of these residues are not well-defined, this conformation is unique among DYNLL-partner complexes and may also have some biological significance. Atomic details of this "backfolding" of myo5a heavy chain peptide were further investigated by MD simulations.

Molecular dynamics simulation of the myo5a peptide in free form and in complex with DYNLL2. To complement the results obtained by NMR spectroscopy and X-ray crystallography MD simulations were completed both with free $M_L$and DYNLL2-M_L complex. These simulations provided more information about the behavior of DYNLL binding motif especially about the N- and C-terminal extensions, because the quality of the electron density map was poor at the N-terminus of the $M_S$in the crystal complex, and any attempts to crystallize DYNLL2 with M_L were unsuccessful. MD is also the best

suited method to study the role of water molecules mediating the interaction between myo5a and DYNLL2.

During MD, the unbound M_L quickly lost its bound (partially β-sheet) geometry observed in the crystal structure and used as an initial structure for the calculation. After the first 250 ns of the MD simulation, a bend→turn→α-helix conformational transition of the D_{-4}-T_1 sequence was observed. After 350 ns, the α-helical secondary structure is stabilized for the terminal 450 ns of the trajectory (**Figure 6 and Movie S1**). Remarkably, the stabilization of the α-helical structure for the D_{-4}-T_1 core region is in accordance with our NMR spectroscopic analysis, namely, that this region has a propensity to fold into an α-helix.

As mentioned above the D_{-4}-T_1 region of the peptide is in an antiparallel β-sheet arrangement in the crystal structure of the DYNLL2-M_S complex. The MD calculation of the complex also confirms this finding and emphasizes that the β-sheet is especially stable at the N_{-2}-M_0core residues, where both copies of the peptide remained tightly H-bound during the entire simulation time (**Figure 6 and Figure S3**). For the second copy (chain D) this region extends to the entire P_{-7}-M_0 region whereas for the first copy (chain B) the appearance of the native-like unbound secondary structure can be observed at the P_{-7}-N_{-2} region . However, this reduction of the antiparallel β-region does not lead to the dissociation of the peptide.The N- and C-terminal parts of the bound peptide do not remain in the extended all-β conformation set as initial conformation (see Methods). Both termini quickly bend over the neighboring parts of the protein (**Movie S2**) and occasionally adopt helical and β-turn conformations during the MD calculations (**Figure 6 and Figure S3**).

DISCUSSION

The unstructured DYNLL2 binding motif of myosin 5a shows a weak helical tendency and folds into a β-sheet upon interaction with DYNLL2. In solution the free M_L shows a very dynamic and highly unstructured nature with a tendency to adopt helices (cca. 10% helicity detected)as

proved both by NMR spectroscopy and MD simulations. Measured relaxation rates are in the average range for the whole molecule, showing local minima in the hydrophobic core at the I_9 residue both for R_2 and R_1. It is known that R_1 average relaxation rates for disordered proteins do not vary dramatically if measured at the same field strength. For the 27-residue long M_L peptide this value is $R_1^{av} = 3.03 \pm 0.54$ s^{-1} at 280K. It was expected that the K_{-6}-D_2 region, including the exon B coded residues and the core motif, should be responsible for binding to DYNLL2. In free M_L, heteronuclear NOE values show the most positive values in the I_9-T_4 region in accordance with increased structural propensity that coincides perfectly with the indicated inherent helical tendency of the K_{-6}-T_4 region indicated by H$^\alpha$ SCS data. The spectral density function at zero frequency reveals regions with restricted motions on the sub-nanosecond time scale for residues that bear a larger than average J(0) value. In this respect, the K_{-6}-K_{-3} segment fulfills this criterion, and again it is in full agreement with the inherent helicity of this region.

IDPs and IDP segments have the tendency of showing mostly inherent helical regions (*60-62*) and the corresponding segments often bear important functional characteristics or contain the binding motifs to partner proteins. A consequence of interaction with the binding partner is the folding of the unstructured region, and formation of all major secondary structure types (helix, sheet, coil) is possible. There are proteins, like the disordered C-terminal domain of the p53 hub protein, which (depending on the binding partner) can fold into all three secondary structure types (*63*). The disorder-to-order transition of the M_L upon binding to DYNLL2 results in the formation of a β-sheet. Proof is obtained from calculated SCS data for both C$^\alpha$ and (C$^\alpha$-C$^\beta$) resonances: the pronounced negative SCS values in the K_{-6}-M_0 region indicate the presence of a β-sheet. On the other hand, chemical shift mapping results show that the K_{-6}-T_1 segment is the main binding site. All our results emphasize that working with a relatively short peptide accommodating the binding sequence it is possible to gather information about the resulting secondary structure of the folded/complexed state (this one being the first example for the LC8 protein family).

Benison *et al.* conducted an NMR study of complex formation between the N-terminal domain of the dynein intermediate chain (IC84-143) and *Drosophila* LC8 (*64, 65*). The IC84-143 segment contains the highly conserved 'TQT box' and binding to LC8 caused broadening below detection limit of the peaks

belonging to the residues involved in binding; a similar broadening effect that was observed for M_L, while the peaks of the residues outside the binding sequence remained unperturbed and maintained the disorder state outside the binding region. On the contrary here we show not only the partial folding of the unbound M_L, but also the structural transition into a β-sheet upon interaction with DYNLL2, an example of structural plasticity in a short linear motif (**Figure 7**).Previously we have shown by kinetic studies that the binding of DYNLL isoforms to their various partners can be best described by a conformational selection model (*18*). The region of the free myo5a peptide contains an inherent α-helix that should be in the unfolded state in order to fit into the binding groove of DYNLL2; this finding is consistent with a conformational selection mechanism for the formation of DYNLL complexes.

The non-canonical recognition motif of myo5a binds into the binding groove of DYNLL2 as other known binding partners. All partners (including the non-canonical myo5a and Pak1) bind to DYNLL by extending one of the β-sheets in the β-sandwich core of DYNLL by a new, antiparallel strand. Side chains of adjacent residues in an antiparallel β-strand point into alternating directions (**Figure 8A**). This arrangement minimizes the possible side chain to side chain contacts in the motif, and maximizes the independence of the individual positions. The surface of DYNLL provides different binding pockets for each position. Close examination of the binding groove reveals that side chains at positions 1, 0, -1 and -3 accomodate into the best defined, deep pockets of DYNLL. Residues in other positions face towards the solvent and/or acomodate into a more permissive, shallow pocket, therby have less interaction with DYNLL. These observations are in good agreement with the results of phage display experiments (*9*) as well as with *in silico* analysis of the contribution of key residues and their context to the global binding energy(*66*).

The backbone of bound polypeptides in all structures occupy almost the same position within the groove of DYNLL (**Figure (8B)**. This mainly results from the antiparallel β-sheet conformation. The hydrogen bonding network necessary for antiparallel β-sheet formation usually starts at position −6 and ends at position 1 (**Figure (8B)**. The DYNLL1-nNOS complex (1CMI) is an exception where the β-sheet extends to the NH group at position 3 (**Figure 8C**). All DYNLL recognition motifs, except in nNOS and

Pak1, have a Thr at position 1 (T_1). As we have already discussed, the hydroxyl group of T_1 of myo5a makes a hydrogen bond with the backbone NH group of F62. For this interaction the side chain of T_1 must turn downwards to the bottom of the groove. The Val innNOS (V_1) in the same position is unable to form hydrogen bonds with the backbone of DYNLL1; that is why the side chain of V_1 turns outwards and forces Y75 and Y77 of DYNLL1 to a distorted position (**Figure 8C**), preventing thus the interaction. These changes caused by V_1 propagate along the backbone of DYNLL, forcing the entire binding groove to a more open conformation (*4*). The width of the binding groove (defined as the distance between 63C to 9C α-carbon atoms of DYNLL (*4*)) is 12.3 Å, 13.4 Å and 14.2 Å, in the case of free DYNLL, LC8-Swallow complex (another "T_1"-type ligand) and DYNLL1-nNOS complex, respectively (*4*). This distance is 13.2 Å in the case of DYNLL2-myo5a and DYNLL2-EML3 complexes (*9*), and indeed both the peptide backbones and the T_1 side chains of these ligands superimpose very well as shown in **Figure 8D**.

Position 0 is the most conserved among the binding motifs. Except myo5a and Pak1, all partners have a Gln in this position. The side-chain of Q_0 caps one of the α-helices of DYNLL by making a hydrogen bond with the backbone NH group of K36 (**Figure 8E**). Both side chain and main chain atoms of K36, E35, and I34 participate in the formation of the well-defined pocket accommodating Q_0. The volume of the pocket would allow only residues that are equal in size or smaller than a Gln in this position. The chemical features of the pocket clearly explain why Glu and Asp, holding a net negative charge are not allowed in this position (would cause an electrostatic repulsion). On the other hand, small and medium sized, uncharged, polar (e.g. Asn) or hydrophobic (e.g. Leu) residues are not preferred either in this position due to the loss of two important intermolecular hydrogen bonds. Such stability loss is somewhat compensated by S_0 in Pak1 by forming a side chain H-bond with the amino-group of K36. However, M_0 in myo5a, fits almost perfectly into the binding pocket and its side chain accommodates the same position as does Q_0 in other partners (**Figure 8E**). Although the sulfur atom can form a weak H-bond with the backbone NH of K36, this residue is clearly suboptimal at this position of the groove. Weaker binding is evidenced by smaller K_d values: for TQT motifs (*e.g.* in Bmf) 1 μM, while for the TMT motif in myo5a the affinity is weaker (9 μM)(*18*).

Thr and Ile are the most common residues in position −1 of the motifs (*9*). They are all accommodated into the deepest pocket formed by Y75, F73, F62, L84 and S64 residues of the binding groove (**Figure 8F**). Interestingly, all crystal structures with bound peptides having a Thr in this position reveal the same interaction pattern including the deeply wedged water molecule and the hydrogen bond network connecting DYNLL and T_{-1} via this water (**Figure 5D**). This is in good agreement with our conclusions based on the crystal structure of DYNLL2-M_S complex and with the results of MD simulations presented here. The contacts of the T_{-1} side chain were detected also by STD NMR experiments. Binding of the bulkier side chain of I_1 in other structures seems to restrict any water molecule to remain in this pocket, most probably by simply limiting the space available for the water. Loss of the hydrogen bond network provided by this water is probably compensated by the favorable entropic effect related to the burial of more hydrophobic surfaces. This is in good agreement with our previous results showing that the entropic contribution is the highest in the case of the nNOS peptide compared to other partners (*18*).

Position −2 shows no obvious amino acid preference. However, *in vitro* evolution showed some preference for the bulky, aromatic side chain of Trp in this position (*9*). Based on our structural comparison, the binding pocket seems to be shallow and relatively permissive here. Side chains in this position face mostly towards the solvent and make only a few contacts with DYNLL. Myo5a seems to be a special case according to the importance of position −2. As shown above, the side-chain of N_{-2} is within hydrogen bonding distance with the backbone carbonyl groups of K36 and S64 residues of DYNLL. Moreover, the carbonyl oxygen of N_{-2} makes a hydrogen bond with the amino group of K36 (**Figure 5D**). As M_0 of myo5a is suboptimal residue one would expect a compensatory mechanism with positive contribution to the binding strength at another site. The observed interactions of N_{-2} may represent this expected mechanism.

DYNLL2 binding and folding of the myosin 5a tail. Cargo binding of myo5a has been shown to be regulated by proteolysis, phosphorylationand Ca^{2+}-binding (*67*). At low levels of Ca^{2+} and in the absence of cargos the motor adopts a closed, triangle−like conformation. In the inactive conformation the heads fold back and directly interact with the globular tail domains (GTD) (*67*). However, if all the

coiled-coil domains and the inter-coiled-coil regions were in an extended conformation, the length of the neck domains would not allow the heads to reach the GTDs. Indeed, the GTDs were shown to interact with the C-terminal part of the proximal coiled-coil region, and this interaction was proved to be necessary for the binding of the heads in the inactive conformation (*68*). It is obvious that the direction of the myosin 5a heavy chains must be totally reversed in this conformation to allow the GTDs to reach the proximal coiled-coil. The tail reversal most likely occurs in a non-coiled-coil region harboring the binding motif of DYNLL2 flanked by the medial and distal coiled-coils. Theoretically, the GTDs can fold back in a symmetric or in an asymmetric fashion. Detailed analysis of pictures obtained by electron microscopy indicates that the rotational symmetry of the dimeric myosin molecule is not preserved in the inactive conformation (*69*). This strongly supports the second model. If we suppose thatthe medial and distal coiled-coils remain intact and they are roughly equal in length, than the inter-coiled-coil region with the DYNLL2 binding site seems to be the only probable "breakpoint" in the inactive conformation in the myo5a tail (**Figure 9**). The resulting asymmetry must influence the conformation of the whole myo5a molecule (at least in the inactive state), because each GTD is oriented differently relative to the proximal coiled-coil.

In this work we have shown by X-ray crystallography, NMR spectroscopy and MD simulations that the unstructured fragment of myo5a tail corresponding to the DYNLL2 binding site and flanking sequences tends to fold upon the binding of the tail light chain. As it was revealed by our crystal structure, the N-terminal extension (Q_{-8}-P_{-7}-K_{-6}) tends to fold back and interact with the free β-edge of the bound sequence itself. STD NMR measurements and MD simulations also showed this tendency. These findings are in good agreement with our theoretical considerations: not only the position relative to the medial and distal coiled-coils, but also the inherent malleability of the unstructured DYNLL2 binding site makes this sequence to the most probable site of chain reversal in the myo5a tail (**Figure 9**)and binding of DYNLL2 might be important role in stabilizing the reversed conformation. Nevertheless, one should keep in mind that the tail light chain could be involved in the regulation only of the myosin 5a isoforms that contain the exon B coded sequence, e.g. in the neuronal-specific isoform. To our knowledge there was only one report addressing the role of the tail light chain in regulation of myo5a (*70*). The authors

concluded that binding of DYNLL2 is apparently not required for adopting the inhibitory folded conformation of the motor protein. However, we noted that a full-length myo5a heavy chain clone lacking exon B was used in those experiments therefore the recombinant protein could not bind the tail light chain. In light of our results it would worth reexamining the role of DYNLL2 in regulation of this myosin motor.

The function of the C-terminal extension (D_2-E_8) seems to be much simpler. Our STD measurements and MD simulations showed that this sequence not only interacts with the side of the canonical binding groove (side of the dimer interface) but also weakly dimerize. Hydrophobic contacts and backbone hydrogen bonds in the region of I_5-L_6-L_7 play key role in the dimerization tendency (**Movie S1**). These results suggest that DYNLL2 can promote the stabilizationof the N-terminal part of the distal coiled-coil by holding the heavy chains of the full-length myosin in close proximity (especially in isoforms where exon D is not present). Such "molecular glue" stabilization function of DYNLL has been previously shown in the case of myo5a (*16*) and other binding partners (*71, 72*).

Interestingly, myo5a heavy chain has two known ubiquitination sites within the N-terminal extension of the DNYLL binding region (K_{-6}, K_{-12}) (found in the PhosphoSite database,(*73*)). Ubiquitination is an important post-translational modification regulating degradation, cellular localization and interactions of many proteins (*74*). The unstructured and flexible DYNLL2 binding site may be easily accessible for ubiquitin ligases. On one hand, binding of DYNLL2 promotes the folding of this region, and the resulting stability (decreased flexibility) may affect the activity of ubiquitin ligases on residues K_{-6} and K_{-12}. On the other hand, ubiquitination may also influence DYNLL binding and/or the conformation of this region. Consequently, both DYNLL binding and also ubiquitination could have important roles in the regulation of myo5a.

CONCLUSION

In this work we found that the DYNLL recognition motif of the myosin 5a motor protein, despite deviations from the consensus motif, accommodates into the same binding groove of this eukaryotic hub

protein as do all its other partners. Moreover the binding motif is extended towards both termini compared to the core consensus sequence. The binding motif of the "tail light chain" (DYNLL2) of myosin 5a is embedded in an intrinsically disordered segment of the tail, between the medial and distal coiled coil domains. Interestingly, the DYNLL recognition segment has a nascent helical tendency in the free form and undergoes an α-helix to β-chain transition (apparently *via* a random coil state) upon complex formation, *i.e.* it is a sequence within the myosin 5a tail that shows high malleability, a general feature of linear motifs and molecular recognition elements in IDPs (*75*). Finally, we suggest that the bound DYNLL could be involved in the stabilization of the folded off state of myosin 5a isoforms (*e.g.* in brain cells) that contain the exon B coded sequence.

ACKNOWLEDGMENT

ACCESSION NUMBER:

Atomic coordinates of the DYNLL2-M_s complex has been deposited with the Protein Data Bank under accession code 4D07.

SUPPLEMENTARY INFORMATION:

Supplementary information contains Result and Figure about the equivalence of the binding positions in the two DYNLL2 grooves as studied by NMR spectroscopy; Results, Method description, Figure and Table about the analysis of stability of crystallographic water positions at the protein-peptide ligand interface; a supporting Table of crystallographic data statistics; supporting Figure of ^{1}H,^{13}C CT HSQC

measurements of the DYNLL2-M_L complex. This material is available free of charge via the Internet at http://pubs.acs.org.

FIGURE LEGENDS:

Figure 1: Bar representation of human myosin 5a (UniProt: Q9Y4I1-3).The schematic view depicts the N-terminal motor domain (red, residues: 1-765), the neck domain (residues: 766-914) with six individual IQ motifs (yellow and green), the tail domain (residues: 915-1880). The latter is divided into five putative coiled-coil ("CC"; orange) sequences (residues: Proximal: 915-1105, Medial1: 1151-1193, Medial2: 1195-1234, Distal1: 1338-1391, Distal2: 1427-1471) interrupted by non-coiled-coil regions (gray; residues: 1106-1150; 1235-1338; 1392-1426) and ends in the C-terminal globular tail domain (blue; residues: 1472-1880). The region of the medial and distal CCs (enlarged) is subjected to alternative splicing. Three short exons (referred to as exon "B", "D", and "F"; residues: 1284-1286, 1321-1347 and 1414-1438, respectively) are expressed alternatively in different isoforms of myosin 5a; the predominant brain-specific isoform contains exon B, while the melanocyte-specific contains isoform exons D and F. Prolines interrupting the medial and distal coiled-coils are also shown (black). Myosin 5a tail sequences used in this work (referred here as M_L and M_S) are also depicted. Numbering of residues is according to UniProt entry Q9Y4I1-3 and is shown between M_L and M_S (orange numbers). Subscripts of the residues refer to the consensus position of the DYNLL binding motifs. Note that four residues at the N-terminus of M_L (gshm) are cloning artefacts.

Figure 2:NMR characterization of the extended DYNLL2 binding peptide (M_L) of myo5a. (A) Secondary chemical shiftvalues for Cα along the amino acid sequence for free (black) and DYNLL2-bound M_L (gray). (**B**) Variation of the R_2/R_1 ratio and (**C**)of the heteronuclear NOE enhancement along the $M_L$amino acid sequence. (**D**) $J(\omega_N)$ variation with $J(0)$ for$M_L$residues.

Figure 3: Chemical shift mapping of complex formation of the M_L peptide with DYNLL2. (**A**)Overlap of the 700 MHz ^{1}H-^{15}N HSQC spectra for the free (black) and DYNLL2 bound (red) M_L at

288 K. **(B)** Combined chemical shift changes (ppm) along the sequence, missing values are due to spectra overlap, or peaks broadened below baseline.

Figure 4:STD-HSQC measurements.(A) Overlap of the 500 MHz ^{1}H-^{15}N HSQC (black) and STD-HSQC (red) spectra at 290 K (protein irradiation at 5500 Hz). (B) Overlap of the 500 MHz ^{1}H-^{13}C CT-HSQC (black for CH, CH$_3$, blue for CH$_2$) and STD-(CT)-HSQC (green for CH, CH$_3$, red for CH$_2$) spectra at 290 K (protein irradiation at -400 Hz).

Figure 5:Structure of the DYNLL2-M$_S$ complex at 1.8 Å resolution. **(A**) Binding groove with M$_S$ peptide (DYNLL2: gray, myo5a: green) **(B)** β-sheet formation (main chain H-bonds: yellow, side chain H-bonds: black) **(C)**: Interactions of Q$_{-8}$-K$_{-3}$: K$_{-3}$, D$_{-4}$ and D$_{-5}$ (corresponding to exon B) are involved in multiple interactions with DYNLL2. Three residues (K$_{-6}$, P$_{-7}$ and Q$_{-8}$) turn back toward the binding groove and make multiple interactions with residues K$_{-3}$-D$_{-5}$. These interactions block the potential hydrogen-bond donor and acceptor groups of the peptide backbone around the -4th position. (cross-eye stereo) **(D)** Interactions of N$_{-2}$-D$_2$: Helix-capping by M$_0$; One water molecule (directly beneath M$_0$; shown here asred sphere) forming hydrogen bonds simultaneously with the peptide and DYNLL2 can be found in the majority of other known crystal structures (cross-eye stereo)

Figure 6:Time-evolution of the secondary structure of M$_L$peptide in free state (A) and bound to DYNLL2 (B and C). The partially β-sheet (DYNLL2-bound) backbone conformation of the M$_S$ crystal structure was used as a starting (t = 0 ns) geometry for all molecular dynamics calculations. The β-sheet quickly disappeared in free solution (**A**) and it is practically not visible in this plot. In the DYNLL2-bound forms the β-sheet remained stable for short (B: chain B) and large (C: chain D) sections of M$_L$, respectively.

Figure 7: Conformational transition of the M$_L$ peptide between the free and the DYNLL-bound form. A large structural transition of M$_L$ (green) could be observed upon complex formation to DYNLL2 (grey) by MD simulation. The transient α-helix of unbound M$_L$ disappeared and a β-strand was developed in both chain B (green) and chain D (red) copies of the peptide. The N-terminal end of the peptide was found to be in a bend/turn region with high curvature (see also Figure 6 for a detailed per-

residue analysis of the secondary structures of both chains). In the case of chain B, the N-terminal end of the peptide bent over the C-terminal region. Only one M_L chain is shown in the free form. The C-terminal ends of the peptides in the complex tend to interact via hydrophobic residues.

Figure 8: Structural comparison of the DYNLL2-M_s complex with other known DNYLL-peptide complexes. (A) Individual binding pockets of DYNLL (positions are numbered according to the arbitrary numbering of residues within bound motifs). All known crystal structures (3P8M, 2XQQ, 2P2T, 3ZKE, 1CMI, 4DS1, 4HT6, 3DVP, 3E2B) were superimposed. Binding motifs are represented by sticks. For simplicity, only the surface of the DYNLL2-M_S complex is shown. (B) Backbones of the bound peptide motifs within the groove of DYNLL show very similar conformation. Surface of the two subunits of DYNLL2 is blue and white. Note that uniquely among the binding motifs the backbone conformation of M_S (green) at position −5 folds back on itself. (C) Binding pocket 1 for M_S (T_1, green) and for nNOS (V_1, orange, 1CMI) is compared. The side chain of T_1 turns downwards and forms a hydrogen-bond with the backbone of DYNLL2. The side chain of V_1 in nNOS protrudes bellow the rings of Y77 and Y75, thereby forces the whole binding groove to open more widely (indicated by orange arrows). (D) Binding pocket 1 for M_S (T_1, green) and for EML3 (T_1, cyan) is compared (2XQQ). Conformation of both the peptide ligands and DYNLL2 are highly similar. (E) The Met (M_0) in M_S (green) and the conserved Gln (Q_0) in EML3 (cyan) adopt similar conformation in the binding pocket of position0. Gln is involved in capping of helix2 (at the backbone NH group of K36) and also forms a hydrogen bond with the side chain of E38 (white and cyan in the DYNLL2-M_s and DYNLL2-EML3 complex, respectively). The sulfur atom of Met can only form a weak interaction with the backbone NH group of K36. The hydrogen bond to the side chain of E38 is missing in this case. (F) Binding pocket −1 for M_S (T_{-1}, green) with a water mediated interaction and for NUP159 (I_1, magenta) is compared (4DS1). Ile is too bulky to allow the incorporation of a water.

Figure 9: Hypothetical model of the asymmetric folding of myosin 5a. Structure of the DYNLL2-M_L complex determined in this work is shown above the medial and distal coiled coils of the tail region

(DYNLL2: surface, green; M_L: cartoon, orange). Blue arrows represent the movement of the motor domains required for the formation of the off-state, where the heads directly interact with the GTDs.
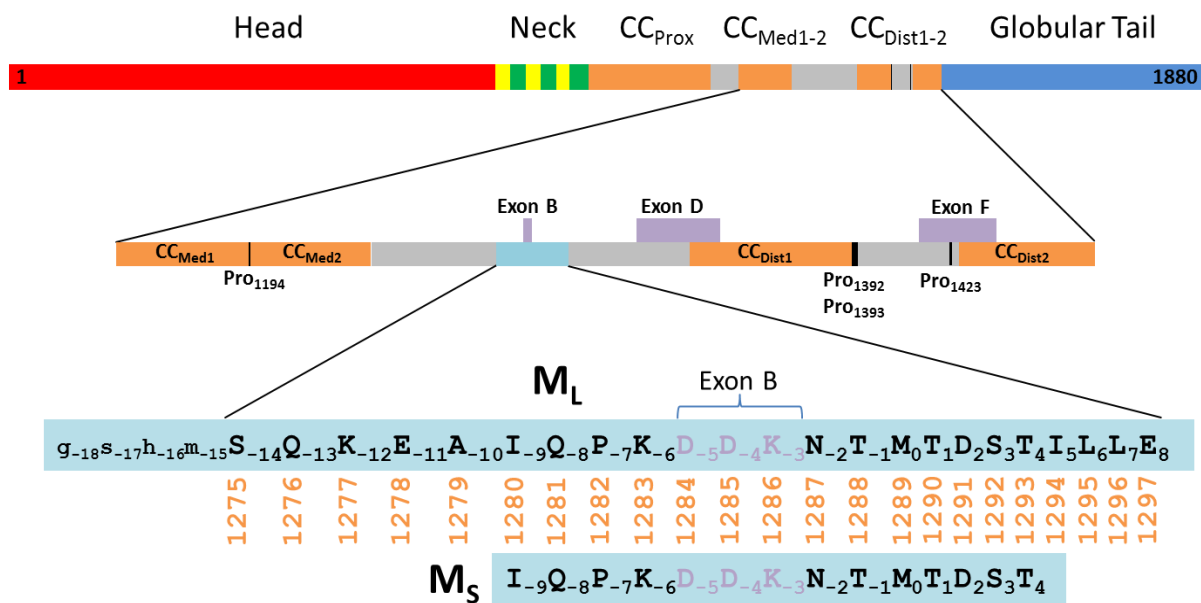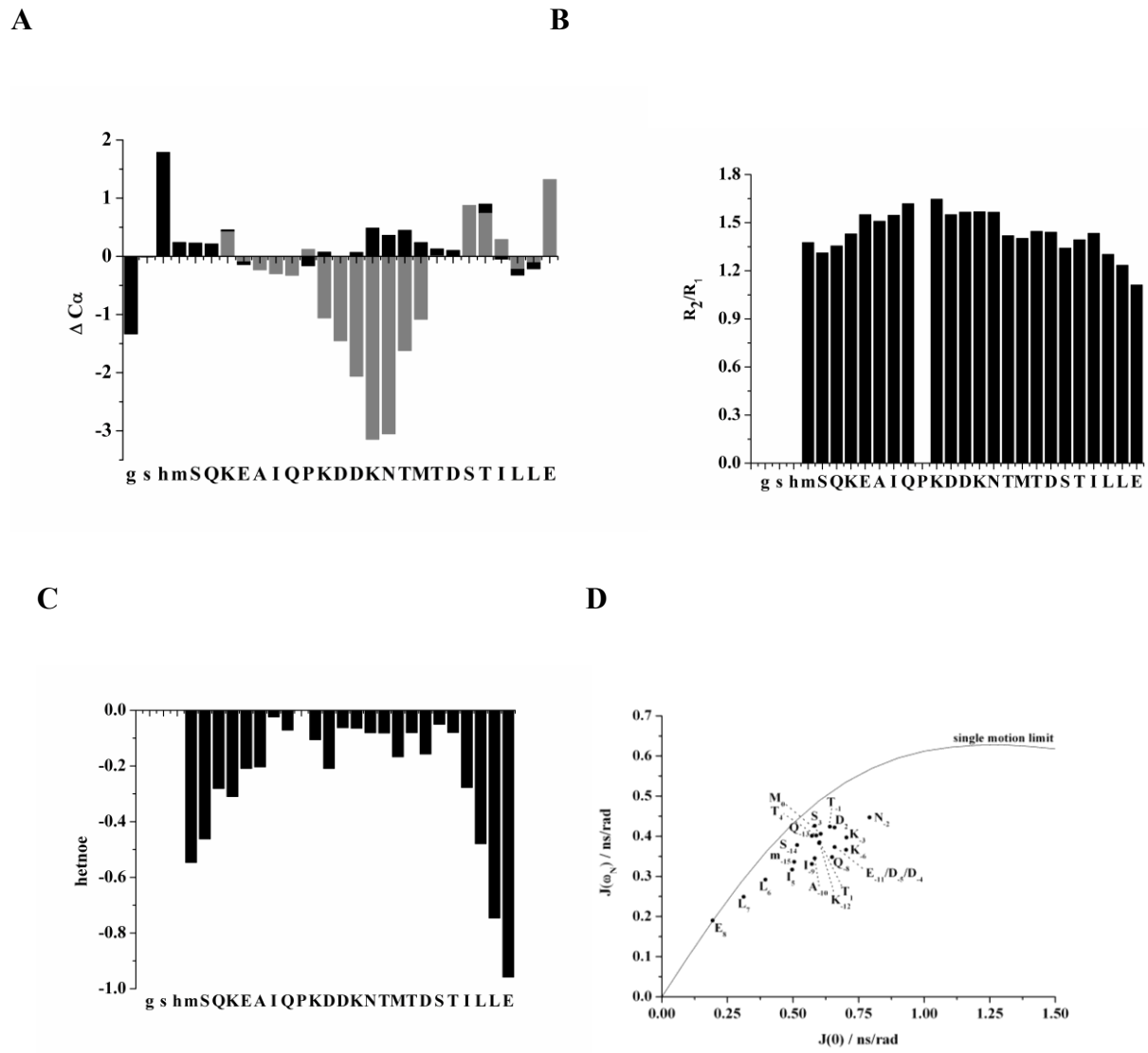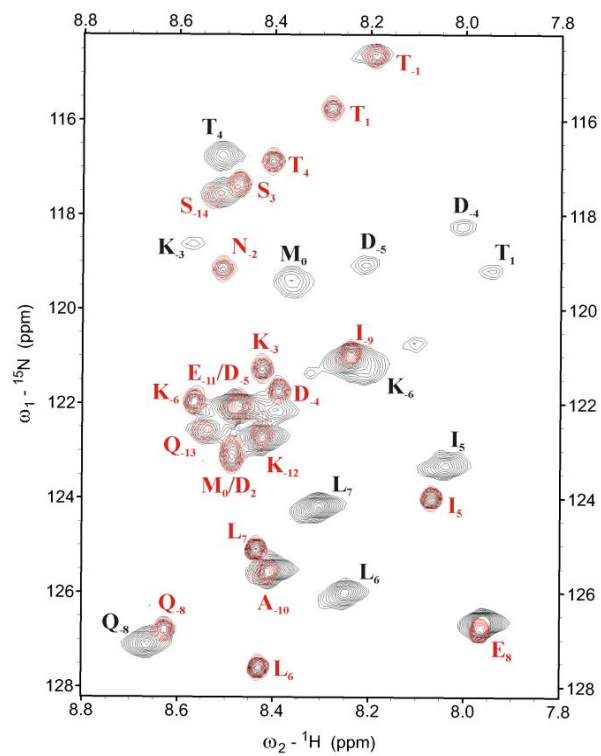
Figure 1.

Figure 2

A
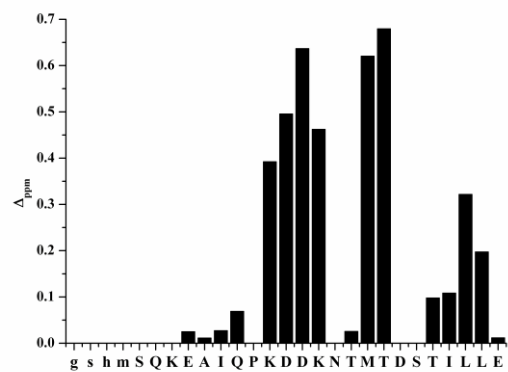

B

C

D

Figure 3

A

B

Figure 4

A

B

31
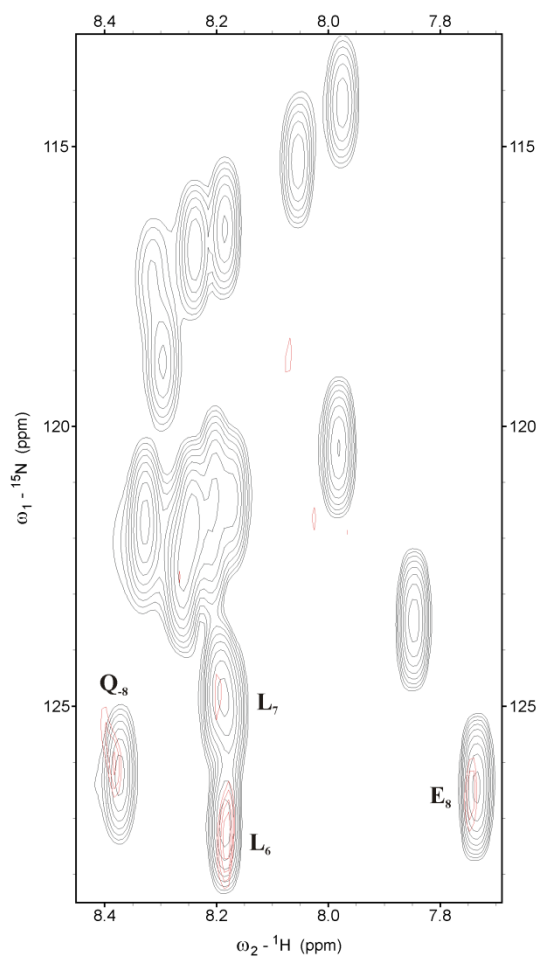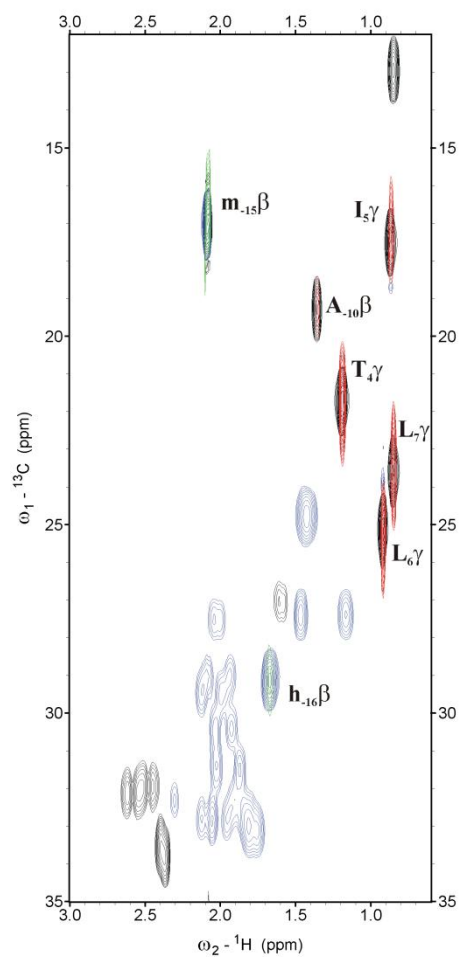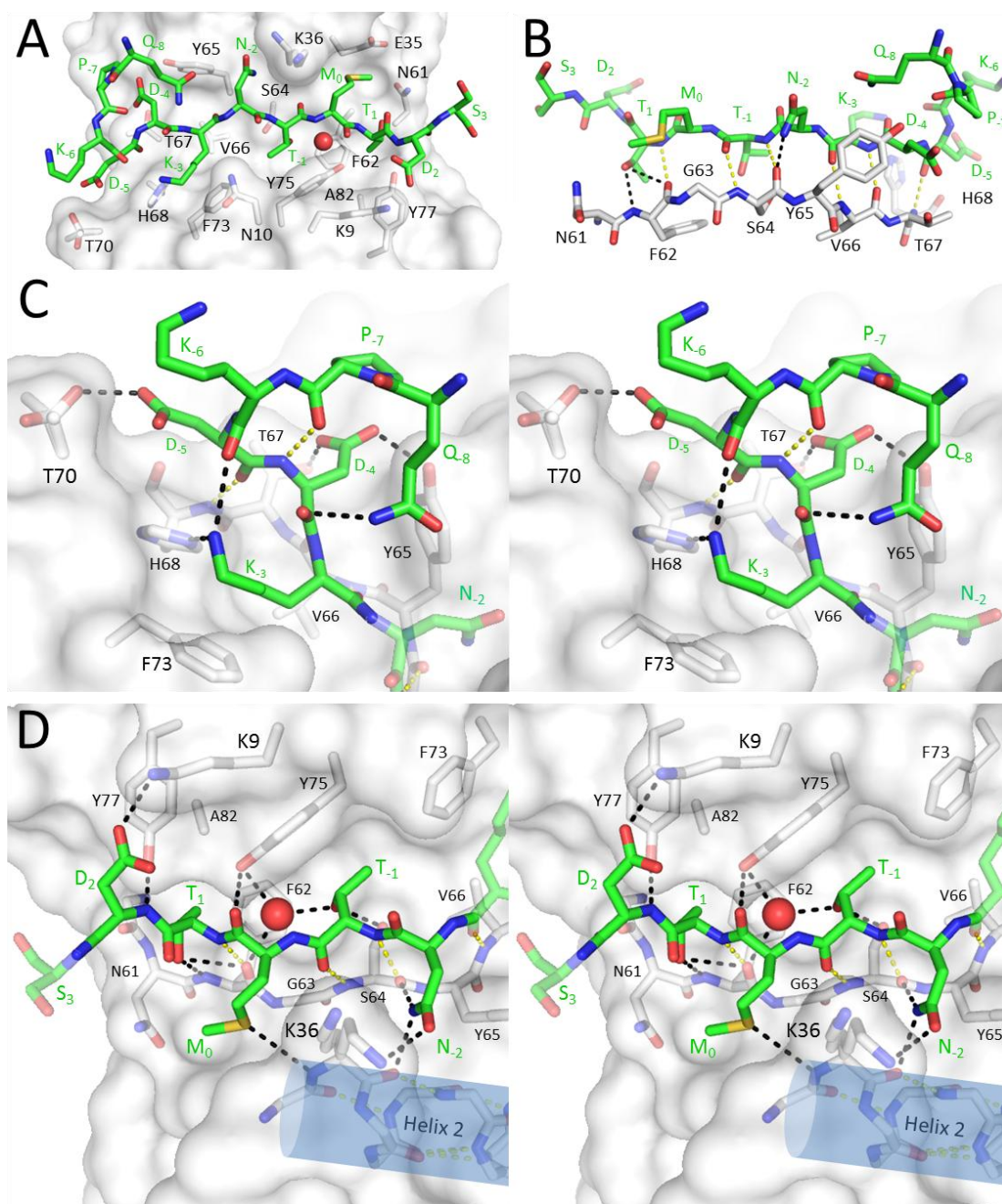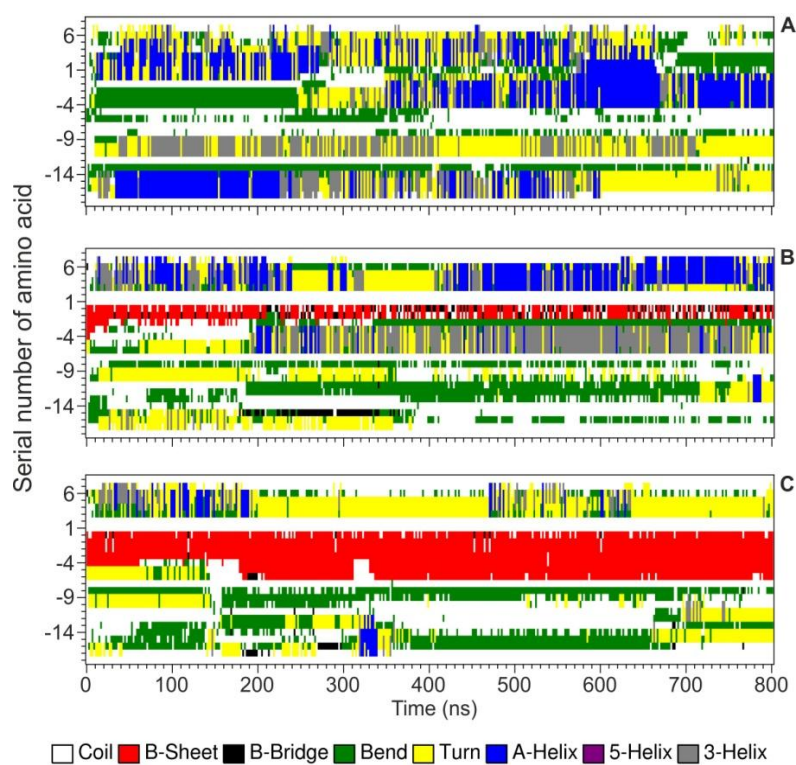
Figure 5

Figure 6.

Figure 7.

Figure 8.

Figure 9.

Head

Neck

Proximal
Coiled-coil

Light chains

Neck

Head

Distal
Coiled-coils
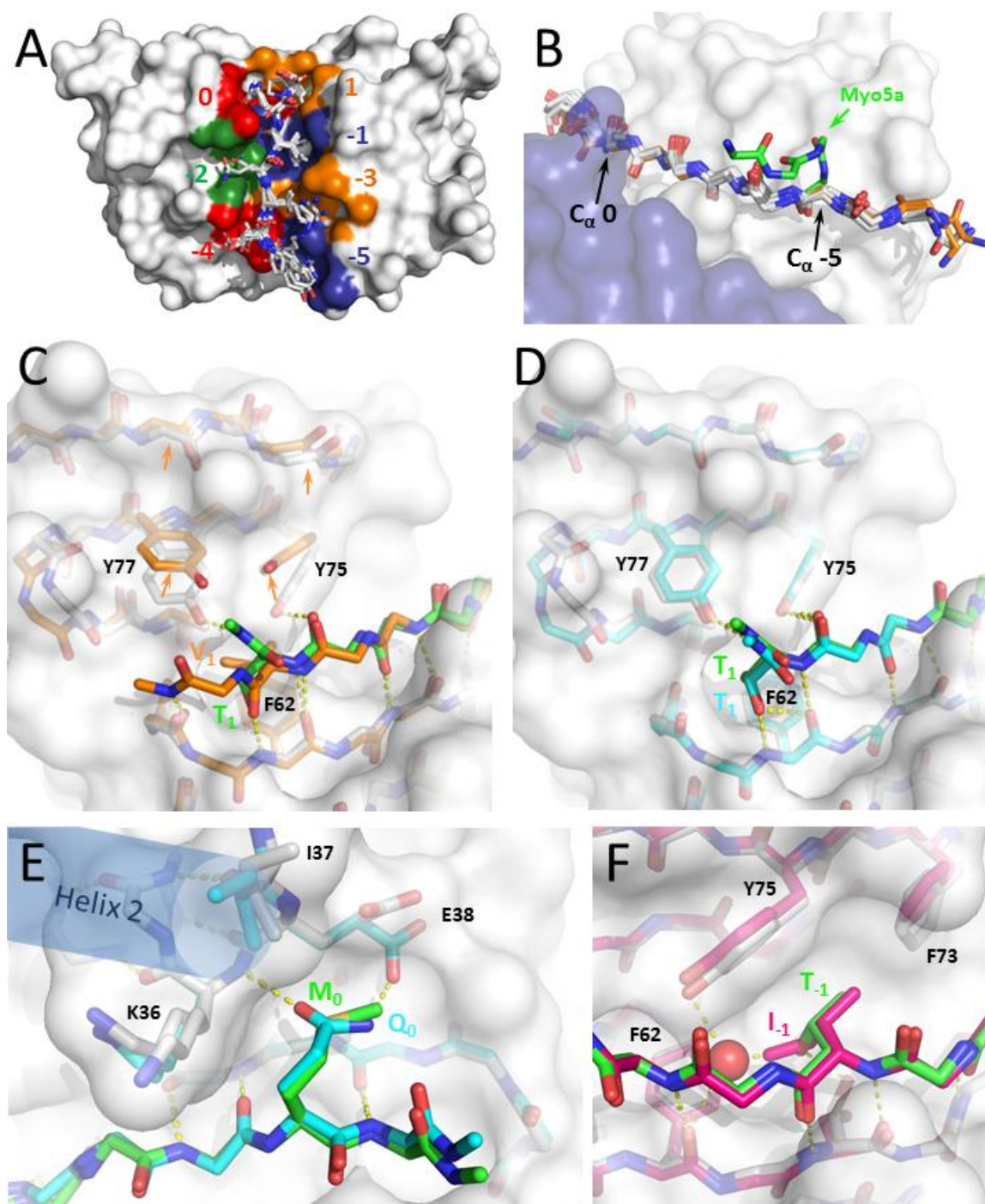
DYNLL2

Medial
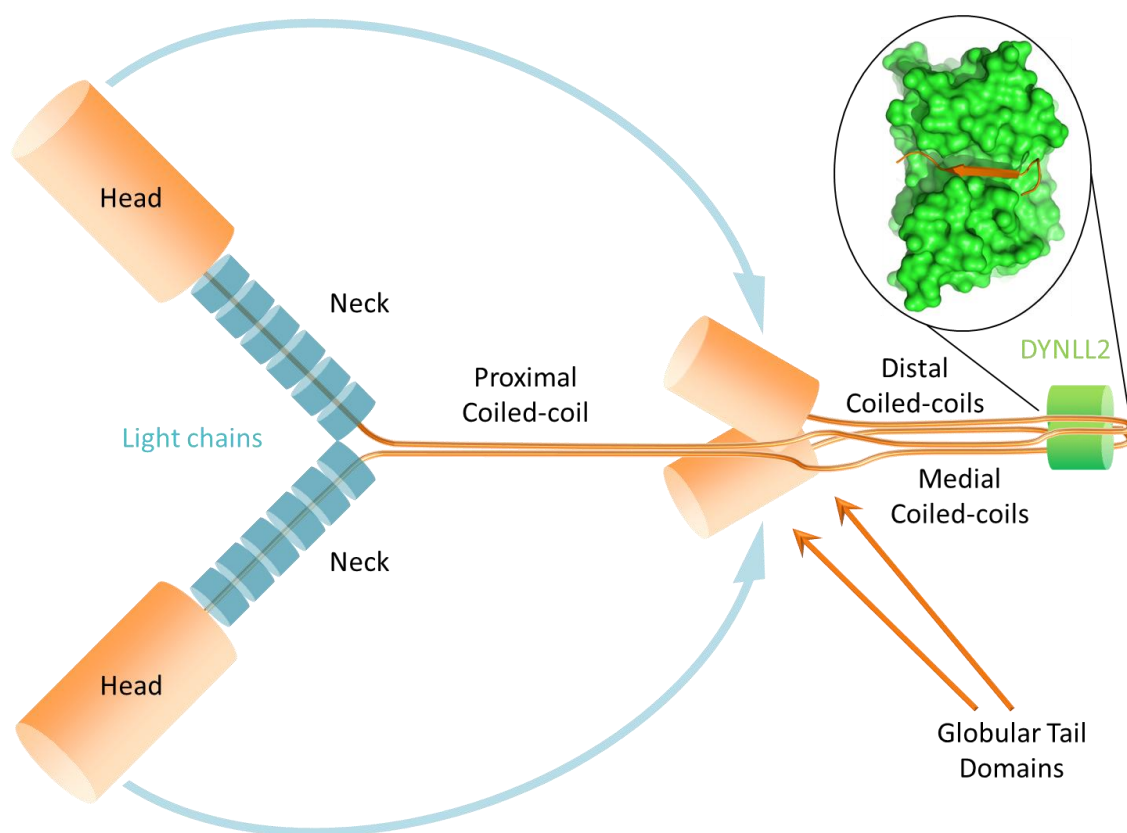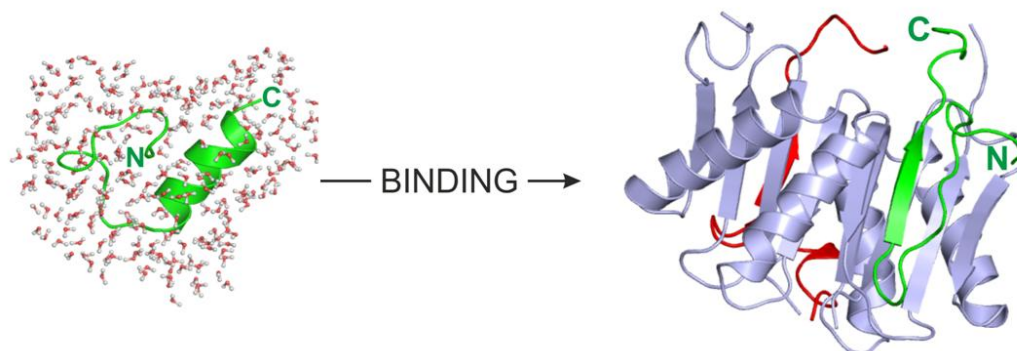Coiled-coils

Globular Tail
Domains
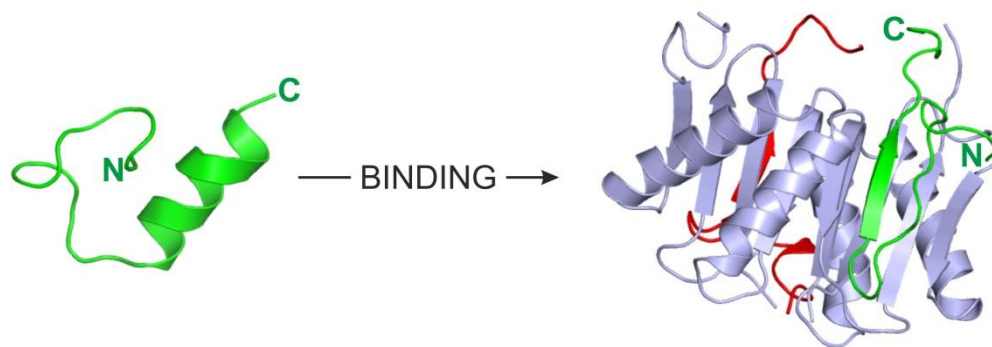
Table of Contents Graphic:

REFERENCES:

1. Rapali, P., Szenes, A., Radnai, L., Bakos, A., Pal, G., and Nyitray, L. (2011) DYNLL/LC8: a light chain subunit of the dynein motor complex and beyond, *Febs J278*, 2980-2996.

2. Barbar, E. (2008) Dynein light chain LC8 is a dimerization hub essential in diverse protein networks, *Biochemistry47*, 503-508.

3. Barbar, E., Kleinman, B., Imhoff, D., Li, M., Hays, T. S., and Hare, M. (2001) Dimerization and folding of LC8, a highly conserved light chain of cytoplasmic dynein, *Biochemistry40*, 1596-1605.

4. Benison, G., Karplus, P. A., and Barbar, E. (2008) The interplay of ligand binding and quaternary structure in the diverse interactions of dynein light chain LC8, *J Mol Biol384*, 954-966.

5. Liang, J., Jaffrey, S. R., Guo, W., Snyder, S. H., and Clardy, J. (1999) Structure of the PIN/LC8 dimer with a bound peptide, *Nat Struct Biol6*, 735-740.

6. Fan, J., Zhang, Q., Tochio, H., Li, M., and Zhang, M. (2001) Structural basis of diverse sequence-dependent target recognition by the 8 kDa dynein light chain, *J Mol Biol306*, 97-108.

7. Dunker, A. K., Silman, I., Uversky, V. N., and Sussman, J. L. (2008) Function and structure of inherently disordered proteins, *Curr Opin Struct Biol18*, 756-764.

8. Neduva, V., Linding, R., Su-Angrand, I., Stark, A., de Masi, F., Gibson, T. J., Lewis, J., Serrano, L., and Russell, R. B. (2005) Systematic discovery of new recognition peptides mediating protein interaction networks, *PLoS Biol3*, e405.

9. Rapali, P., Radnai, L., Suveges, D., Harmat, V., Tolgyesi, F., Wahlgren, W. Y., Katona, G., Nyitray, L., and Pal, G. (2011) Directed evolution reveals the binding motif preference of the LC8/DYNLL hub protein and predicts large numbers of novel binders in the human proteome, *PLoS One6*, e18818.

10. Hammer, J. A., 3rd, and Sellers, J. R. (2012) Walking to work: roles for class V myosins as cargo transporters, *Nat Rev Mol Cell Biol13*, 13-26.

11. Langford, G. M. (2002) Myosin-V, a versatile motor for short-range vesicle transport, *Traffic3*, 859-865.

12. Trybus, K. M. (2008) Myosin V from head to tail, *Cell Mol Life Sci65*, 1378-1389.

13. Lambert, J., Naeyaert, J. M., Callens, T., De Paepe, A., and Messiaen, L. (1998) Human myosin V gene produces different transcripts in a cell type-specific manner, *Biochem Biophys Res Commun252*, 329-333.

14. Roland, J. T., Lapierre, L. A., and Goldenring, J. R. (2009) Alternative splicing in class V myosins determines association with Rab10, *J Biol Chem284*, 1213-1223.

15. Wu, X., Wang, F., Rao, K., Sellers, J. R., and Hammer, J. A., 3rd. (2002) Rab27a is an essential component of melanosome receptor for myosin Va, *Mol Biol Cell13*, 1735-1749.

16. Hodi, Z., Nemeth, A. L., Radnai, L., Hetenyi, C., Schlett, K., Bodor, A., Perczel, A., and Nyitray, L. (2006) Alternatively spliced exon B of myosin Va is essential for binding the tail-associated light chain shared by dynein, *Biochemistry45*, 12582-12595.

17. Wagner, W., Fodor, E., Ginsburg, A., and Hammer, J. A., 3rd. (2006) The binding of DYNLL2 to myosin Va requires alternatively spliced exon B and stabilizes a portion of the myosin's coiled-coil domain, *Biochemistry45*, 11564-11577.

18. Radnai, L., Rapali, P., Hodi, Z., Suveges, D., Molnar, T., Kiss, B., Becsi, B., Erdodi, F., Buday, L., Kardos, J., Kovacs, M., and Nyitray, L. (2010) Affinity, avidity and kinetics of target sequence binding to LC8 dynein light chain isoforms, *J Biol Chem*.

19. King, S. M., and Patel-King, R. S. (1995) The M(r) = 8,000 and 11,000 outer arm dynein light chains from Chlamydomonas flagella have cytoplasmic homologues, *J Biol Chem270*, 11445-11452.

20. Espindola, F. S., Suter, D. M., Partata, L. B., Cao, T., Wolenski, J. S., Cheney, R. E., King, S. M., and Mooseker, M. S. (2000) The light chain composition of chicken brain myosin-Va: calmodulin, myosin-II essential light chains, and 8-kDa dynein light chain/PIN, *Cell Motil Cytoskeleton47*, 269-281.

21. Jaffrey, S. R., and Snyder, S. H. (1996) PIN: an associated protein inhibitor of neuronal nitric oxide synthase, *Science274*, 774-777.

22. Chaudhury, A., He, X. D., and Goyal, R. K. (2011) Myosin Va plays a key role in nitrergic neurotransmission by transporting nNOSalpha to enteric varicosity membrane, *Am J Physiol Gastrointest Liver Physiol301*, G498-507.

23. Vadlamudi, R. K., Bagheri-Yarmand, R., Yang, Z., Balasenthil, S., Nguyen, D., Sahin, A. A., den Hollander, P., and Kumar, R. (2004) Dynein light chain 1, a p21-activated kinase 1-interacting substrate, promotes cancerous phenotypes, *Cancer Cell5*, 575-585.

24. Lu, J., Sun, Q., Chen, X., Wang, H., Hu, Y., and Gu, J. (2005) Identification of dynein light chain 2 as an interaction partner of p21-activated kinase 1, *Biochem Biophys Res Commun331*, 153-158.

25. Kong, X., Gan, H., Hao, Y., Cheng, C., Jiang, J., Hong, Y., Yang, J., Zhu, H., Chi, Y., Yun, X., and Gu, J. (2009) CDK11p58 phosphorylation of PAK1 Ser174 promotes DLC2 binding and roles on cell cycle progression, *J Biochem146*, 417-427.

26. Puthalakath, H., Villunger, A., O'Reilly, L. A., Beaumont, J. G., Coultas, L., Cheney, R. E., Huang, D. C., and Strasser, A. (2001) Bmf: a proapoptotic BH3-only protein regulated by interaction with the myosin V actin motor complex, activated by anoikis, *Science293*, 1829-1832.

27. Puthalakath, H., Huang, D. C., O'Reilly, L. A., King, S. M., and Strasser, A. (1999) The proapoptotic activity of the Bcl-2 family member Bim is regulated by interaction with the dynein motor complex, *Mol Cell3*, 287-296.

28. Fejtova, A., Davydova, D., Bischof, F., Lazarevic, V., Altrock, W. D., Romorini, S., Schone, C., Zuschratter, W., Kreutz, M. R., Garner, C. C., Ziv, N. E., and Gundelfinger, E. D. (2009) Dynein light chain regulates axonal trafficking and synaptic levels of Bassoon, *J Cell Biol185*, 341-355.

29. Naisbitt, S., Valtschanoff, J., Allison, D. W., Sala, C., Kim, E., Craig, A. M., Weinberg, R. J., and Sheng, M. (2000) Interaction of the postsynaptic density-95/guanylate kinase domain-associated protein complex with a light chain of myosin-V and dynein, *J Neurosci20*, 4524-4534.

30. Su, Y., Qiao, W., Guo, T., Tan, J., Li, Z., Chen, Y., Li, X., Li, Y., Zhou, J., and Chen, Q. (2010) Microtubule-dependent retrograde transport of bovine immunodeficiency virus, *Cellular microbiology12*, 1098-1107.

31. Alonso, C., Miskin, J., Hernaez, B., Fernandez-Zapatero, P., Soto, L., Canto, C., Rodriguez-Crespo, I., Dixon, L., and Escribano, J. M. (2001) African swine fever virus protein p54 interacts with the microtubular motor complex through direct binding to light-chain dynein, *Journal of virology75*, 9819-9827.

32. Goddard, T. D., and Kneller, D. G. SPARKY 3, University of California, San Francisco, http://www.cgl.ucsf.edu/home/sparky/.

33. Keller, R. (2004) The Computer Aided Resonance Assignment Tutorial. Ed. Rochus LJ Keller, Editorial Cantina Verlag, Goldau, Swiss.

34. Lefevre, J. F., Dayie, K. T., Peng, J. W., and Wagner, G. (1996) Internal mobility in the partially folded DNA binding and dimerization domains of GAL4: NMR analysis of the N-H spectral density functions, *Biochemistry35*, 2674-2686.

35. Krizova, H., Zidek, L., Stone, M. J., Novotny, M. V., and Sklenar, V. (2004) Temperature-dependent spectral density analysis applied to monitoring backbone dynamics of major urinary protein-I complexed with the pheromone 2- sec-butyl-4,5-dihydrothiazole, *J Biomol NMR28*, 369-384.

36. Kabsch, W. (1993) Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants, *J Appl Crystallogr26*, 795-800.

37. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) Likelihood-enhanced fast translation functions, *Acta Crystallogr D Biol Crystallogr61*, 458-464.

38. Bailey, S. (1994) The Ccp4 Suite - Programs for Protein Crystallography, *Acta Crystallogr D50*, 760-763.

39. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr D Biol Crystallogr53*, 240-255.

40. Lamzin, V. S., and Wilson, K. S. (1993) Automated refinement of protein models, *Acta Crystallogr D Biol Crystallogr49*, 129-147.

41. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr60*, 2126-2132.

42. Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation, *Journal of Chemical Theory and Computation4*, 435-447.

43. Ponder, J. W., and Richards, F. M. (1987) An efficient newton-like method for molecular mechanics energy minimization of large molecules, *Journal of Computational Chemistry8*, 1016-1024.

44. Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations, *J Comput Chem24*, 1999-2012.

45. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *The Journal of Chemical Physics79*, 926-935.

46. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers22*, 2577-2637.

47. Schrodinger, LLC. (2010) The PyMOL Molecular Graphics System, Version 1.3r1.

48. Lupas, A., Van Dyke, M., and Stock, J. (1991) Predicting coiled coils from protein sequences, *Science252*, 1162-1164.

49. Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005) The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins, *J Mol Biol347*, 827-839.

50. Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content, *Bioinformatics21*, 3433-3434.

51. Meszaros, B., Simon, I., and Dosztanyi, Z. (2009) Prediction of protein binding regions in disordered proteins, *PLoS Comput Biol5*, e1000376.

52. Dosztanyi, Z., Meszaros, B., and Simon, I. (2009) ANCHOR: web server for predicting protein binding regions in disordered proteins, *Bioinformatics25*, 2745-2746.

53.   Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects, *J Biomol NMR5*, 67-81.

54.   Braun, D., Wider, G., and Wuethrich, K. (1994) Sequence-Corrected 15N "Random Coil" Chemical Shifts, *Journal of the American Chemical Society116*, 8466-8469.

55.   Mayer, M., and Meyer, B. (2001) Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor, *J Am Chem Soc123*, 6108-6117.

56.   Lightcap, C. M., Sun, S., Lear, J. D., Rodeck, U., Polenova, T., and Williams, J. C. (2008) Biochemical and structural characterization of the Pak1-LC8 interaction, *J Biol Chem283*, 27314-27324.

57.   Benison, G., Karplus, P. A., and Barbar, E. (2007) Structure and dynamics of LC8 complexes with KXTQT-motif peptides: swallow and dynein intermediate chain compete for a common site, *J Mol Biol371*, 457-468.

58.   Williams, J. C., Roulhac, P. L., Roy, A. G., Vallee, R. B., Fitzgerald, M. C., and Hendrickson, W. A. (2007) Structural and thermodynamic characterization of a cytoplasmic dynein light chain-intermediate chain complex, *Proc Natl Acad Sci U S A104*, 10028-10033.

59.   Hall, J., Karplus, P. A., and Barbar, E. (2009) Multivalency in the assembly of intrinsically disordered Dynein intermediate chain, *J Biol Chem284*, 33115-33121.

60.   Kiss, R., Kovacs, D., Tompa, P., and Perczel, A. (2008) Local structural preferences of calpastatin, the intrinsically unstructured protein inhibitor of calpain, *Biochemistry47*, 6936-6945.

61.   Szalaine Agoston, B., Kovacs, D., Tompa, P., and Perczel, A. (2011) Full backbone assignment and dynamics of the intrinsically disordered dehydrin ERD14, *Biomol NMR Assign5*, 189-193.

62.   Csizmok, V., Felli, I. C., Tompa, P., Banci, L., and Bertini, I. (2008) Structural and dynamic characterization of intrinsically disordered human securin by NMR spectroscopy, *J Am Chem Soc130*, 16873-16879.

63.   Oldfield, C. J., Meng, J., Yang, J. Y., Yang, M. Q., Uversky, V. N., and Dunker, A. K. (2008) Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners, *BMC Genomics9 Suppl 1*, S1.

64.   Benison, G., Nyarko, A., and Barbar, E. (2006) Heteronuclear NMR identifies a nascent helix in intrinsically disordered dynein intermediate chain: implications for folding and dimerization, *J Mol Biol362*, 1082-1093.

65.   Benison, G., and Barbar, E. (2009) NMR analysis of dynein light chain dimerization and interactions with diverse ligands, *Methods Enzymol455*, 237-258.

66.   Stein, A., and Aloy, P. (2008) Contextual specificity in peptide-mediated protein interactions, *PLoS One3*, e2524.

67.   Sellers, J. R., Thirumurugan, K., Sakamoto, T., Hammer, J. A., 3rd, and Knight, P. J. (2008) Calcium and cargoes as regulators of myosin 5a activity, *Biochem Biophys Res Commun369*, 176-181.

68.   Li, X. D., Jung, H. S., Mabuchi, K., Craig, R., and Ikebe, M. (2006) The globular tail domain of myosin Va functions as an inhibitor of the myosin Va motor, *J Biol Chem281*, 21789-21798.

69.   Thirumurugan, K., Sakamoto, T., Hammer, J. A., 3rd, Sellers, J. R., and Knight, P. J. (2006) The cargo-binding domain regulates structure and activity of myosin 5, *Nature442*, 212-215.

70.   Krementsov, D. N., Krementsova, E. B., and Trybus, K. M. (2004) Myosin V: regulation by calcium, calmodulin, and the tail domain, *J Cell Biol164*, 877-886.

71.   Nyarko, A., Hare, M., Hays, T. S., and Barbar, E. (2004) The intermediate chain of cytoplasmic dynein is partially disordered and gains structure upon binding to light-chain LC8, *Biochemistry43*, 15595-15603.

72.   Wang, L., Hare, M., Hays, T. S., and Barbar, E. (2004) Dynein light chain LC8 promotes assembly of the coiled-coil domain of swallow protein, *Biochemistry43*, 4611-4620.

73.   Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V., and Sullivan, M. (2012) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse, *Nucleic acids research40*, D261-270.

74.   Komander, D. (2009) The emerging complexity of protein ubiquitination, *Biochemical Society transactions37*, 937-953.

75.   Fuxreiter, M., Tompa, P., and Simon, I. (2007) Local structural disorder imparts plasticity on linear motifs, *Bioinformatics23*, 950-956.