Directed Evolution Reveals the Binding Motif Preference of the LC8/DYNLL Hub Protein and Predicts Large Numbers of Novel Binders in the Human Proteome

Péter Rapali¹, László Radnai¹, Dániel Süveges¹, Veronika Harmat²,³, Ferenc Tölgyesi⁴, Weixiao Y. Wahlgren⁵, Gergely Katona⁵, László Nyitray¹*, Gábor Pál¹*

¹Department of Biochemistry, Eötvös Loránd University, Budapest, Hungary, ²Institute of Chemistry, Eötvös Loránd University, Budapest, Hungary, ³Protein Modeling Research Group, Hungarian Academy of Sciences, Eötvös Loránd University, Budapest, Hungary, ⁴Institute of Biophysics and Radiation Biology, Semmelweis University, Budapest, Hungary, ⁵Department of Chemistry, University of Gothenburg, Gothenburg, Sweden

Abstract

LC8 dynein light chain (DYNLL) is a eukaryotic hub protein that is thought to function as a dimerization engine. Its interacting partners are involved in a wide range of cellular functions. In its dozens of hitherto identified binding partners DYNLL binds to a linear peptide segment. The known segments define a loosely characterized binding motif: [D/S]₄K₂X₂[T/V/I]₃Qₓ[T/V/I]₁[D/E]₂. The motifs are localized in disordered segments of the DYNLL-binding proteins and are often flanked by coiled coil or other potential dimerization domains. Based on a directed evolution approach, here we provide the first quantitative characterization of the binding preference of the DYNLL binding site. We displayed on M13 phage a naive peptide library with seven fully randomized positions around a fixed, naturally conserved glutamine. The peptides were presented in a bivalent manner fused to a leucine zipper mimicking the natural dimer to dimer binding stoichiometry of DYNLL-partner complexes. The phage-selected consensus sequence V₁₅S₁₃R₁₀G₁₀T₁₀Q₁₀T₁₂E₁₂ resembles the natural one, but is extended by an additional N-terminal valine, which increases the affinity of the monomeric peptide twentyfold. Leu-zipper dimerization increases the affinity into the subnanomolar range. By comparing crystal structures of an SRGTQTE-DYNLL and a dimeric VSRGTQTE-DYNLL complex we find that the affinity enhancing valine is accommodated in a binding pocket on DYNLL. Based on the in vitro evolved sequence pattern we predict a large number of novel DYNLL binding partners in the human proteome. Among these EML3, a microtubule-binding protein involved in mitosis contains an exact match of the phage-evolved consensus and binds to DYNLL with nanomolar affinity. These results significantly widen the scope of the human interactome around DYNLL and will certainly shed more light on the biological functions and organizing role of DYNLL in the human and other eukaryotic interactomes.

Citation: Rapali P, Radnai L, Süveges D, Harmat V, Tölgyesi F, et al. (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 ONE 6(4): e18818. doi:10.1371/journal.pone.0018818

Editor: Cameron Neylon, Science and Technology Facilities Council, United Kingdom

Received December 17, 2010; Accepted March 10, 2011; Published April 18, 2011

Copyright: © 2011 Rapali et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The work was supported by the Hungarian Scientific Research Fund (OTKA) NK49812; K81784; NK81950 (LN); and K68408 (GP) (www.otka.hu); as well as by the European Union and the European Social Fund (TÁMOP) 4.2.1.-B/09/KMR-2010-0003 (www.nfu.hu). G.P. is supported by the János Bolyai Research Fellowship (www.mta.hu). W.Y.W. and G.K. are supported by the Swedish Research Council (2008-4248, 2009-350) (www.vr.se) and the Harald and Greta Jeansson Foundation (www.jeanssonstiftelse.se). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: nyitray@elte.hu (LN); palgabor@elte.hu (GP)

Introduction

The LC8 dynein light chain (DYNLL in mammals) is a highly conserved 10 kDa protein that was originally described as a potential cargo binding adapter of both dynein and myosin 5a motor proteins and as an inhibitor of the neuronal NO-synthase enzyme (reviewed in [1]). Subsequent studies demonstrated that it binds to and regulates dozens of other proteins unrelated to the motor proteins and as an inhibitor of the neuronal NO-synthase enzyme (reviewed in [1]). By gene knockout and knockdown experiments DYNLL has been shown to be an essential protein in Drosophila and Caenorhabditis elegans [13,16]. Vertebrate genomes contain two DYNLL paralogs DYNLL1 and DYNLL2 which share 93% sequence identity at the protein level. It is still controversial whether the two isoforms have disparate cellular functions. In some studies DYNLL1 and DYNLL2 were found to bind in vivo specifically to the dynein and myosin 5a complexes, respectively [4,17,18], but other studies do not support this view [19]. The two isoforms have identical in vitro binding characteristics to their interacting proteins studied so far [18,20].

Apo- and ligand-bound structures, determined both by X-ray diffraction and NMR spectroscopic studies, revealed that DYNLL forms a homodimer possessing two identical ligand binding sites. In DYNLL a swapped β-sheet dimer interface is formed. In the complexes two 7-residue peptides bind to parallel grooves at the
edges of the dimer interface expanding the two central β-sheets [18,21,22,23,24,25,26,27]. Since the DYNLL-binding linear peptide motifs are generally located in intrinsically disordered regions of the partner proteins and bind to DYNLL as a short β-strand they could be considered as β-MoRE (molecular recognition element) [20,29]. DYNLL-binding motifs are often flanked by coiled coil or other dimerization domains. DYNLL could function upon complex formation as a “molecular glue” by promoting and stabilizing the dimeric structure and in this way regulating the function of its binding partners [30,31,32].

The diverse DYNLL binding motifs were traditionally divided into two loose consensus sequence classes, [K/R]X₅T₃Q₄T₅, and G₅AₔQ₆V₇D₂ [33,34]. Moreover, a few known partners have non-canonical motifs lacking the conserved Gln such as myosin 5a (D. sD₃R₄G₅K₆A₇T₈M₉T₁₀Dₑ₁) [30,32], p21-activated kinase (Pak1) (R.D. iV₃A₄T₅,S₆P₇I₈) [35] and GRINL1 (E₅T₆R₇E₈L₉G₁₀V₁₁G₂) [36]. In the crystal structure of canonical complexes the side chain of the central glutamine (position 0) caps the N-terminal end of the second α-helix, while the side chains of residues at positions +1, −1 and −3 interact with the hydrophobic interior of the binding groove and some of these also participate in H-bonds formed by structural water molecules [22,23,24,26]. The importance in binding of the conserved glutamine and of the two flanking residues was demonstrated by mutational studies [37]. In the non-canonical Pak1 sequence, a key H-bond network was identified that compensates for the lack of the conserved glutamine [35]. The functional significance of the binding motif classes has not yet been demonstrated yet. Previous studies suggest a hierarchy or continuum of binding affinities (∼1-50 μM) depending on the actual sequence of the motifs [20,33]. On the other hand, different motif classes could achieve roughly the same binding affinities through different thermodynamic mechanisms [20]. It is important to note that many of the binding partners were shown to exist as homodimers therefore DYNLL most likely forms dimer-dimer complexes with its partners [24,30,35]. In fact, binding of bivalent ligands shows a considerable avidity effect; affinity of a dimeric molecule therefore DYNLL1 was hardly detectable (data not shown). DYNLL isoforms are homodimers having two identical peptide binding sites [22,23,24,26]. In order to mimic the natural dimer binding to dimer mechanism [20,24,30,40], we inserted a Leu-zipper motif from GCN4 between the Flag-tagged Bmf peptide and the p3 coat protein. Phage-ELISA using immobilized anti Flag-tag antibody demonstrated successful display in both cases. However, binding to immobilized GST fused DYNLL1 was hardly detectable (data not shown). DYNLL1 isoforms are homodimers having two identical peptide binding sites [22,23,24,26]. In order to mimic the natural dimer binding to dimer mechanism [20,24,30,40], we inserted a Leu-zipper motif from GCN4 between the Flag-tagged Bmf peptide and the p3 coat protein and displayed the peptide in a bivalent format (see Figure 1). A similar dimerizing approach was published by Fuh and colleagues [41,42]. In that study the authors applied competition phage-binding assays and demonstrated the bivalent display through its apparent affinity boosting avidity effect that, as expected, occurred only on surface immobilized targets as opposed to solvent-phase targets. In our case, due to the same effect, Leu-zipper based dimerization resulted in detectable binding to GST fused DYNLL1. Based on simple modeling studies we produced three variants with 4-mer (GGSG), 6-mer (GGSGGS) and 8-mer (GGSGGGSG) linker, respectively, in order to optimize the geometry of the bivalent binding. In binding assays the 4-mer linker was the best one. Since both too low or too high transcription levels decrease display efficiency [41,43], we tried to optimize the expression level of the originally used pTac promoter by inducing it at various IPTG concentrations. As low as 10 μM IPTG concentration had deleterious effect on display (Figure S1) suggesting that even the leakage of the strong pTac promoter might afford higher than optimal transcription level. Therefore we replaced the pTac promoter with the weaker PhoA promoter [42], which increased the display level about 30-fold (Figure S1) allowing for efficient peptide-library display.

The library design fixed the most conserved glutamine at position 0, while in all other positions from −5 to 2 allowed the occurrence of all 20 amino acid residues: X₅X₆X₇X₈X₉X₁₀Q₁₁X₁₂ (Figure 1). Note that we extended the original motif with one extra position (position −5) at the N-terminus. The library contained 2.1×10¹⁰ individual clones. Total randomization of 7 amino acid positions by degenerate (NNK) codons results in a theoretical diversity of 3²×3.4×10¹⁰ variants. Our library size approaches the theoretical one suggesting that most variants were indeed present in the starting repertoire. In the third panning cycles the enrichment was 100 fold. Individual clones from this selection cycle were tested in phage-ELISA experiments for DYNLL1-binding and the DNA of 36 positive clones was sequenced yielding 25 unique sequences (Table S2). From the anti Flag-tag selection 32 ELISA positive clones were sequenced and 30 were unique (Table S3). This set was used for subtracting the effects of display bias from the DYNLL1-selected set.

### Results

**Directed evolution of DYNLL-binding motif from bivalently displayed naïve peptide library**

We combined all 41 hitherto identified canonical DYNLL-binding motifs representing both previously established classes from 33 eukaryotic interacting proteins (Table S1) to define the following 7-mer linear motif: [(D/S)₁,R₃X₅[T/V/I]₁₀,G₁₅[T/V]₁₀]₁₀/D/E₁₁. Published affinities of the natural binding peptides are in the micromolar range (0.75 μM< Kd<50 μM) [20,30,39].

We aimed to elucidate the thermodynamically driven characteristic amino acid preferences of individual motif positions. To do so, we applied directed evolution by selecting DYNLL1 binders from a phage displayed peptide library. Before building the library we determined the optimal display format by displaying the strongest known binding peptide (Kd=0.75 μM) from Bmf [20] on M13 phage. The Bmf peptide (E.D₃K₅A₇T₈Q₁₀T₁₂) was fused to an N-terminal Flag-tag and displayed monovalently on the p3 and also multivalently on the p8 coat protein. Phage-ELISA using immobilized anti Flag-tag antibody demonstrated successful display in both cases. However, binding to immobilized GST fused DYNLL1 was hardly detectable (data not shown). DYNLL1 isoforms are homodimers having two identical peptide binding sites [22,23,24,26]. In order to mimic the natural dimer binding to dimer mechanism [20,24,30,40], we inserted a Leu-zipper motif from GCN4 between the Flag-tagged Bmf peptide and the p3 coat protein and displayed the peptide in a bivalent format (see Figure 1). A similar dimerizing approach was published by Fuh and colleagues [41,42]. In that study the authors applied competition phage-binding assays and demonstrated the bivalent display through its apparent affinity boosting avidity effect that, as expected, occurred only on surface immobilized targets as opposed to solvent-phase targets. In our case, due to the same effect, Leu-zipper based dimerization resulted in detectable binding to GST fused DYNLL1. Based on simple modeling studies we produced three variants with 4-mer (GGSG), 6-mer (GGSGGS) and 8-mer (GGSGGGSG) linker, respectively, in order to optimize the geometry of the bivalent binding. In binding assays the 4-mer linker was the best one. Since both too low or too high transcription levels decrease display efficiency [41,43], we tried to optimize the expression level of the originally used pTac promoter by inducing it at various IPTG concentrations. As low as 10 μM IPTG concentration had deleterious effect on display (Figure S1) suggesting that even the leakage of the strong pTac promoter might afford higher than optimal transcription level. Therefore we replaced the pTac promoter with the weaker PhoA promoter [42], which increased the display level about 30-fold (Figure S1) allowing for efficient peptide-library display.

The library design fixed the most conserved glutamine at position 0, while in all other positions from −5 to 2 allowed the occurrence of all 20 amino acid residues: X₅X₆X₇X₈X₉X₁₀Q₁₁X₁₂ (Figure 1). Note that we extended the original motif with one extra position (position −5) at the N-terminus. The library contained 2.1×10¹⁰ individual clones. Total randomization of 7 amino acid positions by degenerate (NNK) codons results in a theoretical diversity of 3²×3.4×10¹⁰ variants. Our library size approaches the theoretical one suggesting that most variants were indeed present in the starting repertoire. In the third panning cycles the enrichment was 100 fold. Individual clones from this selection cycle were tested in phage-ELISA experiments for DYNLL1-binding and the DNA of 36 positive clones was sequenced yielding 25 unique sequences (Table S2). From the anti Flag-tag selection 32 ELISA positive clones were sequenced and 30 were unique (Table S3). This set was used for subtracting the effects of display bias from the DYNLL1-selected set.
The logo of the *in vitro* evolved pattern is compared to the logo of 41 known naturally evolved DYNNL1 binding peptides (listed in Table S2). Generally, the two logos are similar, but there are a few characteristic differences. In the −1 and +1 positions both logos show strong conservation and marked preference for threonine. At position +2 there is a lower level conservation and slight preference for a negative charge (Glu, Asp) in both sets. There are similarities at the moderately conserved positions −2 and −3 as well. At position −2 mostly small residues are selected (Gly, Ala), while at position −3 there is a preference for positively charged residues (Arg, Lys). The characteristic differences appear at positions −4 and −5. In the natural binding peptides at position −4 a negatively charged Asp is the most frequent amino acid, which is absent from the *in vitro* selected pool. The most striking difference emerged at position −5, where the known natural binding peptides show lack of preference, while the *in vitro* evolved set presents significant non-randomness. Here mostly apolar residues (Val, Met, Ile, Leu) or polar residues with long aliphatic side chain portion were selected (Arg, Lys), with the most preferred residue being Val.

**Binding properties of the *in vitro* evolved consensus peptide – the significance of Val**

Binding parameters of the *in vitro* evolved Ac-VSRGTQTE and naturally evolved Ac-DKSTQTD consensus peptides were determined by isothermal titration calorimetry (ITC). Synthetic peptides were used and N-terminal acetylation was applied to better mimic a non-terminal peptide motif by avoiding the positively charged N-terminus. In order to elucidate the particular contribution of Val, to the binding affinity, several control peptides were produced. The *in vitro* evolved Val, was replaced with a glycine (Gly, ) to assess side chain function. In a further
truncated variant the entire amino acid residue was replaced with an acetate group (Ac-5) to assess the function of the amide group at position -5. The same variations of Val-5, Gly-5 or Ac-5 were also introduced in the framework of the naturally evolved consensus peptide. DYNNL1 binding of all six variants was analyzed by ITC. The results are illustrated in Figure 3 and the binding parameters are summarized in Table 1. In all measurements the stoichiometry of binding was 1:1 (2 peptides binding to one homodimer DYNNL1). The binding was enthalpy-driven and was accompanied with unfavorable entropy change. Ac-5 versions of both the natural and the in vitro evolved consensus peptide had binding affinities in the micromolar range. Replacing the acetyl group with Gly-5 led to a tenfold affinity increase with both peptides. This was due to a favorable change in binding enthalpy contribution suggesting that the main chain at position -5 participates in stabilizing molecular interactions. Replacing Gly-5 with a Val

Figure 3. Thermodynamic binding properties of natural and in vitro evolved motifs determined by isothermal titration calorimetry. Data for natural consensus (Ac-DKSTQTD) and its two control peptides and data for phage selected consensus (Ac-VSRGTQTE) and its two control peptides are shown in A and B, respectively.

doi:10.1371/journal.pone.0018818.g003
Peptide dimerization increases binding affinity

Homodimer structure of DYNLL1 allows for simultaneous binding of two peptides resulting in an avidity effect provided that the peptides are presented in dimer format [20,24]. This effect was utilized in the *in vitro* evolution process by displaying the peptide library in a Leu-zipper fused bivalent format. We tested how the same bivalency affects binding properties of the peptides. Dimerization of the Gly\_5 containing peptides enhanced their affinity compared to the monomeric versions (see in Table 1). In case of the naturally evolved consensus the improvement was over 500-fold. For the *in vitro* evolved consensus it was over 500-fold.

Peptide dimerization increases binding affinity

Table 1. Thermodynamic analysis of DYNLL-binding interactions.

| Partner                  | $K_d$ (µM) | $\Delta H$ (kJ mol$^{-1}$) | $-\Delta S$ (kJ mol$^{-1}$) |
|--------------------------|------------|---------------------------|-----------------------------|
| Ac-DKSTQTD               | 5.6        | -48.8                     | 18.9                        |
| Ac-GDKSTQTD              | 0.56       | -70.3                     | 34.4                        |
| Ac-VDKSTQTD              | 0.07       | -65.4                     | 24.8                        |
| dimeric-GSVDKSTQTD\*     | < 0.007    | <76.2                     | <29.4                       |
| dimeric-GSVDKSTQTD\*     | < 0.009    | <68                       | <42.0                       |
| Ac-SRGQTQTE              | 15.2       | -48.6                     | 21.0                        |
| Ac-GSRGTQTE              | 1.64       | -69.7                     | 36.5                        |
| Ac-VSRGTQTE              | 0.08       | -75.6                     | 35.0                        |
| dimeric-GSVSRGTQTE\*     | < 0.003    | -69.6                     | 20.6                        |
| dimeric-GSVSRGTQTE\*     | < 0.007    | -64.9                     | 17.6                        |
| EML3 (8-94)              | 0.05       | -32.4                     | 9.465                       |

*Data reach the dynamic range limit of the ITC method.

doi:10.1371/journal.pone.0018818.t001

Structural significance of position 5

For crystallographic studies we used DYNLL2 because its binding site and binding properties are identical to those of DYNLL1 [20], but it lacks a free surface exposed cysteine and hence is chemically more stable during crystallization. In order to understand the positive contribution of Val\_5 to DYNLL binding we crystallized and analyzed the DYNLL2 complex formed with the monomeric Ac-SRGQTQTE peptide and with the Leu zipper dimerized GS-VRGTQTE peptide (where the first two residues originate from the recombinant construct) (Table S4). We also made multiple attempts to grow crystals of acceptable quality from the Ac-VSRGTQTE peptide-DYNLL2 complex. Initially, this complex was prepared similarly to the Ac-SRGQTQTE peptide-DYNLL2 complex (see Experimental Procedures), but it showed markedly different behavior. After 5 minutes thousands of tiny crystal groups composed of dozens of small needles appeared. Several thousand of conditions were tested, but even the best crystals diffracted only around 10 Å.

Crystal structure of the Ac-SRGQTQTE-DYNLL2 complex was solved and refined at 1.31 Å resolution (Figure 4A). The asymmetric unit contains four chains forming two dimers. In the resulting electron density map the bound Ac-SRGQTQTE peptide could be clearly identified. All residues were in the most favored Ramachandran region except for Asn51, which has a positive $\phi$ angle, as in all other DYNLL structures [18,21,22,23,24,25,26,27]. Similar to the human DYNLL/Swa peptide complex (PDB entry 3e2b), each dimer contains two five-stranded $\beta$ sheets, with four strands from one subunit and a fifth strand crossed over from the other subunit. One side of each sheet is flanked by two helices while the other side forms the dimer interface. Both Ac-SRGQTQTE peptides bind identically to their symmetry related binding sites forming a sixth $\beta$ strand in the grooves created at the dimer interface (Figure 4F). The side chains of residues Thr\_1, Thr\_3 and Arg\_5 of the peptide project toward the interior of DYNLL2, into a deep hydrophobic groove lined by aromatic DYNLL residues Phe73, Tyr75, Tyr77, Phe62 and Tyr56 (Figure 4D).

Crystal structure of the complex formed between DYNLL2 and Leu zipper dimerized GS-VRGTQTE peptide complex was solved and refined to 2.9 Å resolution (Figure 4B). Overall conformation of the DYNLL2 dimer and the binding region highly resembles to that of the Ac-SRGQTQTE complex structure. Individual DYNLL2 dimers can be fitted with an RMSD in the range of 0.38–0.66 Å for the DYNLL2 backbone atoms and with an RMSD in the range of 0.42–0.61 Å for the peptide (Figure 4C). A few side chains outside of the motif segment were not resolved in the electron density map perhaps due to inherent flexibility. The binding region was well resolved due to conformational stabilization by molecular contacts. The DYNLL2 dimer and the interactions within the two binding grooves are symmetric. In contrast, the coiled coil region of the Leu-zipper diverges from the symmetry axis of the complex due to asymmetric crystal contacts of its two $\alpha$-helices. Different backbone conformations of the four-residue linkers in the two peptides as well as their elevated thermal factors indicate that the linkers are flexible (Figure 4B) suggesting that the Leu-zipper does not interfere with the binding interaction. The bound SRGTQTE segment of the peptide preserves the 5 backbone hydrogen bonds of the $\beta$-sheet interaction and the 11 additional hydrogen bonds of the high resolution complex structure (Figure 4F). Favorable interactions of the additional Val\_5 residue highlight direct and indirect structural basis of its contribution to increased binding affinity. The Val\_5 residue extends the $\beta$-sheet establishing two backbone hydrogen bonds (Figure 4G), while its side chain is accommodated into a shallow pocket formed partially by the His68 imidazole ring, which is also involved in binding the Arg\_3 side chain through an H-bond (Figure 4E). The recombinant construct derived Gly\_5 and Ser\_6 residues do not interact directly with DYNLL2.

In the crystal of the dimer-to-dimer complex, due to crystal packing, neighboring GS-VRGTQTE-DYNLL2 units form antiparallel $\beta$-sheet structures. We speculate that the same antiparallel $\beta$-sheet structures form in case of the Ac-SRGQTQTE-DYNLL2 complex resulting in the observed almost instantaneous crystal formation. If this is the case, presence of the Leu-zippers would slow down the association of the resulting long
linear “rods”, and that is why crystallization of dimerized GS\textsubscript{SVSRTQTQTE} proceeds at a much lower rate.

Predicting novel DYNLL binding partners based on the phage-selected sequence set

It was previously shown that display bias-normalized affinity-selected amino acid frequencies correlate with binding energy contributions of the selected amino acid residues [46,47, 48,49,50,51]. This finding together with the assumption that amino acids in a linear motif contribute independently to the binding affinity, led to the design of a simple scoring system (see Materials and Methods). The general logic of binding motif prediction is illustrated as a flowchart in Figure 5. When an octapeptide segment is aligned with the phage-selected sequence logo, at each position the given amino acid type of the peptide gets scored based on the display-bias normalized frequency of the same amino acid type in the selected pool. The eight positional score values are summed up to give the final score (see Materials and methods). This system is suitable for the 7 fully randomized positions but is incompatible with the fixed Gln at position zero. Giving zero score at this position for the lack of a Gln would have underestimated such motifs compromising the entire scoring system. Therefore we analyzed only those segments that do have a Gln at position zero and we did not give a score to this position. This way the minimal score, zero is given to octapeptides that out of the fixed Gln do not contain any residue present in the selected set. The maximal score, 367 is given to the octapeptide representing the consensus of the selected pool.

We restricted our search to cytoplasmic and predictably disordered regions of the human proteome since known DYNLL

Figure 4. Crystal structure of Ac-SRGTQTE – DYNLL2 (A, C, D and F) and Leu-zipper dimerized VSRGTQTE – DYNLL2 complex (B, C, E and G). Peptides bound to DYNLL adopt similar conformation, see alignment in C. As shown in B the Leu-zipper dimerized VSRGTQTE peptides lay into two parallel binding grooves on DYNLL2 forming a dimer-dimer complex. Colors from red to blue correspond to high and low B-factors, respectively (from 131 to 48). The GGSG linker between the binding motif and the Leu-zipper appears to be flexible. Side chain interactions between the peptides and the binding groove of DYNLL2 are shown in D and E. Side chains of DYNLL2 having at least one atom within 4 Å distance from the peptide are shown as sticks. F and G are rotated versions (by approximately 90° along the axes of the binding motifs) of D and E, respectively, and show that the peptides bind to DYNLL2 in a β-strand conformation. Interestingly, the interaction is partially mediated by H-bridges of several buried structural waters represented as red spheres in D and E. Comparing F and G shows that the backbone of Val\textsubscript{5} forms one additional hydrogen bond, while E and G illustrate that its side chain interacts with DYNLL2 His68. Details of this interaction are highlighted by orange areas.

doi:10.1371/journal.pone.0018818.g004
binding motifs are limited to these sequence classes. The density distribution of the scores from 0 to 367 shows a nearly continuous, exponentially decaying profile (Figure 6A). The score distribution of known binding motifs (Table S1) shows that: i) the two motif families, \([K/R]_{-3}X_{-2}T_{-1}Q_{0}T_{1}\) and \(G_{-2}I_{-1}Q_{0}V_{1}D_{2}\) are separated from each other, the former having higher scores suggesting that members of this family are closer to the thermodynamic binding optimum; and ii) members of the \([K/R]_{-3}X_{-2}T_{-1}Q_{0}T_{1}\) family are dispersed in the high score region. Considering these as inner controls their distribution indicates that the scoring system is appropriate to predict novel partners.

A fundamental question was how to set the cutoff score value in order to define a set of predicted novel binders depleted in false positives. The global decaying trend of the probability density function of scores is dictated by simple statistical rules. Starting with an arbitrary sequence pattern for which no binding protein exists in the human proteome the same global trend of scores would apply. We argue that in case of a sequence pattern for which a corresponding binding protein does exist, natural selection should drift the proportion of high-score motifs away from the stochastically expected values. Following this logic we used a bootstrap analysis to determine the statistically driven baseline of the probability distribution as follows. We produced one thousand bootstrap analysis to determine the statistically driven baseline of stochastically expected values. Following this logic we used a

doi:10.1371/journal.pone.0018818.g005

220 the normalized density values were very close to one, but above that score a statistically significant amount of high-score motifs was detected (Figure 6B). Thus, we used the score value of 220 as a statistically verified cutoff to define our predicted dataset of DYNLL binders. Proteins with motif(s) having a score 220 or higher are considered to be potential DYNLL binders. Normalized density values above score 220 indicates positive natural selection that evolved a large number of DYNLL-binding partners, which is expected for a hub protein. Above the threshold level there are 242 motifs provided by 219 human proteins (listed in Table S3). Among these we find 8 out of the 10 human DYNLL binding motifs for which the binding site had already been identified. Interestingly, the highest possible score i.e. the phage-selected consensus sequence is present only in the human proteome in a microtubule associated protein called EML3 [52], which has not yet been identified as a DYNLL interacting partner. Our preliminary ITC studies verified that a recombinant fragment of EML3 (residues 8-94) containing a coiled coil domain and the in vitro evolved consensus motif binds to DYNLL with a \(K_d\) of 50 nM (Figure S2).

**Discussion**

Most biological processes manifest through complex networks of protein-protein interactions. In order to understand how a network functions, one needs to identify all individual interactions and determine their affinity and specificity. DYNLLs are hub proteins recognizing a short linear motif, a \(\beta\)-MoRE in their growing number of newly identified binding partners [2,57]. As residue conservation indicates functional importance, comparison of natural binder sequences could in principle elucidate the preference of their target binding site. However, known natural sequences usually represent only a subset of all existing binders biased the observed sequence trend. Moreover, natural evolution can select sequences based on several functions in addition to binding, which complicates interpretation of an observed pattern. A systematic study of the binding epitope through single amino acid replacements can test the functional importance of individual residues, but a complete analysis of even a small binding site would require an impractically large number of mutants. If the function relies on cooperating residues, mutation combinations might require astronomical number of variants. Directed evolution combines combinatorial mutagenesis with functional selection and easily handles the above problems [53]. Diversity of our phage display library covered almost all possible mutation combinations representing a near complete, unbiased starting variant set. Our *in vitro* binding-selection combined with a powerful display bias normalization revealed characteristic sequence features, but we have to keep in mind that sequence patterns obtained this way are the results of a thermodynamics-driven binding-selection as opposed to a complex natural selection. Based on the phage-selected binding pattern we applied a simple scoring system combined with an untraditional baseline correction approach to identify novel natural binding partners in the human proteome.

Our library design covered both natural DYNLL-binding sequence classes. Importantly, the 25 unique phage-selected binders demonstrated that the two class characteristics can freely mix. Thus, the existence of the two natural classes is not due to a simple binding selection. It is either due to functional selections other than DYNLL binding, or could simply be a sampling error due to the limited number of known natural DYNLL binders. The same conclusion was drawn from the *in vitro* thermodynamic characterization of target sequences binding to DYNLL [20].
ITC measurements show that dimerization of the peptides leads to an at least 100 fold increase in binding affinity. For successful phage selection it was necessary to start with a bivalently displayed peptide library. Nevertheless, the resulted sequence pattern is relevant for the monovalent interaction as well. Monomeric form of the consensus binding motif (VSRGTQTE, \( K_D = 0.08 \, \text{mM} \)) binds tenfold tighter than the strongest known natural DYNLL-binding sequence of the Bmf protein (EDKATQTL, \( K_D = 0.75 \, \text{mM} \)) \([20]\). The natural and the \textit{in vitro} evolved sequence patterns had strong similarities at positions \(2, 21\) and \(+1\) localizing the TQT segment as a binding epicenter. Additional similarities in preferences were found at positions \(-3, -2\) and \(+2\) suggesting that natural selection at these positions is also governed mostly by DYNLL-binding. However, compared to the TQT motif there is a relatively low level of conservation at these positions in both sets indicating less stringent requirement for a particular stereochemistry.

We identified two positions with characteristic differences between naturally and \textit{in vitro} evolved preferences. In nature Asp is the most frequent residue at position \(-4\), but it is absent from the \textit{in vitro} selected population, where the second most frequent residue in nature, Ser was selected. The anti Flag-tag antibody selected clone set showed that peptides having Asp within the octamer segment can be readily produced and displayed on phage.

Figure 6. Statistical analysis for threshold score determination. (A) Density distributions of scores. Dashed baseline is derived from average values from 1000 randomized score sets (see Methods) while continuous line represents the \textit{in vitro} selected scores. In the high score region (inset), the phage-selected set yields higher frequencies than the randomized baseline set. (Error bars show standard deviation of frequencies in the randomized sets). (B) The ratio of the density distribution functions of the scores. Above a threshold score value of 220 predicted DYNLL-binding motif frequencies exceed stochastic frequencies obtained from the randomized set, therefore their ratio exceeds one.

doi:10.1371/journal.pone.0018818.g006
There...
proteins listed in the predicted motif sets. The more of the predicted interactions are verified the better we will understand how DYNLL participates in organizing diverse protein networks and we will also learn more about how these networks function.

Materials and Methods

All chemicals unless otherwise specified were purchased from Sigma-Aldrich. DNA modifying enzymes were from Fermentas and New England Biolabs.

Expression vector constructions

For bacterial GST-fused expression the DYNLL1 gene was cloned into the pGEX4T-1 expression vector (Amersham) between BamHI and NdeI restriction sites using PCR. For His-tagged bacterial expression DYNLL1 and DYNLL2 (residues: 1–89, Uniprot accession numbers: P63167 and Q96FJ2, respectively) were cloned into the pET-15b (Novagen) expression vector as previously described [20]. The leucine zipper from yeast GCN4 peptide (residues: 250–281) was cloned between BamHI and BglII sites into a modified pET expression vector which allows N-terminal GST fusion and cleavage by TEV protease.

Protein expression, purification and peptide synthesis

6xHis-tagged DYNLL1 and DYNLL2; the N-terminal GST-tagged GCN4 leucine zipper-fused DYNLL-binding peptides and the N-terminal GST-tagged DYNLL1 were expressed in E. coli BL21(DE3) Star cells (Novagen). The cells were allowed to grow in LB media up to OD600 = 1.5 at 37°C and induced overnight with 0.5 mM IPTG at 18°C. After affinity chromatography using GST-Bind™ Resin (Novagen) or Profinity™ IMAC Ni-Charged Resin (Bio-RAD), the N-terminal GST or 6xHis tags were removed by proteolysis. In case of the GCN4 leucine zipper-fused DYNLL-binding peptides TEV protease was used. The DYNLL1, DYNLL2 and GST-tagged DYNLL1 proteins were further purified by anion exchange chromatography using HiTrapQ ion-exchange column (Amersham) as previously described [30].

Phagemid vector construction

The pS1602a phagemid [67] displaying human growth hormone (hGH) on the p3 coat protein of the M13 phage under the control of a pTac promoter was generously provided by Genentech. This vector contains an NsiI site between the signal peptide and the hGH coding sequence. Between the hGH and the p3 protein coding segments there is a glycine-serine (G/S) linker coding segment, which does not contain unique restriction enzyme cleavage site. A unique KpnI site was introduced to the beginning of the G/S linker by Kunkel mutagenesis [68] resulting the phagemid, pG2B. Then, in several cloning steps using synthetic adaptors, primers and PCR reactions, the pTFBL-p3 vector was produced. This vector contains a pTac promoter, a malE signal peptide and a p3 coding segment from pS1602a [67]. Between the signal peptide and the p3 segment it contains a FLAG-tag (DYKDDDD), a 5 residue G/S linker, an 8 residue (EDKATQTL) DLC-binding peptide from the Bmf protein (residues: 66–73; Uniprot accession number: Q96LC9) [5], a 4 residue G/S linker and a Leu-zipper from the yeast GCN4 protein (residues: 250–281, Uniprot accession number: P03069). The Leu-zipper segment was obtained by a PCR reaction from yeast cDNA library [69]. In a subsequent step the pTac promoter was replaced with a PhoA promoter by replacing a cassette between the EcoRI and NsiI sites of the pTFBL-p3 vector with a cassette between the same sites from the phGHR(1–238) vector [42], which was a generous gift from Genentech. This final step resulted in the vector pPFBL-p3 which is illustrated in Figure 1A.
Library construction

The library was produced based on previously described protocols [70]. In the pPFBL-p3 vector the Bmf octapeptide coding segment was replaced by 8 consecutive TAA stop codons by Kunkel mutagenesis [68] resulting in the pPF-STOP-L-p3 vector. This vector served as a stop template for the Kunkel mutagenesis, when stop codons were replaced with the following segment: NNK NNK NNK NNK CAG NNK NNK. The NNK codons represent a set of 32 codons covering all 20 amino acid residues. CAG codes for glutamine. The resulted pPF-Lib-p3 library construct is illustrated in Figure 1B and C.

Phage display

All steps of the phage display selection and analysis of individual peptide-displaying clones were carried out as previously described [70]. At first DYNLL1 was immobilized onto MaxiSorp plates (NUNC), but the immobilization efficiency was poor. The amount of properly folded immobilized homodimer DYNLL1 was increased by using the GST fused form. The increased yield might have been due to two factors. Larger proteins immobilize better on the MaxiSorp plate and GST is a homodimer, which as a tag, appears to stabilize the homodimer state of DYNLL1 [71]. To prevent selection of GST-binding peptides, the peptide-phae solution was supplemented with 30 µg/ml recombinant GST. Two independent selections on two different targets were performed. Nunc MaxiSorp ELISA plates were coated with GST-DYNLL1 (3 µg/ml) or anti-Flag tag antibody (2 µg/ml; SIGMA, F3165), respectively. Three selection rounds were carried out on each target separately as described. The eluted phage population was amplified in XL1 Blue cells superinfected with helper phage. Binding properties of individual peptide-phage clones were tested in a phage-ELISA format as described [70].

Sequence analysis

The genes of individual peptide-phage clones producing an ELISA signal on their target 3-fold above background (measured on BSA or casein containing wells) were sequenced by the Big Dye Terminator v3.1 cycle Sequencing Kit (Applied Biosystems). To eliminate the effects of display bias, the amino acid frequencies determined for DYNLL1-binding peptide-phage population were normalized to data from the anti-Flag tag selected population. Nunc MaxiSorp ELISA plates were coated with GST-DYNLL1 (3 µg/ml) or anti-Flag tag antibody (2 µg/ml; SIGMA, F3165), respectively. Three selection rounds were carried out on each target separately as described. The eluted phage population was amplified in XL1 Blue cells superinfected with helper phage. Binding properties of individual peptide-phage clones were tested in a phage-ELISA format as described [70].

Isothermal titration calorimetry

All peptides and DYNLL1 were dialyzed overnight in PBS buffer supplemented with 3 mM 2-mercaptoethanol (pH 7.4). A typical binding experiment involved 15 µM DYNLL1 in the cell and 30–40 injections varying between of 3–5 µl peptide solution. For each experiment the thermodynamic parameters were determined at 299 K. We used a VP-ITC instrument from Microcal and the data were fitted by the software package Origin 5.0 (OriginLab) using the simple A + B to AB binding model. The signal from an initial 2 µl injection was eliminated. The equilibration time between injections was 900 sec.

Crystallographic studies

DYNLL2 was dialyzed twice against TBS (20 mM TRIS-HCl, 150 mM NaCl, 3 mM NaN₃, 3 mM DTT, pH 7.6) and concentrated using Amicon Ultra-4 Centrifugal Filter Unit with UltraCel-5 membrane (Millipore). The Ac-SRGQTQTE peptide and the dimerized-VSRGTQTE peptide were dissolved in the same buffer and subsequently complexed with DYNLL2. In the final solutions of the complexes the Ac-SRGQTQTE concentration was 2.0 mM, the dimerized-VSRGTQTE concentration was 1.3 mM, while the concentration of dimeric DYNLL2 was 1.5 mM and 1.2 mM, respectively. Crystals were grown using the hanging drop method at 293 K with a reservoir solution of 31% PEG4000, 0.4 M CH₃COONaH₂O, 0.1 M CH₃COONa pH 4.6 in the case of the Ac-SRGQTQTE complex and 20% PEG8000, 0.2 M MgCl₂, 0.1 M TRIS pH 7.0 in the case of the dimerized-VSRGTQTE complex. Drops were composed of 2 µl reservoir solution and 2 µl protein solution. The crystals were soaked in reservoir solution plus 20% glycerol for 1 minute and subsequently flash frozen in liquid nitrogen.

X-ray diffraction data were collected from a crystal of the Ac-SRGQTQTE complex at 100 K with an ADSC Q315R CCD detector at ID29 of the ESRF (λ = 0.93 Å). Data were indexed, integrated and scaled to a resolution of 1.3 Å using XDS and XSCEL [72]. The space group is P2₁2₁2₁ with unit cell dimensions of a = 35.6 Å, b = 64.0 Å, c = 151.8 Å. Data collection of the crystal of the dimerized GSVSRGTQTE complex was carried out using a Rigaku R-AXIS IV++ detector with Cu Kα radiation (λ = 1.5418 Å) focused by Osmic confocal optics. Data were indexed, integrated and scaled to 2.9 Å resolution using the CrystalClear software from Rigaku. The space group is P2₁2₁2₁ with unit cell dimensions a = 53.8 Å, b = 68.4 Å, c = 101.7 Å.

The structure of the DYNLL2/Ac-SRGQTQTE peptide complex was solved by molecular replacement using the program PHASER [73] of CCP4 6.1.2 Program Suite [74,75]. As search model, the structure of 1CMI was used, thereafter automated model building was carried out with Arp/wArp [76] using amino acid sequences of both DYNLL2 monomer and the synthetic peptide Ac-SRGQTQTE. The model was systematically improved using iterative cycles of manual rebuilding with the program Coot [77] and refinement with Refmac5 of CCP4 6.1.2 Program Suite [74,75]. At the final steps of refinement restrained anisotropic temperature factors for the protein atoms were introduced and this step reduced R_ract and R_free by 2.4% and 1.5% respectively. During the automated model building, 419 water molecules were modeled and the final refined structure had overall R_cryst = 12.1% and R_free = 15.6%. The structure of the DYNLL2/Leu-zipper dimerized GSVSRGTQTE peptide complex was solved using the MOLREP program off the CCP4 package [78]. The DYNLL2/Ac-SRGQTQTE complex structure was used as search model. Model building was carried out using the program Coot [77]. The model was refined with Refmac5 [75] using restrained maximum-likelihood refinement and TLS refinement [79]. During refinement non-crystallographic restraints were added to the DYNLL2 molecules of the dimer as well as to the bound segment and coiled coil region of the engineered peptide. The final model contains 1995 atoms and 40 water molecules and it has overall R_ract = 25.0% and R_free = 29.5%. The stereochemistry of the structures was assessed with WHATCHECK [80] and PROCHECK [81]. The structures were deposited with the PDB under reference codes 2XQQ and 3P8M, respectively.
Crystallographic data and refinement statistics are shown in Table S4.

**Binding partner prediction**

For binding partner prediction a simple scoring matrix was used containing display-normalized position-specific percentages of each amino acid type of the selected sequence set. The score equals the summed up occurrences along the eight residue motif (Eq. 1).

\[ S = \sum_{i=1}^{8} p_{i}^N - 100 \]  

(1)

S is the score of the eight residue motif, \( p_{i}^N \) is the percentage occurrence of amino acid N in the \( i \)th position in the selected sequence set. Subtraction of 100 stands for eliminating the score contribution of the nonrandomized Gln at position 0. The input dataset was a filtered subset of the annotated human protein sequence set from UniProt [82]. Secreced proteins, transmembrane and extracellular domains of membrane proteins were excluded. From the remaining intracellular set only predicted disordered regions were used. A score value was assigned to each eight residue segment if at least one residue of the segment was predictedly unstructured and if the segment contained a Gln at position 0. For disorder prediction we used IUPred [83] as proposed by a recent comparative study [84]. The prediction program was implemented in Perl and is freely available with the scoring matrix upon request. (For an overview of binding site prediction see Figure 5.)

We determined the threshold as follows: we generated 1000 scoring matrices by keeping the original phage selected matrix scores at each position but randomized the positions themselves. Using the same input sequence set, we repeated the above described scoring with all randomized matrices. To calculate the probability density function, hits were summed up for ten score unit windows. Density distribution function values generated by the real phage selected scoring matrix were normalized by the averages of corresponding values obtained by the randomized matrices. Threshold was defined as a score value above which the normalized distribution function values exceed one.

**Supporting Information**

**Figure S1 Peptide display level.** (A) Effect of IPTG induction on the display level of the IPTG inducible pTac promoter driven pTBL-p3 construct. (B) Replacing of the pTac promoter with phoA promoter results in ~30-fold increase in display level. (DOC)

**Table S1** 41 known DYNLL binding motifs from 33 proteins. (DOC)

**Table S2** 25 non-identical (based on DNA level) phage selected sequences. (DOC)

**Table S3** 30 non-identical (based on DNA level) Flag-tag selected sequences. (DOC)

**Table S4** Crystallographic data and refinement statistics. (DOC)

**Table S5** List of predicted DYNLL binding motifs with scores higher than the threshold level (≥220). Sequences are sorted in three classes according to the reliability of the prediction: A) most probable, C) least probable interacting partners. For a detailed description of the classification rules see the main text. (DOC)

**Acknowledgments**

We thank Miklós Sahin-Toth for critical reading of the manuscript, Zoltán Gáspári for modeling the optimal SG linker length and for valuable discussions, Katálin Kekesi for MS analysis, Richard Chazal for part of protein crystallization, Csaba Hetényi for modeling studies and András Patthy for peptide synthesis. The source code of IUPred was a kind gift for their expert assistance. The sequence analysis was done on the “Atlasz” HPC of Eötvös Loránd University.

**Author Contributions**

Conceived and designed the experiments: FT GK LN GP. Performed the experiments: PR LR DS VH WW. Analyzed the data: PR LR DS VH WW GK. Wrote the paper: PR LR DS VH GK LN GP.

**References**

1. Pfister KK, Shah PR, Hummerich H, Russ A, Cotton J, et al. (2006) Genetic analysis of the cytoplasmic dynein subunit families. PLoS Genet 2: e1.

2. Barbar E (2008) Dynein light chain LC8 is a dimerization hub essential in diverse protein networks. Biochemistry 47: 503–508.

3. Hodi Z, Rapali F, Radnai L, Molnár T, Szemes A, et al. (2007) The LC8 family of dynein light chains: Multifunctional chaperone-like proteins. FEBS J 274: 106–106.

4. Pathalakath H, Huang DC, O’Reilly LA, King SM, Strasser A (1999) The proapoptotic activity of the Bcl-2 family member Bim is regulated by interaction with LC8 is required for Pak1 nuclear import and is indispensable for zebrafish development. PLoS One 4: e6025.

5. Pathalakath H, Villunger A, O’Reilly LA, Beaumont JG, Coutts I, et al. (2001) Bmf: a proapoptotic BH3-only protein regulated by interaction with the myosin V actin motor complex, activated by anisakis. Science 293: 1029–1032.

6. Lo KW, Kan HM, Chan LN, Xu WG, Wang KP, et al. (2005) The 8kDa dynein light chain binds to p53-binding protein 1 and mediates DNA damage-induced p53 nuclear accumulation. J Biol Chem 280: 8172–8179.

7. Herzog RP, Andresson U, Scarpulla RC (2000) Dynein light chain interacts with NRF-1 and EWG, structurally and functionally related transcription factors from humans and drosophila. J Cell Sci 113(Pt 23): 4263–4273.

8. Stelter P, Kunze R, Flemming D, Hopfner D, Diepholz M, et al. (2007) Molecular basis for the functional interaction of dynein light chain with the nuclear-pore complex. Nat Cell Biol 9: 780–786.

9. Lightcap GM, Kari G, Arias-Romero LE, Chernoff J, Rodeck U, et al. (2009) Interaction with LC8 is required for Pak1 nuclear import and is indispensable for zebrafish development. PLoS One 4: e6025.

10. Jacob Y, Radzha H, Coccaldi PE, Tordo N (2000) Cytoplasmic dynein LC8 interacts with lysis virus phosphoprotein. J Virol 74: 10217–10222.

11. Vadlamudi RK, Bagheri-Yarmand R, Yang Z, Balasenthil S, Nguyen D, et al. (2004) Dynein light chain 1, a p21-activated kinase 1-interacting substrate, promotes cancerous phenotypes. Cancer Cell 5: 575–585.

12. Fajtova A, Dasylova D, Bischof F, Lazarevic V, Altrock WD, et al. (2009) Dynein light chain regulates axonal trafficking and synaptic levels of Bassoon. J Cell Biol 185: 341–355.
13. Naishbitt S, Valschonaf O, Allison DW, Sala G, Kim E, et al. (2000) Interaction of the postsynaptic density-95/guanylate kinase domain-associated protein complex with a light chain of myosin-V and dynin. J Neurosci 20: 4524–4534.

14. Raykova KS, den Hollander P, Manavathi B, Tahdher AH, Song C, et al. (2006) Essential role of KIBRA in co-activator function of dynin light chain 1 in mammalian cells. J Biol Chem 281: 19092–19099.

15. Dick T, Ray K, Salz HK, Chia W (1996) Cytoplasmic dynein (dcl1) mutations cause morphological defects and apoptotic cell death in Drosophila melanogaster. Mol Cell Biol 16: 1966–1977.

16. Kamath RS, Fraser AG, Dong Y, Poulin G, Darbion R, et al. (2003) Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature 421: 231–237.

17. Espindola FS, Suter DM, Partata LB, Cao T, Wolenski JS, et al. (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 Cytoskeleton 47: 269–291.

18. Day CL, Puthalakath H, Skea G, Strasser A, Barsukov I, et al. (2004) Localization of dynein light chains 1 and 2 and their pro-apoptotic ligands. Biochem J 377: 597–605.

19. Lo KW, Kogoy JM, Kaoou BA, Kang SM, Fisher KK (2007) Interaction of the DYNLT (TCTEX1/RPS) light chains and the intermediate chains reveals novel intersubunit regulation during assembly of the dynemic complex. J Biol Chem 282: 58671–58678.

20. Radlau R, Rapali P, Hodi Z, Szweges D, Molnár T, et al. (2010) Affinity, avidity and binding kinetics of target sequence binding to LC8 dynein light chain isoforms. J Biol Chem.

21. Tochio H, Oishi S, Zhang Q, Li M, Zhang M (1998) Solution structure of a tetrameric 8-kDa dynein light chain with a bound peptide. Nat Struct Biol 5: 735–740.

22. Fan J, Zhang Q, Tochio H, Li M, Zhang M (2001) Structural basis of diverse sequence-dependent target recognition by the 8-kDa dynein light chain. J Mol Biol 306: 97–109.

23. Williams JC, Roulhac PL, Roy AG, Valle RB, Fitzgerald MC, et al. (2007) Structural and thermodynamic characterization of a cytoplasmic dynein light chain-intermediate chain complex. Proc Natl Acad Sci U S A 104: 10942–10947.

24. Benison G, Karplus PA, Barbar E (2008) The interplay of ligand binding and quaternary structure in the diverse interactions of dynein light chain LC8. J Biol Chem 283: 4004–4013.

25. Espindola FS, Suter DM, Partata LB, Cao T, Wolenski JS, et al. (2001) Identification of novel cellular proteins that bind to the LC8 dynein light chain. J Biol Chem 276: 41497–41499.

26. Naisbitt S, Valtschanoff J, Allison DW, Sala C, Kim E, et al. (2000) Dimerization status and functions. J Biol Chem 283: 4004–4013.

27. Wang W, Lo KW, Kan HM, Fan JS, Zhang M (2003) Structure of the LC8 dimer with a bound peptide. Nat Struct Biol 6: 735–740.

28. Mohan A, Oldfield CJ, Radivojac P, Vacic V, Cortese MS, et al. (2006) Analysis of the molecular recognition features, MoRFs, and their binding partners. J Proteome Res 5: 2351–2366.

29. Hodi Z, Nemeth AL, Radnai I, Herényi C, Schlett K, et al. (2006) Alternatively spliced exon B of myosin Va is essential for binding the tail-associated light chain showing dynamic properties. Biochemistry 45: 11564–11577.

30. Nyarko A, Hare M, Hays TS, Barbar E (2004) The intermediate chain of cytoplasmic dynein is partially disordered and gains structure upon binding to the intermediate chain. J Proteome Res 6: 2351–2366.

31. Naisbitt S, Valtschanoff J, Allison DW, Sala C, Kim E, et al. (2000) Cytoplasmic dynein promotes assembly of the coiled-coil domain of swallow protein. Biochemistry 49: 4161–4162.

32. Chen CV, Sudhir SS, Fu H (2004) Bivalent antibody phage display mimics natural immunoglobulin. J Immunol Methods 294: 119–132.

33. Wang L, Hare M, Hays TS, Barbar E (2004) Dynein light chain LC8 promotes assembly of the coiled-coil domain of swallow protein. Biochemistry 43: 14059–14066.

34. Sparks AB, Rider JE, Hoffman NG, Fowlkes DM, Brugge JS, et al. (1996) Phage display of ligand specificity by a human protein inhibitor of neuronal nitric oxide synthase. Nat Struct Biol 3: 965–969.

35. Sidhu SS, Fairbrother WJ, Deshayes K (2003) Exploring protein-protein interactions with phage display. Curr Opin Struct Biol 13: 14–25.

36. Paoluzzi S, Castagnoli L, Lauro I, Saleni AE, Cordella L, et al. (1998) Recognition specificity of individual EH domains of mammalian and yeast. EMBO J 17: 6531–6538.

37. Pal G, Ultsch MH, Clark KP, Currell B, Kossiakoff AA, et al. (2005) Comprehensive mapping of protein-protein interactions with rapid combinatorial scanning. J Biol Chem 281: 22387–22395.

38. Kim E, Lomeli H, Lin J, Kugiyama S, Naisbitt S, et al. (2007) Localizing the specificity profiles of phage-displayed peptide libraries. Nat Protoc 2: 1368–1386.

39. Tonikian R, Zhang Y, Boone C, Sudhir SS, et al. (2007) Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries. Nat Protoc 2: 1368–1386.

40. Dente L, Vertani C, Zucconl A, Pellicer G, Lanfrancone L, et al. (1997) Modified phage peptide libraries as a tool to study specificity of phosphorylation and recognition of tyrosine containing peptides. J Mol Biol 269: 694–703.

41. Feng S, Kazahara C, Rickels RJ, Schreiber SL, et al. (1995) Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands. Proc Natl Acad Sci U S A 92: 12295–12299.

42. Zhang Y, Appleton BA, Wu P, Wismann C, Sudhir SS (2007) Functional and structural analysis of the ligand specificity of the HtrA2/Omi PDZ domain. Protein Sci 16: 1738–1740.

43. Tonikian R, Zhang Y, Boone C, Sudhir SS (2007) Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries. Nat Protoc 2: 1368–1386.

44. Tegha-Dunghu J, Neumann B, Reber S, Krause R, Erle H, et al. (2008) EML1 is a nuclear microtubule-binding protein required for the correct alignment of chromosomes in metaphase. J Cell Sci 121: 1718–1726.

45. Walk SL, Li M, Zhang M (2003) Structural basis for the diverse interactions of dynein light chain LC8. J Biol Chem 283: 4004–4013.

46. Laura RP, Witt AS, Held HA, Gerstner R, Deshayes K, et al. (2002) The Erbin family. J Biol Chem 277: 2282–2288.

47. Tonikian R, Zhang Y, Boone C, Sudhir SS (2007) Function and divergent ligand specificity among PDZ domains of the LAP and zonula occludens (ZO) families. J Biol Chem 282: 22950–22955.

48. Zhang Y, Appleton BA, Wu P, Wismann C, Sudhir SS (2007) Structural and functional analysis of the ligand specificity of the HtrA2/Omi PDZ domain. Protein Sci 16: 1738–1740.

49. Tonikian R, Zhang Y, Boone C, Sudhir SS (2007) Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries. Nat Protoc 2: 1368–1386.

50. Dente L, Vertani C, Zucconl A, Pellicer G, Lanfrancone L, et al. (1997) Modified phage peptide libraries as a tool to study specificity of phosphorylation and recognition of tyrosine containing peptides. J Mol Biol 269: 694–703.

51. Wang L, Hare M, Hays TS, Barbar E (2004) Dynein light chain LC8 promotes assembly of the coiled-coil domain of swallow protein. Biochemistry 43: 4161–4162.

52. Pal G, Ultsch MH, Clark KP, Currell B, Kossiakoff AA, et al. (2005) Comprehensive mapping of protein-protein interactions with rapid combinatorial scanning. J Biol Chem 281: 22387–22395.

53. Tonikian R, Zhang Y, Boone C, Sudhir SS (2007) Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries. Nat Protoc 2: 1368–1386.
68. Kunkel TA (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc Natl Acad Sci U S A 82: 488–492.
69. Li Y, Brown JH, Rehetnikova L, Blazsek A, Farkas L, et al. (2003) Visualization of an unstable coiled coil from the scallop myosin rod. Nature 424: 341–345.
70. Sidhu SS, Lowman HB, Cunningham GC, Wells JA (2009) Phage display for selection of novel binding peptides. Methods Enzymol 328: 333–363.
71. Lyon RP, Hill JJ, Atkins WM (2003) Novel class of bivalent glutathione S-transferase inhibitors. Biochemistry 42: 10410–10420.
72. Kabach W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J Appl Cryst 26: 795–800.
73. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. Acta Crystallogr D Biol Crystallogr 61: 458–464.
74. Collaborative Computational Project N (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50: 760–763.
75. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53: 240–255.
76. Lamzin VS, Wilson KS (1993) Automated refinement of protein models. Acta Crystallogr D Biol Crystallogr 49: 129–147.
77. Emley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60: 2126–2132.
78. Vagin A, Teplyakov A (1997) Molecular replacement with MOLREP. Acta Crystallogr D Biol Crystallogr 66: 22–23.
79. Wion MLagNM M (2000) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Cryst 57: 122–133.
80. Hooft RW, Vriend G, Sander C, Abola EE (1996) Errors in protein structures. Nature 381: 272.
81. Laskowski RA, Moss DS, Thornton JM (1993) Main-chain bond lengths and bond angles in protein structures. J Mol Biol 231: 1049–1067.
82. UniProt T (2010) The Universal Protein Resource (UniProt) in 2010. Nucleic Acids Res 38: D142–148.
83. Dosztanyi Z, Csizmok V, Tompa P, Simon I (2003) The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. J Mol Biol 347: 827–839.
84. Szappanos B, Suveges D, Nyitray L, Perczel A, Gaspari Z (2010) Folded-unfolded cross-predictions and protein evolution: the case study of coiled-coils. FEBS Lett 584: 1623–1627.