The FEBS Journal

FEBS PRESS science publishing by scientists

Structural determinants governing S100A4-induced isoform-selective disassembly of nonmuscle myosin II filaments

Bence Kiss[1], Lajos Kalmár[2,3], László Nyitray[1] and Gábor Pál[1]

1 Department of Biochemistry, Eötvös Loránd University, Budapest, Hungary
2 Institute of Enzymology, Research Centre for Natural Sciences of the Hungarian Academy of Sciences, Budapest, Hungary
3 Department of Veterinary Medicine, University of Cambridge, Cambridgeshire, UK

Keywords
directed protein evolution; fuzzy complex; paralog specificity; phage display; protein–protein interaction

Correspondence
G. Pál and L. Nyitray, Department of Biochemistry, Eötvös Loránd University, 1117, Pázmány Péter sétány 1/C, Budapest H-1117, Hungary
Fax: +36 1 382 2172
Tel: +36 1 381 2171
E-mail: gabor.pal@ttk.elte.hu and nyitray@elte.hu

(Received 4 February 2016, revised 21 March 2016, accepted 30 March 2016)

doi:10.1111/febs.13728

The Ca^{2+}-binding protein S100A4 interacts with the C terminus of nonmuscle myosin IIA (NMIIA) causing filament disassembly, which is correlated with an increased metastatic potential of tumor cells. Despite high sequence similarity of the three NMII isoforms, S100A4 discriminates against binding to NMIIB. We searched for structural determinants of this selectivity. Based on paralog scanning using phage display, we identified a single position as major determinant of isoform selectivity. Reciprocal single amino acid replacements showed that at position 1907 (NMIIA numbering), the NMIIA/NMIIC-specific alanine provides about 60-fold higher affinity than the NMIIB-specific asparagine. The structural background of this can be explained in part by a communication between the two consecutive α-helical binding segments. This communication is completely abolished by the Ala-to-Asn substitution. Mutual swapping of the disordered tailpieces only slightly affects the affinity of the NMII chimeras. Interestingly, we found that the tailpiece and position 1907 act in a nonadditive fashion. Finally, we also found that the higher stability of the C-terminal coiled-coil region of NMIIB also discriminates against interaction with S100A4. Our results clearly show that the isoform-selective binding of S100A4 is determined at multiple levels in the structure of the three NMII isoforms and the corresponding functional elements of NMII act synergistically with one another resulting in a complex interaction network. The experimental and *in silico* results suggest two divergent evolutionary pathways: NMIIA and NMIIB evolved to possess S100A4-dependent and -independent regulations, respectively.

Introduction

Nonmuscle myosin II (NMII) motor proteins have central roles in cell adhesion, migration, and cytokinesis. These modular proteins consist of an N-terminal head or motor domain, a neck region that binds the essential and the regulatory light chain, and a coiled-coil tail domain that ends in a nonhelical (i.e., disordered) tailpiece region. NMII hexamers, consisting of two heavy chains and four light chains, self-assemble into filaments via their tail region enabling actin cross-linking and force generation by the motor domains [1]. In mammals, there are three paralog NMII heavy chain genes (*MYH9*, *MYH10*, and

Abbreviations
ACD, assembly competence domain; CD spectroscopy, circular dichroism spectroscopy; EDTA, ethylenediamine-tetraacetic acid; NMII, nonmuscle myosin II; NTA, nitrilotriacetic acid; RX-DMD, discrete molecular dynamics combined with replica exchange simulation; SPR, surface plasmon resonance; TCEP, tris(2-carboxyethyl)phosphine.

MYH14) encoding NMIIA, NMIIB, and NMIIC isoforms, respectively. The different heavy chains do not form hetero coiled-coils; however, the coiled-coil tails of different NMIIs coassemble into heterotypic filaments *in vivo* [2,3]. The NMII isoforms have both overlapping and distinct function and localization, determined mostly by characteristic differences in the motor and tail domains, respectively [4]. For example, distinct features of the enzymatic cycle of the NMIIB motor domain enables this isoform to exert a more prolonged tension on actin filaments, which is essential for smooth muscle contraction and cytokinesis [5–8]. However, it is the C-terminal end of the coiled-coil tail and the disordered tailpiece that determine the localization and filament morphology of the isoforms [9–11]. The assembly competence domain (ACD), which is essential for filament formation in all conventional myosin motors, resides near the C-terminal end of the coiled-coil domain. Besides the properties of the coiled-coil sequence, extrinsic factors, such as reversible phosphorylation and protein–protein interactions, also regulate filament formation. Phosphorylation of numerous sites within the disordered tailpiece by protein kinase C (PKC) and casein kinase 2 (CK2) is considered to be a general mechanism to regulate the filament assembly/disassembly and cellular localization of NMII isoforms [12,13]. In contrast, the Ca^{2+}-dependent S100A4 protein (metastasin, Fsp1, mts1) isoform specifically regulates the NMIIA filaments [14,15]. Overproduction of S100A4 results in increased motility and migration of cultured cells [16,17], hence it contributes to the metastatic phenotype of several cancer types [18,19].

Two parallel studies led to the structure determination of the S100A4–NMIIA complex [20,21]. The structural models revealed that the S100A4 homodimer binds one NMII heavy chain fragment in an asymmetric manner, and S100A4 interacts not only with the C terminus of the coiled-coil region but also with the N-terminal part of the disordered tailpiece. Simply based on sequence similarity of the three isoforms, one would assume that the most diverse tailpiece region should be the major source of isoform-selective interaction with S100A4 [20,21]. This suggestion was also supported by the fact that this region is markedly more hydrophobic in NMIIA than in NMIIB and NMIIC. However, binding studies with various NMIIA fragments, some containing the entire S100A4-binding site while others missing either the N-terminal helical or the C-terminal disordered binding segment, suggested a more important role for the helical segment. It was shown that (in terms of dissociation constant values), the helical-binding segment provides about three orders of magnitude higher contribution to the overall binding energy than the disordered tailpiece [20]. Additionally, we found that NMIIC binds to S100A4 with similar affinity than NMIIA despite the two isoforms having rather different disordered tailpieces [20]. Finally, it is the disordered segment of the NMIIA peptide that has the most diverse conformation in existing S100A4–NMIIA complex structures further supporting the assumption that this element is not a key determinant of the interaction [20–22].

Sequence comparison of the high-affinity α-helical binding regions of the three isoforms and analysis of the 3D structures highlight several positions that face S100A4 and carry different residues in the low-affinity NMIIB versus the high-affinity NMIIA and NMIIC isoforms. In the latter isoforms, these residues are Ala1907, Met1910, and Asn1911 (NMIIA numbering), which are substituted by Asn, Leu, and Ser, respectively, in the NMIIB isoform (Fig. 1A). As these positions reside in the high-affinity segment and their variation correlates with binding affinity of the respective isoforms, we assumed that at least some of them might contribute to isoform selectivity.

To investigate the functional significance of the above-mentioned sequence diversity on the isoform-selective S100A4–NMII interactions, we used an unbiased high-throughput directed protein evolution method, phage display. By performing paralog scanning phage display experiments followed by binding studies on individual point mutants, we reveal that despite the complex multi-segmented nature of the S100A4–NMIIA contact surface, a single amino acid residue has the utmost importance in isoform selectivity of NMIIs. Namely, Ala1907 in NMIIA and NMIIC is a key residue promoting isoform-selective binding to S100A4, while Asn at the same position in NMIIB is a key component that hinders such binding.

Based on our comprehensive multi-level analyses, we suggest a new mechanistic model that explains the importance of position 1907; assigns a role for the weaker contributing disordered tailpiece; highlights how the higher C-terminal coiled-coil stability of NMIIB hinders its binding to S100A4, and describes the interplay of these three structural elements in determining isoform-specificity. Moreover, based on our atomistic-level structural dynamics analyses of the whole molecular recognition motif, we also propose a model for a 'fuzzy complex' [23], where transient structural elements and helical propensities play important roles in the interaction.

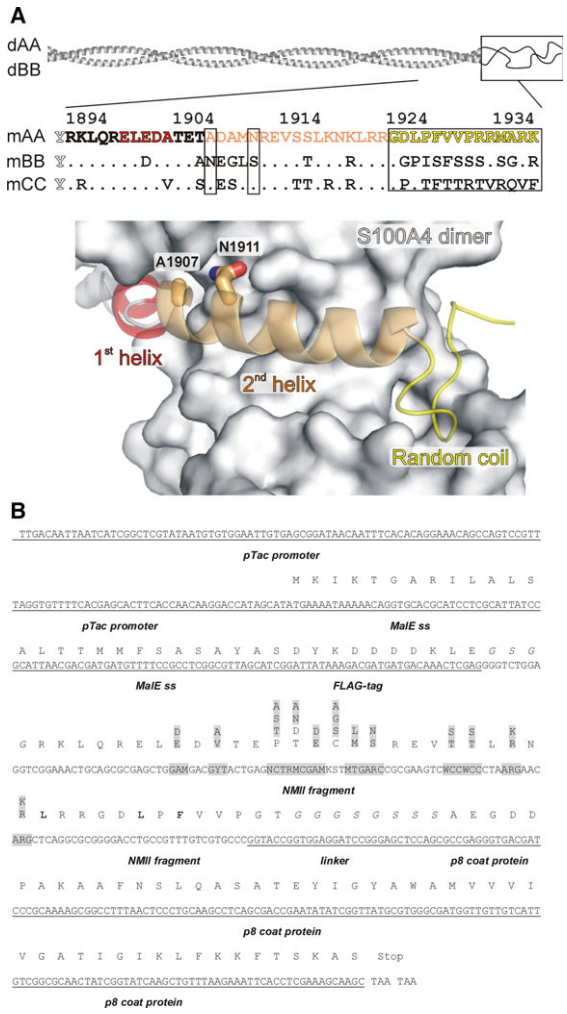


Fig. 1. Modular organization of the S100A4-binding site of NMII proteins and the NMII phage library construct. (A) *dAA* (NMIIA S1712-E1960) and *dBB* (NMIIB S1720-E1976) fragments form stable coiled-coils at room temperature and assemble to paracrystals in an ionic strength-dependent manner. *mAA*, *mBB*, and *mCC* are monomeric peptides representing the S100A4-binding site. An N-terminal tyrosine (Y) was added to the monomeric fragments to facilitate spectrophotometric concentration determination. Frames highlight the two residues and the disordered region that were swapped among NMII isoforms to generate chimera variants for binding studies. NMII-binding segments, referred to in the text as first helix (red), second helix (orange), and the fuzzy tailpiece (yellow), are depicted using the crystal structure of S100A4–*mAA* complex (PDB ID: 3ZWH). (B) An NMII fragment (residues 1894–1931 of NMIIA numbering) flanked by an N-terminal epitope tag (FLAG-tag) was fused to the p8 coat protein through a Gly/Ser linker. The MalE signal sequence (MalE ss) targets the variants to the periplasm before getting incorporated in the phage coat. The fusion gene is controlled by the pTac promoter. Paralog-scanned positions in the NMII fragment are highlighted with gray. Residues without gray background do not occur in any isoforms, but they are necessitated by the nature of the genetic code.

Results

Combinatorial paralog scanning by phage display

Paralog scanning combined by phage display is a powerful approach to locate amino acid residues that are responsible for paralog-specific functions [24]. We constructed an NMII-phage library (Fig. 1B) displaying the Arg1894–Pro1931 NMIIA segment. We decided to leave out the last six residues of the previously crystallized NMIIA fragment Arg1894–Lys1937 [20], as this positively charged region dramatically decreased display efficiency of the peptide. The NMII library was designed to include all residues that differ in the isoforms between positions 1901 and 1920 (Fig. 1B) and to produce all the corresponding chimeras. In this segment, there are 11 positions in which 2, and 1 position in which three different amino acids are presented by the three isoforms, altogether defining a set of 6144 sequences. Nevertheless, the structure of the genetic code and the traditional monomer-based oligonucleotide synthesis together necessitated representation of four different residues instead of two (i.e., two extra residue types not present in NMII isoforms) at 2 positions and four different residues instead of three (i.e., one extra residue type not present in NMII isoforms) at 1 position, so the library design led to 32 768 different variants. An NMII-phage library of 10^9 clones was produced and selected separately either on GST-S100A4 or on anti-FLAG-tag antibody, the latter one providing a readout and normalization for potential display bias. Altogether 29-29 clones from the second cycle of the two parallel selections were sequenced (Fig. 2A). The GST-S100A4-selected sequence pattern is illustrated as a display-normalized sequence logo [25] in Fig. 2B. It is well established that normalized amino acid frequencies obtained from the binding-selected population are reliable predictors of the binding energy contributions of individual residues [26–31].

Paralog scanning fully corroborated the previous assumption we made based on our published S100A4–NMIIA crystal structure that Ala1907 might contribute to isoform-specific S100A4 binding [20]. The NMIIA-(and NMIIC-) specific alanine was 54-fold more frequently selected than the NMIIB-specific asparagine (Fig. 2B). In the case of position 1911, a clear, but less dramatic, effect was observed, as the frequency of the NMIIA- and NMIIC-specific asparagine was fourfold higher than that of the NMIIB-specific serine. Importantly, at the remaining 10 paralog-scanned positions, no amino acid preference occurred. Such an 'inert' position was 1910, where the stronger binding NMIIA and NMIIC isoforms carry a

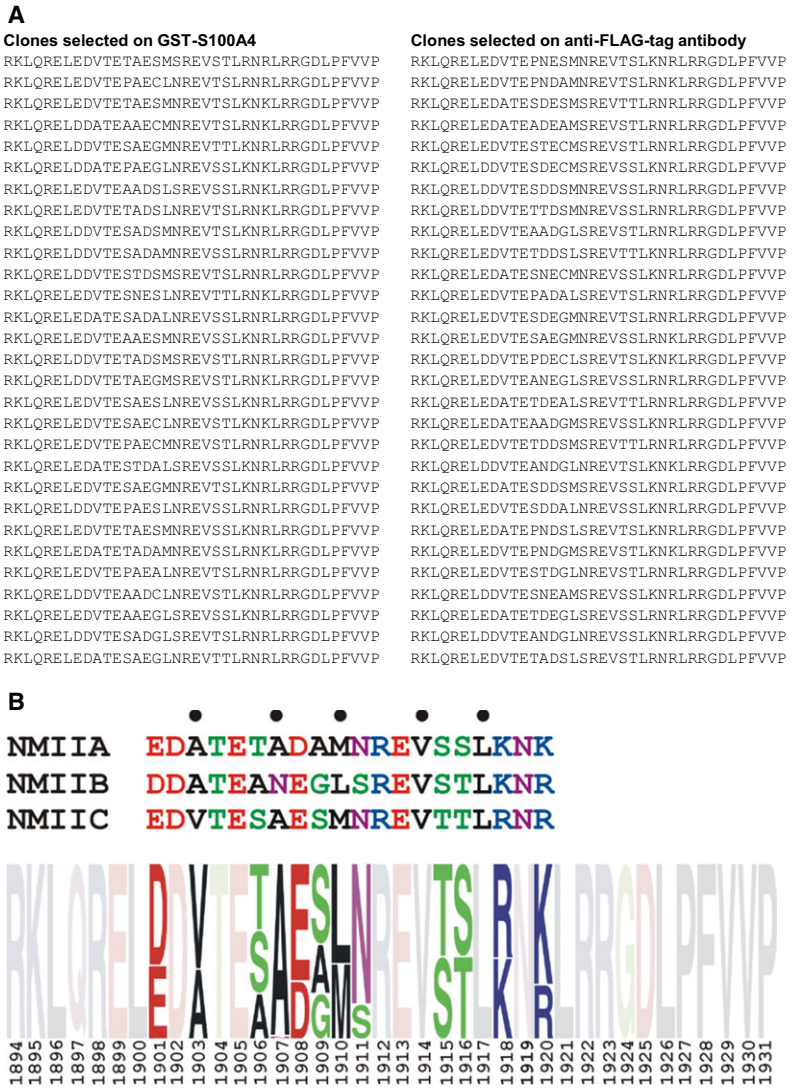


Fig. 2. Result of the NMII-phage library selection. (A) From the selection of NMII-phage library depicted in Fig. 1B, 29-29 unique S100A4 and anti-FLAG-tag antibody binders were sequenced. These sequences were used to generate display bias normalized (see in 'Materials and methods') sequence logo (B), where letter heights indicate normalized amino acid frequencies. Faded letters indicate positions that were not varied. Note that amino acids that had to be included in the starting library only because of the nature of the genetic code were omitted from logo generation. Amino acid sequence of the three NMII isoforms corresponding to the paralog-scanned region is shown on the top of the LOGO. The black dots indicate the *a* and *d* positions in coiled-coil heptad repeat. The corresponding residues face inward the S100A4–NMIIA complex.

methionine, but it was not positively selected over the NMIIB-specific leucine. Moreover, one could assume that in this hydrophobic binding interface, the larger NMIIC-specific valine in position 1903 is more favorable than the small NMIIA- and NMIIB-specific alanine, but these two residues were equally selected.

Interaction of S100A4 with monomeric NMII chimeras and their point mutants

To quantitatively assess how the Ala/Asn diversity at position 1907 (NMIIA numbering) affects S100A4 binding, we produced the corresponding swapped variants: NMIIA and NMIIC peptides containing Asn named $mA^{Ala\rightarrow Asn}A$ and $mC^{Ala\rightarrow Asn}C$, respectively, and an NMIIB peptide containing Ala named $mB^{Asn\rightarrow Ala}B$. In this nomenclature, the letter '*m*' stands for 'monomeric'; '*A*', *B*', and '*C*' refer to NMIIA, IIB, and IIC isoform, respectively. For example, in '*AA*', the first and second letters refer to the helical-binding region and the C-terminal disordered tailpiece-binding region of NMIIA isoform, respectively.

Results of binding assays performed by surface plasmon resonance (SPR) were in good accordance with the paralog scanning phage display experiment. The affinity of the $mB^{Asn\rightarrow Ala}B$ variant was 55-fold higher than that of *mBB*, while $mA^{Ala\rightarrow Asn}A$ and $mC^{Ala\rightarrow Asn}C$ bound approximately 70-fold weaker to S100A4 than their corresponding parent molecules (Fig. 3, Table 1). The affinity of *mAA* and *mCC* is three orders of magnitude higher than that of mBB, where difference is clearly governed by the alterations in the dissociation rate constants (Table 1).

We also tested the effect of the Ser1911Asn mutation on the affinity of *mBB* and $mB^{Asn\rightarrow Ala}B$ peptides. As we could not detect any affinity change in either case, the reciprocal mutants of *mAA* and *mCC* were not produced. Therefore, in the following experiments, we focused only on the single mutants at position 1907.

To assess the importance of the disordered tail in binding to S100A4, the peptide segments 1924–1937 were mutually swapped among the three isoforms yielding six chimera variants: *mAB, mAC, mBA, mBC, mCA*, and *mCB*. As above, the name reflects the new combination of the two structural elements in monomeric format.

Swapping the disordered segment had a relatively small effect on the S100A4–NMII peptide interactions suggesting that compared to the α-helical segment, the tailpiece element has a much lower contribution to the overall binding energy. This is also manifested in the form of a weaker contribution to isoform selectivity. Generally, the fragments comprising the NMIIA tailpiece bound about twofold stronger to S100A4

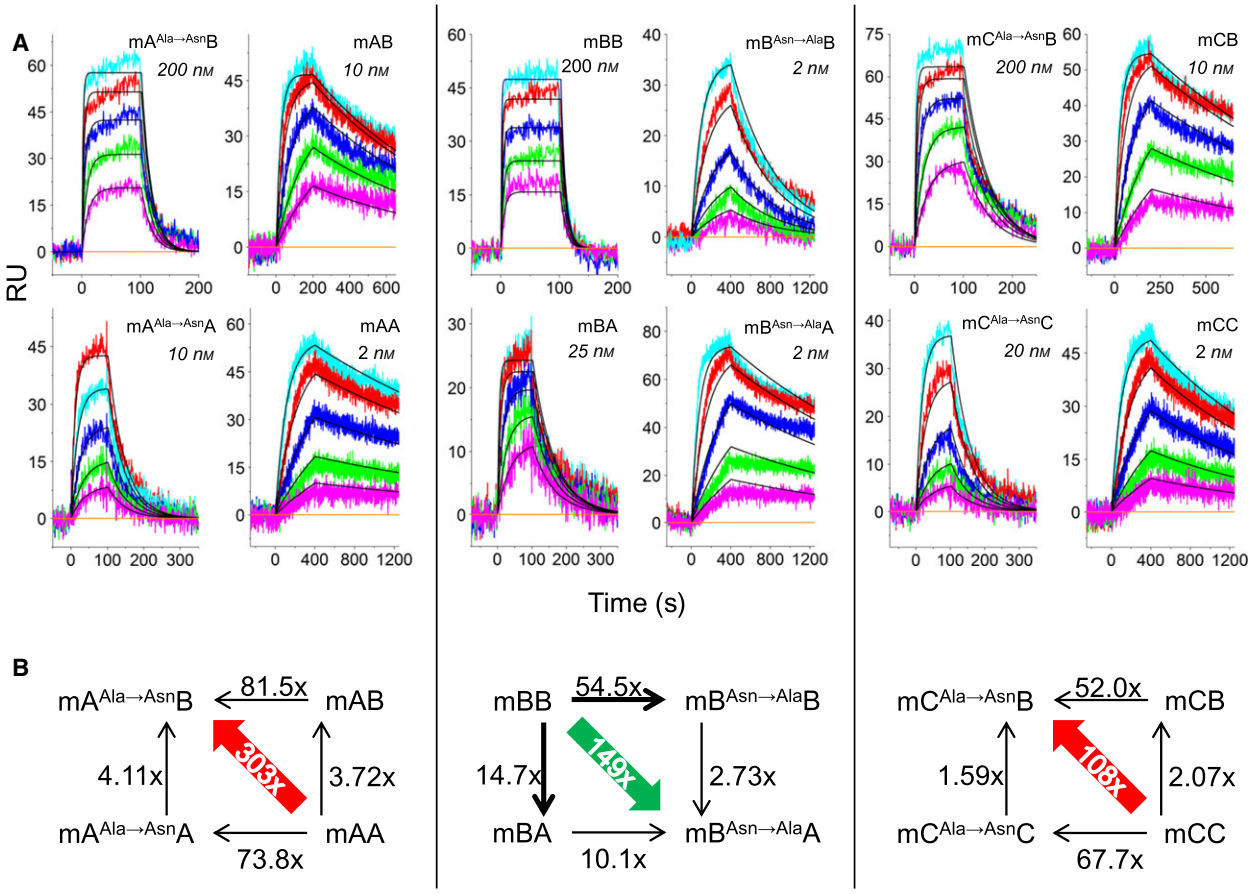


Fig. 3. Interaction of monomeric NMII fragments with S100A4. Monomeric NMII chimeras are referred to as m*AB*, a monomeric fragment that comprises the NMIIA helical region followed by the disordered tailpiece of NMIIB. Single mutants are depicted as, e.g., $mA^{Ala\rightarrow Asn}A$, a monomeric NMIIA fragment that comprises the NMIIB-specific asparagine in position 1907. (A) Monomeric NMIIA and NMIIB variants were injected to His_6-S100A4 immobilized on Tris-NTA sensor chip as described in 'Materials and methods'. Sensorgrams derived from five different analyte injections are indicated with colored lines, while black lines represent the global fit of the experimental data to a 1 : 1 Langmuir model. The highest analyte concentrations from the twofold dilution series are indicated in each panel. (B) Visual representation of the effects of single-residue mutations or tail swapping on the affinities of monomeric NMII isoforms. Black arrows always point from the wild-type to the mutant form. Note that variants with alanine at position 1907 are placed always on the right side, while variants with NMIIB tailpiece are always placed at the top of the boxes. Consequently, the effect of identical module swapping always resides on identical edge of the squares. Bolded arrow indicates a case when the effect of the corresponding mutation exceeds that of the same mutation shown as a nonbolded parallel arrow at the opposite edge of the box. Such situation suggests cooperativity of the functional modules. Red diagonal arrows highlight negative, while green arrow highlights positive cumulative effect of combined mutations on the affinities of monomeric NMII fragments. Note that these diagonal arrows just like all other arrows point from the wild-type to the mutant state.

Table 1. Analysis of surface plasmon resonance data for NMII binding to S100A4.

NMII fragment	k_{on} ($\mu M^{-1} \cdot s^{-1}$)	±SE[a]	k_{off} (s^{-1})	±SE[a]	K_d (nM)
mAA	4.21E+00	1.77E-02	3.79E-04	3.42E-06	0.09
dAA	3.86E-01	3.16E-03	8.10E-04	6.17E-06	2.10
$mA^{Ala\rightarrow Asn}A$	3.29E+00	2.56E-02	2.19E-02	1.78E-04	6.65
mAB	3.87E+00	2.05E-02	1.29E-03	8.52E-06	0.34
dAB	1.86E-01	1.68E-03	1.65E-03	8.63E-06	8.90
$mA^{Ala\rightarrow Asn}B$	1.54E+00	8.03E-02	1.67E-01	6.64E-03	109.00
mAC	4.78E+00	2.38E-02	1.11E-03	4.52E-06	0.23
mBB	3.93E+00	1.06E-01	1.21E-01	2.62E-03	30.80
$mB^{Asn\rightarrow Ala}B$	3.97E+00	4.13E-02	2.24E-03	1.15E-05	0.57
mBA	7.53E+00	1.49E-01	1.58E-02	1.62E-04	2.10
$mB^{Asn\rightarrow Ala}A$	5.32E+00	4.36E-02	1.14E-03	6.69E-06	0.21
$dB^{Asn\rightarrow Ala}A$	1.41E-01	1.24E-03	1.13E-03	6.80E-06	8.00
mBC	1.66E+00	8.30E-02	3.62E-02	9.23E-04	21.80
mCC	4.62E+00	2.01E-02	6.58E-04	3.66E-06	0.14
$mC^{Ala\rightarrow Asn}C$	1.83E+00	3.77E-02	1.76E-02	1.58E-04	9.61
mCA	3.90E+00	1.54E-02	2.80E-04	3.13E-06	0.07
mCB	3.06E+00	1.43E-02	9.00E-04	7.45E-06	0.29
$mC^{Ala\rightarrow Asn}B$	1.29E+00	2.42E-02	1.97E-02	2.02E-04	15.30

[a] Values represent the standard error of global fitting. Note that the affinity of the monomeric NMII fragments to S100A4 is governed by the dissociation rate constants.

than those having the NMIIC tailpiece, and approximately fourfold stronger than those having a NMIIB tailpiece.

The only notable exception is *mBA*, which bound to S100A4 14.7-fold stronger than *mBB* (Fig. 3B, Table 1). As the Ala/Asn diversity at position 1907 was found to be the major selectivity determinant of the helical segment, we tested if the observed nonadditive effect can be mapped to this position. Therefore, we constructed three additional NMII variants, $mA^{Ala\rightarrow Asn}B$, $mC^{Ala\rightarrow Asn}B$, and $mB^{Asn\rightarrow Ala}A$, and determined their binding constants to S100A4. We found that S100A4 bound $mA^{Ala\rightarrow Asn}A$ 4.1-fold stronger than $mA^{Ala\rightarrow Asn}B$, effect that is very similar to that of the *mAA*/*mAB* pair, where the NMIIA tailpiece provided a 3.7-fold stronger binding. Essentially, the same functional independence of position 1907 and the tailpiece was detected in the context of the NMIIC helical segment. S100A4 bound $mC^{Ala\rightarrow Asn}C$ 1.6-fold stronger than $mC^{Ala\rightarrow Asn}B$, effect that is very similar to that of the *mCC*/*mCB* pair, where the NMIIC tailpiece increases affinity 2.1-fold compared to NMIIB tail.

In contrast, while in the context of the *mBA*/*mBB* pair, the NMIIA tail provides a 14.7-fold stronger binding over the NMIIB tailpiece (Fig. 3B, Table 1), the affinity of $mB^{Asn\rightarrow Ala}A$ is only 2.7-fold higher than that of $mB^{Asn\rightarrow Ala}B$, which represents a 5.4-fold nonadditive effect.

In summary, the unusually high energetic contribution of NMIIA tailpiece observed in the context of NMIIB helical-binding segment cannot be mapped simply to the Asn/Ala diversity at position 1907 as it does not show up in the Ala-to-Asn mutants of the corresponding NMIIA and NMIIC variants. The fact that the observed nonadditivity manifests only in the context of NMIIB suggests that it is related to some other, perhaps overall characteristics of the NMIIB helical segment.

It seems that only in the context of NMIIB, the presence of Asn at position 1907 enhances the contribution of the NMIIA tailpiece to the binding energy 5.4-fold compared to when Ala occurs at this position. Alternatively, the complex stabilizing effect of Ala in position 1907 is 5.4-fold greater when the disordered binding region of NMIIB is present compared to when it is replaced with the equivalent region of NMIIA, but again, only in the context of NMIIB.

The simplest description to this phenomenon is that within the context of the NMIIB helical binding segment, Ala1907 and the NMIIA tailpiece act with negative cooperativity, i.e., they lower each other's binding energy contribution. The most straightforward mechanistic explanation is that there is an optimal conformation of the helical part for Ala1907 to bind to S100A4 and there is also an optimal conformation for the NMIIA tailpiece to bind to S100A4, but in the context of the NMIIB helix, these optimal conformations cannot coexist in the same complex, or if they do coexist, they exert a tension on the structure. The other side of the coin is that in the presence of the inherently

weaker binder NMIIB tailpiece, Ala1907 can bind in optimal local helical conformation, while when the inherently weaker binder Asn1907 is present, the NMIIA tailpiece can bind in optimal conformation.

MD simulations reveal an allosteric communication that tunes isoform selectivity

One might consider that amino acid substitution at position 1907 requires some conformational change of the helical binding site upon formation of the NMIIA–S100A4 complex, and the accompanied free enthalpy change is different for the two helical segments of the binding motif. To reveal more details about the structural dynamics of the peptides, we used the discrete molecular dynamics (DMD) method [32] to map the conformation ensemble of *mAA*, *mBB*, *mCC*, $mA^{Ala\rightarrow Asn}A$, $mB^{Asn\rightarrow Ala}B$, *and* $mC^{Ala\rightarrow Asn}C$. DMD combined with replica exchange simulation (RX-DMD) was previously used to predict helix content in experimentally analyzed disordered proteins or protein regions, and was able to detect the position and propensity of preformed structural elements with high fidelity [33].

Conformational ensembles of the two isoforms showed only a few small differences (Fig. 4). Helix propensity was relatively low but significant in *mAA, mBB*, *and mCC* (Fig. 4A,C,E).

In contrast to the crystal structure PDB ID: 3ZWH, in the crystal structure of *mAA* complexed with a truncated form of S100A4 (PDB ID: 4CFQ), only the second helix (Ala1907–Arg1923 in PDB ID: 3ZWH) of *mAA* has strictly defined conformation. The first helix (Glu1899–Ala1903) is invisible, while the C-terminal tailpiece segment adopts two alternative conformations [22]. Although the two 3D structures differ in the crystallized S100A4 variant, as well as in crystal packing, the identical position of NMIIA segment Asp1907–Arg1923 and alterations in the conformation of the other binding segments imply a binding mechanism as follows. The helical binding region might bind to S100A4 in two consecutive steps. In the first step, the second helical segment would bind to S100A4 yielding an intermediate state in which the second helix is already at its final position, while the first helix and the tailpiece provide fuzziness to the complex.

To simulate this state, we folded up the second helical-binding segment (positions 1910–1923, leaving more flexibility in its N-terminal end) to an α-helix and put a constraint on it to maintain the helical structure during the DMD simulation. We found that fixing the second helix in the complex almost doubles the helix propensity of the first helix and this way initiates its folding in *mAA* (Fig. 4A). In *mBB*, no such effect was detected; the helix propensity of the first helix was insensitive on fixing the longer second helix (Fig. 4C). In *mCC*, we also found some communication between the second and first helix, but it was less significant compared to *mAA* (Fig. 4E). Importantly, the single Ala-to-Asn substitution in $mA^{Ala\rightarrow Asn}A$ located in the linker between the two helices resulted in the same phenomenon observed in *mBB*: the Ala-to-Asn replacement completely abolished the allosteric effect (Fig. 4B). However, the reciprocal Asn-to-Ala substitution in *mBB* had only a slight effect on helix propensity of the first binding segment (Fig. 4D), which became apparent only when the full spectrum (all temperatures) of the RX-DMD simulations were compared. In $mC^{Ala\rightarrow Asn}C$, we also observed only a slight effect on helix propensity. Upon the Ala-to-Asn substitution, the first helix showed almost the same level of helicity as the wild-type.

The high-resolution NMIIA–S100A4 structures, the results of the DMD simulation, and the context dependence of the binding energy change upon the Asn-to-Ala replacement suggest that Ala1907 has two functions in isoform NMIIA. One is binding to a shallow hydrophobic binding site on S100A4, where it appears to fit better than an Asn. The other is an allosteric effect: Ala1907 manages to increase helical propensity of the first α-helical segment and in this way increases its binding affinity.

While based on the DMD simulation, an Ala substitution in NMIIB does not seem to perform this second job, but it could still fulfill the first one. We can approximate the direct binding energy contribution of Ala1907 by comparing affinities of *mBA* and $mB^{Asn\rightarrow Ala}A$, which represents the contribution of Ala in the context of an isotype NMIIB helix that does not increase the helical propensity of the first helical segment. In this context, Ala increases affinity 10.1-fold. Note that this variant is our closest control for *mAA*, as it also has an '*A*' nonhelical tailpiece. In the context of isoform NMIIA, the same Ala increases affinity 73.8-fold, so the contribution of the helix propensity increasing allosteric effect in NMIIA is approximately 73.8/10.1 = 7.3-fold.

In summary, we were able to detect a short range intramolecular allostery within the peptide through which the first and the second helices communicate in S100A4-complexed *mAA*, and showed that Ala1907 of NMIIA has an important role in this allosteric effect. The direct binding energy contribution and the context-dependent allosteric contribution of Ala1907 are apparently in the same magnitude.

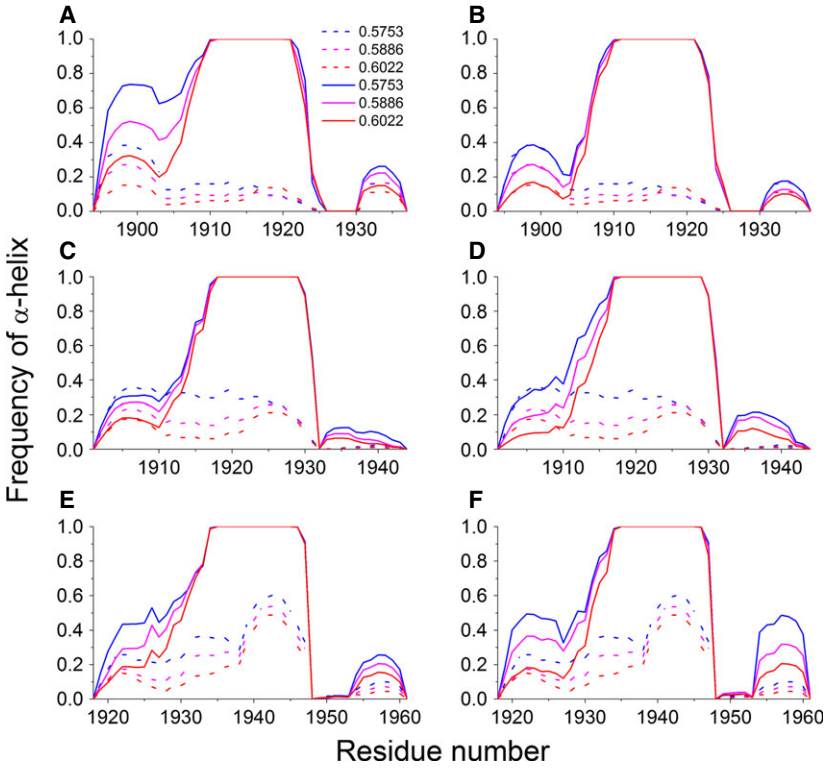


Fig. 4. Helical propensity differences of NMII fragments calculated from RX-DMD simulations. Frequencies were calculated from 4500 structures of each indicated temperature boxes, where 0.5886 corresponds to room temperature. Originally, RX-DMD simulations were run in 12 temperature boxes (see 'Materials and methods'), but for clarity, only the room temperature and the adjacent temperature boxes are plotted. Helix propensities of the peptides with (solid line) and without (dashed line) helical constrain on the middle region are shown on (A) *mAA*, (B) $mA^{Ala\rightarrow Asn}A$, (C) *mBB*, (D) $mB^{Asn\rightarrow Ala}B$, (E) *mCC*, and (F) $mC^{Ala\rightarrow Asn}C$. Same temperatures are plotted with the same color.

Interaction of S100A4 with dimeric NMII chimeras and their point mutants

Starting with coiled-coil forming NMIIA (*dAA*, Ser1712–Glu1960) and NMIIB (*dBB*, Ser1720–Glu1976) fragments, where '*d*' denotes dimeric, we reconstructed the same 1907 Ala-to-Asn point mutants already described for the monomers, and produced similar tail swapping chimeras as well (Fig. 1A). The only difference in this respect was that these fragments contained the full-length disordered tailpieces. The corresponding 259-residue long NMIIC fragment was difficult to keep in solution even at high ionic strength, thus the binding experiments were performed only with the NMIIA and NMIIB variants. When NMII coiled-coils were injected over the interaction surface, high nonspecific binding to the chip was observed (the same applied when we tried to use S100A4 as analyte), thus we needed to keep the NMII concentration relatively low. Consequently, we could measure only the relatively high, approximately nanomolar affinity (*dAA*, *dAB*, and $dB^{Asn\rightarrow Ala}A$) interactions (Fig. 5, Table 1).

Even the limited amount of data provided very useful information showing that all of the three coiled-coil variants bind to S100A4 with an association rate constant (k_{on}) 10- to 40-fold lower than their respective monomeric versions. This finding is in line with our hypothesis published previously [20] that the coiled-coil formation must compete with S100A4 binding. Namely, the S100A4-binding site on NMII proteins is in part buried in the coiled-coil structure resulting in its reduced accessibility to S100A4, which lowers the association rate constant of the interaction. Moreover, the k_{on} value of *dAA* is fourfold higher than that of $dB^{Asn\rightarrow Ala}A$ in spite of the fact that k_{on} values of the two monomeric versions of these forms are almost the same (Table 1). This fourfold difference in the k_{on} values of the two dimeric forms comprising different types of helical (and coiled-coil) segments suggests that the C-terminal coiled-coil stability of NMIIB should be higher than that of NMIIA hindering complex formation and this way contributing to isoform selectivity.

Stability of NMII coiled-coil fragments

To test whether there are coiled-coil stability differences between the isoforms, we determined the thermal stability of the *dAA* and *dBB* fragments by heat denaturation experiments using circular dichroism spectroscopy. In good accordance with our expectation, melting temperature of *dBB* was higher than that of *dAA* (T_M = 49.1 ± 0.1 °C and 34.7 ± 0.1 °C, respectively) (Fig. 6A).

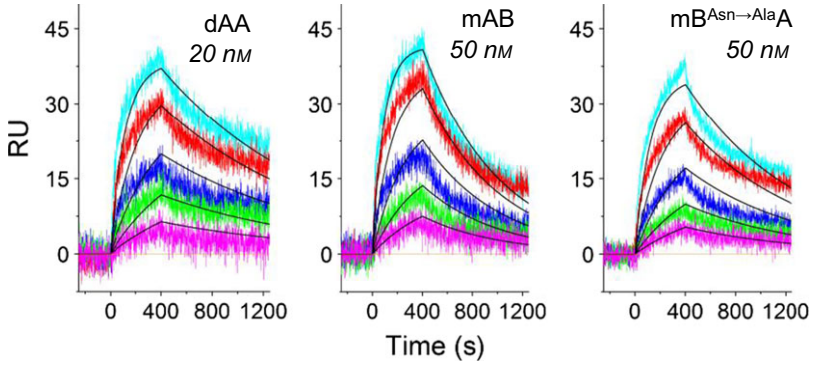


Fig. 5. Interaction of dimeric NMII fragments with S100A4. Dimeric NMII chimeras are referred to as *dAB*, a dimeric fragment that comprises the NMIIA coiled-coil region followed by the disordered tailpiece of NMIIB. The $dB^{Asn→Ala}A$ variant comprises the NMIIA-specific alanine in position 1907 (NMIIA numbering). The dimeric NMII variants were injected to His_6-S100A4 immobilized on Tris-NTA sensor chip as described in 'Materials and methods'. Sensorgrams derived from five different analyte injections are indicated with colored lines, while black lines represent the global fit of the experimental data to a 1 : 1 Langmuir model. The highest analyte concentrations from the twofold dilution series are indicated in each panel.

We performed RX-DMD simulations on the coiled-coil regions of dimeric *dAA* and *dBB* to localize regions (residue differences) responsible for the different stability. We found significant differences between the dissociation profiles of the two coiled-coil dimers. While only the ends and a short inner region (1767–1788) of *dBB* showed increased dissociation propensity upon heating, in *dAA* we detected two, more extended unstable regions (1766–1801, 1823–1848). While the positions of the first unstable region in the two isoforms partly overlap and their sequence is also highly similar, the second unstable region is unique to *dAA* (Fig. 6B–D). We found some evolutionarily conserved differences between NMIIA and NMIIB on the sequence level in the region 1747–1839 (Fig. 7) further corroborating that this segment might be responsible for the characteristically different coiled-coil stability of the two isoforms. Although (because of poor solubility) we could not obtain experimental results on *dCC*, we estimated its stability by RX-DMD simulations. Interestingly, it was even less stable than *dAA*.

Interaction of S100A4 with NMII paracrystals

Finally, we tested the effect of single-residue replacement and/or tailpiece swapping on S100A4-induced NMII filament disassembly. As it was previously described, S100A4 is unable to disassemble *dBB* paracrystals [15]. Importantly, we show here that both tailpiece swapping and the Asn-to-Ala mutation markedly increased the sensitivity of the corresponding variant *dBB* paracrystals to S100A4 binding (Fig. 8B). Notably, S100A4 disintegrated $dB^{Asn→Ala}A$ assemblies with the same efficiency as it disassembled the *dAA* and *dAB* filaments (Fig. 8).

S100A4 apparently bound to $dA^{Ala→Asn}A$ assemblies with similar affinity as to $dB^{Asn→Ala}B$, and the affinities of $dA^{Ala→Asn}B$ and *dBA* to S100A4 were also similar (Table 2). This is in contrast with corresponding data determined for the monomeric fragments, where S100A4 bound $mB^{Asn→Ala}B$ 12-fold stronger than $mA^{Ala→Asn}A$, and it bound *mBA* 13-fold stronger than $mA^{Ala→Asn}B$ (Table 1). The fact that preferential binding to $mB^{Asn→Ala}B$ and *mBA* is not manifested in the context of the dimeric format can be explained by the fact that the stability of the NMIIB coiled-coil is higher than that of the NMIIA protein. However, it was previously described that the structure and stability of paracrystal-like assemblies formed by different NMII isoforms also varies, which may also contribute to the isoform selectivity of the NMII–S100A4 interactions [9].

Discussion

Over the last decade, the role of intrinsically disordered regions (IDRs) and their main functional elements (known as short linear motifs – SliMs or molecular recognition elements – MoREs) has been widely recognized in regulatory processes [34]. After gene duplication, paralogs can retain their basic function, but in their rapidly evolving IDRs, it can easily acquire novel short functional elements that could alter localization and/or regulation [35]. In the case of nonmuscle myosins, the intrinsically disordered nonhelical tailpiece plays an essential role in differentiating

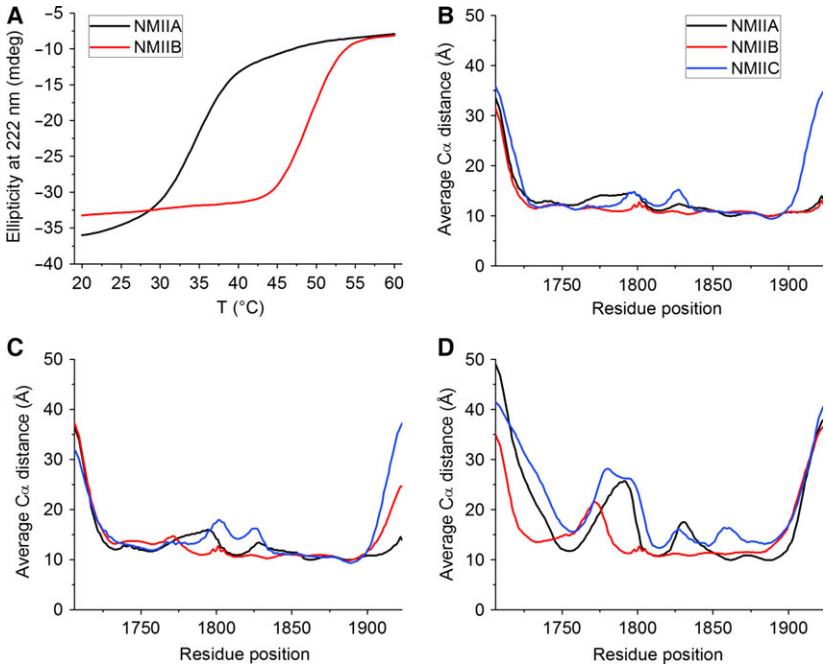


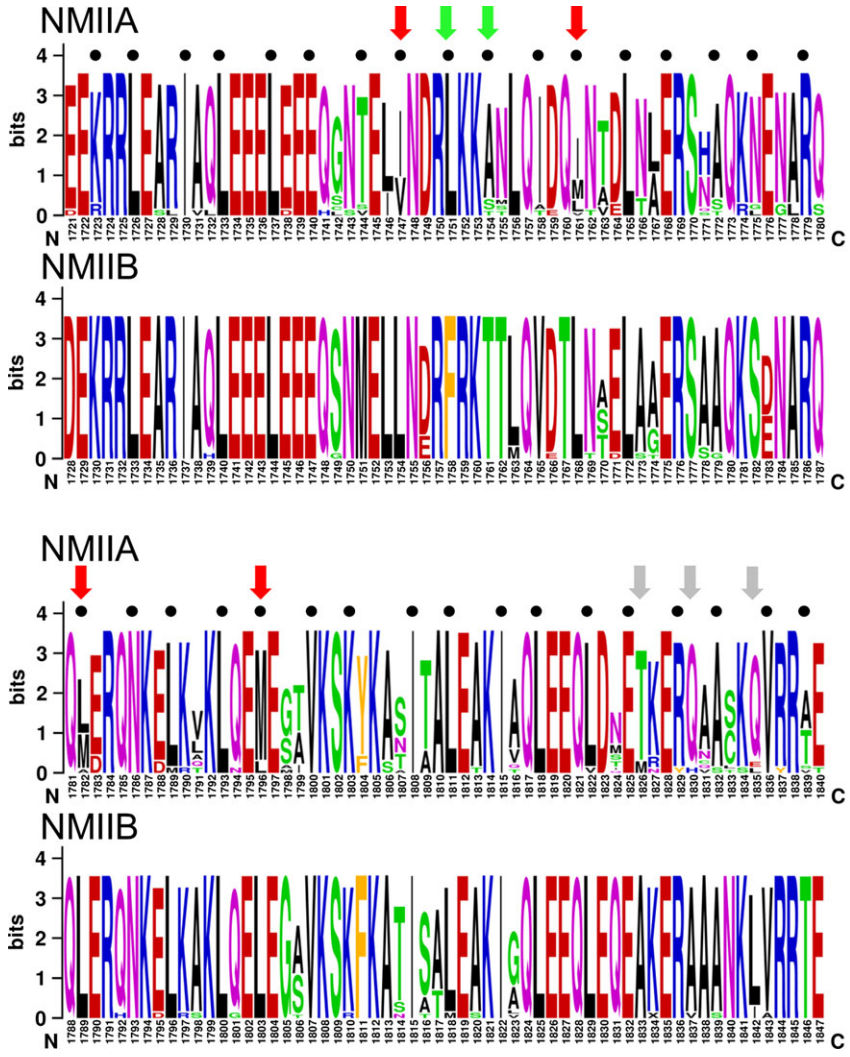
Fig. 6. Stability of dimeric NMII fragments. (A) Thermal unfolding of 5 μM *dAA* and *dBB* was followed with circular dichroism spectroscopy at 222 nm. (B–D) Average distances between identical Cα atoms of the two chains of the coiled-coil structure. Averages were calculated using 5000 structures from temperature boxes 0.6161 (B), 0.6303 (C), 0.6448 (D) of RX-DMD simulations (originally eight boxes) on NMIIA-, NMIIB-, and NMIIC-terminal coiled-coil region. We used a seven-residue long sliding window to smooth waves caused by the helical structure. Residue positions are denoted according to NMIIA numbering.

the regulation and localization of NMII isoforms [9–11]. While the charge pattern of the disordered tailpiece is conserved, the length and the amino acid composition notably vary. The proximal, positively charged region contains numerous phosphorylatable Ser and Thr residues in NMIIB and NMIIC [36]. In contrast, these residues are excluded from NMIIA. Around the S100A4-binding region (Gln1897–Ala1935), Ser1943 is the closest phosphorylatable residue within the tailpiece. As the two sites are seven residues away from each other, the phosphorylation event unlikely interferes with S100A4 binding. On the other hand, NMIIA is rather abundant in hydrophobic residues at the proximal tailpiece (Fig. 1). This phenomenon implies an elegant evolutionary scenario representing two characteristic divisions of SLiMs [35]: The same part of an intrinsically disordered region divergently evolved to (a) a ligand motif in NMIIA to interact with the hydrophobic pocket of S100A4 or (b) to a post-translational modification motif in NMIIB and NMIIC to be phosphorylated.

Nevertheless, it is important to keep in mind the following considerations.

Firstly, the whole S100A4-binding site also includes the C-terminal coiled-coil region and therefore it is much more extended than a regular SLiM [35]. Secondly, the helical-binding part has a greater contribution to the binding energy than the disordered segment [20]. Thirdly, both NMIIC and NMIIA bind to S100A4 with subnanomolar affinity, while SLiMs usually enable low-affinity, transient interactions [35]. To sum up these notions, while the disordered tail can be considered as an IDR/SLiM, the nature of this binding region precludes it to be an autonomous determinant of isoform selectivity. On the other hand, it should also be noted that in the context of the myosin filaments, the tailpiece is the only binding motif region that is freely accessible for S100A4 to make the first transient and likely fuzzy interaction in all NMII isoforms [23]. This fact in itself brought up the possibility that the whole binding motif has an inner modular architecture and if so, the elements of this architecture might act synergistically. We decided to conduct a thorough analysis to dissect this question.

Here, we showed that isoform-selective binding and filament disassembling activity of S100A4 are indeed encoded at multiple levels in the sequence and structure of the NMII isoforms. Moreover, we found that the corresponding functional elements of NMII act synergistically resulting in a complex interaction network. In this network, we identified four elements: a key residue at position 1907; stability of the coiled-coil segment; nature of the disordered tailpiece; and structural plasticity of the helical binding segment, and showed that all these play an interrelated role. On one hand, we used SPR for measuring binding kinetics and affinity values as this method is sensitive and delivers highly reproducible results. However, we emphasize that absolute values delivered by SPR are not expected to be identical with the ones prevailing inside the cell.



Species	NMIIA	NMIIB
Xenopus laevis	B7ZS76	Q04834
Meleagris gallopavo	G1NJB2	G3UUP1
Ficedula albicollis	U3KF79	U3K3Y7
Anas platyrhynchos	U3IIB9	U3J7L2
Gallus gallus	P14105	Q789A6
Anolis carolinensis	G1KSP5	G1KU94
Oryctolagus cuniculus	G1SL68	G1TC33
Loxodonta africana	G3UIR6	G3SWZ8
Callithrix jacchus	U3DST1	F7I9C3
Mus musculus	Q8VDD5	Q61879
Rattus norvegicus	Q62812	G3V9Y1
Bos taurus	F1MQ37	Q27991
Canis familiaris	Q258K2	F1PVV7
Macaca mulatta	H9YUC5	H9ENN2
Nomascus leucogenys	G1RWQ0	G1RG61
Homo sapiens	P35579	P35580
Otolemur garnettii	H0WTE2	H0X188
Danio rerio	F8W3L6	I3ISA3
Takifugu rubripes	H2RW07	H2TKH5

Fig. 7. Sequence logo representation of the C-terminal coiled-coil segments of various vertebrate NMIIA and NMIIB proteins. The corresponding input sequences were retrieved from the UniProt database. The inserted table contains the species list and the UniProt codes of the input proteins. Black dots indicate 'a' and 'd' coiled-coil positions. Colored arrows mark 'a' and 'd' positions that show isoform-specific conservation. The four red arrows point to positions that in NMIIB carry an optimal, while in NMIIA, a suboptimal residue for coiled-coil stability. The two green arrows highlight positions where NMIIA carries an optimal, while NMIIB a suboptimal residue for stability. We note that between the heptads 1793–1799 and 1808–1814, there are eight (instead of seven) positions necessitating the existence of a skip residue. Moreover, between positions 1823 and 1835, the two logos detect evolutionary conserved differences in the physicochemical nature of residues. At positions 1826, 1830, and 1835 (NMIIA numbering) indicated by gray arrows, NMIIA orthologs carry Thr, Gln, and Gln residues, while NMIIB orthologs Ala, Ala, and Leu, respectively. Because of the more hydrophobic nature of this NMIIB segment, correct assignment of the coiled-coil register is problematic hence and the functional significance of the difference is unclear.

On the other hand, the ratios of the measured values corresponding to isoform-selective contributions of individual residues or structural segments should provide interpretable information for the *in vivo* situation.

The individual observed binding energy contributions and the detected nonadditive effects, when all combined, point to a consistent mechanistic model that explains how and why S100A4 preferentially disassembles filaments of the NMIIA isoform (Fig. 9). The tail domain of all NMII isoforms contains a long coiled-coil segment followed by a C-terminal disordered tailpiece. While the larger part of the S100A4-binding site on NMII is buried in the coiled-coil structure and becomes available only for short times when the two strands temporarily separate, the tailpiece is always available for binding. Out of the

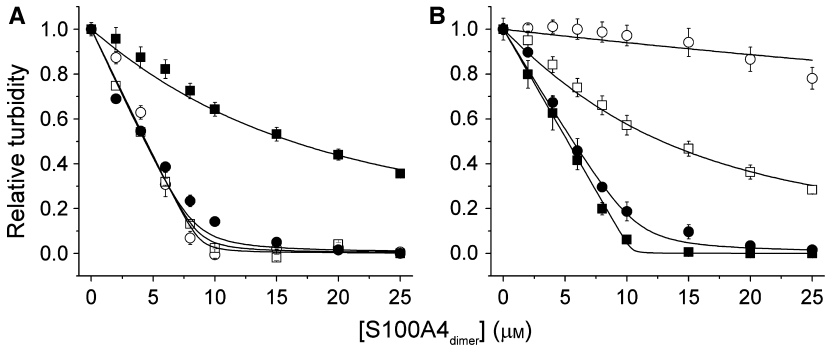


Fig. 8. S100A4-induced disassembly of NMII paracrystals. (A) About 10 μM (monomeric heavy chain concentration) *dAA* (open circle), *dAB* (open square), $dA^{Ala\rightarrow Asn}A$ (filled circle), and $mA^{Ala\rightarrow Asn}B$ (filled square), and (B) about 10 μM *dBB* (open circle), *dBA* (open square), $dB^{Asn\rightarrow Ala}B$ (filled circle), and $dB^{Asn\rightarrow Ala}A$ (filled square) was titrated with S100A4 (indicated on the abscissa as dimer). The data points represent the mean ± SE of three independent experiments.

Table 2. Interaction of the S100A4 homodimer with NMII paracrystals.

NMII fragment	K_d (nM)	±SE[a]
dAA	39.7	17.8
dAB	93.4	54.9
$dA^{Ala\rightarrow Asn}A$	191	135
$dA^{Ala\rightarrow Asn}B$	11 200	538
dBB	148 000	39 700
dBA	7850	446
$dB^{Asn\rightarrow Ala}B$	238	239
$dB^{Asn\rightarrow Ala}A$	4.9	4.6

[a] Values represent the standard error of fitting. Note that the K_d values in the submicromolar range are ill-defined due to the relatively high concentration (10 μM) of the dimeric NMII fragments.

three isoforms, NMIIA displays the inherently strongest S100A4-binding tailpiece. It is highly conceivable that filament disintegration starts with an interaction between S100A4 and the freely available tailpiece of NMIIA. The tailpiece can bind in several different conformations to form a fuzzy complex, but eventually it accommodates an optimal, higher affinity intermediate state, which facilitates efficient binding of the second helical-binding segment (residues Ala1907–Arg1923). As this segment participates in a coiled-coil structure, its binding to S100A4 requires coiled-coil melting.

This could happen essentially through two ways implying two interaction mechanisms. The coiled-coil could go through spontaneous transient unzipping exposing the second helical segment for S100A4, which points to a conformational selection mechanism. Alternatively, S100A4 could actively unzip the coiled-coil by using part of the binding energy, which would invoke an induced fit mechanism, i.e., a two-step binding. It is becoming widely accepted that these two mechanisms are not mutually exclusive [37]. Linear topology of the S100A4–NMII interaction suggests that there might be a smooth continuum between the two pure mechanisms. It is conceivable that sometimes the coiled-coil opens up spontaneously such that the entire S100A4-binding site becomes available, while other times only some parts become accessible and in these cases, the rest is opened up by S100A4 in an induced fit mechanism. At present, we have no direct evidence to decide between these pure mechanisms or to estimate their respective contribution to a mixture of the two. Nevertheless, as we found that monomer NMII fragments tend to bind S100A4 significantly faster than their respective dimer variants suggests that conformational selection plays an important role in the binding mechanism. Whatever mechanism leads to the binding of the second helix, based on comparative affinity values and MD simulations, we suggest that in NMIIA, this binding step in turn facilitates the folding of the first helical-binding segment (residues Glu1899–Ala1903). This way S100A4 gets even closer to the ACD, which could eventually lead to filament disassembly.

The proposed mechanism is congruent with all the quantitative data we measured and also with the DMD simulations; moreover, it also shows why NMIIA is preferred over NMIIB by S100A4. Firstly, the tailpiece of NMIIB is an inherently weaker binder than its NMIIA counterpart; secondly, Asn at 1914 (1907 in NMIIA numbering) in the second helical-binding segment provides about two orders of magnitude lower affinity than Ala; thirdly, the formation and fixation of the second long helical-binding region does not increase helical propensity in the first helical-binding segment as Asn breaks the allosteric

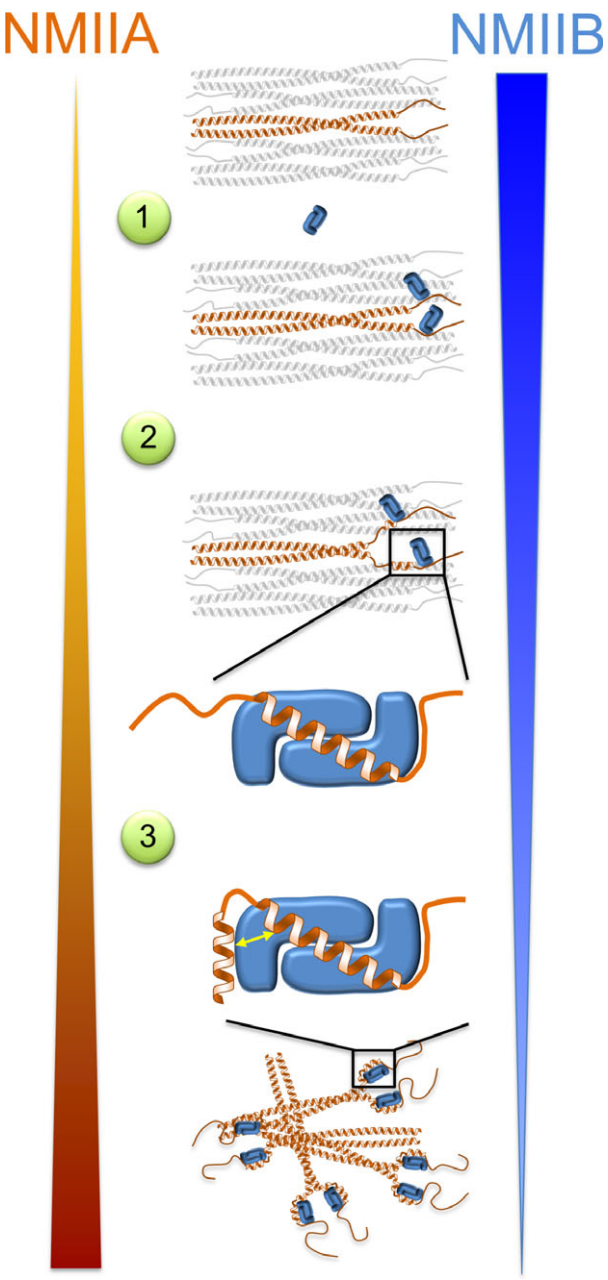


Fig. 9. Mechanism of S100A4 binding to NMIIA and NMIIB. The complex binding mechanism of NMII isoforms is going through several bidirectional steps, in which the two proteins NMIIA and NMIIB behave differently. The initiation of the interaction (1) depends on the hydrophobic nature of the disordered tailpiece. Binding of the helical motif to the Ca^{2+}-activated S100A4 dimer (blue) requires the partial dissociation of the myosin heavy chain dimer (gray and orange), thus the rate of this step is affected by the stability of the C-terminal coiled-coil (2). The lifetime of the S100A4–NMII complex is mainly determined by a single position at the N-terminal end of the second helical-binding segment (position 1907 in NMIIA) (2, 3). The formation of the fully bound form, i.e., the binding of the first helix is further enhanced by the allosteric communication (yellow arrow) between the preformed second and the nascent first helical-binding segments (3). In contrast to NMIIB, NMIIA has (a) a hydrophobic disordered binding segment; (b) moderate coiled-coil stability, (c) a favorable Ala in position 1907, and (d) an allosteric communication between the helical-binding segments. As a consequence, S100A4 prefers NMIIA and discriminates against NMIIB binding and filament dissociation.

communication between the structural elements of the binding motif; and fourthly, the coiled-coil of NMIIB is inherently more stable than that of NMIIA, which also hinders binding of this isoform by S100A4 (Fig. 9).

In the case of NMIIC, we have limited data to claim whether it is a relevant S100A4-binding partner or not. While *mCC* binds S100A4 with similar affinity than *mAA* and we could detect a similar allosteric communication between the binding helices, the less hydrophobic nature of its random coil-binding region and/or its phosphorylation might restrain S100A4 binding to NMIIC filaments *in vivo*. Based on the DMD simulations, the coiled-coil stability of *dCC* is more similar to *dAA* than to *dBB*; however, further *in vitro* experiments are needed for an extensive characterization of the NMIIC coiled-coil structure.

It is now well accepted that the molecular mechanism of an interaction is determined by and can be understood through the structure and conformational dynamics of the interacting partners. To this end, the long linear binding motif on the NMIIA tail domain appears to have a modular organization. This architecture can easily code for a directed translational transposition of binding energy along the NMII molecule gradually approaching key filament stabilizing positions. We are certainly aware that our mechanistic model is based mostly on indirect evidence, but we also note that it is consistent and is in great accordance with our and others' experimental findings [20,21,38]. Moreover, all the suggested steps can be tested by independent techniques to prove or confute our model. Nevertheless, the ultimate answer could be delivered only by direct detection of the individual binding steps.

It is also intriguing how the machinery behind the complex molecular event of filament disassembly evolved. The linear nature and proposed modular inner architecture of the S100A4-binding site on NMIIA appears to be an optimal format for efficient modular evolution. Only comprehensive bioinformatics studies that compare statistically relevant numbers of ortholog and paralog sequences could reveal whether the proposed segmental architecture is indeed detectable and if so, how and when the NMIIA function

emerged through evolution. Unfortunately, at the time of writing, the available sequences are too few and sparse to conduct such studies.

Finally, we discuss the possible functional significance of the isoform-specific S100A4 binding to NMII motors. It is conceivable that it represents one of the mechanisms by which cells differentially regulate isoforms-specific NMII functions, perhaps most importantly, the dynamic turnover and rearrangement of the bipolar myosin minifilaments. Two recent studies have shown that when the contractile system assembles, at first, the various NMII isoforms coassemble into heterotypic filament [2,3]. Later, however, in polarized cells, homotypic NMIIA and NMIIB filaments are generated due to an unidentified sorting mechanism [39]. We suggest that isoform-specific S100A4 binding could initiate such sorting mechanism as follows: induced by a local increase in Ca^{2+}, S100A4 can promote heterotypic filament disassembly and restrict NMIIA to reincorporate into new filaments, which eventually would result in homotypic NMIIB filaments. Importantly, tailpiece phosphorylation could be responsible for a second, and in a functional sense, reciprocal type of isoform-specific filament turnover mechanism. As discussed above, the phosphorylation pattern of the tailpiece is isoform-specific: certain kinases (e.g., PKC, CK2) could specifically phosphorylate NMIIB (or NMIIC) tailpiece [36]. In an analogous way to the S100A4/NMIIA pair, phosphorylation of NMIIB (and NMIIC) could lead to heterotypic filament disassembly and dynamic rearrangements, but in this case, it would result in homotypic NMIIA filaments. The validity of the above mechanisms remains to be tested experimentally.

We hypothesize that the need for NMII isoform-specific functions, such as different rates of filament turnover for NMIIA and NMIIB could have been an effective driving force during vertebrate evolution that resulted in paralog-specific conservation of: (a) the disordered tailpieces; (b) key residues (such as Ala/Asn in position 1907) in the helical-binding region, and (c) C-terminal coiled-coil stability.

Materials and methods

Production and purification of human S100A4 and NMII fragments

His-tagged human S100A4 (UniProt accession no. P26447) was expressed in *Escherichia coli* BL21 (DE3) cells, and purified by Ni^{2+}-affinity chromatography. After cleavage of the His_6-tag with tobacco etch virus (TEV) protease, samples were applied to a phenyl sepharose column, washed thoroughly and eluted with 5 mM EGTA-containing buffer. The purified S100A4 proteins were dialyzed against 20 mM Hepes pH 7.5, 20 mM NaCl, 0.1 mM tris(2-carboxyethyl)phosphine (TCEP) buffer, concentrated, pooled, and stored at −70 °C. For SPR measurements, the uncleaved His_6-S100A4 was applied to the phenyl sepharose column. For producing GST-tagged S100A4, the sequence encoding S100A4 was cloned into a pGEX4TI-based vector pETARA using *Bam*HI and *Xho*I sites. The construct was expressed in *E. coli* BL21 (DE3) cells. The cells were allowed to grow in LB media up to $OD_{600} = 0.8$ at 37 °C and induced for 4 h with 0.3 mM IPTG at 28 °C. The fusion protein was purified using glutathione sepharose 4B GST affinity resin (GE Healthcare, Little Chalfont, Buckinghamshire, UK) in 20 mM Hepes pH 7.5, 150 mM NaCl, 0.1 mM TCEP. After a thorough wash, GST-S100A4 was eluted by the washing buffer complemented with 10 mM reduced glutathione and then stored at −70 °C until further use.

Coding regions of NMIIA (UniProt accession no. P35579), NMIIB (UniProt accession no. P35580), and NMIIC (UniProt accession no. Q7Z406) heavy chain fragments were obtained from HEK cell mRNA by reverse transcriptase-based PCR. NMII point mutants and chimeras were generated by the megaprimer method. The monomeric NMII fragments (1894–1937 in NMIIA) were expressed with an N-terminal His_{10}-ubiquitin fusion in *E. coli* BL21 cells, and purified by Ni^{2+}-affinity chromatography. His-tagged ubiquitin was removed by yeast ubiquitin hydrolase, while it was dialyzed against buffer containing 20 mM Tris pH 8.0, 100 mM NaCl, and 0.2 mM DTT. The completely processed sample was applied to Ni^{2+}-affinity column. The peptide went into the flow-through fraction, and was finally purified by reverse-phase HPLC on a Jupiter 300 C18 column (Phenomenex, Torrance, CA, USA). The His_6-tagged dimeric NMII fragments (S1712-E1960 in NMIIIA) were expressed in *E. coli* BL21 cells, and purified by Ni^{2+}-affinity chromatography. The elution fraction was digested with TEV protease and dialyzed against 20 mM Hepes pH 7.5. The precipitated NMII fragments were sedimented by centrifugation, washed extensively with 20 mM Hepes pH 7.5 and dissolved in 20 mM Hepes pH 7.5, 1 M NaCl, and 1 mM TCEP and then pooled and stored at −70 °C until use.

Construction of NMII-phage library

The $NMIIA^{1894–1931}$-based peptide library is based on the pKS-Tag-SGCI-p8 phagemid vector, which was constructed from pBluescript II KS(-) (Stratagene, La Jolla, CA, USA), pMal-p2X (New England Biolabs, Ipswich, MA, USA), and the M13KO7 helper phage. The vector encodes a periplasmic signal; a FLAG epitope followed by a monovalently displayed NMIIA module and the p8 coat protein

(Fig. 1B). The FLAG-tag allows for assessing display bias. The library was produced in two successive mutagenesis steps [40]. First, pKS-Tag-NMIIA$^{1894–1931}$-p8 was used as the template to produce pNMIIA$^{1894–1931}$-STOP in which all codons to be scanned were replaced with stop codons (underlined) using the primer 5′CTGCAGCG CGAGCTGTAAGACTAAACTGAGTAATAATAATAA TAATAACGCGAAGTCTAA TAACTATAAAACTAAC TCAGGCGCGGG3′. Then pNMIIA$^{1894–1931}$-STOP was used as template for combinatorial mutagenesis. Stop codons were replaced with degenerate triplets encoding all amino acid types occurring in the three NMII isoforms, using the mutagenesis primer 5′CTGCAGCGCGAG CTGGAMGACGYTACTGAGNCTRMCGAMKSTMT GARCCGCGAAGTCWCCWCCCTAARGAACARGCT CAGGCGCGGG3′. The phagemid library was electroporated into *E. coli* SS320 to generate phage libraries [40].

Selection of NMII-phage library

MaxiSorp plates were coated with GST-S100A4 or anti-FLAG-tag antibody. The protein concentration was 30 μg·mL^{-1} for GST-S100A4 and 2 μg·mL^{-1} for the antibody. The phage library was incubated in the target protein-coated wells for 2 h at room temperature in 'binding buffer' (20 mM Tris pH 7.5, 150 mM NaCl, 0.1 mM TCEP, and 1 mM $CaCl_2$). Phages were eluted by the binding buffer complemented with 10 mM ethylenediamine-tetraacetic acid (EDTA). Two selection rounds were carried out on each target, and the binding properties of individual NMII-phage clones were tested by phage ELISA [40].

Sequence analysis

NMII-phage clones producing an ELISA signal on GST-S100A4 threefold above background (which was measured on BSA-containing wells) were sequenced by the Big Dye Terminator v3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA, USA). To eliminate the effects of display bias, GST-S100A4-selected amino acid frequencies were normalized to data from the anti-FLAG-tag selected NMII-phage population. A dataset of 100 sequences representing the normalized amino acid frequencies at each randomized position was generated and used as the input set for sequence logo generation by the WEBLOGO program [25].

Surface plasmon resonance

The SPR measurements were carried out on a ProteOn XPR36 protein interaction array system (Bio-Rad, Hercules, CA, USA) using HTG sensor chip. The running buffer contained 20 mM Hepes pH 7.5, 300 mM NaCl (or 500 mM in the case of dimeric NMII fragments), 1 mM $CaCl_2$, 0.1 mM TCEP, and 0.005% Tween-20. The immobilization of His_6-S100A4 resulted in 300–600 RU ligand density, which was highly stable as a consequence of the avid binding of S100A4 derived from the Tris-nitrilotriacetic acid (NTA) functional groups and the homodimer structure of the protein. Analytes were injected into the chip at five different concentrations simultaneously at a flow rate of 60 μL·min^{-1}. In the sixth analyte channel, running buffer was injected for double referencing. After each analyte injection, the surface was regenerated with 500 μM EGTA-containing buffer, which did not strip off the Ni^{2+}, but caused the rapid and effective analyte dissociation from S100A4. The double referenced data were global fitted to the 1 : 1 Langmuir binding model.

Circular dichroism spectroscopy

Circular dichroism (CD) measurements were carried out on a Jasco J-715 spectropolarimeter (Jasco, Easton, MD, USA). Thermostability studies were made in a buffer containing 10 mM Hepes pH 7.5, 500 mM NaCl, 0.1 mM TCEP with a temperature ramp of 1 °C·min^{-1}. To calculate the melting temperatures, experimental data were fitted to the Boltzmann sigmoid equation. The ellipticity of 5 μM *dAA* and *dBB* fragments was monitored at 222 nm with 0.1 cm path length.

Filament disassembly assay

About 10 μM NMII coiled-coil fragments (monomeric concentration) was titrated with S100A4 in triplicates in 20 mM Hepes pH 7.5, 100 mM NaCl, 1 mM $CaCl_2$, 2 mM $MgCl_2$, and 0.1 mM TCEP-containing buffer. Turbidity was measured at 320 nm using 384-well transparent microplates (Greiner #781162) with Synergy H4 multi-mode microplate reader (BioTek, Winooski, VT, USA). The titration data were fitted to the quadratic binding equation.

Discrete molecular dynamics combined with replica exchange method (RX-DMD)

To examine the structural dynamics of monomeric NMII fragments, RX-DMD simulations were started with fully extended or partially folded (constrained second helical sequence in NMIIA: MNREVSSLKNKLRR, and the corresponding region in NMIIB and NMIIC) peptides in 12 different temperature boxes (from 0.5252 to 0.6749). The production run was 1 million steps long (~2 ns) and the exchange of replicas was permitted every 1000 steps. About 5000 structures were extracted from every simulations (from every 200 steps), and were separated according to the temperature box they were in. We used DSSP program [41] to map helix propensity within the peptide and counted the

frequency of being in an α-helical structure for every residues. To filter out frequency bias caused by the initial extended structure, we omitted the first 500 structures (100 000 steps) from the DSSP analysis. In coiled-coil stability experiments, we first used CCBuilder [42] to build up stable dimeric coiled-coils from segment 1706–1923 in NMIIA and from the homologous sequence of NMIIB and NMIIC. The completely folded dimers were put into a production run in RX-DMD in eight different temperature boxes (from 0.5497 to 0.6448) for 1 million steps with the same exchange permissions as before. After creating 5000 structures from each replica, average CA distances were calculated for every identical residue pairs on the two chains, and were plotted using a seven-residue wide sliding window (to smooth waves caused by the helical structure).

Acknowledgements

We thank Mr. Gergő Gógl for critical reading of the manuscript. This work was supported by the Hungarian Scientific Research Fund (OTKA) grants K108437 and NK100769 to LN and GP, respectively, as well as by the National Development Agency Grant KMOP-4.2.1/B-10-2011 and the MedInProt program of the Hungarian Academy of Sciences.

Author contributions

BK designed and performed the experimental studies and analyzed their data. LK designed and performed the *in silico* studies and analyzed their data. LN oversaw the research and contributed to writing. GP designed the directed evolution studies and contributed to writing. All authors reviewed the results and approved the final version of the manuscript.

References

1 Vicente-Manzanares M, Ma X, Adelstein RS & Horwitz AR (2009) Non-muscle myosin II takes centre stage in cell adhesion and migration. *Nat Rev Mol Cell Biol* **10**, 778–790.
2 Beach JR, Shao L, Remmert K, Li D, Betzig E & Hammer JA III (2014) Nonmuscle myosin II isoforms coassemble in living cells. *Curr Biol* **24**, 1160–1166.
3 Shutova MS, Spessott WA, Giraudo CG & Svitkina T (2014) Endogenous species of mammalian nonmuscle myosin IIA and IIB include activated monomers and heteropolymers. *Curr Biol* **24**, 1958–1968.
4 Wang A, Ma X, Conti MA & Adelstein RS (2011) Distinct and redundant roles of the non-muscle myosin II isoforms and functional domains. *Biochem Soc Trans* **39**, 1131–1135.
5 Kovacs M, Wang F, Hu A, Zhang Y & Sellers JR (2003) Functional divergence of human cytoplasmic myosin II: kinetic characterization of the non-muscle IIA isoform. *J Biol Chem* **278**, 38132–38140.
6 Wang F, Kovacs M, Hu A, Limouze J, Harvey EV & Sellers JR (2003) Kinetic mechanism of non-muscle myosin IIB: functional adaptations for tension generation and maintenance. *J Biol Chem* **278**, 27439–27448.
7 Yuen SL, Ogut O & Brozovich FV (2009) Nonmuscle myosin is regulated during smooth muscle contraction. *Am J Physiol Heart Circ Physiol* **297**, H191–H199.
8 Ma X, Kovacs M, Conti MA, Wang A, Zhang Y, Sellers JR & Adelstein RS (2012) Nonmuscle myosin II exerts tension but does not translocate actin in vertebrate cytokinesis. *Proc Natl Acad Sci USA* **109**, 4509–4514.
9 Ronen D & Ravid S (2009) Myosin II tailpiece determines its paracrystal structure, filament assembly properties, and cellular localization. *J Biol Chem* **284**, 24948–24957.
10 Sandquist JC & Means AR (2008) The C-terminal tail region of nonmuscle myosin II directs isoform-specific distribution in migrating cells. *Mol Biol Cell* **19**, 5156–5167.
11 Vicente-Manzanares M, Koach MA, Whitmore L, Lamers ML & Horwitz AF (2008) Segregation and activation of myosin IIB creates a rear in migrating cells. *J Cell Biol* **183**, 543–554.
12 Murakami N, Chauhan VP & Elzinga M (1998) Two nonmuscle myosin II heavy chain isoforms expressed in rabbit brains: filament forming properties, the effects of phosphorylation by protein kinase C and casein kinase II, and location of the phosphorylation sites. *Biochemistry* **37**, 1989–2003.
13 Rosenberg M & Ravid S (2006) Protein kinase Cgamma regulates myosin IIB phosphorylation, cellular localization, and filament assembly. *Mol Biol Cell* **17**, 1364–1374.
14 Ford HL, Silver DL, Kachar B, Sellers JR & Zain SB (1997) Effect of Mts1 on the structure and activity of nonmuscle myosin II. *Biochemistry* **36**, 16321–16327.
15 Li ZH, Spektor A, Varlamova O & Bresnick AR (2003) Mts1 regulates the assembly of nonmuscle myosin-IIA. *Biochemistry* **42**, 14258–14266.
16 Li ZH & Bresnick AR (2006) The S100A4 metastasis factor regulates cellular motility via a direct interaction with myosin-IIA. *Cancer Res* **66**, 5173–5180.
17 Kim EJ & Helfman DM (2003) Characterization of the metastasis-associated protein, S100A4. Roles of calcium binding and dimerization in cellular localization and interaction with myosin. *J Biol Chem* **278**, 30063–30073.
18 Schneider M, Hansen JL & Sheikh SP (2008) S100A4: a common mediator of epithelial-mesenchymal transition, fibrosis and regeneration in diseases? *J Mol Med (Berl)* **86**, 507–522.

19 Boye K & Maelandsmo GM (2010) S100A4 and metastasis: a small actor playing many roles. *Am J Pathol* **176**, 528–535.
20 Kiss B, Duelli A, Radnai L, Kekesi KA, Katona G & Nyitray L (2012) Crystal structure of the S100A4-nonmuscle myosin IIA tail fragment complex reveals an asymmetric target binding mechanism. *Proc Natl Acad Sci USA* **109**, 6048–6053.
21 Elliott PR, Irvine AF, Jung HS, Tozawa K, Pastok MW, Picone R, Badyal SK, Basran J, Rudland PS, Barraclough R *et al*. (2012) Asymmetric mode of Ca(2)(+)-S100A4 interaction with nonmuscle myosin IIA generates nanomolar affinity required for filament remodeling. *Structure* **20**, 654–666.
22 Duelli A, Kiss B, Lundholm I, Bodor A, Petoukhov MV, Svergun DI, Nyitray L & Katona G (2014) The C-terminal random coil region tunes the Ca(2)(+)-binding affinity of S100A4 through conformational activation. *PLoS One* **9**, e97654.
23 Fuxreiter M & Tompa P (2012) Fuzzy complexes: a more stochastic view of protein function. *Adv Exp Med Biol* **725**, 1–14.
24 Szenthe B, Patthy A, Gaspari Z, Kekesi AK, Graf L & Pal G (2007) When the surface tells what lies beneath: combinatorial phage-display mutagenesis reveals complex networks of surface-core interactions in the pacifastin protease inhibitor family. *J Mol Biol* **370**, 63–79.
25 Crooks GE, Hon G, Chandonia JM & Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* **14**, 1188–1190.
26 Weiss GA, Watanabe CK, Zhong A, Goddard A & Sidhu SS (2000) Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proc Natl Acad Sci USA* **97**, 8950–8954.
27 Pal G, Fong SY, Kossiakoff AA & Sidhu SS (2005) Alternative views of functional protein binding epitopes obtained by combinatorial shotgun scanning mutagenesis. *Protein Sci* **14**, 2405–2413.
28 Pal G, Kossiakoff AA & Sidhu SS (2003) The functional binding epitope of a high affinity variant of human growth hormone mapped by shotgun alanine-scanning mutagenesis: insights into the mechanisms responsible for improved affinity. *J Mol Biol* **332**, 195–204.
29 Pal G, Kouadio JL, Artis DR, Kossiakoff AA & Sidhu SS (2006) Comprehensive and quantitative mapping of energy landscapes for protein-protein interactions by rapid combinatorial scanning. *J Biol Chem* **281**, 22378–22385.
30 Pal G, Ultsch MH, Clark KP, Currell B, Kossiakoff AA & Sidhu SS (2005) Intramolecular cooperativity in a protein binding site assessed by combinatorial shotgun scanning mutagenesis. *J Mol Biol* **347**, 489–494.
31 Rapali P, Radnai L, Suveges D, Harmat V, Tolgyesi F, Wahlgren WY, Katona G, Nyitray L & 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 One* **6**, e18818.
32 Shirvanyants D, Ding F, Tsao D, Ramachandran S & Dokholyan NV (2012) Discrete molecular dynamics: an efficient and versatile simulation method for fine protein characterization. *J Phys Chem B* **116**, 8375–8382.
33 Szollosi D, Horvath T, Han KH, Dokholyan NV, Tompa P, Kalmar L & Hegedus T (2014) Discrete molecular dynamics can predict helical prestructured motifs in disordered proteins. *PLoS One* **9**, e95795.
34 Tompa P, Schad E, Tantos A & Kalmar L (2015) Intrinsically disordered proteins: emerging interaction specialists. *Curr Opin Struct Biol* **35**, 49–59.
35 Van Roey K, Uyar B, Weatheritt RJ, Dinkel H, Seiler M, Budd A, Gibson TJ & Davey NE (2014) Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation. *Chem Rev* **114**, 6733–6778.
36 Dulyaninova NG & Bresnick AR (2013) The heavy chain has its day: regulation of myosin-II assembly. *Bioarchitecture* **3**, 77–85.
37 Csermely P, Palotai R & Nussinov R (2010) Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends Biochem Sci* **35**, 539–546.
38 Badyal SK, Basran J, Bhanji N, Kim JH, Chavda AP, Jung HS, Craig R, Elliott PR, Irvine AF, Barsukov IL *et al*. (2011) Mechanism of the Ca(2)+-dependent interaction between S100A4 and tail fragments of nonmuscle myosin heavy chain IIA. *J Mol Biol* **405**, 1004–1026.
39 Beach JR & Hammer JA III (2015) Myosin II isoform co-assembly and differential regulation in mammalian systems. *Exp Cell Res* **334**, 2–9.
40 Sidhu SS, Lowman HB, Cunningham BC & Wells JA (2000) Phage display for selection of novel binding peptides. *Methods Enzymol* **328**, 333–363.
41 Kabsch W & Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637.
42 Wood CW, Bruning M, Ibarra AA, Bartlett GJ, Thomson AR, Sessions RB, Brady RL & Woolfson DN (2014) CCBuilder: an interactive web-based tool for building, designing and assessing coiled-coil protein assemblies. *Bioinformatics* **30**, 3029–3035.