Conformational plasticity and evolutionary analysis of the myotilin tandem Ig domains

Vid Puž1, Miha Pavšič1, Brigita Lenarcič2,2 & Kristina Djinović-Carugo1,3

Myotilin is a component of the sarcomere where it plays an important role in organisation and maintenance of Z-disk integrity. This involves direct binding to F-actin and filamin C, a function mediated by its Ig domain pair. While the structures of these two individual domains are known, information about their relative orientation and flexibility remains limited. We set out to characterise the Ig domain pair of myotilin with emphasis on its molecular structure, dynamics and phylogeny. First, sequence conservation analysis of myotilin shed light on the molecular basis of myotilinopathies and revealed several motifs in Ig domains found also in I-band proteins. In particular, a highly conserved Glu344 mapping to Ig domain linker, was identified as a critical component of the inter-domain hinge mechanism. Next, SAXS and molecular dynamics revealed that Ig domain pair exists as a multi-conformation species with dynamic exchange between extended and compact orientations. Mutation of AKE motif to AAA further confirmed its impact on inter-domain flexibility. We hypothesise that the conformational plasticity of the Ig domain pair in its unbound form is part of the binding partner recognition mechanism.

Movement is vital to all living organisms, from the transport of single molecules in cells to the movement of the entire organism. In vertebrates, skeletal muscle is essential for all voluntary movements such as walking, running, swimming or flying. Involuntary movements, such as those occurring in cardiac muscle resulting in contraction of the heart (beating), or those occurring in smooth muscle resulting in peristalsis are fundamental to the viability of the organism.

Muscles are composed of repetitive isotropic (I) and anisotropic (A) bands, giving them a striped appearance. Here, the smallest periodic unit, termed sarcomere, is composed of actin and myosin filaments which together with other accessory proteins form the sarcomeric cytoskeleton1,2. The precise ultra-structural order of these filaments is of utmost importance for converting the molecular interactions produced by actin and myosin in each sarcomere into efficient contraction at the macroscopic level. Titin and nebulin (and its smaller cardiac variant nebulette) act as a general blueprint for sarcomere assembly and length, due to their outstanding size and modular architecture. Sarcomere is on both ends delimited with Z-discs, which act as passive transducers of the generated tension. This is also the spot of actin filament crosslinking, primarily by the α-actinin, however other signalling, stretch-sensing and actin-binding proteins are involved, for example myotilin, filamin C, ZASP/cypher/oracle, FATZ/myozenin/calsarcin and many others3.

Myotilin, the founding member of the myotilin/myopalladin/palladin family was initially identified and characterised as α-actinin binding protein4. Both myopalladin and myotilin have restricted expression profile, with the highest expression in the skeletal and cardiac muscle4,5, in contrast to palladin, which shows much wider expression pattern both in skeletal and non-skeletal tissues, depending on the expressed isoform6,7. Myotilin is comprised of two C2-type Ig domains and a unique serine-rich N-terminal region4, while palladin and myopalladin host up to five Ig domains, where the two most C-terminal domains share high homology with the Ig domains of myotilin and two most N-terminal domains are mostly related to the N2B region of titin8 (Fig. 1a). Ig-like domains are required for the myotilin dimerisation4,9 and interaction with other binding partners such as G-
F-actin\textsuperscript{8,9}, filamin C\textsuperscript{10} and FATZ\textsuperscript{11}. With the motif\textsuperscript{494ESEEL} at the very C-terminal tail of the molecule myotilin is able to bind to the PDZ domain of protein ZASP, similarly to the C-terminal motif of α-actinin and FATZ family of proteins\textsuperscript{12}. At the N-terminal part, encompassing residues 89–105, myotilin binds to the calmodulin-like domain of α-actinin-2 (ref. 13). More specifically, via canonical motif 1-4-5-8 (numbers correspond to sequence positions of hydrophobic amino acid residues) myotilin is able to interact with EF3-4 hands of α-actinin, in a
similar way as it was shown for titin Z-repeats14–17, palladin13, 18 and more recently for neck peptide of α-actinin-2 in its auto-inhibited form19.

Myotilin is thought to play a role in anchoring and stabilising actin filaments at the Z-disc, and is involved in the organisation and maintenance of Z-disk integrity22. Its involvement in actin crosslinking, bundling in concert with α-actinin-2 and its interactions with important components of Z-disc, suggest that it plays a role in structural stabilisation and assembly of the Z-disc22. During the early stages of myofibrillogenesis, myotilin is present in the pre-myofibrils (Z-bodies) together with actin, α-actinin-2, filamin C, FATZ and ZASP21. Recent results reported that it is highly mobile not only in the Z-bodies but also in the Z-bands and possesses ability of dynamic exchange with the cytoplasmic pool of proteins, indicating important role in initial organisation and maintenance of Z-disc22.

The N-terminal part of molecule, which is predicted to be intrinsically unstructured (Fig. 1b) also represents a «mutational hotspot», where missense mutations (all but two residing between residues 29 and 124) have been shown to cause muscle disorders such as limb-girdle muscular dystrophy 1A, myofibrillar myopathy, and spheroid body myopathy20, 23–26 (Fig. 1a), collectively termed myotilinopathies24, 27. Missense disease mutations in myotilin were shown to have no impact on the localisation and dynamics in the Z-disc, however degradation rates were slower for the mutant forms of myotilin, compared to the wild-type28. It is important to note that members of the family may have overlapping functions, such as binding to the same partner and can probably compensate for each other, as it was shown with the myo V′′ mice, which did not show any abnormalities or differences, compared to the normal mice29. Although several properties of myotilin as described above propose its role in the Z-disc, its primary function still remains unknown.

It is well established that function of several sarcomeric proteins critically depends on the flexibility/rigidity of their tandem Ig domain repeats. For example, flexibility of tandem Ig repeats in filamin directly affects its interactions with other proteins30, 31. In titin, inter-domain flexibility is essential for its mechanical functions and stability32–38. More specifically, titin Z1Z2 tandem domains, localized at the Z-disc, possess an adaptable dynamics for recruitment of the binding partner telethonin31. However, no such data exists for the myotilin tandem Ig pair. Since myotilin is a member of tandem Ig domain containing proteins, homologous to titin, it can be expected that the inter-domain flexibility could have an impact on its interaction with the diverse proteins and, in turn, the mechanical properties of the sarcomeric cytoskeleton. To gain insight into its dynamic properties we used several complementary computational and experimental approaches. First, we applied molecular evolution approach to analyse myotilin conservation across homologous sequences with the aim to identify region of flexibility. Next, we examined structural properties and dynamics of myotilin Ig domains in solution by using SAXS measurements and molecular dynamics simulations.

Results
Evolutionary analysis of myotilin. Molecular evolution analysis has already proved to be a valuable tool for identification of evolutionary relationship between Ig domains of different sarcomeric proteins39–41. To gain insight into myotilin amino acid sequence conservation we assembled a dataset of 85 homologous sequences of Vertebrates subphyla ranging from Chondrichthyes to Mammalia classes and constructed a multiple sequence alignment. Myotilin homologues of the evolutionarily distant species of the Chordata phylum could not be identified here; the highest sequence similarity was observed for Ig domains of titin and titin-like proteins, which were not included in the analysis. For comparison, a sequence alignment using only one representative species of each class (i.e. excluding redundant sequences from Mammalia and Aves classes) was constructed and showed similar results. The Bio-NJ tree calculated using maximum likelihood approach showed the expected evolutionary relationship between species with a good branch support for main branches indicating a reliable sequence alignment (Supplementary Fig. S1). Interestingly, detailed inspection of the alignment showed insertion of 11 residues at the N-terminal part in the sequences from the Aves class which was not observed in the other analysed sequences (Supplementary Data S1).

Per residue conservation scores clearly demonstrate that Ig domains 1 and 2, together with the short region (residues 160–190), represent the most conserved part of the molecule (Fig. 1c), which is in line with their functional role in dimerisation and interactions with other proteins. As expected, mapping of conservation score to the Ig domains’ structure revealed the highest conservation for residues of the hydrophobic core, specifically the tryptophan residues in the centre of each of the domains (W283 in Ig1 and W382 in Ig2), which are critical for the stability of the immunoglobulin fold and form together with other hydrophobic residues a folding nucleus42.

On the other hand, surface residues are less conserved with the exception of three important motifs: PPxf, NxxG and PxP. The PPxf motif (x = any residue, f = a hydrophobic residue, most often Phe) is located at the N-terminus of both Ig domains and has already been found in Ig domains of other proteins from the I-band43. Interestingly, conserved proline residues, first identified at the border of fibronectin type III domains of tenascin, were shown to prevent non-specific self-association and aggregation of their cognate proteins44. We speculate that similar effects could be also present in myotilin. The NxxG motif (x = any residue) present at the β-hairpin turn connecting the F and G strands, is a central part of the N-terminal network of conserved residues. Similarly to the PPxf motif, it is found in the N-conserved type of I-band Ig domains44. Together with the PxP motif, located at the BC loop on both domains, these motifs form a conserved N-terminal network of residues (Fig. 1d). Overall, these results demonstrate that throughout the course of evolution certain structurally and/or functionally important features of Ig domains were retained while other parts, particularly solvent-exposed regions, are more varied to accommodate different binding partners.

Next, inspection of alignment of the linker region within tandem Ig domain repeat of myotilin sequence (141AKEHKR346) and corresponding repeats in myopalladin (1185AKEVKK1170) and palladin (1227AKEAHK1232) revealed that Ala, Lys and Gln residues are highly conserved within this region with an overall basic character. Interestingly, highly conserved Glu is also characteristic for I-band linker regions where it sometimes merges with
PPIX to form the EPPxF motif. Mutational studies of this motif within the titin Ig65–70 domain repeats suggested that it is not a determinant of inter-domain orientation33, however latest studies of tandem titin Ig67–68 repeats indicate that it is important for titin global dynamics38. While myotilin lacks the combined EPPxF motif, the PPIX motif could nevertheless be involved in inter-domain dynamics. Even more, similar inter-domain dynamics are expected for domain pairs Ig1–2 (myotilin) and Ig4–Ig5 (myopalladin/palladin) due to their high sequence similarity and linker length. Linker regions within other Ig domain pairs of myopalladin/palladin are longer (76 and 44 residues for Ig1–2 and Ig3–4, respectively) indicating a higher degree of flexibility.

In contrast to the two Ig domains of myotilin, the N-terminal (residues 1–250) and the C-terminal part (residues 441–498) display lower conservation scores. While those two regions contain functionally important sites (binding sites for α-actinin-2 on N-terminus, and ZASP binding site at the very C-terminus (Fig. 1a), detailed structural analysis indicates that they are intrinsically disordered (Fig. 1b). Link between intrinsic disorder and higher sequence variation is in line with the observation that disordered regions typically evolve faster (i.e., have higher mutation rate) than the structured domains45. This was also observed for myopalladin and palladin (Supplementary Fig. S1) where unstructured parts are mainly located between the Ig domains. Still, the structure of such regions could be induced/stabilised upon binding of another protein. A representative example is the formation of an α-helical structure at the otherwise unstructured N-terminal of palladin upon binding to EF34 hands of α-actinin-2 (ref. 13).

### Analysis of disease-causing mutations in myotilin.

Several mutations within the myotilin gene, mostly mapping to the intrinsically disordered N-terminal part, result in Z-disc alterations and polymorphous aggregate formation, which is associated with different myopathies20, 23–26. Therefore, we performed detailed inspection of the disease-linked mutations to gain an insight into their possible common denominator. Conservation scores of the mutation hotspots showed no significant differences compared to the other regions; average conservation score for all mutations was 0.673, where 0 present fully variable and 1 fully conserved score. Majority of the mutations are substitutions from polar/charged residues to a hydrophobic residue with the exception of K36E, Q74K and R405K (Fig. 1a). The newly formed hydrophobic clusters could together with the nearby hydrophobic rich region (HRR, Fig. 1a) of the molecule promote self-association and formation of aggregates. Since almost all mutations are located within the intrinsically disordered region we speculate that mutations could positively affect disorder-to-order transition and thereby enabling further progression of aggregation. To test this hypothesis, we compared the disorder tendency plots for wt myotilin and the disease-related mutants (Supplementary Fig. S2). Mutations resulting in introduction of hydrophobic residues displayed reduced disorder while other mutations (K36E, Q74K and R405K) did not show significant differences. Similar effect of mutations on the disorder-to-order transition was also predicted for the large dataset of disordered proteins, and more specifically for the mutation R243W of the disordered region of p63 (ref. 46). Moreover, in γ-C-crystallin the mutation R168W, directly linked to congenital cataracts, increased aggregation and precipitation propensity, possibly due to the increased exposure of a hydrophobic patch47.

### Ig domain pair displays conformational plasticity.

SAXS analysis of tandem Ig domains. To examine structural properties of the tandem Ig domains of myotilin in solution, small-angle X-ray scattering (SAXS) was employed. SAXS data of MYOTIg1–2 construct encompassing the two Ig domains was measured using highly purified samples of myotilin Ig domains at different concentrations. Slightly increased radius of gyration ($R_g$) values. Even more, when compared to bovine serum albumin (BSA)49, a compact globular protein, myotilin shows a less globular shape which is indicated by Kratky plot curve broadening and the maxima shifted to higher $s$ × $R_g$ values. Even more, when compared

| **Name**   | $I_0$ (Guinier) [a.u.] | $I_0$ (P(r)) [a.u.] | $R_g$ (Guinier) [nm] | $R_g$ (P(r)) [nm] | $D_{max}$ [nm] | V [Porod] [kDa] | $M_r$ (calc.) [kDa] | $M_r$ (Porod) [kDa] |
|------------|------------------------|----------------------|----------------------|-------------------|----------------|-----------------|-------------------|-------------------|
| MYOT Ig1-2 | 15.34 ± 0.06           | 15.37                | 2.57 ± 0.10          | 2.63              | 9.37           | 29.14           | 22.3              | 16.3              | 16.4              |

Table 1. Parameters derived from the SAXS data for MYOTIg1–2.
to the Ig domain pair (Ig20–21) of filamin A, both in its compact and slightly more extended/flexible migfilin peptide-bound form, myotilin Ig domain pair is significantly more flexible, similarly as Ig domain pairs Ig12–13 and Ig22–23 of filamin A (Fig. 2b). Here, increased flexibility is correlated to the extended orientation of the tandem Ig domains with less frequent/stable inter-domain contacts. Therefore, in MYOTIg1–2 being the most flexible of the shown examples, the extended SAXS envelope corresponds to a nearly coaxial domain orientation on average where the two Ig domains do not form stable contacts.

To gain detailed insight into the relative orientation and dynamics of the two domains of the Ig1–2 pair, both rigid body and flexible/ensemble fitting approaches were used and assessed with regard to $\chi^2$ and inter-domain clashes. For rigid body fitting, the $ab$ initio three-dimensional low-resolution molecular envelope was reconstructed using DAMMIF. Next, the most probable averaged model was calculated with the DAMAVER set of programs, showing a “dumbbell” shape with dimensions of approximately 10 × 4.5 nm. Subsequently, rigid body modelling using CORAL was performed, showing relatively good chi value ($\chi^2 = 2.957$), however with a slight mismatch at the higher angles (Fig. 2c). The superposition of rigid body model with the $ab$ initio reconstructed molecular envelope of MYOTIg1–2 is presented in the inset.

Based on the Kratky plot analysis, we further explored flexibility of the tandem Ig domains using ensemble fitting. For ensemble fitting, ensemble optimisation method (EOM) and MultiFoXS approaches were employed to fit the theoretical scattering intensities (calculated from a pool of models with different inter-domain orientations) to the experimental scattering data. EOM results indicate various inter-domain orientations which are represented by four models ranging from more compact to a more extended arrangement, fitting the experimental scattering with a low chi value ($\chi^2 = 1.481$). Radius of gyration distribution plot shows three peaks with maxima ranging from 2.3 nm (compact), 2.7 nm (average extended), and 3.0 nm representing fully extended

Figure 2. SAXS analysis of myotilin Ig domains. (a) $P(r)$ vs. $r$ plot for MYOTIg1-2, showing bimodal profile, with the $D_{max} = 9.37$ nm. Curve is showing smooth decay to $P(r)=0$ at the $D_{max}$ and good fit to the experimental data presented in the inset, using GNOM program ($s = 4\pi\sin(\theta)/\lambda$). (b) Comparison of dimensionless Kratky plots of MYOTIg1-2 with different proteins as an examples of globular protein (BSA), tandem Ig domains of filamin A with compact conformation (Ig20–21), and the same construct with the addition of migfilin peptide (Ig20–21 + migfilin), which increases flexibility of the Ig domains. MYOTIg1–2 shows less compact conformation with the degree of flexibility as indicated by Kratky plot broadening and maxima shift and is similar to Ig22–23 of filamin A. Molecular envelopes of the analysed proteins, with the available high-resolution structures are presented above graph. (c) Fit of the calculated rigid body model by CORAL to experimental scattering with goodness of the fit expressed in $\chi^2$. Superposition of the rigid body model with the averaged $ab$ initio reconstructed molecular envelope of MYOTIg1–2 is presented in the inset.
species (Supplementary Fig. S4). Similarly, end-to-end distance and $D_{\text{max}}$ distribution plots show bimodal profiles, thereby additionally confirming flexibility of the Ig1–2 system.

However, EOM-produced models using separate Ig domains for fitting have some clashes between Ig domains and the dummy atoms of the linker. To overcome this problem, we prepared a model of Ig1–2 together with the linker (I-TASSER) and employed MultiFoXS server to explore the multiple conformations by conformational sampling with rapidly exploring random trees (RRT) search. Here, the linker residues were assigned as flexible. This significantly improved the fitting to the experimental data for 4-state model, showing lowered chi value ($\chi^2 = 1.120$) in comparison with the rigid body modelling (for comparison see Fig. 2c and Fig. 3a, inset). Radius of gyration distribution plot for different best scoring N-state models revealed a trimodal $R_g$ peak distribution with maxima at 2.2, 2.6 and 2.8 nm (Fig. 3a), which is in excellent agreement with the data calculated by the EOM method. Graphical representation of the calculated multi-state model confirmed the presence of the different populations of tandem Ig domain arrangements (Fig. 3b). A similar distribution was observed for the semi-extended Ig domain pair 6–7 of filamin C49. Taken together, these results reveal that tandem Ig domains of myotilin display a certain degree of flexibility and exist in different relative orientations, most likely due to the flexible nature of the linker between the two domains.

**Molecular modelling and dynamics.** In order to confirm flexibility and dynamics of Ig domains of myotilin shown by SAXS measurements, we employed molecular dynamics simulations. First, we produced models of tandem Ig domains of myotilin using I-TASSER, with the Z-score as a selection criteria. Comparatively, best model generated in the Swiss-model repository showed almost identical structure with the RMSD value of 1.16 Å between 191 equivalent Cα atoms and almost coaxial orientation. For the molecular dynamics simulations the best model generated by I-TASSER was used.

Residues with highest degree of flexibility were predicted using a coarse-grained dynamics approach implemented in CABS-Flex server. Here, high root mean square fluctuation (RMSF) value is indicative of high residue flexibility. For several other proteins, results of such simulations have already been proved to be in good agreement with NMR measurements and all-atom MD simulations aimed to determine flexible regions52. The residue fluctuation profile of Ig1–2 domain pair of myotilin shows that highest flexibility is associated with the linker region between Ig1 and Ig2 while residues of the Ig domains’ hydrophobic core have far lower RMSF values and are hence less flexible (Fig. 4a). Fluctuations in N-to-C-terminus distance of Ig1–2 domain pair during the coarse-grained simulation indicate a transition from fully-extended (starting model) to a more compact domain arrangement.

Detailed analysis of fluctuations in domain arrangement was analysed using two 10 ns all-atom MD simulations. Here, the Ig1–2 domain pair moved from an extended to a more compact arrangement, followed again by a transition to regain a locally more extended form. In case of simulation 1, we observed almost complete regain to extended orientation (Fig. 4b), whereas simulation 2 showed significantly more compact domain orientation at the end of simulation. More extended orientation and limited flexibility in simulation 1 could originate from the observed formation of short α-helix of linker residues. These residues showed preference for the α-helical structure, as was corroborated by structure prediction with I-TASSER and PEP-FOLD3, showing helical structure of all produced models.

During simulation 2 the maximum relative rotation angle between the domains reached a value of 104° with the domain closure of 54.5 % as analysed by the DynDom server (here, values of 0 and 100 % correspond to fully extended and fully compact forms, respectively). Further analysis of the extended and compact conformations from the all-atom MD simulations revealed that the residue stretch $^{345}$EHKR$^{347}$ represents a hinge, where the
main chain bond rotation results in the domain torsion (Fig. 4c). This hinge region is also involved in the formation of the α-helical region observed during molecular dynamics simulations. Interestingly, the Glu residue within this stretch is also conserved in the linker regions of tandem domains of titin where it was suggested to be involved in intra- and inter-domain stabilisation, however no structural data is available to unequivocally confirm this hypothesis. Our data demonstrate that the corresponding Glu residue of myotilin (Glu344), in concert with other hinge residues, plays an important role in the hinge mechanism of myotilin, where its main dihedral \( \psi \) angle changes by 54° as measured by the DynDom server. Moreover, formation of intra-domain interactions via hydrogen bonds was observed between Lys343 and Ile375 and between Pro376/Ala429 and Glu344 (Fig. 4c). Remarkably, Pro376 and Ala429, located within the Ig2 domain, are part of the highly conserved motifs PxP and NxxG. Additional MD simulations were performed with the in silico mutated AKE motif to AAA, where we observed increased flexibility and less extended structure of the mutant, suggesting that the AKE motif is important for the stabilisation of the inter-domain orientation. Additionally, distribution of end-to-end distances differs for wt and mutant (Fig. 4d). Distribution of mutant is significantly broader (median = 80.2 Å, SD = 5.5 Å), compared to the wt (median = 84.6 Å, SD = 3.2 Å).

Therefore, our data clearly demonstrates a connection between conserved motifs and myotilin dynamics on a level of intra-domain contacts involving residues within the conserved motifs.
Discussion

To date, great effort has been put in studies of mechanical and structural properties of sarcomeric proteins located within the I-band and Z-disk, particularly with regard to their single and poly-tandem Ig domain regions which are known to be able to adopt different relative orientations and in turn affect the flexibility and elasticity of their cognate proteins. In particular, the main focus of the research has been elastic I-band region of titin and inter-domain interactions of filamins. Results have shown essential impact of titin domain unfolding, stability and inter-domain mobility on mechanical properties directly correlated to muscle contraction and elasticity\(^{39}\). Studies of the titin N-terminal Z1Z2, located within the Z-disc\(^{34}\), have revealed important implications of inter-domain dynamics in unbound state for interaction with binding partner telethonin\(^{35}\). Also, compact domain-domain interactions of filamins\(^{30, 31, 49, 54}\), as well as inter-molecular interaction between lg24 domains leading to dimerisation\(^{35, 58}\), were found to be functionally important. For example, relative orientation and inter-domain contacts of filamin lg domains 16–24 (Rod 2 region) play an important role in ligand binding regulation by the \(\beta\) sheet augmentation mechanism\(^{30}\). Additional example of such multi-domain protein is myomesin, located at the M-band. Crystal structures of its poly-tandem Ig9–13 domains revealed structured, \(\alpha\)-helical linkers between domains, which restrict domain flexibility. These domain-connecting linkers possess a spring-like mechanism – they unfold upon applied tension and again refold when tension is reduced\(^{37, 38}\).

Here, we have focused on a small family of Z-disc proteins, composed of myotilin, myopalladin and palladin. Myotilin is the smallest member of the family with the two Ig domains, compared to a modular proteins myopalladin and palladin, which possess up to five Ig domains. To date, only a limited structural and mechanano-dynamical data exist for this family; while structures of single myotilin Ig1 (ref. \(^{59}\)) and Ig2 and palladin Ig3 (ref. \(^{60}\)) and Ig4 domains have been determined, no information about the tandem domain orientations and dynamics is available. SAXS data, which allowed us to produce averaged low-resolution molecular envelope of tandem myotilin Ig domains, revealed «dumbbell» shape with an almost extended domain conformation. Flexibility analysis using dimensionless Kratky plot suggests a certain degree of conformational plasticity, which was confirmed using ensemble optimisation method (EOM) and MultiFoXS approaches. Use of these methods significantly improved fitting to the experimental data, compared to the rigid body modelling, and showed that myotilin Ig domains in solution can be described as a multi-conformation species that adopt various relative orientations resulting in a compact or more extended forms. However, these states are not equally populated since in the predominant form (50 %) the Ig domain pair adopts a semi-extended orientation. Interestingly, a preferential orientation has also been reported for the Z1Z2 domain pair of titin with implications in selective recruitment of the binding partner telethonin\(^{34}\). Since myotilin Ig domains present an important interaction site for prominent proteins of the sarcoplasm, as well as inter-domain interactions of filamins\(^{30, 31, 49, 54}\), have revealed important implications of inter-domain dynamics in unbound state for interaction with binding partner telethonin\(^{35}\). Also, compact domain-domain interactions of filamins\(^{30, 31, 49, 54}\), as well as inter-molecular interaction between lg24 domains leading to dimerisation\(^{35, 58}\), were found to be functionally important. For example, relative orientation and inter-domain contacts of filamin lg domains 16–24 (Rod 2 region) play an important role in ligand binding regulation by the \(\beta\) sheet augmentation mechanism\(^{30}\). Additional example of such multi-domain protein is myomesin, located at the M-band. Crystal structures of its poly-tandem Ig9–13 domains revealed structured, \(\alpha\)-helical linkers between domains, which restrict domain flexibility. These domain-connecting linkers possess a spring-like mechanism – they unfold upon applied tension and again refold when tension is reduced\(^{37, 38}\).

While previous studies showed that myotilin is able to form dimers via Ig domains\(^{4, 8, 9}\), we did not observe species with molecular weight corresponding to the mass of the dimer (calculated from Ig values), which indicates that the concentration effect described above was not a consequence of oligomerisation. Different construct design used in our and in previously reported studies suggests that regions outside of the Ig1–2 domain pair promote dimer formation and/or that the conditions used do not favour the formation of a stable dimer. For example, study of Salmikangas et. al showed that construct encompassing residues 229–441 is able to dimerise\(^{8}\). On the other hand, another study revealed that smallest fragment of myotilin with the ability to dimerise is 345–498 (ref. \(^{9}\)).

Molecular dynamics simulations using coarse-grained and all-atom approaches corroborated these findings and further showed that the transition of extended form to the more compact one is reversible. Specifically, we identified the highly conserved Glu344, located in the middle of the linker, as a central point of the torsion mechanism involving the rotation of dihedral angle \(\psi\) around the main chain. Presence of this conserved acidic residue within the linker sequence was already reported for I-band Ig domains, but its involvement in intra- or inter-domain interactions was not clearly shown\(^{35}\). In case of myotilin Ig domains, we noticed that during simulation Lys343 and Glu344 formed H-bonds with the Ile375 and Pro376/Ala429, respectively, within the N-terminal PnP and NxxG motifs on Ig2 domain, indicating impact of these conserved features on intra-domain stabilising interactions. Moreover, we observed that linker residues during simulation showed tendency to form \(\alpha\)-helical structure, therefore restricting flexibility and consequently affecting Ig domain dynamics.

Molecular evolution analysis of the myotilin revealed that Ig domains are the most conserved part and present the only structured region of the molecule, consistent with the disorder prediction. When only Ig domains were analysed, mainly hydrophobic residues were conserved, along with the PPxf and NxxG motifs, whereas other solvent accessible residues showed high variability, implying that these residues co-evolved in respect to the binding partners. Linker residues between homologous Ig domains of the family were highly conserved and it can be speculated that they share similar mechanisms of intra-domain orientation changes in contrast to other Ig domain pairs in palladin and myopalladin, which have different linker length and amino acid composition.

Our molecular evolution analysis additionally allowed us to shed light on the molecular basis of myotillinopathies. Based on our results, we propose that aggregation through newly exposed hydrophobic patches plays an important role in pathological mechanism, probably in concert with the slower degradation and changed turnover as already indicated in previous studies\(^{38}\).

In conclusion, our results clearly show that Ig domain pair of myotilin is flexible, however with a general preference for a semi-extended inter-domain orientation with possible implications in binding partner recognition by analogy to other known examples of tandem Ig domain-containing proteins. Also, due to similar linker length and high sequence conservation, similar tandem Ig domain dynamics is expected for myopalladin and palladin, which is in line with their partially overlapping roles.
**Methods**

**Cloning, protein expression and purification.** DNA fragment encoding Ig1–2 domains of human myotilin (residues 250–444; Uniprot accession code Q9UBF9) was cloned into pETM-14 vector (EMBL) to create fusion construct with the plasmid-encoded N-terminal His$_6$-tag and a HRV-3C protease cleavage site at the 5’ end (MYOTIg1–2). Correct assembly was verified by DNA sequencing (GATC Biotech). For protein expression *E. coli* strain BL31(DE3) was used. Briefly, cells were grown at 37 °C in LB media supplemented with ampicillin (100 μg/mL) to an OD$_{600}$ of approximately 0.6. After cooling to 20 °C recombinant protein expression was induced by addition of IPTG to a final concentration of 0.5 mM. After 16 h cells were collected by centrifugation, resuspended in lysis buffer (20 mM HEPES, 500 mM NaCl, 10 mM imidazole, pH 7.4), and lysed by sonication. Soluble fraction was obtained by centrifugation (20 000 rpm, 4 °C) and applied onto HiTrap IMAC FF column (GE Healthcare). After washing (20 mM HEPES, 500 mM NaCl, 10 mM imidazole, 5 % glycerol, pH 7.4) the bound proteins were eluted by gradually raising imidazole concentration to 500 mM. His$_6$-tag was removed by incubation of the pooled eluates containing MYOTIg1–2 with GST-tagged HRV-3C protease (prepared in-house; mass ratio MYOTIg1–2/HRV-3C 100:1) during overnight dialysis at 4 °C against the dialysis buffer (50 mM Tris, 150 mM NaCl, 1 mM DTT, pH 7.4). The His$_6$-tag free MYOTIg1–2 was recovered as a flow-through after applying the cleavage mixture onto HiTrap IMAC FF and GSTrap FF columns to remove uncleaved MYOTIg1–2 and HRV-3C protease, respectively. Final purification step was size-exclusion chromatography on a Superdex 75 10/300 column (GE Healthcare) equilibrated in buffer 20 mM HEPES, 150 mM NaCl, 5 % glycerol, 1 mM DTT, pH 7.4. The purity of the final samples was assessed by SDS-PAGE (Supplementary Fig. S3).

**Evolutionary analysis.** For evolutionary analysis, two sets of data were prepared: (1) complete myotilin sequences (Supplementary Table S1) and (2) sequences of Ig domains of myopalladin and palladin (Supplementary Table S2) from Vertebrates subphylla. Sequences were retrieved from the NCBI nr database via Protein Blast and Ensembl genome browser using either myotilin (Q9UBF9) or Ig domain sequences of palladin (Q8WX93) and myopalladin (Q86TC9). Following manual data inspection and filtering, the sequences were aligned and phylogenetic trees were constructed using maximum likelihood based method implemented in the MEGA 7 software. Input parameters for constructing phylogenetic trees were: LG substitutional model, Bio-NJ as initial tree, Nearest-Neighbour-Interchange (NNI) as tree inference heuristic method and branch support calculated with 100 bootstrap replicates. Calculated trees were graphically edited using FigTree software (http://tree.bio.ed.ac.uk/software/figtree/). Conservation scores for each myotilin residue in the multiple sequence alignment (MSA) were calculated using the ConSurf server. Pre-calculated MSA for Ig1 consisted of 301 homologues and 152 homologues for Ig2. Calculated conservation rates were projected onto both Ig structures and coloured according to the used colouring scheme.

**Small-angle X-ray scattering (SAXS).** SAXS data for purified construct MYOTIg1–2 was collected at ESRF beamline BM29 BioSaxs (Grenoble, France) equipped with the Pilatus 1M detector. Samples were measured at concentrations of 3.67, 7.00 and 14.12 mg/mL in buffer 20 mM Na$_2$-HEPES, 150 mM NaCl, 5 % glycerol, 1 mM DTT, pH 7.4, in two independent measurements. The scattering intensity (I) was measured in the range from 0.03 < s < 5.0 Å$^{-1}$, where s = 4π sinθ/λ, and λ = 1 Å, at the distance between sample and detector of 2.867 m. Background scattering was subtracted and data reduced, normalised according to the measured concentration and extrapolated to infinite dilution using the two lowest measured concentrations using PRIMUS module of the ATSAS software package. Forward scattering (I$_0$) and radius of gyration (R$_g$) were obtained by fitting the linear Guinier region of the data. Pair distribution function P(r) with the corresponding maximum particle size parameter (D$_{max}$) was determined using GNOM program. The dimensionless Kratky plots were generated by the use of data from SASBDB for BSA (accession code: SASDBT4), filamin A Ig20-Ig21 (accession code: SASDAH8), filamin A Ig20-Ig21 (PDB accession code: SASDAH8) and filamin A Ig22-Ig23 (accession code: SASDAU3).

For reconstruction of theoretical molecular envelope *ab initio* modelling was performed 20 times using the program DAMMIF, where scattering from the calculated envelopes were fitted against the experimental scattering and evaluated by the chi values. The most typical envelope was selected by comparing the normalised spatial discrepancy (NSD) values between the pair of envelopes (Supplementary Table S3) and later averaged by DAMAVER set of programs. In rigid body modelling, high resolution structures of myotilin Ig1 (PDB accession code: 2KDG) and Ig2 (PDB accession code: 2KQK) were used to fit SAXS scattering data using program CORAL. Linker residues were designated as dummy atoms. *Ab initio* calculated envelope was superposed to the rigid body model using SUPCOMB program.

Flexibility analysis was performed using EOM2.1 program with enforced P1 symmetry while the linker between Ig domains was again designated as dummy atoms. Conformations consistent with scattering data were selected from the pool of 50 000 models using a genetic algorithm. The flexibility analysis was independently repeated three times; all runs gave comparable results. For comparative analysis MultiFoXS server was used where I-TASSER structure was used as the input model; residues Leu341, Ala342, Lys343, Glu344, His345, Lys346, Arg347 were designated as flexible residues. More than 10 000 conformations were calculated and scored with the rapidly exploring random trees (RRT) search to extract only those conformations, which adequately describe observed system. All structure figures were prepared using PyMOL (The PyMOL Molecular Graphics System, version 1.3, Schrödinger, LLC).

**Modelling, coarse-grained and all-atom molecular dynamics.** Models of tandem Ig domains were calculated using I-TASSER and the best model was selected on the basis of its Z-score. This model was then
used to analyse structure fluctuations as well as in molecular dynamics simulations. Same model was used for the
in silico mutagenesis, which was performed in PyMOL, to obtain model with the mutated AKE motif to AAA.

Fluctuations of relative orientation between the tandem Ig domains were predicted using CABS-Flex server
which implements structure coarse-graining and Monte Carlo dynamics sampling scheme. RMSF values of the
2000 frames were analysed in terms of distance between Ig domains (defined as distance from C_{\alpha} of N-terminal
to C_{\alpha} of C-terminal amino acid residue) vs. number of frames.

All-atom molecular dynamics (MD) simulations were performed using NAMD 2.1 program suite with
CHARMM-27 force-field. Initial structure was solvated and neutralised (by adding Na+ or Cl- ions) and the
 topology was generated using psfgen. Equilibration was done at 300 K using constant temperature Langevin
dynamics, constant pressure via Nose-Hoover Langevin piston and Particle Mesh Edward for full-system peri-
odic electrostatics. Energy minimisation was done for 500 steps of conjugate gradients. Time-step was set to 2 fs.
The length of the final production run was 10 ns. MD trajectories were analysed using VMD where end-to-end
distances were measured. Further trajectory analysis was done using the DynDom server to reveal hinge resi-
dues which could participate in the torsion mechanism of conformational switching and to calculate rotation
angle of inter-domain bending. Inter- and intra-domain interactions via hydrogen bonding were analysed and
graphically presented using Cytoscape coupled with UCSF Chimera and structureViz programs.

References

1. Gautel, M. & Djinovic-Carugo, K. The sarcomeric cytoskeleton: from molecules to motion. J. Exp. Biol. 219, 135–145 (2016).
2. Au, Y. The muscle ultrastructure: a structural perspective of the sarcomere. Cell. Mol. Life Sci. 61, 3016–3033 (2004).
3. Luther, P. K. The vertebrate muscle Z-disc: sarcomere anchor for structure and signalling. J Muscle Res Cell Motil 30, 171–185 (2009).
4. Salminen, P. et al. Myotilin, a novel sarcomeric protein with two Ig-like domains, is encoded by a candidate gene for limb-girdle
muscular dystrophy. Hum. Mol. Genet. 8, 1329–1336 (1999).
5. Bang, M. L. et al. Myopalladin, a novel 145-kilodalton sarcomeric protein with multiple roles in Z-disc and I-band protein
assemblies. J Cell Biol 153, 413–427 (2001).
6. Parast, M. M. & Otey, C. A. Characterization of palladin, a novel protein localized to stress fibers and cell adhesions. J Cell Biol 150,
643–656 (2000).
7. Mykkänen, O. M. et al. Characterization of human palladin, a microfilament-associated protein. Mol. Biol. Cell 12, 3060–3073
(2001).
8. Salmikangas, P. et al. Myotilin, the limb-girdle muscular dystrophy 1A (LGMD1A) protein, cross-links actin filaments and controls
sarcomere assembly. Hum. Mol. Genet. 12, 189–203 (2003).
9. Nandelstädth, von P. et al. Actin-organizing properties of the muscular dystrophy protein myotilin. Exp. Cell Res. 310, 131–139
(2005).
10. van der Ven, P. F. M. et al. Indications for a novel muscular dystrophy pathway, gamma-filamin, the muscle-specific filamin isoform,
interacts with myotilin. J Cell Biol 151, 235–248 (2000).
11. Gontier, Y. et al. The Z-disc proteins myotilin and FATZ-1 interact with each other and are connected to the sarcolemma via muscle-
specific filamins. J. Cell Biol. 118, 3739–3749 (2005).
12. Nandelstädth, von P. et al. A class III PDZ binding motif in the myotilin and FATZ families binds enigma family proteins: a common
link for Z-disc myopaties. Mol. Cell. Biol. 29, 822–834 (2009).
13. Beck, M. R., Otey, C. A. & Campbell, S. L. Structural Characterization of the Interactions between Palladin and \( \alpha \)-actinin. J. Mol.
Biol. 413, 712–725 (2011).
14. Young, P., Ferguson, C., Bañuelos, S. & Gaštel, M. Molecular structure of the sarcomeric Z-disc: two types of titin interactions lead
to an asymmetrical sorting of alpha-actinin. EMBO J 17, 1614–1624 (1998).
15. Young, P. & Gautel, M. The interaction of titin and alpha-actinin is controlled by a phospholipid-regulated intramolecular pseudoligand mechanism. EMBO J 19, 6331–6340 (2000).
16. Joseph, C. et al. A structural characterization of the interactions between titin Z-repeats and the alpha-actinin C-terminal domain.
Biochemistry 40, 4957–4965 (2001).
17. Atkinson, R. A. et al. Ca2+–independent binding of an EF-hand domain to a novel motif in the alpha-actinin-titin complex. Nat.
Struct. Biol. 8, 853–857 (2001).
18. Rönty, M., Taivainen, A., Moza, M., Otey, C. A. & Carpen, O. Molecular analysis of the interaction between palladin and \( \alpha \)-actinin. FEBS Letters 566, 30–34 (2004).
19. Ribeiro, E. A. et al. The structure and regulation of human muscle \( \alpha \)-actinin. Cell 159, 1447–1460 (2014).
20. Otey, C. A., Dixon, R., Stack, C. & Goicoechea, S. M. Cytoplasmic Ig-domain proteins: Cytoskeletal regulators with a role in human
disease. Cell Motil. Cytoskeleton 66, 618–634 (2009).
21. Wang, J. et al. Dynamics of Z-band based proteins in developing skeletal muscle cells. Cell Motil. Cytoskeleton 61, 34–48 (2005).
22. Wang, J., Dube, D. K., Mittal, B., Sanger, J. M. & Sanger, J. W. Myotilin dynamics in cardiac and skeletal muscle cells. Cytoskeleton (Hoboken) 68, 661–670 (2011).
23. Hauser, M. A. et al. Myotilin is mutated in limb girdle muscular dystrophy 1A. Hum. Mol. Genet. 9, 2141–2147 (2000).
24. Selcen, D. & Engel, A. G. Mutations in myotilin cause myofibrillar myopathy. Neurology 62, 1363–1371 (2004).
25. Selcen, D. Myofibrillar myopathies. Neuromuscul. Disord. 21, 161–171 (2011).
26. Foroud, T. et al. A mutation in myotilin causes spheroid body myopathy. Neurology 65, 1936–1940 (2005).
27. Olivi, M., Goldfarb, L. G., Shatunov, A., Fischer, D. & Ferrer, I. Myotilinopathy: refining the clinical and myopathological phenotype.
Brain 128, 2315–2326 (2005).
28. Nandelstädth, von P., Solymani, R., Baumann, M. & Carpen, O. Analysis of myotilin turnover provides mechanistic insight into the
role of myotilinopathy-causing mutations. Biochem. J. 436, 113–121 (2011).
29. Moza, M. et al. Targeted deletion of the muscular dystrophy gene myotilin does not perturb muscle structure or function in mouse.
Mol. Biol. Cell 27, 244–252 (2007).
30. Lad, Y. et al. Structure of three tandem filamin domains reveals auto-inhibition of ligand binding. EMBO J. 26, 3993–4004 (2007).
31. Sepulda, J. et al. Flexible Structure of Peptide-Bound Filamin A Mechanosensor Domain Pair 20–21. PLoS ONE 10, e0136969 (2015).
32. Tchkurovbova, L. & Trinick, J. Properties of titin immunoglobulin and fibronectin-3 domains. J. Biol. Chem. 279, 46351–46354
(2004).
33. Marino, M. et al. Poly-Ig tandems from I-band titin share extended domain arrangements irrespective of the distinct features of their
modular constituents. J Muscle Res Cell Motil 26, 355–365 (2005).
34. Marino, M. et al. The Ig doublet Z1Z2: a model system for the hybrid analysis of conformational dynamics in Ig tandems from titin.
Structure 14, 1437–1447 (2006).
35. Zou, P. et al. Palindromic assembly of the giant muscle protein titin in the sarcomeric Z-disc. Nature 439, 229–233 (2006).
36. Lee, E. H., Gao, M., Pinotsis, N., Wilmans, M. & Schulten, K. Mechanical strength of the titin Z1Z2-telethonin complex. Structure 14,
497–509 (2006).
37. Lee, E. H., Hsin, J., Castelmuir, vonE., Mayans, O. & Schulten, K. Tertiary and secondary structure elasticity of a six-Ig titin chain. *Biophys. J.* 98, 1085–1095 (2010).
38. Lukman, S., Grant, G. H. & Bui, J. M. Unraveling evolutionary constraints: a heterogeneous conservation in dynamics of the titin Ig domains. *FEBS Letters* 584, 1235–1239 (2010).
39. Higgins, D. G., Labett, S., Gautel, M. & Gibson, T. J. The evolution of titin and related giant muscle proteins. *J. Mol. Evol.* 38, 395–404 (1994).
40. Kenny, P. A., Liston, E. M. & Higgins, D. G. Molecular evolution of immunoglobulin and fibronectin domains in titin and related muscle proteins. *Gene* 232, 11–23 (1999).
41. Light, S., Sagit, R., Ithychanda, S. S., Qin, J. & Elofsson, A. The evolution of filamin – A protein domain repeat perspective. *J. Struct. Biol.* 179, 289–298 (2012).
42. Clarke, J., Kota, E., Fowler, S. B. & Hamill, S. J. Folding studies of immunoglobulin-like β-sandwich proteins suggest that they share a common folding pathway. *Structure* 7, 1145–1153 (1999).
43. Witt, C. C. et al. A Survey of the Primary Structure and the Interspecies Conservation of I-Band Titin's Elastic Elements in Vertebrates. *J. Struct. Biol.* 122, 206–215 (1998).
44. Steward, A., Adhya, S. & Clarke, J. Sequence conservation in Ig-like domains: the role of highly conserved proline residues in the fibronectin type III superfamily. *J. Mol. Biol.* 318, 935–940 (2002).
45. van der Lee, R. et al. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* 114, 6589–6631 (2014).
46. Vacic, V. et al. Disease-associated mutations disrupt functionally important regions of intrinsic protein disorder. *PLoS Comput Biol* 8, e1002709 (2012).
47. Talla, V., Narayanan, C., Srinivasan, N. & Balasubramanian, D. Mutation causing self-aggregation in human gamma-c-crystallin leading to congenital cataract. *Invest. Ophthalmol. Vis. Sci.* 47, 5212–5217 (2006).
48. Skou, S., Gillilan, R. E. & Ando, N. Synchrotron-based small-angle X-ray scattering of proteins in solution. *Nature Protocols* 9, 1727–1739 (2014).
49. Sethi, R. & Ylänne, J. Small-angle X-ray scattering reveals compact domain-domain interactions in the N-terminal region of filamin C. *PLoS ONE* 9, e107457 (2014).
50. Valrentini, E., Kikhtney, A. G., Previtali, G., Jeffries, C. M. & Svergun, D. I. SASBDB, a repository for biological small-angle scattering data. *Nucl. Acids Res.* 43, D357–63 (2015).
51. Heikkinen, O. K. & et al. Atomic structures of two novel immunoglobulin-like domain pairs in the actin cross-linking protein filamin. *J. Biol. Chem.* 284, 25450–25458 (2009).
52. Jamroz, M., Kolinski, A. & Kmiecik, S. CABS-flex predictions of protein flexibility compared with NMR ensembles. *Bioinformatics* 30, 2150–2154 (2014).
53. Labeit, S., Kolmerer, B. & Linke, W. A. The giant protein titin. Emerging roles in physiology and pathophysiology. *Circ. Res.* 80, 290–294 (1997).
54. Sethi, R. et al. A Novel Structural Unit in the N-Terminal Region of Filamins. *J. Biol. Chem.* jbc.M113.537456. doi:10.1074/jbc.M113.537456 (2014).
55. Pudas, K., Kietma, T.-R., Butler, P. J. G., Stewart, M. & Ylänne, J. Structural Basis for Vertebrate Filamin Dimerization. *Structure* 13, 111–119 (2005).
56. Sjekloča, L. et al. Crystal structure of human filamin C domain 23 and small angle scattering model for filamin C 23–24 dimer. *J. Mol. Biol.* 368, 1011–1023 (2007).
57. Pinotis, N., Lange, S., Periard, J.-C., Svergun, D. I. & Wilmanns, M. Molecular basis of the C-terminal tail-to-tail assembly of the sarcomeric filament protein myomesin. *EMBO J.* 27, 253–264 (2008).
58. Pinotis, N. et al. Superhelical architecture of the myosin filament-linking protein myomesin with unusual elastic properties. *PLoS Biol.* 10, e1001261 (2012).
59. Heikkinen, O. et al. Solution structure of the first immunoglobulin domain of human myotilin. *J Biomed NMR* 44, 107–112 (2009).
60. Beck, M. R. et al. Structure and function of palladin’s actin binding domain. *J. Mol. Biol.* 425, 3325–3337 (2013).
61. Caspermeyer, J. MEVA Evolutionary Software Re-Engineered to Handle Today’s Big Data Demands. *Mol. Biol. Evol.* 33, 1887–1887 (2016).
62. Valdar, W. S. J. Scoring residue conservation. *Proteins: Structure, Function, and Bioinformatics* 48, 227–241 (2002).
63. Kozlowski, L. P. & Bujnicki, J. M. MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins. *BMC Bioinformatics* 13, 111 (2012).
64. Goldenberg, O., Erez, E., Nimrod, G. & Ben-Tal, N. The ConSurf-DR: pre-calculated evolutionary conservation profiles of protein structures. *Nucl. Acids Res.* 37, D123–7 (2009).
65. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. I. & Svergun, D. I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* 36, 1277–1282 (2003).
66. Petoukhov, M. V. et al. New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr* 45, 342–350 (2012).
67. Svergun, D. I. IUCr. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* 25, 495–503 (1992).
68. Franke, D. & Svergun, D. I. IUCr. DAMMiC: a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr* 42, 342–346 (2009).
69. Volkov, V. V. & Svergun, D. I. IUCr. Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Crystallogr* 36, 860–864 (2003).
70. Konarev, P. V., Petoukhov, M. V. & Svergun, D. I. Rapid automated superposition of shapes and macromolecular models using spherical harmonics. *J Appl Crystallogr* 49, 953–960 (2016).
71. Tria, G., Mertens, H. D. T., Kachala, M. & Svergun, D. I. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* 2, 207–217 (2015).
72. Schneiderman-Duhovny, D., Hamnel, M., Tainer, J. A. & Salz, A. FoXS, FoXS Dock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. *Nucl. Acids Res.* 44, W424–9 (2016).
73. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9, 40 (2008).
74. Jamroz, M., Kolinski, A. & Kmiecik, S. CABS-flex: Server for fast simulation of protein structure fluctuations. *Nucl. Acids Res.* 41, W427–31 (2013).
75. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* 26, 1781–1802 (2005).
76. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* 14, 33–8–27–8 (1996).
77. Hayward, S. & Lee, R. A. Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *J. Mol. Graph. Model.* 21, 181–183 (2002).
78. Shannon, P. & et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504 (2003).
79. Petterson, E. E. et al. UCSF Chimera–a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25, 1605–1612 (2004).
80. Morris, J. H., Huang, C. C., Babbit, P. C. & Ferrin, T. E. structureViz: Linking Cytoscape and UCSF Chimera. *Bioinformatics* 23, 2345–2347 (2007).
Acknowledgements
This work was supported by the Slovenian Research Agency young researcher grant (No. 35337) and research program P1–0140. We thank staff of SAXS beamline at ESRF in Grenoble for their excellent supports. We are grateful to Euripides de Almeida Ribeiro Jr. (University of Vienna) for data collection and excellent assistance during SAXS measurements. KDC research was supported by a Marie Curie Initial Training Network: MUZIC (N°238423), Austrian Science Fund (FWF) Projects I525 and I1593, P22276, P19060 and W1221, Federal Ministry of Economy, Family and Youth through the initiative “Laura Bassi Centres of Expertise”, funding the Centre of Optimized Structural Studies, N°253275 and by the University of Vienna.

Author Contributions
The project was designed by K.D.C., B.L. and V.P. V.P. prepared protein samples, performed SAXS, molecular dynamics simulations and evolutionary analysis. V.P., M.P. and K.D.C. wrote the manuscript with all authors contributing to editing the manuscript and supporting the conclusions.

Additional Information
Supplementary information accompanies this paper at doi:10.1038/s41598-017-03323-6

Competing Interests: The authors declare that they have no competing interests.

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

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

© The Author(s) 2017