Actin-Induced Structure in the Beta-Thymosin Family of Intrinsically Disordered Proteins

Xue, B.; Robinson, Robert Charles

2016

Xue, B., & Robinson, R. C. (2016). Actin-Induced Structure in the Beta-Thymosin Family of Intrinsically Disordered Proteins. *Vitamins and Hormones*, 102, 55-71.

<https://hdl.handle.net/10356/83977>

<https://doi.org/10.1016/bs.vh.2016.04.007>

© 2016 Elsevier. This is the author created version of a work that has been peer reviewed and accepted for publication by *Vitamins and Hormones*, Elsevier. It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [<http://dx.doi.org/10.1016/bs.vh.2016.04.007>].

Downloaded on 26 Apr 2023 21:53:29 SGT

Actin-induced structure in the Beta-thymosin family of intrinsically disordered proteins

Bo Xue^a and Robert C. Robinson^{a,b,c,d,e,f}

^aInstitute of Molecular and Cell Biology, A*STAR (Agency for Science, Technology and Research), Biopolis, Singapore 138673.

^bDepartment of Biochemistry, National University of Singapore, 8 Medical Drive, Singapore 117597.

^cDepartment of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597.

^dNTU Institute of Structural Biology, Nanyang Technological University, 59 Nanyang Drive, Singapore 636921.

^eSchool of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551.

^fLee Kong Chan School of Medicine, 50 Nanyang Avenue, Singapore 639798.

Corresponding author:

Bo Xue, Institute of Molecular and Cell Biology, Proteos, 61 Biopolis Drive, Singapore 138673; Tel: +65-6586-9831; Fax: +65-6779-1117; E-mail: bxue@imcb.a-star.edu.sg

Abstract

Thymosin β 4 (T β 4) is a 43-amino acid signature motif peptide that defines the beta-thymosin (β T) family of proteins. β Ts are intrinsically unstructured in their free states, and undergo disorder-to-order transitions in carrying out their biological functions. This property poses challenges in determining their 3D structures, mainly favoring structural studies on the complexes formed between β Ts and their interaction partners. One of β Ts' primary binding partners is monomeric actin, a major component of the cytoskeleton in eukaryotic cells. T β 4's role in this system is to maintain the highly concentrated pool of monomeric actin that can be accessed through profilin by actin filament nucleating machineries. Here, we give an account of the structures of β Ts that have been illuminated by nuclear magnetic resonance (NMR) and X-ray crystallography. NMR has been the method of choice for probing regions that have intrinsic conformational preference within the largely disordered β Ts in their native states in solution. X-ray crystallography has demonstrated at atomic detail how β Ts interact with actin. Detailed analysis of these structures highlights the disorder-to-order transition of T β 4 in binding to actin and its isoform specificity.

Keywords:

Beta-thymosin, thymosin β 4, intrinsically disordered/unstructured proteins, actin, structure, crystallography, NMR, protein hybrid,

Introduction

Beta-thymosins (β Ts) are a family of proteins that are comprised of one or more of a ~43 amino acid signature motif (β T repeat) that was first defined for the prototypical single motif protein, thymosin β 4 (T β 4). T β 4 was discovered in the 1970s in an effort to extract hormones from thymus tissue (Goldstein, 2007). In the early 1990s, T β 4 was found to bind and sequester monomeric actin, a major component of the cytoskeleton in eukaryotic cells (Safer et al., 1991; Safer et al., 1990). Around that time, the first nuclear magnetic resonance (NMR) structural studies on isolated T β 4 began, and later on the T β 4:actin complex (Czisch et al., 1993; Domanski et al., 2004; Zarbock et al., 1990). X-ray crystallography studies were only possible once strategies were developed to stabilize the intrinsically disordered T β 4 in its complex with actin (Aguda et al., 2006; Hertzog et al., 2004; Irobi et al., 2004; Xue et al., 2014). The accumulation of these structural data, together with the discovery of homologs through protein and nucleotide sequencing, allowed for a precise definition of the features of the β T repeat: it has a consensus LKKT motif at the center of the sequence, which is flanked by two moderately conserved segments that tend to weakly form amphipathic α -helices in isolation, yet become stabilized when bound to actin. Since similar LKKT motifs, N-terminal α -helices and actin-binding properties can also be found in the Wiskott-Aldrich syndrome protein (WASP) homology domain 2 (WH2)-containing proteins, β Ts and WH2s are often grouped in the same functional superfamily (Paunola et al., 2002), whether or not they arose from a common ancestor (Edwards, 2004).

β T repeats appear within proteins either singly or as multiple repeats. T β 4, and close homologs, such as T β 10 and T β 15, contain one β T repeat, whereas other members contain from 2 to 27 such repeats (Xue and Robinson, 2013). One widely studied multi- β T repeat protein is the three β T repeat *Drosophila* ciboulot, which was the first protein to yield an actin-bound β T repeat structure by X-ray crystallography (Hertzog et al., 2004). Generally, single- and multi-repeat β Ts appear to be mutually exclusive in their phylogenetic distributions (Van Troys et al., 2007), with few exceptions (Dehal et al., 2002; Dhaese et al., 2009). Single- and multi-repeat β Ts also show differences in their actin-binding functionality: all of the studied single-repeat β Ts act like T β 4 in inhibiting their bound actin from forming or joining a filament, whereas multi-repeat β Ts, such as ciboulot and the four-repeat TetraThymosin β , often allow their bound actin to join the fast-growing, barbed end of a preexisting actin filament, a profilin-like activity (Hertzog et al., 2004; Van Troys et al., 2004). Recent studies have attributed the functional difference to the stability of the C-terminal helices within the β T repeats (Didry et al., 2012; Xue et al., 2014).

From a structural point of view, β Ts belong to a class of proteins termed intrinsically disordered/unstructured proteins (IDPs/IUPs) (Andreeva et al., 2014; Oldfield and Dunker, 2014; van der Lee et al., 2014). Members of this class of proteins contain regions, or in the case of β Ts, the entire protein, that lack ordered structure in their native states, but undergo disorder-to-order transitions in carrying out their biological functions (Uversky, 2002). A large collection of IDPs/IUPs can be found in DisProt, the curated database of disordered proteins (Sickmeier et al., 2007), among which is T β 4 and other actin-binding proteins, such as WASP, tropomodulin, and troponin. Despite being generally unstructured, IDPs/IUPs do have regions with intrinsic conformational preferences, which are further classified into two motifs by SCOP2 (Andreeva et al., 2014), flexible disordered helical regions and extended binding regions. We give a brief historical account of the structural studies on β Ts and present a detailed analysis of the interaction between T β 4 and actin, which reveals that β Ts contain both of the IDPs/IUPs motifs.

Beta-thymosins in solution

T β 4 in mixed organic-aqueous solvent

Due to the small size and unstructured nature of β Ts, solution NMR has been the method of choice to study their isolated structures. The first β T NMR structure determined was that of bovine T β 4, which is identical in amino acid sequence to human T β 4 (Zarbock et al., 1990). The experiments were conducted in solutions of 60% (v/v) trifluoroethanol-d₃ or 50% (v/v) hexafluoroisopropyl-d₂ in water, because no obvious structure could be detected by circular dichroism and ¹H NMR spectra for T β 4 in aqueous solutions. The presence of alcohols in the solution promoted the formation of ordered helical structures in T β 4 at residues 4-16 and 30-40. Due to insufficient nuclear Overhauser enhancement (NOEs) signals for residues 17 and 18, the relative orientation of these two helices could not be determined. In addition to the helices, there was also a loop region between residues 24-29 of T β 4.

T β 4 in aqueous solution

The NMR study was subsequently continued to probe the influences of temperature and buffer pH on the structure of T β 4 in aqueous solutions (Czisch et al., 1993). The measurements were performed at 1, 4, 14 and 29°C for samples at pH 3.0, and at 1 and 14°C for samples at pH 6.5, which were the conditions favorable for minimizing amide exchange. Temperature was found to have a significant effect on the secondary structure T β 4. Residues 5-16 had a tendency to form an α -helix, which was stable at up to 14°C and became destabilized at higher temperatures. Residues 31-37 also showed a propensity to adopt a helical conformation at 1 and 4 °C, albeit less well defined than residues 5-16; however, residues 31-37 became completely random at 14°C. In agreement, the helical content percentages determined from the circular dichroism spectra were 15, 9 and 7% for 4, 9 and 20 °C, respectively. The first four residues at the N-terminus of T β 4 were always disordered irrespective of the temperature.

T β 10 in mixed organic-aqueous solvent

Following these studies on bovine T β 4, the NMR study was extended to bovine T β 10 which is a close homolog of T β 4 and was originally called T β 9 (Stoll et al., 1997). Two helical regions at residues 4-27 and 32-41 of T β 10 were detected in a solution of 40% (v/v) 1,1,1,3,3,3-hexafluoro-2-propanol-d₂ in water, which were separated by a poorly defined loop region between residues 28-31 (Fig. 1A). The first helix of T β 10 was, however, twice the length of its counterpart in T β 4. Similar to T β 4, the lack of medium or long-range NOEs for the loop region impeded the determination of the relative orientation of the two helices in T β 10. Moreover, the N-terminus of T β 10 also showed random-coil structure.

T β 4 mutants in mixed organic-aqueous solvent

To further correlate T β 4's structure with its actin binding-activity, NMR spectra were recorded in 60% (v/v) trifluoroethanol for wild-type T β 4 and its mutants (Simenel et al., 2000). Consistent with the previously reported T β 4 structures in alcohols, the wild-type conformation under these conditions consisted of two α -helices: residues 5-16, of which Lys11 to Lys16 form a more stable core; and the second helix from residues 31-39. Several point mutations were made in the first helix and in the LKKT motif (residues 17-22), the two main regions known to be responsible for actin binding. One such mutation, S15A, caused the first helix to become extended at its C-terminal end, to include the LKKT motif leading to a 50% reduction in its actin-binding activity. Another mutation, K11P, disrupted the formation of the first helix, resulting in complete loss of actin binding. These data

demonstrated that the presence of the first helix and its correct orientation in relation to the LKKT motif are essential for the T β 4:actin interaction.

Thus, these structural data indicate that β Ts are largely unstructured in solution, but have a latent propensity to form two regions of helix, which can be induced by aqueous-organic mixed solvents or low temperatures.

Beta-thymosins in complex with actin

NMR-based T β 4:actin model

Chemical cross-linking data had established that T β 4 adopts an extended conformation when bound to actin, with its N-terminal region in contact to subdomain 1 of actin and its C-terminal region in close vicinity to His40 and Gln41 of actin subdomain 2 (De La Cruz et al., 2000; Reichert et al., 1996; Safer et al., 1997). The first high-resolution model for the structural basis of the T β 4:actin interaction was constructed from NMR data using selective labeling of T β 4 (Domanski et al., 2004). The relatively large size of the T β 4:actin complex (47 kDa) had precluded NMR spectroscopy studies of the whole complex. However, a complete assignment of amide protons and nitrogen resonances for T β 4 was successfully achieved, while bound to actin, using a combination of selective labeling of Leu, Thr, and Lys of T β 4 in a three-dimensional ^1H - ^{15}N NOESY-HSQC experiment. The actin-bound model (PDB ID: 1UY5, Fig. 1B) was constrained in its positioning on actin by the distance restraints arising from the crosslinking data. This model predicts that the entire length of T β 4 is engaged in binding to actin, causing the T β 4 backbone to adopt a conformation, with the two terminal helices flanking the extended region. Cleavage of the actin DNase I binding loop by subtilisin did not modify the signal arising from T β 4 in the complex, suggesting that the T β 4 backbone is not in the vicinity of actin residues 42-47. These data established that T β 4 terminal segments form helices that are induced and/or stabilized through binding to actin.

X-ray structures

C-terminal segment of T β 4

Further breakthroughs in understanding the T β 4:actin interaction came with novel strategies to stabilize the complex for X-ray crystallography. Through creating a hybrid protein between domain 1 of gelsolin (G1), a protein that severs and caps actin filaments (Nag et al., 2013), and the C-terminal segment of T β 4 (hereafter referred to as G1-T β 4), the structure of the G1-T β 4:actin complex was elucidated at 2 Å resolution (PDB ID: 1T44, Fig. 1C) (Irobi et al., 2004). This structure revealed that the C-terminal segment of T β 4 forms a helix which bridges, and narrows, the gap between subdomains 2 and 4 of actin, in contrast to the NMR model where it only contacts subdomain 2. The new conformation provided an explanation for the biochemical evidence that Lys38 of T β 4, located at the T β 4 C-terminus, could be crosslinked to Gln41 in the DNase I loop of actin (Safer et al, 1997). In the G1-T β 4:actin structure, the flexible actin DNase I loop (residues 39-50) is, like in most other actin structures, disordered. The distance between the mainchain Ca carbon atoms of T β 4's Lys38 and actin's Pro38, which is the closest residue to Gln41 that was resolved in the G1-T β 4:actin structure, is 15.1 Å. Given the length of the side chains of Lys (6.4 Å) and Gln (4.4 Å), the flexibility of actin's DNase I loop, and the unstructured nature of T β 4 before becoming fully bound to actin, it is reasonable to envisage that these two residues may be accessed by a zero-length crosslinker. Therefore, this crystal structure revealed the binding geometry of T β 4's C-terminal segment to actin that induces and/or stabilizes the helical conformation of T β 4.

N-terminal segment of ciboulot's first β T repeat

In the same year as the determination of the G1-T β 4:actin X-ray structure (2004), Hertzog et al. determined the X-ray structure of a 25-residue peptide from the three-repeat β -thymosin protein, ciboulot, in complex with actin at 2.5 Å (PDB ID 1SQK, Fig. 1D) (Hertzog et al., 2004). The peptide in the structure corresponds to the N-terminal segment of ciboulot's first β T repeat (Asp10 to Ser34, hereafter referred to as Cib1N), which has moderate sequence similarity to T β 4. This X-ray structure is of considerable significance because it was the first to reveal how the N-terminal helix and the central LKKT motif of the β T/WH2 superfamily of proteins interact with actin, and the mode of interaction was later confirmed repeatedly by other complex structures of actin-bound WH2 motifs (Chereau et al., 2005; Lee et al., 2007). The N-terminal α -helix of Cib1N in the structure was extended relative to later β T/WH2:actin structures. Later it became apparent that the helix was extended by stabilization through crystal packing, and that the actual actin-bound helix is one turn shorter (Didry et al., 2012). Nevertheless, this work demonstrated that actin binding induces and/or stabilizes the β T/WH2 N-terminal helix and revealed the binding geometry of this helix and the conserved LKKT motif.

C-terminal segment of ciboulot's second β T repeat

Utilizing a similar strategy employed for the G1-T β 4 hybrid, a hybrid protein comprised of G1 and ciboulot's second and third β T repeats (hereafter referred to as G1-Cib23), was used in an attempt to determine the structure of this hybrid bound to actin (PDB ID: 2FF6, Fig. 1E) (Aguda et al., 2006). Although the hybrid contained nearly half of ciboulot (Glu72 to Ala129), the structure revealed only a short stretch of 11 residues from the C-terminal segment of ciboulot's second β -thymosin repeat (Glu72 to Lys82). This visible portion closely resembled T β 4, in that it bound to the same place on the surface of actin. The last visible ciboulot residue, Lys82, corresponded to Lys31 in T β 4, the first residue of the C-terminal helix. The absence of an ordered C-terminal helix in the G1-Cib23:actin structure may be attributed to the sequence differences between T β 4 and ciboulot: Ser31 and Thr34, in T β 4, are Asp81 and Ala84 in ciboulot, respectively. These substitutions were reasoned to cause destabilization of the helix, leading to partial loss of the actin binding by the C-terminal segment of ciboulot's second β T repeat. In agreement with such a hypothesis, the second repeat in isolation is the weakest in binding to actin when compared with the first and the third ciboulot repeats (Chereau et al., 2005). Thus, this work established that actin stabilization of the C-terminal helix is β T/WH2 module specific.

Structure-based models of T β 4

A structure-based model for the full-length T β 4:actin complex was constructed by combining the G1-T β 4:actin and Cib1N:actin structures (Irobi et al., 2004). The model relied on the conserved central LKKT motif common to the β T/WH2 superfamily. By simply superimposing the actins from these two structures, the N-terminal segment of ciboulot's first β T repeat, from the Cib1N:actin structure, naturally connected, via its LKKT motif, to the C-terminal segment of T β 4 in the G1-T β 4:actin structure. Ciboulot residues were then substituted with the corresponding residues from T β 4 to yield a model of full-length T β 4. Overall, this model is of good credibility, as its construction was based on X-ray structures and involved minimum structural manipulations. Certain details are, however, expected to be less reliable, due to the sequence differences in these two proteins. The model was later further refined to take into account the presence of a helix-breaking proline (Pro4) at the N-terminus of T β 4 (Fig. 1F) (Xue et al., 2007). This proline residue effectively restricted the length of T β 4's N-terminal helix, making the helix sterically more compatible with the formation of a ternary complex composed of actin, T β 4 and profilin (Xue et al., 2014; Yarmola et al., 2001). Sequence alignment

indicates that this proline is generally absent in multi-repeat β Ts and WH2 motifs, but is common among the single-repeat β Ts (Edwards, 2012; Paunola et al., 2002), implying similar functional roles of this residue in other single-repeat β Ts.

Full-length and N-terminal segment of T β 4

Recently, we reported two novel X-ray structures of T β 4:actin complexes, which provided structural information on the actin interactions of full-length T β 4 and, independently, its N-terminal segment (Xue et al., 2014). The full-length T β 4:actin structure was made possible by creating a hybrid protein consisting of full-length human T β 4 fused to *Pichia* actin (PDB ID: 4PL7, Fig. 1G; hereafter referred to as T β 4FL). In this structure there was interpretable electron density for the majority of T β 4 (Asp5 to Ala40), but a lack of electron density for the mobile termini, residues 1-4 and 41-43, respectively. The second structure was a complex between rabbit skeletal actin and a hybrid peptide comprising T β 4's N-terminal segment prior to the LKKT motif (residues 1-16), the lysine-rich region of cordon-bleu (residues 1095-1106), and full-length T β 4 (residues 1-43). Clear electron density was apparent for residues 4-25 for the full-length T β 4 (residues 34-55 in the hybrid, hereafter referred to as T β 4N), which were unambiguously assigned (PDB ID: 4PL8, Fig. 1H).

In good agreement with the NMR studies and the structure-based models, full-length T β 4 in the T β 4FL:actin structure interacts extensively with actin through its two terminal helices and throughout the connecting linker region. The validity of the structure could be crosschecked with the segments of T β 4 in the G1-T β 4:actin and T β 4N:actin structures. Despite the differences in hybrid design and crystallization conditions, the full-length T β 4 in T β 4FL:actin corresponded well with its independently determined N- and C-terminal segments taken from T β 4N:actin and G1-T β 4:actin, respectively (Fig. 2). Not only did the main chains of T β 4 in these structures superimpose precisely, but many side chains also adopted nearly identical conformations, suggesting their critical roles in anchoring of T β 4 to the surface of actin. The first ordered residue in T β 4FL is Asp5, which lies at the beginning of the N-terminal helix, in agreement with the NMR measurements in aqueous solution at 14°C (Czisch et al., 1993). The N-terminus in T β 4N:actin was extended by one residue to Pro4; a residue that has a minimal contact with actin. Thus, these structures defined the full length T β 4 interaction with actin and demonstrated that many of the residues that could be induced to adopt helical conformations in aqueous-organic mixed solvents or at low temperatures are also functional as helices in interacting with actin.

T β 4:actin interface analysis

Here we present details of the T β 4:actin interface based on the analysis of the three available T β 4:actin structures using PISA (Krissinel and Henrick, 2007). Of the 36 ordered T β 4 residues in the T β 4FL:actin structure, Met6, Ile9, Phe12, Leu17, Pro27, Leu28, Pro29 and Ile34 are the major contributors of hydrophobic interactions, whereas Asp5, Glu8, Glu10, Lys14, Lys18, Thr20, Thr22, Gln23, Lys25, Asn26, Lys31, Glu37 and Lys38 engage actin through hydrogen bonds and salt bridges (Fig. 3A); only eight residues (Ala7, Lys11, Asp13, Ser15, Glu32, Glu35, Gln39 and Ala40) have no contribution in actin binding, judged by the absence of buried surface area for these residues. However, as pointed out earlier (Xue et al., 2007), some of the non-contacting residues, for example Asp13 and Ser15, contribute indirectly by stabilizing the actin-bound conformation of T β 4.

An interesting question arose when we considered the T β 4-bound actins in these structures, which are isoforms from either rabbit skeletal muscle or from the yeast, *Pichia pastoris*. Measured tissue distributions of T β 4 in mouse indicates that its concentration in muscle is much lower when compared

with some other tissues such as spleen, brain, liver and lung (Hannappel, 2007). Likewise, EST data, which have been shown to correlate with protein concentrations in human and mouse (Xue and Robinson, 2013), show similar patterns of T β 4 enrichment in mammals and birds (interested readers may check the EST profiles for the individual entries under <http://www.ncbi.nlm.nih.gov/unigene/?term=TMSB4X> for more details). Therefore, the most biologically relevant binding partners for T β 4 in mammals are the cytoplasmic β - and γ 1-actins. *Pichia* actin is cytoplasmic. It is, however, only ~85% identical in sequence to mammalian cytoplasmic actins. Moreover, the *Pichia* actin-T β 4 interaction is non-endogenous, as no β T family members are found in yeasts.

To address the above-question, we performed a multiple sequence alignment of *Pichia*, rabbit skeletal muscle, and human cytoplasmic β - and γ 1-actins. We then compiled all the interface residues according to the PISA analysis of the three T β 4:actin structures, and indicated them alongside the aligned sequences (Fig. 4). Particular attention was paid to the residues involved in forming hydrogen bonds and salt bridges with T β 4, because these interactions are critical in engaging the relatively hydrophilic T β 4 peptide and in defining the specificity between T β 4 and actin. There are altogether 59 interface residues on actin. Using the two human cytoplasmic actins as references, which have identical interfaces, the numbers of non-identical interface residues are only 2 and 7 for the rabbit skeletal muscle and the *Pichia* actins, respectively, indicating the atomic details on T β 4:actin interaction obtained from the rabbit skeletal muscle and *Pichia* isoforms are indeed highly relevant to the mammalian cytoplasmic actins. Moreover, the variations in these two isoforms do not completely overlap, but somewhat complement each other. For the two variations in the rabbit skeletal muscle isoform, M16L and T201V, the former is non-existent in the *Pichia* isoform, whereas the latter becomes a more similar T201S substitution. Conversely, 6 out of the 7 variations found in the *Pichia* isoform are non-existent in the rabbit skeletal muscle isoform. This complementarity allowed us to examine the interactions between T β 4 and human cytoplasmic actin by combining the structural information from the three known structures.

We took the T β 4FL:actin structure as the starting model, and focused on the non-identical interface residues on actin that were involved in hydrogen bonding with T β 4. One region of interest is at the barbed-face of actin, in which Gly350 in *Pichia* actin corresponds to Ser350 in both rabbit skeletal muscle and human cytoplasmic actins. As shown by the comparison between the T β 4N:actin and the T β 4FL:actin structures (Figs. 3F vs. 3E), the presence of the hydroxyl group on the side chain of Ser350 gave rise to one extra hydrogen bond to T β 4's Glu8. Another region of interest lies at the opposite end, the pointed-face of actin, in which *Pichia* actin's Thr199, Ser201 and Ser203 are replaced by a serine and two threonines in human cytoplasmic actins. Given the similarity between serines and threonines, these variations are not expected to change the hydrogen bonding patterns observed in the T β 4FL:actin structure for this region (Fig. 3B). In fact, a threonine at residue 203 had been observed in the G1-T β 4:actin structure to form a hydrogen bond to T β 4's Glu37 (Fig. 3C). It should be pointed out that the overall hydrogen-bonding pattern for the muscle isoform is rather different in this region. The likely cause for the difference is the muscle specific residue, Val201, which eliminates the possibility of a 2.5 Å hydrogen bond observed between *Pichia* actin's Ser201 and T β 4's Lys38, leading to complete disengagement of Lys38 in G1-T β 4:actin. A third region of interest is at residue 28 of actin, which harbors a histidine in *Pichia*, or an arginine in muscle and cytoplasmic actins (Fig. 3D). Despite having longer side chain and greater hydrogen bonding potential, the arginine in T β 4N:actin did not provide an extra hydrogen bond to T β 4, because it was tied up by participating in an intra-molecular hydrogen bond to Asp24 of actin. Taken together, the interactions

between mammalian cytoplasmic actins and T β 4 appear to be captured by the *Pichia* actin isoform structure despite the non-existence of T β 4 in this species.

Conclusion

Tremendous progress has been made during the past two and half decades in elucidating the structures of β T family proteins, and in particular that of the family's archetypal member, T β 4. A body of evidence has emerged, suggesting that two regions within the largely unstructured β T repeat have a higher propensity towards helix formation, which can be induced when the β T repeat executes its biological function in binding and sequestering monomeric actin. How well the C-terminal helix interacts with actin determines whether the β T repeat acts as a suppressor or facilitator of actin polymerization at the fast growing barbed end of actin filaments. Advances in genomics have brought forth a large number of new members into the β T family, which await further structural and functional characterization.

In addition to sequestering actin, T β 4 also participates in numerous biological functions such as stimulating wound healing and tissue repair (Malinda et al., 1999), suppressing inflammation (Young et al., 1999), and enhancing cardiomyocyte regeneration (Qian et al., 2012; Smart et al., 2011). However, the mechanisms behind many of T β 4's hormonal activities remain elusive (Goldstein et al., 2005). Once the pathways and receptors involved in these processes have been identified, new structural data will be required to compare how closely the disorder-to-order transitions mirror those determined for actin binding, and to help drive drug development in these areas.

Acknowledgements

BX and RCR would like to thank the Agency for Science, Technology and Research (A*STAR), Singapore for support.

References

- Aguda, A. H., Xue, B., Irobi, E., Preat, T., and Robinson, R. C. (2006). The structural basis of actin interaction with multiple WH2/beta-thymosin motif-containing proteins. *Structure* **14**, 469-476.
- Andreeva, A., Howorth, D., Chothia, C., Kulesha, E., and Murzin, A. G. (2014). SCOP2 prototype: a new approach to protein structure mining. *Nucleic Acids Res* **42**, D310-314.
- Chereau, D., Kerff, F., Graceffa, P., Grabarek, Z., Langsetmo, K., and Dominguez, R. (2005). Actin-bound structures of Wiskott-Aldrich syndrome protein (WASP)-homology domain 2 and the implications for filament assembly. *Proc Natl Acad Sci U S A* **102**, 16644-16649.
- Czisch, M., Schleicher, M., Horger, S., Voelter, W., and Holak, T. A. (1993). Conformation of thymosin beta 4 in water determined by NMR spectroscopy. *Eur J Biochem* **218**, 335-344.
- De La Cruz, E. M., Ostap, E. M., Brundage, R. A., Reddy, K. S., Sweeney, H. L., and Safer, D. (2000). Thymosin-beta(4) changes the conformation and dynamics of actin monomers. *Biophys J* **78**, 2516-2527.
- Dehal, P., Satou, Y., Campbell, R. K., et al. (2002). The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* **298**, 2157-2167.

- Dhaese, S., Vandepoele, K., Waterschoot, D., Vanloo, B., Vandekerckhove, J., Ampe, C., and Van Troys, M. (2009). The mouse thymosin beta15 gene family displays unique complexity and encodes a functional thymosin repeat. *J Mol Biol* **387**, 809-825.
- Didry, D., Cantrelle, F. X., Husson, C., et al. (2012). How a single residue in individual beta-thymosin/WH2 domains controls their functions in actin assembly. *EMBO J* **31**, 1000-1013.
- Domanski, M., Hertzog, M., Coutant, J., Gutsche-Perelroizen, I., Bontems, F., Carlier, M. F., Guittet, E., and van Heijenoort, C. (2004). Coupling of folding and binding of thymosin beta4 upon interaction with monomeric actin monitored by nuclear magnetic resonance. *J Biol Chem* **279**, 23637-23645.
- Edwards, J. (2004). Are beta-thymosins WH2 domains? *FEBS Lett* **573**, 231-232; author reply 233.
- Edwards, J. (2012). Single-domain β -thymosins: the family history. *Ann N Y Acad Sci* **1269**, 7-15.
- Goldstein, A. L. (2007). History of the discovery of the thymosins. *Ann N Y Acad Sci* **1112**, 1-13.
- Goldstein, A. L., Hannappel, E., and Kleinman, H. K. (2005). Thymosin beta4: actin-sequestering protein moonlights to repair injured tissues. *Trends Mol Med* **11**, 421-429.
- Hannappel, E. (2007). Beta-thymosins. *Ann N Y Acad Sci* **1112**, 21-37.
- Hertzog, M., van Heijenoort, C., Didry, D., et al. (2004). The beta-thymosin/WH2 domain: structural basis for the switch from inhibition to promotion of actin assembly. *Cell* **117**, 611-623.
- Irobi, E., Aguda, A. H., Larsson, M., Guerin, C., Yin, H. L., Burtnick, L. D., Blanchoin, L., and Robinson, R. C. (2004). Structural basis of actin sequestration by thymosin-beta4: implications for WH2 proteins. *EMBO J* **23**, 3599-3608.
- Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797.
- Lee, S. H., Kerff, F., Chereau, D., Ferron, F., Klug, A., and Dominguez, R. (2007). Structural basis for the actin-binding function of missing-in-metastasis. *Structure* **15**, 145-155.
- Malinda, K. M., Sidhu, G. S., Mani, H., Banaudha, K., Maheshwari, R. K., Goldstein, A. L., and Kleinman, H. K. (1999). Thymosin beta 4 accelerates wound healing. *J Invest Dermatol* **113**, 364-368.
- Nag, S., Larsson, M., Robinson, R. C., and Burtnick, L. D. (2013). Gelsolin: the tail of a molecular gymnast. *Cytoskeleton* **70**, 360-384.
- Oldfield, C. J., and Dunker, A. K. (2014). Intrinsically Disordered Proteins and Intrinsically Disordered Protein Regions. *Annual Review of Biochemistry* **83**, 553-584.
- Paunola, E., Mattila, P. K., and Lappalainen, P. (2002). WH2 domain: a small, versatile adapter for actin monomers. *FEBS Lett* **513**, 92-97.
- Qian, L., Huang, Y., Spencer, C. I., Foley, A., Vedantham, V., Liu, L., Conway, S. J., Fu, J.-d., and Srivastava, D. (2012). In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. *Nature* **485**, 593-598.
- Reichert, A., Heintz, D., Echner, H., Voelter, W., and Faulstich, H. (1996). Identification of contact sites in the actin-thymosin beta 4 complex by distance-dependent thiol cross-linking. *J Biol Chem* **271**, 1301-1308.
- Safer, D., Elzinga, M., and Nachmias, V. T. (1991). Thymosin beta 4 and Fx, an actin-sequestering peptide, are indistinguishable. *J Biol Chem* **266**, 4029-4032.
- Safer, D., Golla, R., and Nachmias, V. T. (1990). Isolation of a 5-kilodalton actin-sequestering peptide from human blood platelets. *Proc Natl Acad Sci U S A* **87**, 2536-2540.
- Safer, D., Sosnick, T. R., and Elzinga, M. (1997). Thymosin beta 4 binds actin in an extended conformation and contacts both the barbed and pointed ends. *Biochemistry* **36**, 5806-5816.
- Sickmeier, M., Hamilton, J. A., LeGall, T., et al. (2007). DisProt: the Database of Disordered Proteins. *Nucleic Acids Res* **35**, D786-793.
- Simenel, C., Van Troys, M., Vandekerckhove, J., Ampe, C., and Delepierre, M. (2000). Structural requirements for thymosin beta4 in its contact with actin. An NMR-analysis of thymosin beta4 mutants in solution and correlation with their biological activity. *Eur J Biochem* **267**, 3530-3538.

- Smart, N., Bollini, S., Dube, K. N., et al. (2011). De novo cardiomyocytes from within the activated adult heart after injury. *Nature* **474**, 640-644.
- Stoll, R., Voelter, W., and Holak, T. A. (1997). Conformation of thymosin beta 9 in water/fluoroalcohol solution determined by NMR spectroscopy. *Biopolymers* **41**, 623-634.
- Uversky, V. N. (2002). Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* **11**, 739-756.
- van der Lee, R., Buljan, M., Lang, B., et al. (2014). Classification of intrinsically disordered regions and proteins. *Chem Rev* **114**, 6589-6631.
- Van Troys, M., Dhaese, S., Vandekerckhove, J., and Ampe, C. (2007). Multirepeat β -thymosins. Actin-Monomer-Binding Proteins. Lappalainen, P. (Ed). New York, Springer, 71-81.
- Van Troys, M., Ono, K., Dewitte, D., Jonckheere, V., De Ruyck, N., Vandekerckhove, J., Ono, S., and Ampe, C. (2004). TetraThymosinbeta is required for actin dynamics in *Caenorhabditis elegans* and acts via functionally different actin-binding repeats. *Mol Biol Cell* **15**, 4735-4748.
- Xue, B., Aguda, A. H., and Robinson, R. C. (2007). Models of the actin-bound forms of the beta-thymosins. *Ann N Y Acad Sci* **1112**, 56-66.
- Xue, B., Leyrat, C., Grimes, J. M., and Robinson, R. C. (2014). Structural basis of thymosin- β 4/profilin exchange leading to actin filament polymerization. *Proc Natl Acad Sci U S A* 201412271.
- Xue, B., and Robinson, R. C. (2013). Guardians of the actin monomer. *Eur J Biochem* **92**, 316-332.
- Yarmola, E. G., Parikh, S., and Bubb, M. R. (2001). Formation and implications of a ternary complex of profilin, thymosin beta 4, and actin. *J Biol Chem* **276**, 45555-45563.
- Young, J. D., Lawrence, A. J., MacLean, A. G., Leung, B. P., McInnes, I. B., Canas, B., Pappin, D. J., and Stevenson, R. D. (1999). Thymosin beta 4 sulfoxide is an anti-inflammatory agent generated by monocytes in the presence of glucocorticoids. *Nat Med* **5**, 1424-1427.
- Zarbock, J., Oschkinat, H., Hannappel, E., Kalbacher, H., Voelter, W., and Holak, T. A. (1990). Solution conformation of thymosin beta 4: a nuclear magnetic resonance and simulated annealing study. *Biochemistry* **29**, 7814-7821.

Figures

Fig. 1

Figure 1 Structures of beta-thymosins. The structures were solved by X-ray crystallography unless stated otherwise. Beta-thymosins are shown as red cartoons; actin is shown as a grey surface, with its bound nucleotide shown as green sticks; gelsolin domain 1 (G1) in two of the structures is shown as a blue cartoon. PDB IDs are given following the abbreviations of the structures, whose meanings can be found in the main text. (A) Solution NMR structure of T β 10 (1HJ0). (B) NMR-based model of T β 4:actin (1UY5). (C) Cib1N:actin (1SQK). (D) G1-T β 4:actin (1T44). (E) G1-Cib23:actin (2FF6). (F) Structure-based model of T β 4:actin. (G) T β 4FL:actin (4PL7). (H) T β 4N:actin (4PL8). This and the subsequent structural figures were prepared with the PyMOL Molecular Graphics System (DeLano Scientific LLC).

Fig. 2

Figure 2 Superimposition of the three known Tβ4:actin structures. Tβ4s in Tβ4FL:actin, Tβ4N:actin and G1-Tβ4:actin are colored in green, yellow and violet, respectively. Actin from Tβ4FL:actin is shown as a grey surface.

Fig. 3

Figure 3 Detailed interactions between Tβ4 and actin. (A) The Tβ4:actin interface based on the Tβ4FL:actin structure. Tβ4 is shown as green sticks except for its buried hydrophobic residues, which are colored in chocolate. Actin is shown as a grey surface, with its interface residues shown as light blue sticks. Dashed lines in magenta represent hydrogen bonds and/or salt bridges. (B-E) Differences between the *Pichia* and rabbit skeletal muscle isoforms on the pointed-face (B vs. C), side (D), and the barbed-face (E vs. F) of actin. Tβ4 residues from Tβ4FL, Tβ4N and G1-Tβ4 are colored in green, yellow and violet, respectively. Non-identical actin residues from the *Pichia* and the muscle isoforms are colored in light blue and orange, respectively.

Fig. 4