Structural Studies on the LINC Complex and Fic-1

by

Xuanzong Guo

B.S. Biotechnology
Zhejiang University, China, 2012

SUBMITTED TO THE DEPARTMENT OF BIOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN BIOLOGY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

FEBRUARY 2015

Signature of Author: Signature redacted
Department of Biology
January, 2015

Certified by: Signature redacted
Thomas U. Schwartz
Associate Professor of Biology
Thesis Supervisor

Accepted by: Signature redacted
Amy Keating
Associate Professor of Biology
Chairman, Committee for Graduate Students

Structural Studies on the LINC Complex and Fic-1

by

Xuanzong Guo

Submitted to the Department of Biology on January 16, 2015 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Biology

ABSTRACT

LINC complexes span the nuclear envelope and connect the nucleoskeleton to the cytoskeleton. In 2012, our lab solved the first LINC complex structure, that of SUN domain of human SUN2 bound with KASH1 or KASH2 peptides. In this project testes-specific human SUN proteins (SUN3, SPAG4, and SUN5) were compared to ubiquitously-expressed SUN2. Secondly, fission and budding yeast LINC complexes differ from human ones and were analyzed as well. I was able to confirm SUN-KASH interaction in human and yeast. For structural analysis I explored various expression strategies.

Fic-1 is a *C. elegans* Fic-domain protein with diverse cellular functions. As a subfamily III Fic enzyme, Fic-1 may reveal valuable insights into Fic enzyme mechanisms from its structure. After trying different knowledge-informed constructs and crystal optimization, small Fic-1 crystals were obtained, which diffracted X-rays to ~ 7 Å. With modest additional effort diffraction-quality crystals should be achievable.

Thesis Supervisor: Thomas U. Schwartz
Title: Associate Professor of Biology

1. LINC Complex

1.1 Background

LINC complexes connect nucleoskeleton and cytoskeleton, participating in processes as nuclear anchorage and migration, telomere clustering, and mechanotransduction[1]. They span the nuclear envelope and consist of SUN proteins and KASH proteins. SUN proteins cross the inner nuclear envelope and extend into the perinuclear space as a coiled coil with trimerizing SUN domain at the C-terminal end. KASH proteins cross the outer nuclear membrane and connect SUN domains to the cytoskeleton. In 2012, our lab solved the structure of SUN domain of human SUN2 with KASH1 or KASH2 peptides, which provide molecular insight into how LINC complexes generate stable, force-bearing connection[2].

The SUN domain folds into a beta-sandwich structure, with an N-terminal helix that forms a three-stranded coiled coil with two other SUN domains, effectively generating a SUN homotrimer[3]. KASH peptides bind to a groove between adjacent SUN domains and are stabilized by the ~20 amino-acid KASH lid that extends from the beta-sandwich. Two highly conserved cysteines from the SUN domains form a disulfide bond, generating a ~10 residue loop structure that coordinates a metal ion (Fig. 1).

LINC complexes are conserved from yeast to human. At least one SUN domain protein has been identified in both budding and fission yeast, two are found in worms and fruit flies, and mammals posses at least five family members[4]. Both budding and fission yeast possess two KASH proteins, worms three, and fruit flies two. Six KASH domain proteins have been described in vertebrates. KASH mutations have been associated with autism, hearing loss, cancer, muscular dystrophy and other diseases[5].

1.2 Results

1.2.1 Purification of stoichiometric SUN-KASH complexes.

In fission yeast there are 2 SUN proteins (Sad1 and *Uncharacterized*) and 2 KASH proteins (Kms1 and Kms2). In constructs made by Dr. Brian Sosa, SUN proteins and KASH proteins were cloned into the two expression cassettes of the pETDuet-1 vector. SUN proteins were N-terminally tagged with 6xHis, trimerizing GCN4, and a 3C protease recognition site. KASH proteins were N-terminally tagged with MBP and a 3C protease recognition site. MBP was used for tracking KASH peptides on SDS-PAGE and for solubilizing the complex. Before crystallization, tags were removed using 3C protease.

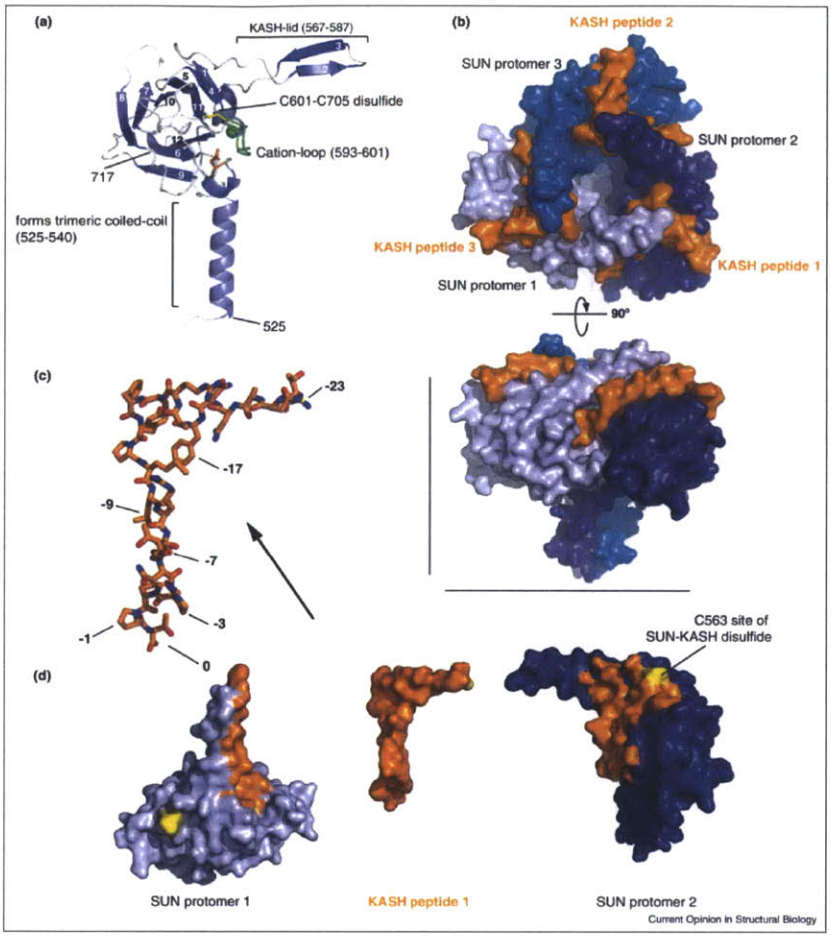

Fig. 1. Structure of SUN domain of SUN2 with KASH2 peptide. (A) Overview of a human SUN2 protomer isolated from its Nesprin-2 binding partners in the trimeric SUN–KASH complex. The protein is organized around a compact β-sandwich core, decorated with features important for function (labeled). Bound cation depicted as a green sphere. (B) View from the ONM facing the bottom of the trimeric SUN2 arrangement (blue colors) with three individual KASH peptides (orange) bound. (C) Side view of the SUN2–KASH2 complex. It is easy to recognize how deeply the three KASH peptides are buried in clefts formed between neighboring SUN2 protomers. (D) Explosion view of the KASH peptide interacting with neighboring SUN domains in the SUN2 trimer. Areas on the SUN domains in close contact with the KASH peptide are highlighted in orange. Note the L-shaped, extended conformation of the bound KASH peptide. Important residues for SUN interaction are labeled in the zoomed, stick representation of KASH2.

BS108	Sad1 (283-514); Kms1 (584-607)
BS109	Sad1 (283-514); Kms2 (432-457)
BS231	*Uncharacterized* (391-601); Kms1 (584-607)
BS232	*Uncharacterized* (391-601); Kms2 (432-457)

The two SUN proteins and the two KASH proteins bind to each other during Ni column elution (Fig. 2A), proving that the last 24 amino acids of Kms1 and last 26 amino acids of Kms2 are sufficient for SUN binding. However, when KASH peptides were truncated to last 9 amino acids of Kms1 and last 10 amino acids of Kms2, SUN-KASH binding was abolished (Fig. 2B). Further experiment showed that the last 17 amino acids of Kms1 are sufficient for binding to Sad1.

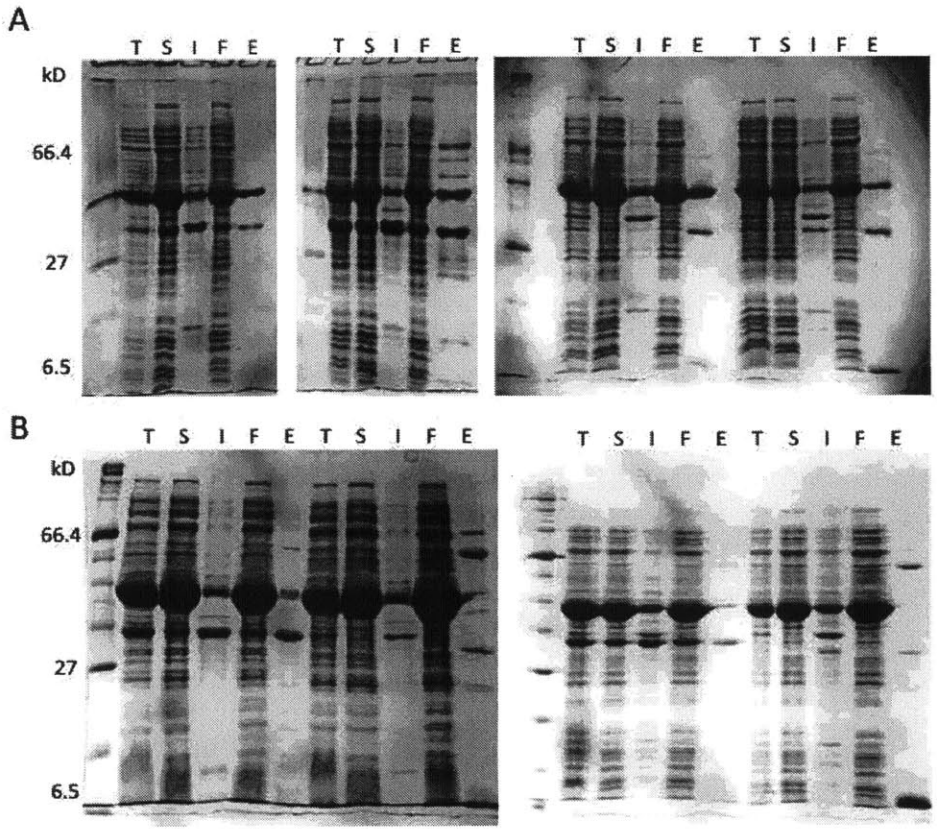

Fig. 2. Ni column purification of BS108, BS109, BS231, BS232, XG1, XG2, XG3, and XG4. (A) Ni column purification of BS108, BS109, BS231, and BS232 (left to right). In the elution fraction, the upper band is MBP-KASH, and the lower band is His-SUN. (B) Ni column purification of XG1, XG3, XG2, and XG4 (left to right).

XG1	Sad1 (283-514); Kms1 (599-607)
XG2	Sad1 (283-514); Kms2 (448-457)
XG3	*Uncharacterized* (391-601); Kms1 (599-607)
XG4	*Uncharacterized* (391-601); Kms2 (448-457)
XG5	Sad1 (283-514); Kms1 (591-607)
XG6	Sad1 (283-514); Kms2 (441-457)
XG7	*Uncharacterized* (391-601); Kms1 (591-607)
XG8	*Uncharacterized* (391-601); Kms2 (441-457)

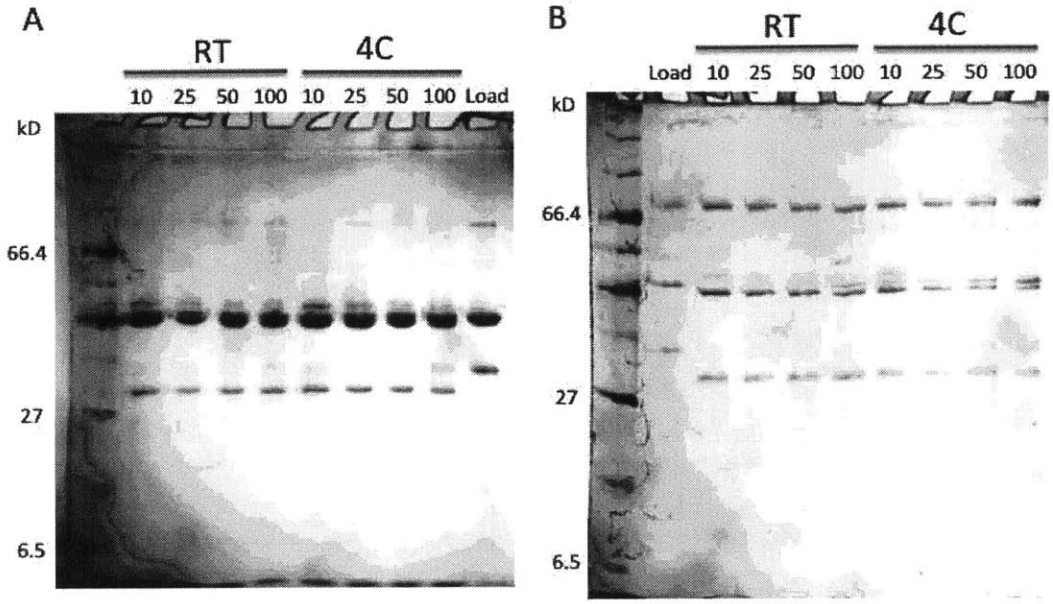

Fig. 3. 3C cutting test of BS108 and BS109. (A) 3C cutting test of BS108. RT denotes at room temperature, 4C for 4 degrees Celsius. The number above each lane represents the mass ratio of substrate to enzyme, e.g., 10 means that for 1 mass of enzyme, 10 masses of substrate were put into the reaction. The amount of substrate was kept constant for comparison. (B) 3C cutting test of BS109.

After LINC complexes were purified from the cell lysate, a problem arose as MBP tag turned out to be resistant to cutting in all existing constructs (Fig. 3&4). I speculated that the way that SUN and KASH bind precludes accessibility of 3C protease. To solve this problem, I used different strategies.

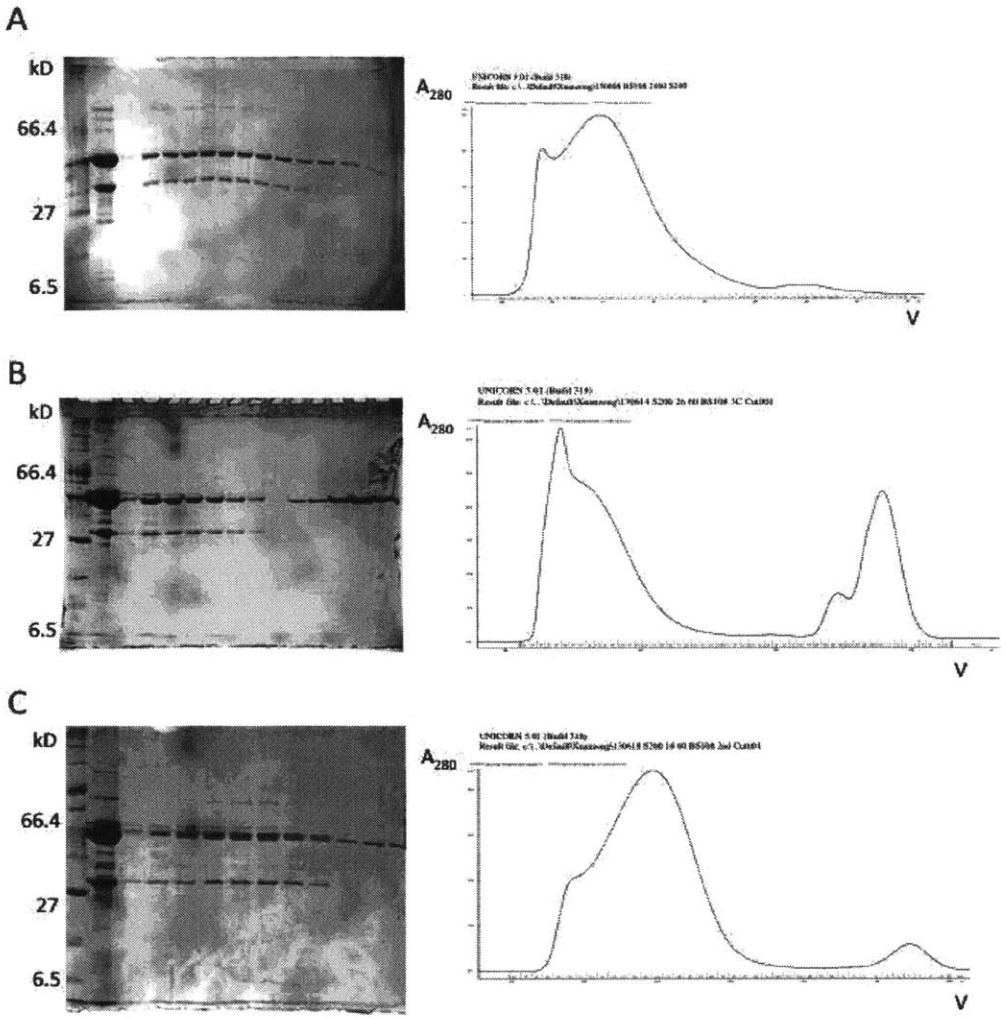

Fig. 4. Gel filtration of uncut, once-cut, and twice-cut samples of BS108. (A) On the left is an SDS-PAGE analysis of elution profile of uncut BS108 sample gel filtration; on the right is the chromatogram of the gel filtration. (B) Taking major peak fraction from (A) and do a 1:10 (enzyme:substrate) cut. (C) Taking major peak fraction from (B) and do another 1:10 cut.

Replacing MBP with sfGFP[6] (super folder green fluorescent protein, an engineered version of GFP that enhances folding when expressed in bacteria) didn't solve the problem, as GFP seemed to come off autonomously in Ni elution and the suspected GFP could not be separated from SUN-KASH complex via gel filtration (Fig. 5). Detergents Tween 20, Triton X-100, NP-40, and DM were tested as a means to improve 3C cutting efficiency, but to no avail (Fig. 6). While adding an N-terminal linker to the 3C cleavage site didn't improve cutting either, extending the C-terminal linker slightly improved cutting (Fig. 7). This supports the idea that binding of KASH peptides blocks the protease from accessing the cleavage site. However, further increasing the length of the linker did not enhance cutting. Using a much smaller tag GB1 does not bring complete cutting of KASH peptides either (Fig. 8). Changing the cut site to SUMO brought about expression issues: SUMO-tagged Kms1 did not express well, and Sad1 became insoluble (Fig. 9). Leaving KASH peptides untagged, on the other hand, rendered SUN proteins completely insoluble (Fig. 10).

XG9	Sad1 (283-514); GFP:Kms1 (584-607)
XG10	Sad1 (283-514); GFP:Kms2 (432-457)
XG11	*Uncharacterized* (391-601); GFP:Kms1 (584-607)
XG12	*Uncharacterized* (391-601); GFP:Kms2 (432-457)
XG15	Sad1 (283-514); LINK:3C:Kms1 (584-607)
XG16	Sad1 (283-514); LINK:3C:Kms2 (432-457)
XG17	*Uncharacterized* (391-601); LINK:3C:Kms1 (584-607)
XG18	*Uncharacterized* (391-601); LINK:3C:Kms2 (432-457)
XG19	Sad1 (283-514); 3C:LINK:Kms1 (584-607)
XG20	Sad1 (283-514); 3C:LINK:Kms2 (432-457)
XG21	*Uncharacterized* (391-601); 3C:LINK:Kms1 (584-607)
XG22	*Uncharacterized* (391-601); 3C:LINK:Kms2 (432-457)
XG27	Sad1 (283-514); Kms1 (584-607)
XG28	Sad1 (283-514); Kms2 (432-457)
XG29	*Uncharacterized* (391-601); Kms1 (584-607)
XG30	*Uncharacterized* (391-601); Kms2 (432-457)
XG31	Sad1 (283-514); GB1:Kms1 (584-607)
XG32	Sad1 (283-514); GB1:Kms2 (432-457)
XG33	*Uncharacterized* (391-601); GB1:Kms1 (584-607)
XG34	*Uncharacterized* (391-601); GB1:Kms2 (432-457)
XG101	Sad1 (283-514); SUMO:Kms1 (584-607)

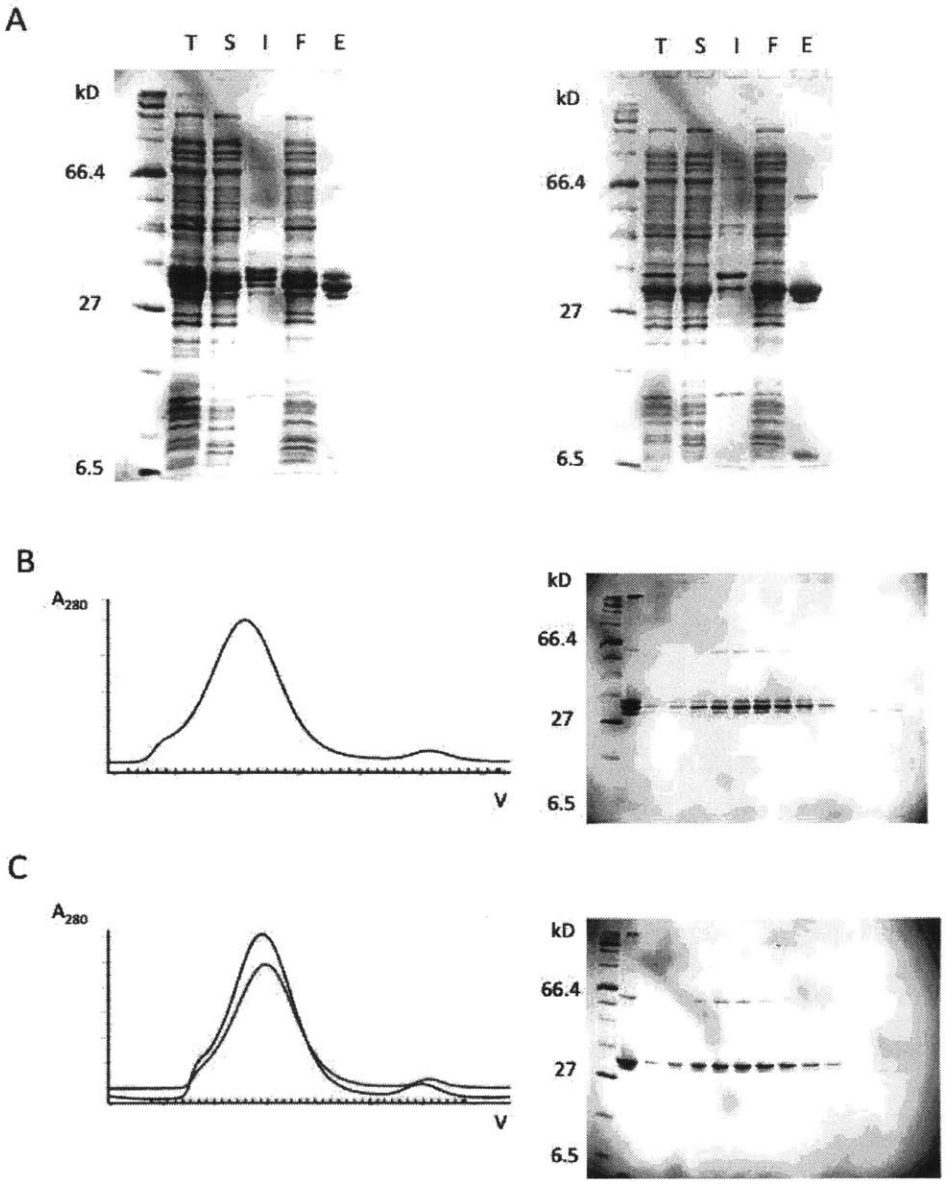

Fig. 5. Ni column and gel filtration purification of GFP-tagged constructs XG9 and XG11. (A) Ni column purification of XG9 and XG11. (B) Gel filtration purification of XG9. (C) Gel filtration purification of XG11.

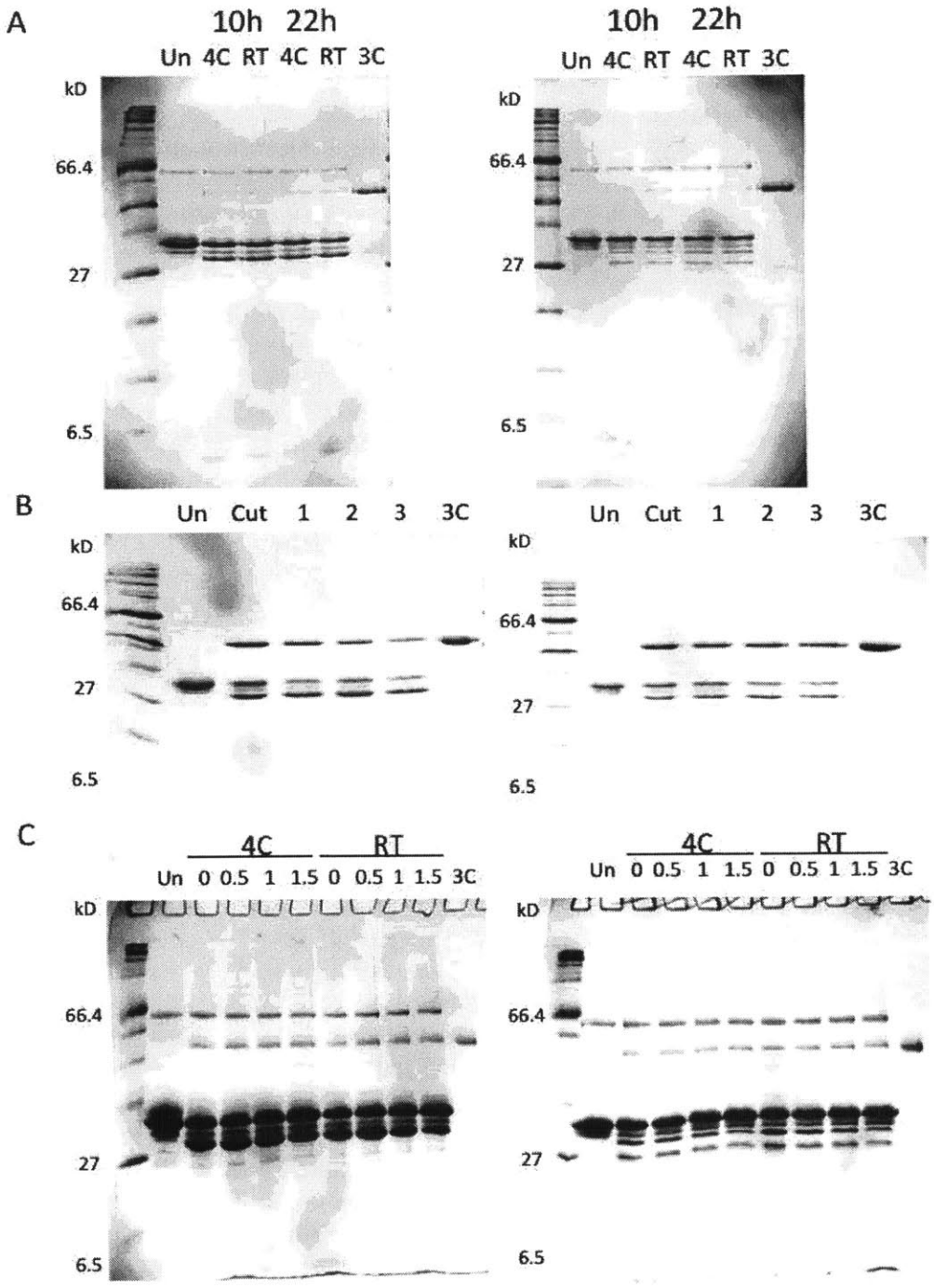

Fig. 6. 3C cutting test of XG9 and XG11. (A) Time and temperature on cutting. substrate:enzyme (mass ratio) = 20. (B) Detergent on 6-hour cutting at room temperature with 1:20 enzyme. 1: 0.1% Tween 20. 2: 0.1% NP 40. 3: 0.1% Triton X-100. (C) Detergent DM on 6-hour cutting at room temperature with 1:20 enzyme. 1: 0.3% DM.

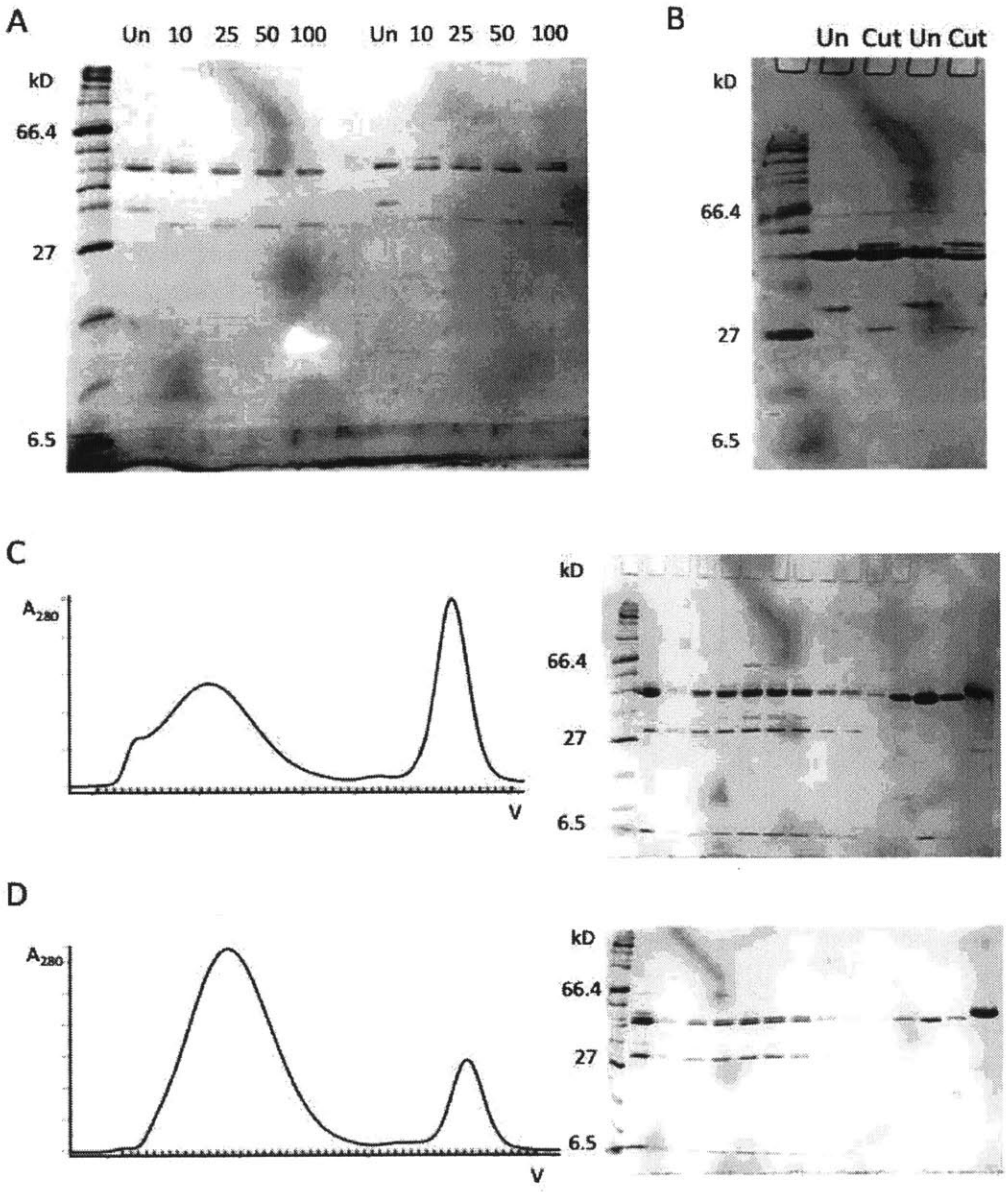

Fig. 7. 3C cutting test and gel filtration purification of linker-extending constructs. (A) 3C cutting test of XG15 (left) and XG19 (right). Number denotes mass ratio of substrate to enzyme. (B) 3C cutting test of XG17 (left) and XG21 (right). (C) S200 26 60 gel filtration purification of cut XG19. (D) S200 26 60 gel filtration purification of re-cut XG19 pooled from major peak in (C).

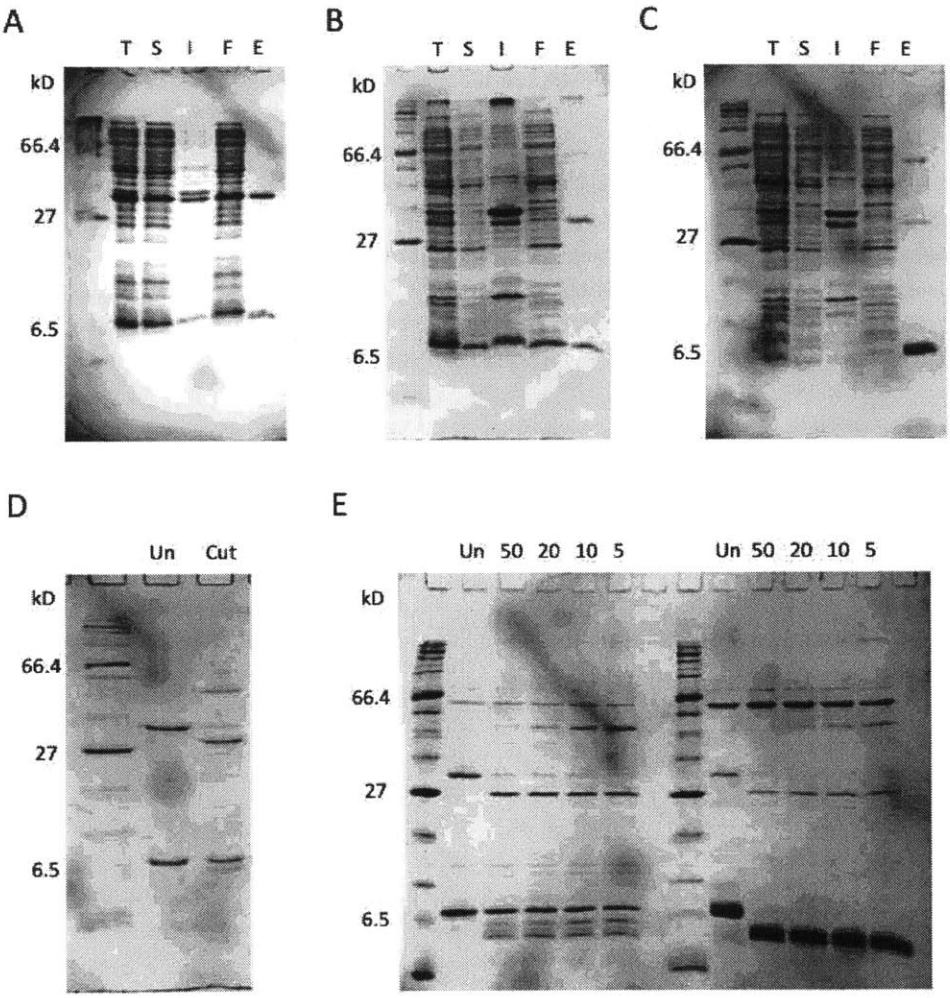

Fig. 8. Ni column purification and 3C test cuts of GB1-tagged KASH constructs. (A) Ni column purification of XG31. (B) Ni column purification of XG33. (C) Ni column purification of XG34. (D) Cutting test of XG31 with 1:10 mass ratio of substrate to enzyme. (E) Cutting test of XG33 and XG34. The number above the lane indicates mass ratio of substrate to enzyme.

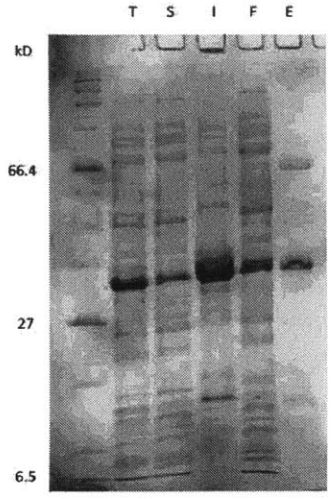

Fig. 9. Ni column purification of XG101.

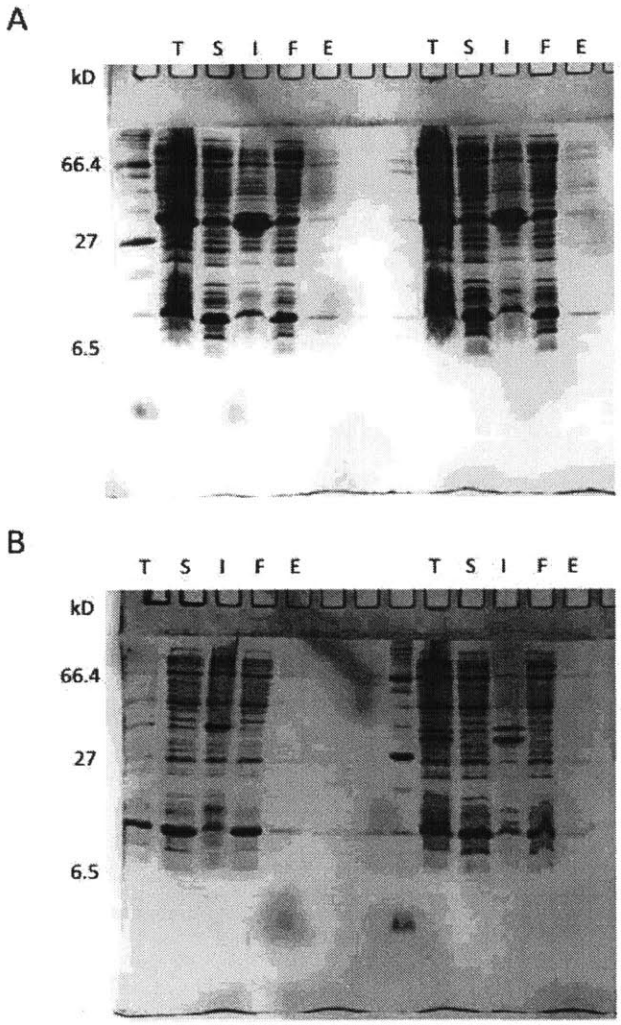

Fig. 10. Ni column purification of XG27, XG28, XG29, and XG30. (A) Ni column purification of XG27 and XG28. (B) Ni column purification of XG29 and XG30.

1.2.2 Apo-SUN proteins are insoluble under tested conditions

For apo-SUN structures of Sad1 and *Uncharacterized*, expressing SUN domains alone rendered them completely insoluble (Fig. 11A). Deleting the trimerizing GCN4 (in case that Sad1 and *Uncharacterized* do not trimerize) did not help (Fig. 11B).

XG23	Sad1 (283-514)
XG24	*Uncharacterized* (391-601)
XG58	Sad1 (283-514) w/o GCN4
XG59	*Uncharacterized* (391-601) w/o GCN4

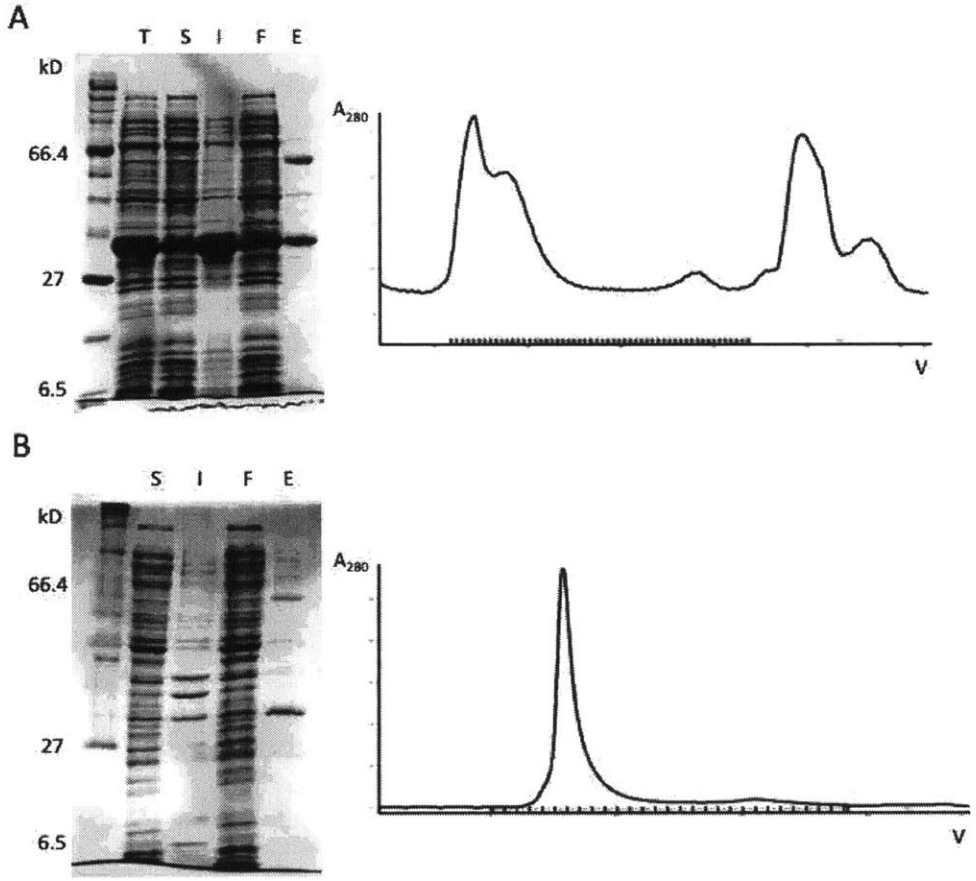

Fig. 11. Apo-SUN constructs with or without GCN4 trimerizing domain. (A) Ni column purification of XG23 and gel filtration purification of cut XG23. (B) Ni column purification of XG58 and gel filtration purification of cut XG58.

In human, there are 5 SUN proteins (SUN1, SUN2, SUN3, SPAG4, and SUN5) and 6 KASH proteins (Nesprin-1, Nesprin-2, Nesprin-3, Nesprin-4, KASH5, and LRMP). SUN1 and SUN2 are ubiquitously expressed, while SUN3, SPAG4, and SUN5 are only expressed in testes. Since the structure of SUN2-KASH1/2 has been solved by the lab, focus has been given to the testes specific SUN proteins, i.e., SUN3, SPAG4, and SUN5.

In contrast to apo-SUN2 and despite great conservation, apo-SUN3, -SPAG4, and -SUN5 were insoluble (Fig. 12). Based on results from XG42 (SUN3; KASH5), GB1 seemed to be able to solubilize SUN domains, but proved hard to cut as well (Fig. 13). A new strategy for co-expression was co-transformation of SUN-bearing and KASH-bearing plasmids with selection by two antibiotics. GCN4 was absent from SUN constructs and KASH peptides were tagged with SUMO. Expression of KASH peptides alone was successful, but during co-expression, KASH peptides seemed not expressed and SUN was insoluble (XG82, XG83, XG84, XG85) (Fig. 14). Cloning SUN and KASH into the same plasmid yielded similar results (Fig. 15).

XG35	SUN3 (161-357)
XG36	SPAG4 (235-437)
XG37	SUN5 (176-379)

XG102	SUMO:Nesprin-1 (8769-8797); SUN3 (161-357)

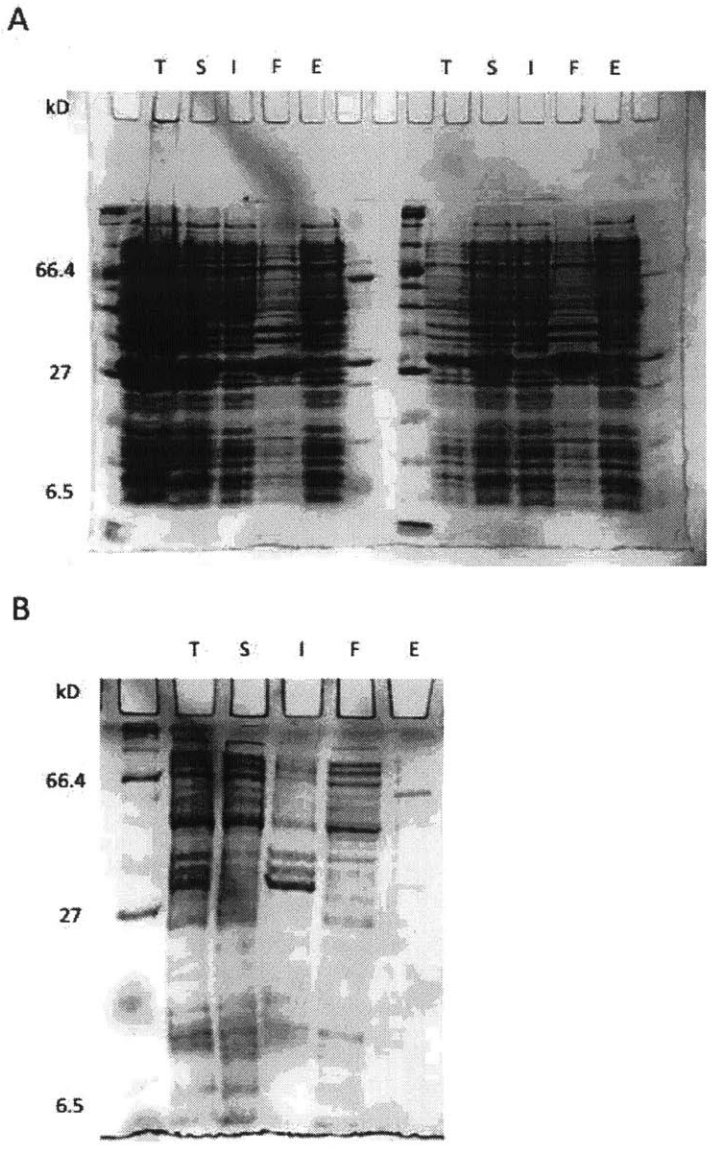

Fig. 12. Ni column purification of SUN3, SPAG4, and SUN5. (A) Ni column purification of XG35 (SUN3, left) and XG37 (SUN5, right). (B) Ni column purification of XG36 (SPAG4).

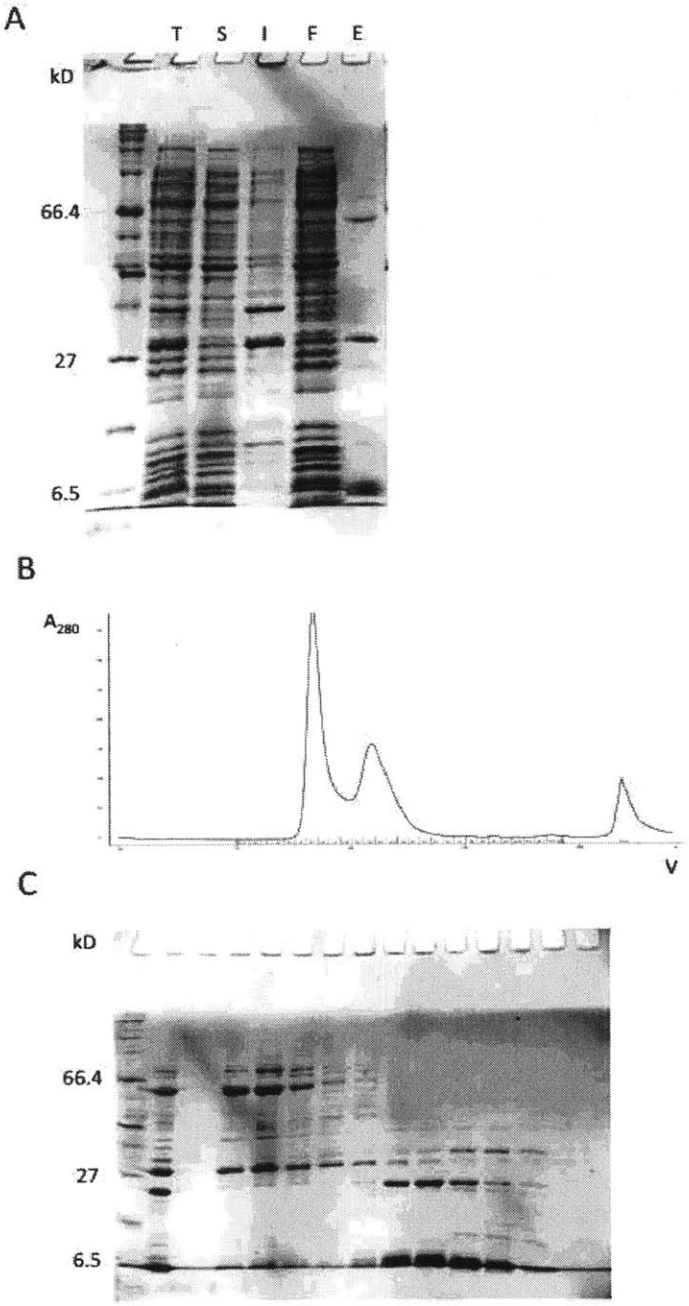

Fig. 13. Ni column and gel filtration purification of XG42. (A) Ni column purification of XG42. (B) Chromatogram of S75 10 300 gel filtration purification of XG42. (C) SDS-PAGE analysis of S75 10 300 gel filtration purification of XG42.

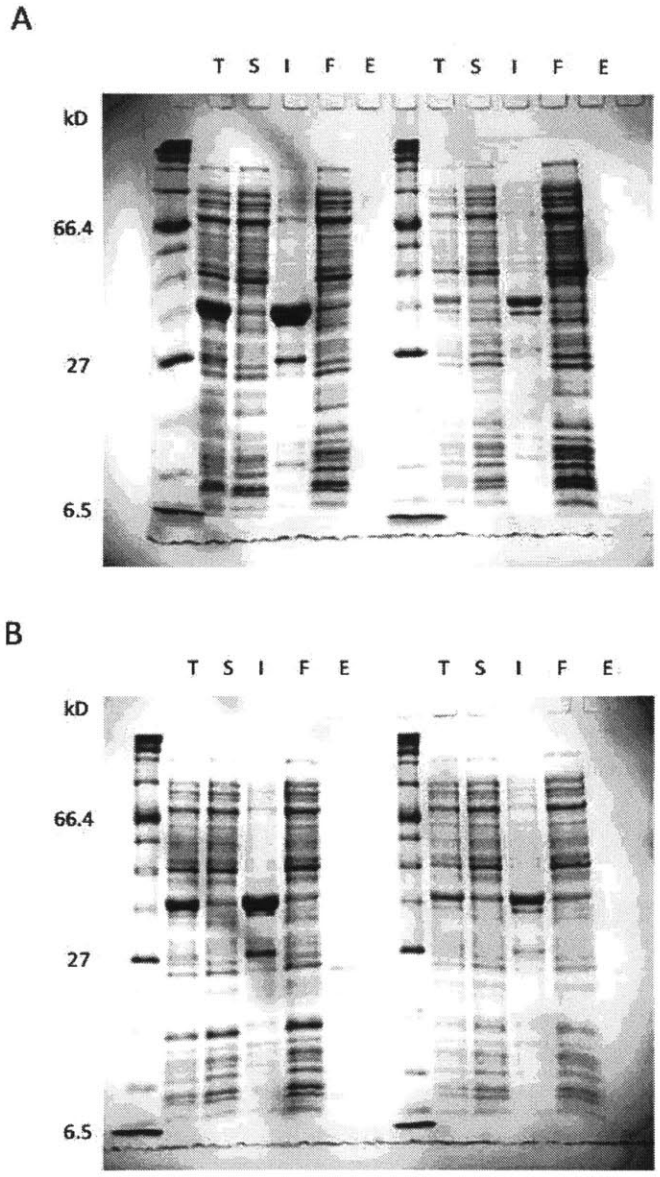

Fig. 14. Ni column purification of XG82, XG83, XG84, and XG85. (A) Ni column purification of XG82 and XG83. (B) Ni column purification of XG84 and XG85.

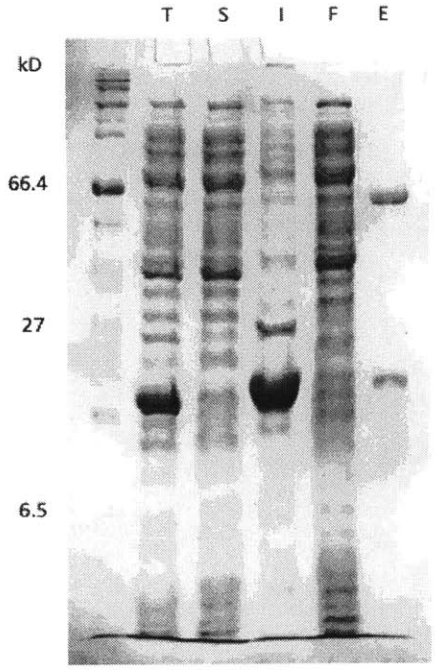

Fig. 15. Ni column purification of XG102.

2. Fic-1

2.1 Background

Fic enzymes are found across from all domains of life. They participate in eukaryotic signal transduction, drug tolerance, bacterial pathogenicity, and bacterial stress response by post-translationally modifying proteins through AMPylation, UMPylation, phosphorylation, or phosphocholination[7]. Fic enzymes transfer parts of a nucleotide cofactor onto target proteins, which include eukaryotic GTPases, unstructured protein segments, and bacterial enzymes[7, 8, 9]. Flies lacking Fic enzymes are viable and fertile, but blind[10]. Fic proteins share a conserved active site, a loop supported by six alpha-helices (Fig. 16). With emerging structures of Fic enzymes, it is now possible to understand how a conserved active site can recognize different substrates and perform various functions[11, 12, 13, 14]. One way that Fic enzymes can perform various functions was understood by the study of Fic protein Doc, which uses an inverted substrate to phosphorylate and inactivate EF-Tu[15].

Based on domain complexity, Fic enzymes are classified into three subfamilies[7]. Subfamily I has minimal Fic enzyme features and is represented by toxin Doc from phage P1. Subfamily II has insertions at various positions of the minimal Fic domain and is represented by *E. coli* Fic enzymes. Subfamily III includes HYPE enzymes typical of animal Fic proteins, and they usually contain a transmembrane helix and several tetratricopeptide repeats (TPR).

An inhibitory strategy is employed to regulate Fic enzymes, i.e., an intermolecular or intramolecular alpha helix (S/T)XXXE(G/N) binds to the Fic active site in competition with the nucleotide[16]. The glutamate in the inhibitory helix is highly conserved and interacts with the second arginine of the conserved Fic motif. Truncation of the inhibitory glutamate allows ATP to bind in a hydrolysis competent orientation[17]. However, how Fic enzymes are activated is currently unknown.

Currently, almost all Fic enzymes with known structures belong to subfamilies I and II. In December 2014, the first eukaryotic Fic enzyme structure (HYPE from human) was published, and it was found that the TPR domain serves as a dimerization domain[18].

C. elegans Fic-1 (508 aa) belongs to subfamily III and its has a N-terminal transmembrane helix (aa45-63), two TPR repeats (aa147-214), an inhibitory helix (aa270-275) and a Fic domain (aa326-461). Fic-1 mutant E274G is constitutively active.

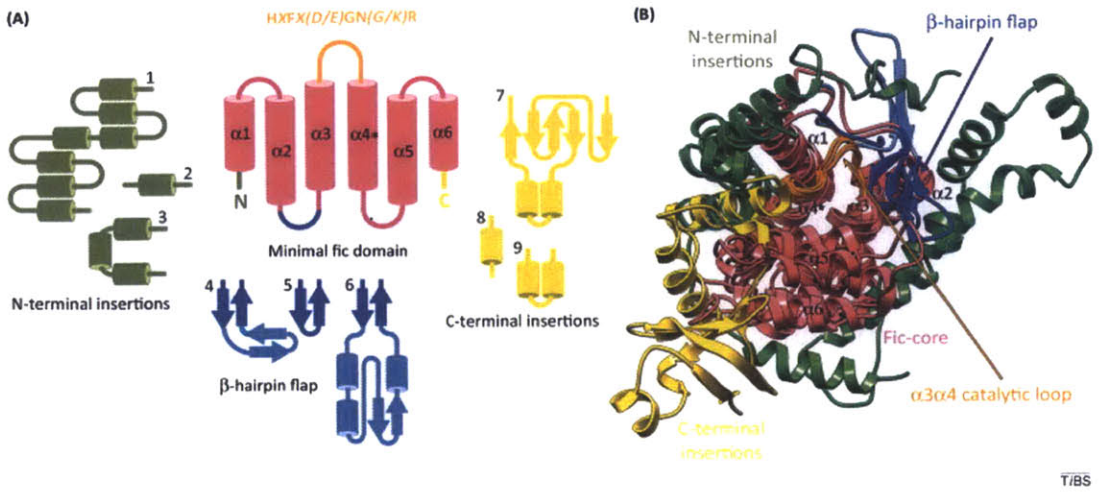

Fig. 16. The Fic core and functional extensions. (A) A minimal Fic domain, as observed in the toxin Doc (coloured raspberry). The six α-helices constituting the domain are labelled in black and the regulatory helix α4 is marked with an *. The three major classes of functional extensions, and their grafting sites at the N terminus, C terminus, and α1α2 loop (or β-hairpin flap), are coloured green, yellow, and blue, respectively. Additional structural elements from the β-hairpin flap are light blue. Examples of specific extensions are shown to the left, right, and below the minimal Fic domain: (1) N-terminal lobe of $IbpA_2$ (immunoglobulin binding protein A of*Histophilus somni*); (2) CDP–choline binding domain from AnkX (ankyrin repeat-containing protein; *Legionella pneumophila*); (3) N-terminal α helix of NmFic (*Neisseria meningitidis*); (4) BepA (*Bartonella henselae*) β-hairpin flap with an additional β-hairpin insertion; (5) NmFic β-hairpin flap; (6) AnkX insert domain protruding from the β-hairpin flap; (7) BepA C-terminal domain; (8) NmFic C-terminal α helix; and (9) AnkX ankyrin tandem repeat. (B) Superposition of four representative Fic enzymes represented as cartoons, Doc (death on curing; bacteriophage P1), NmFic (*N. meningitidis*), BepA (*B. henselae*), and $IbpA_2$ (*H. somni*). The different structural elements of the enzymes are coloured according to panel (A). (from Garcia-Pino et al., 2014)

2.2 Results

2.2.1 A few truncation constructs of Fic-1 yielded well-behaving proteins.

To find a crystal-yielding Fic-1, I made a series of truncation constructs. Fic-1 (65-508) WT and E274G were derived from the complete sequence after the transmembrane helix. They were perfectly soluble, and appeared as a single sharp monodisperse peak in gel filtration (Fig. 17). Oddly, wild-type construct can only be concentrated to 1.56 mg/ml while mutant construct to 25 mg/ml. Fic-1 (134-508) WT and E274G were made based on limited proteolysis. Both turned out perfectly soluble, and on gel filtration appeared as a single monodisperse peak (Fig. 18).

Interestingly, wild-type can be concentrated to 70 mg/ml, while mutant to 7.5 mg/ml. Fic-1 (145-508) WT and E274G included the TPR domain in addition to the Fic domain, since the two domains were adjacent in homology modeling and could be interacting. Both were quite soluble, but in gel filtration they either appeared as a double peak (WT) or a broad trailing peak (mutant) (Fig. 19), suggesting oligomerization, conformational heterogeneity, or both. Fic-1 (234-508) WT and E274G were made based on secondary structure prediction, and they included a short stretch of alpha-helix that seemed to be part of the domain. Fic-1 (234-508) WT and E274G turned out to be mostly insoluble in Ni purification (Fig. 20). In a subsequent gel filtration, the constructs showed inhomogeneous behavior, and therefore were not used further. Fic-1 (258-508) WT and E274G behaved well in biochemical essays (Dr. Matthias Truttmann) and were quite soluble and essentially pure after Ni column. On gel filtration, they both appeared as a trailing peak suggesting multiple conformations or oligomeric states (Fig. 21). They were concentrated to 20 mg/ml.

To sum up, both wild-type and mutant constructs of Fic-1 (65-508), Fic-1 (134-508), and Fic-1 (258-508) behaved well during purification and were worth pursuing for crystallization.

M1	Fic-1 (1-508)
M4	Fic-1 (1-508) E274G
XG76	Fic-1 (65-508)
XG77	Fic-1 (65-508) E274G
XG99	Fic-1 (134-508)
XG100	Fic-1 (134-508) E274G
XG78	Fic-1 (145-508)
XG79	Fic-1 (145-508) E274G
XG56	Fic-1 (234-508)
XG57	Fic-1 (234-508) E274G
M3	Fic-1 (258-508)
M2	Fic-1 (258-508) E274G

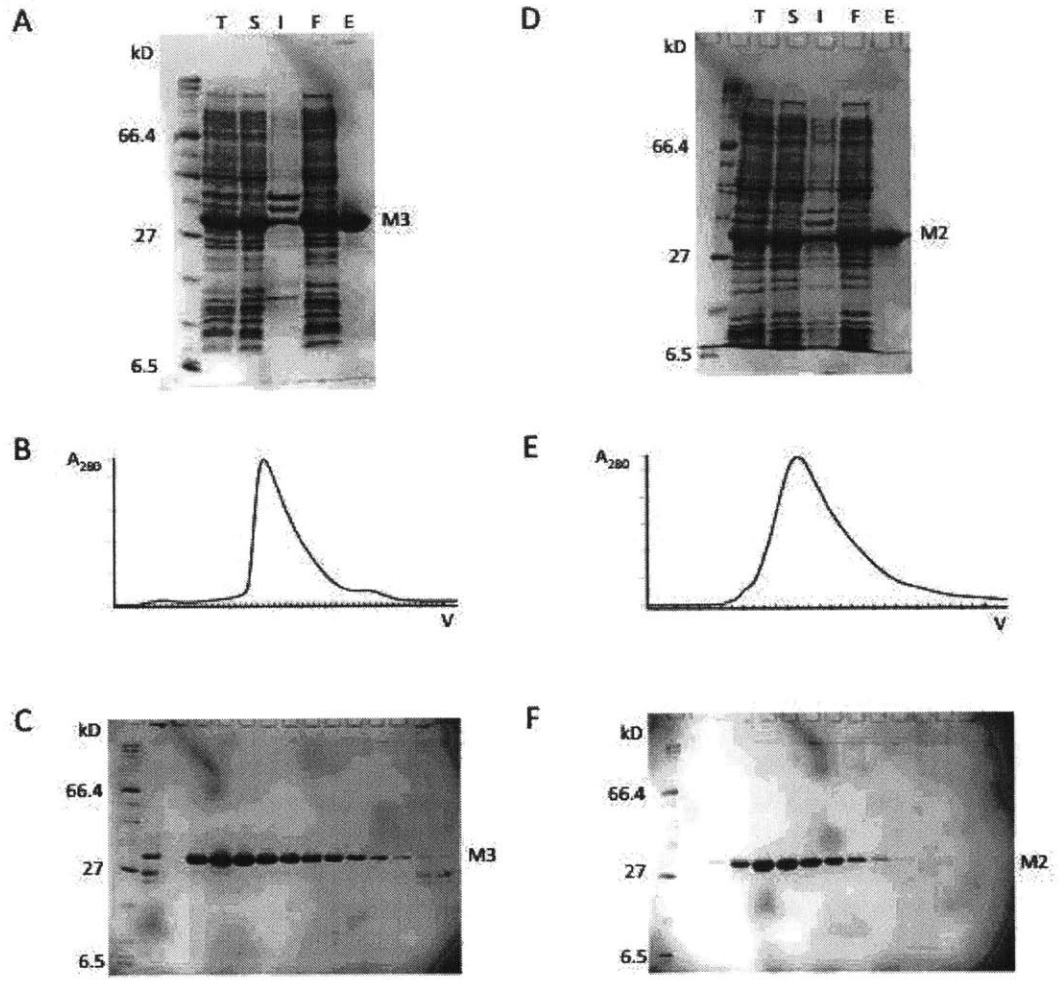

Fig. 17. Ni Column and gel filtration purification of M3 and M2. (A) Ni column purification of M3. T stands for total fraction from cell lysis, S for soluble fraction after centrifugation, I for insoluble fraction, F for flow-through fraction from Ni column, and E for elution fraction from Ni column. (B) Chromatogram of M3 on S200 26 60. The blue curve is the sample absorption at 280 nm over elution volume. (C) SDS-PAGE analysis of peaks in (B). The first lane after marker is sample before gel filtration. (D) Ni column purification of M2. (E) Chromatogram of M3 on S200 10 300. (F) SDS-PAGE analysis of peaks in (E).

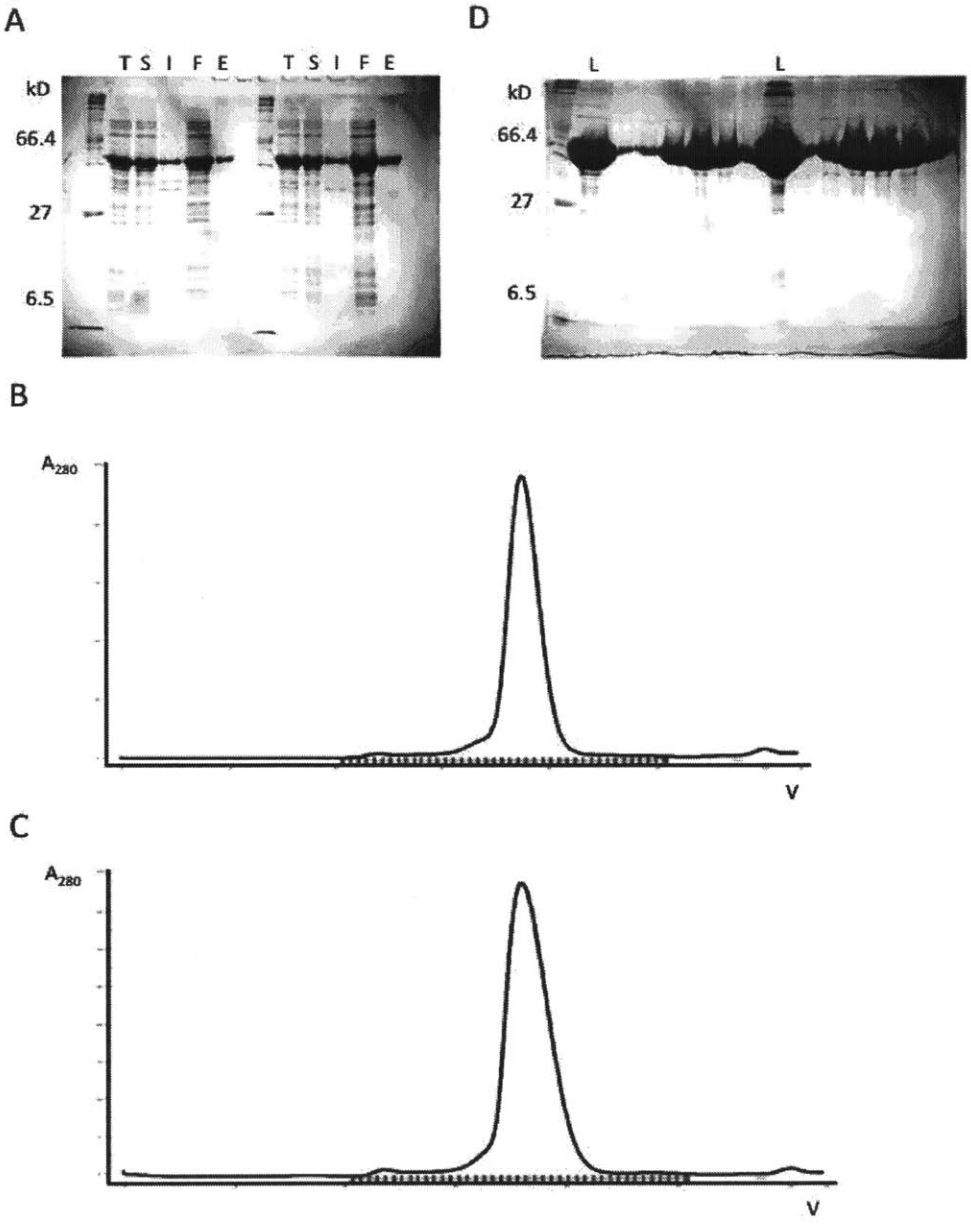

Fig. 18. Ni column and gel filtration purification of XG76 and XG77. (A) Ni column purification of XG76 (left) and XG77 (right). (B) S200 16 60 gel filtration chromatogram of XG76. (C) S200 16 60 gel filtration chromatogram of XG77. (D) SDS-PAGE analysis of gel filtration elution profile of XG76 (left) and XG77 (right).

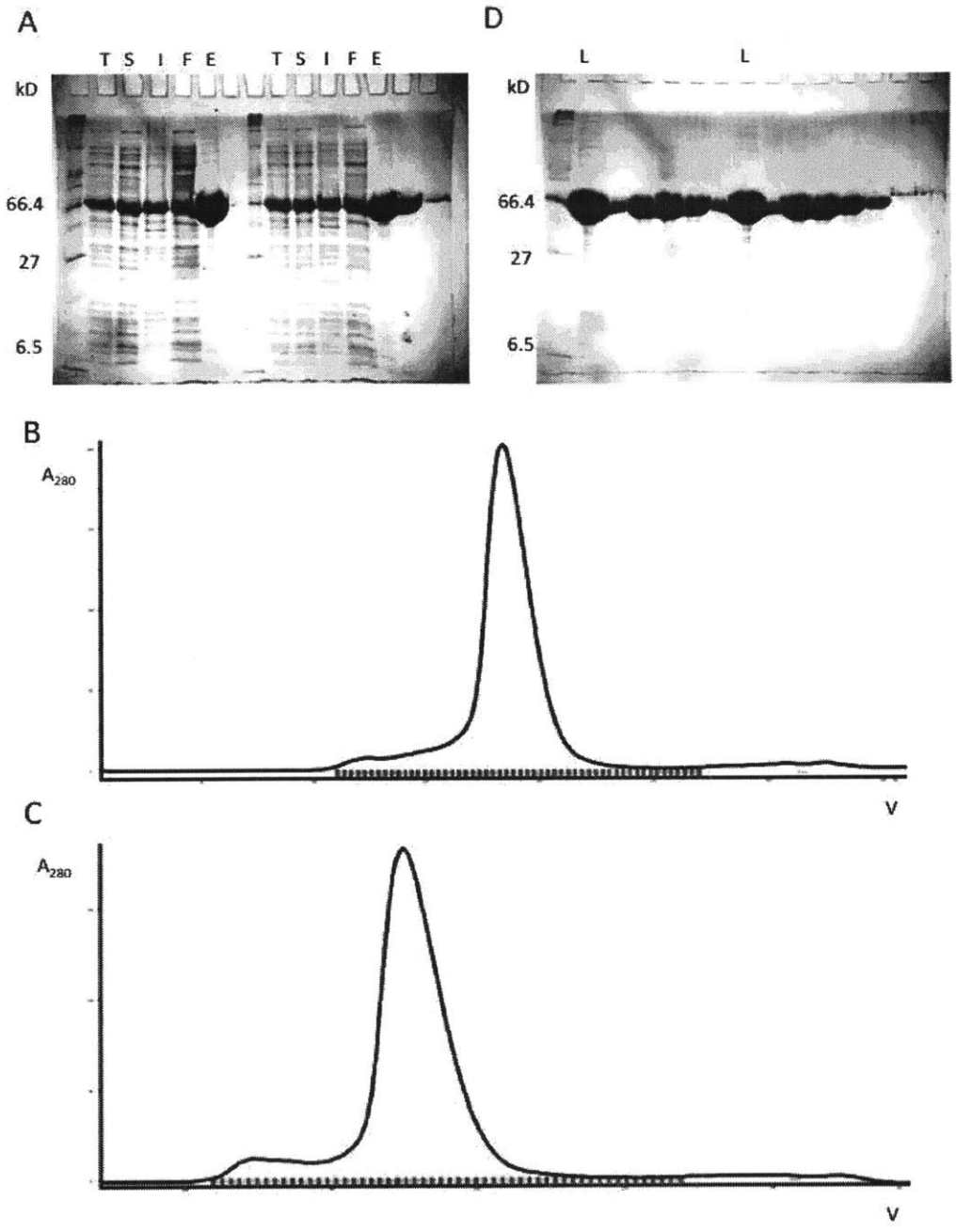

Fig. 19. Ni column and gel filtration purification of XG99 and XG100. (A) Ni column purification of XG99 (left) and XG100 (right). (B) Elution profile of S200 26 60 gel filtration of XG99. (C) Elution profile of S200 26 60 gel filtration of XG100. (D) SDS-PAGE analysis of gel filtration of XG99 (left) and XG100 (right).

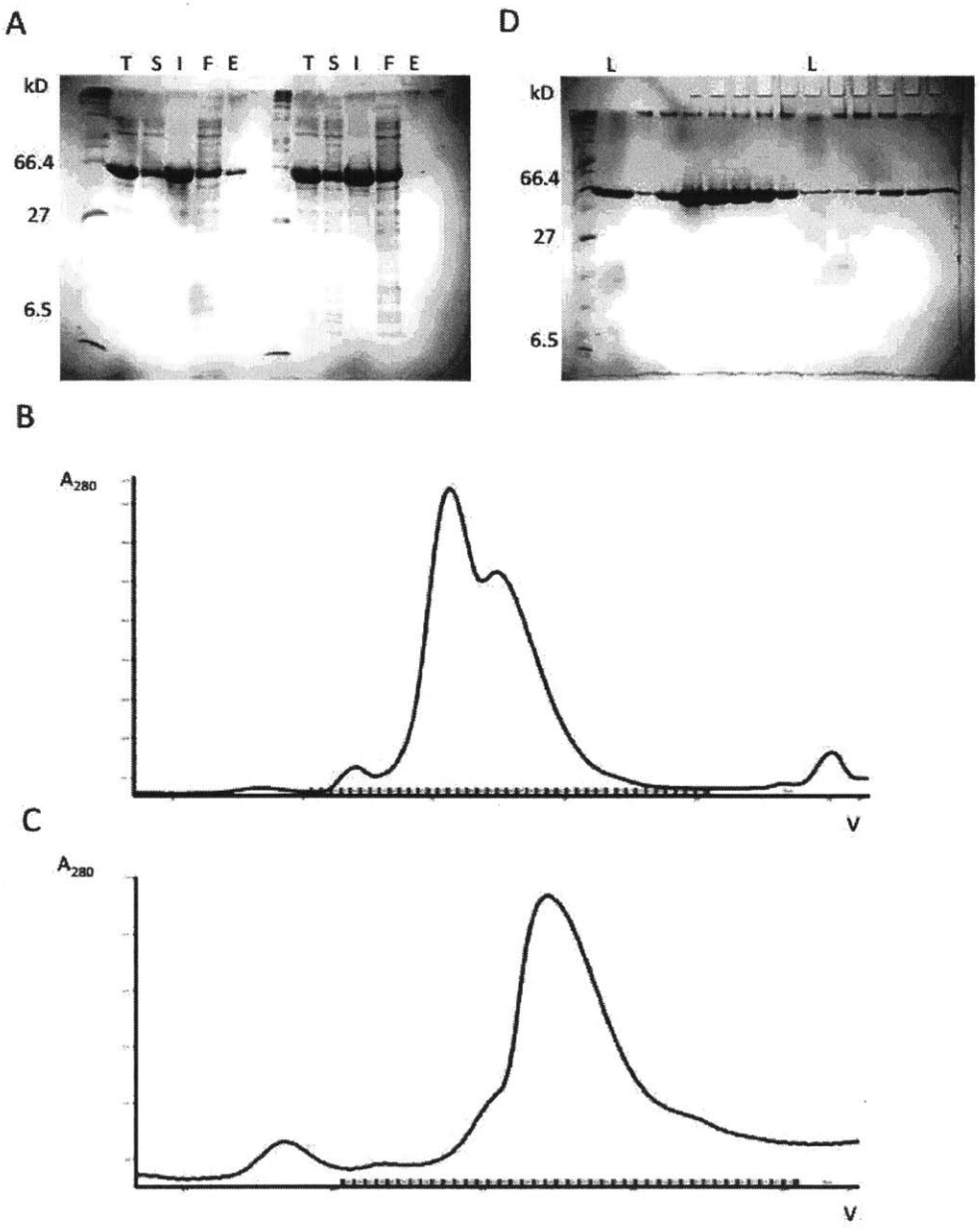

Fig. 20. Ni column and gel filtration purification of XG78 and XG79. (A) Ni column purification of XG78 (left) and XG79 (right). (B) S200 16 60 gel filtration chromatogram of XG78. (C) S200 16 60 gel filtration chromatogram of XG79. (D) SDS-PAGE analysis of gel filtration elution profile of XG78 (left) and XG79 (right).

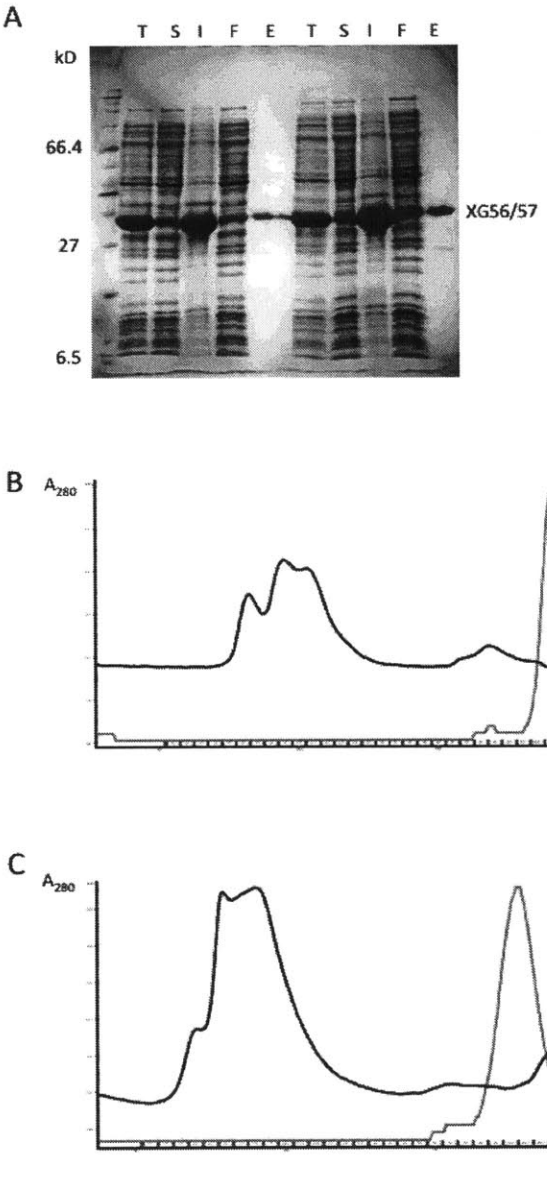

Fig. 21. Ni column and gel filtration purification of XG56 and XG57. (A) Ni column purification of XG56 (left) and XG57 (right). (B) Chromatogram of XG56 on S75 10 300. The dark blue curve is the sample absorption at 280 nm over elution volume, and light blue conductance. (C) Chromatogram of XG57 on S75 10 300.

2.2.2 Limited proteolysis suggested that TPR and Fic domains were packed.

Limited proteolysis is a powerful way to find stable folds within proteins, thus a limited proteolysis analysis of Fic-1 (65-508) was performed.

First, the kind and optimal amount of proteases to perform the assay with was determined. Trypsin, chymotrypsin, and V8 proteases were tested in serial dilutions to partially cut purified XG76, and the digests were analyzed on a gel. 81-fold diluted trypsin stock (0.1 OD) proved sufficient to produce a stable fragment pattern, while for chymotrypsin, 3-fold dilution of stock (0.1 OD), and V8 protease, stock (Fig. 22).

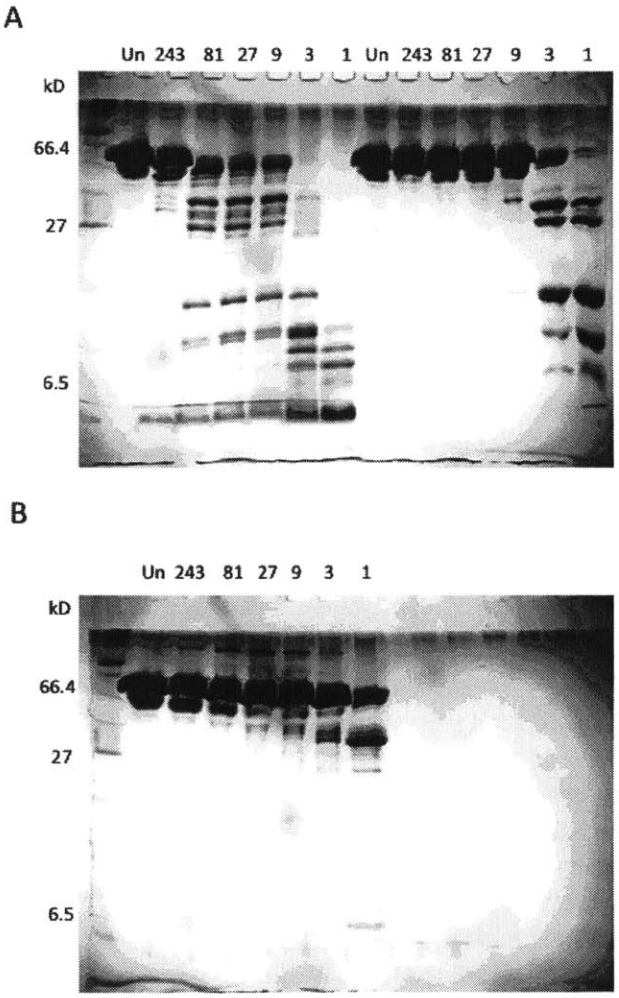

Fig. 22. Limited proteolysis test of trypsin, chymotrypsin, and V8 on XG76. (A) Limited proteolysis test of trypsin (left) and chymotrypsin (right) on XG76. The number indicates the dilution of the stock enzyme. The stock enzyme has OD of 0.1. (B) Limited proteolysis test of V8 on XG76.

Next, large-scale digestion of XG76 with trypsin and chymotrypsin was performed separately in optimized amounts. Digestion products were separated over a Superdex S200 10/300 gel filtration column and the peaks were analyzed by SDS-PAGE. While the trypsin digest was uninformative, the chymotrypsin digest appeared as a single peak on gel filtration and yet as multiple bands on the gel, indicating that Fic-1 fragments remained in a stable structure despite the cuts (Fig. 23).

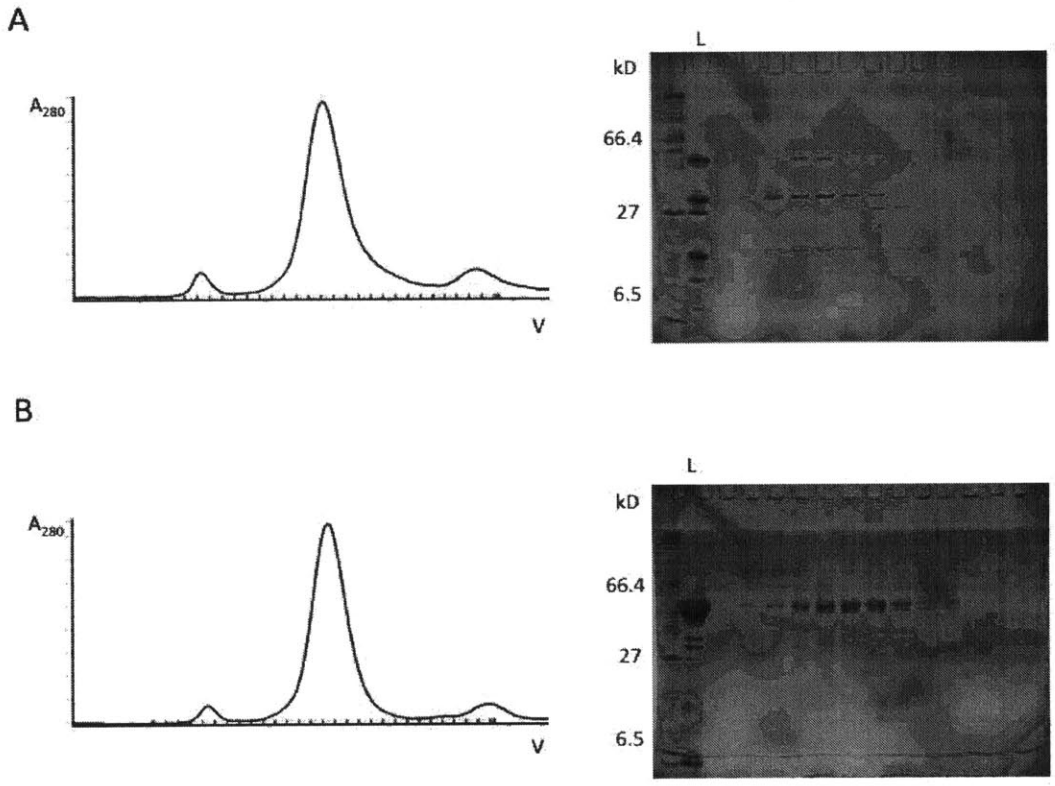

Fig. 23. Gel filtration separation of XG76 digested with chymotrypsin or trypsin. (A) Gel filtration elution profile and its SDS-PAGE analysis of XG76 digested with chymotrypsin. (B) Gel filtration elution profile and its SDS-PAGE analysis of XG76 digested with trypsin.

The peak fraction from the chymotrypsin digest was analyzed by mass spectroscopy (Fig. 24). Matching the fragment mass information with potential cut sites revealed the cut site location. There were three major cut sites, and the following conclusions were drawn.

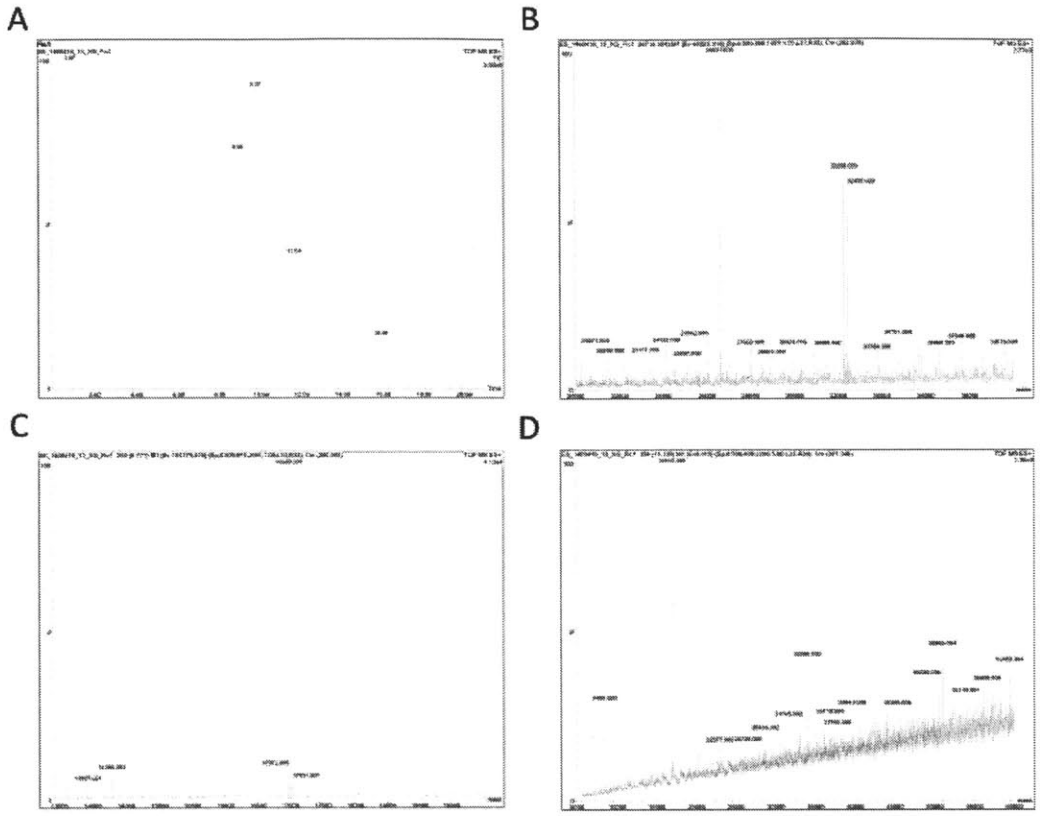

Fig. 24. Mass spectroscopy of a gel filtration fraction of chymotrypsin digestion of XG76. (A) Chromatogram of the mass spec run on XG76. (B) De-convoluted spectrum from the first peak in (A). (C) De-convoluted spectrum from the second peak in (A). (D) De-convoluted spectrum from the later peaks in (A).

a) Loop from aa340 to aa368 (GNADPVEAGRIRTTQVYVGRFTPVSPEYV) in the Fic domain was exposed on the surface and could probably be deleted.
b) Between TPR and Fic domain, there were two potential sites that went uncut, meaning they were not exposed.
c) There was a cut site at aa119, indicating exposure, but sites before aa119 went uncut, meaning they were buried. Thus, there was probably another domain before TPR.

Therefore, a good truncation site would be at around aa119. With consideration to the local amino acid characteristics, constructs Fic-1 (134-508) WT and E274G were designed.

2.2.3 Crystal hits from the well-behaving Fic-1 constructs.

For Fic-1 (258-508), proteins precipitated in most of the drops in the first set of crystallization trials, which were Index HT (Hampton Research), AmSO4 (Qiagen), and Protein Complex (Qiagen) screens with proteins in 5 mg/ml, 10 mg/ml and 20 mg/ml concentrations in a 96-well sitting drop. The protein concentration was reduced to 1 mg/ml and 2 mg/ml, but there was still abundant precipitation without any useful leads. The same happened to Fic-1 (258-508) E274G, with protein concentrations of 2 mg/ml and 5.2 mg/ml in the same three screens

Since ATP is a substrate of Fic-1, adding ATP or its analogs can possibly stabilize the structure of Fic-1. Fic-1 (258-508) WT and E274G were incubated with a 10-fold molar excess of AMPPNP (non-hydrolyzable ATP) on ice for 30 minutes, then the screens were set up with Fic-1 (258-508) E274G at 2 mg/ml and 5.2 mg/ml, and Fic-1 (258-508) at 1.5 mg/ml, 3 mg/ml, and 6 mg/ml. The amount of precipitation was reasonable, but still no crystals formed.

Fic-1 (258-508) was set up at 0.39 mg/ml, 0.78 mg/ml, and 1.56 mg/ml, and Fic-1 (258-508) E274G at 6.3 mg/ml, 12.5 mg/ml, and 25 mg/ml. No crystal was found, even after lowering Fic-1 (258-508) E274G concentration to 1 mg/ml, 2 mg/ml, and 4.7 mg/ml.

Fic-1 (145-508) was set up at 17 mg/ml, 35 mg/ml, and 70 mg/ml, and Fic-1 (258-508) E274G at 1.8 mg/ml, 2.7 mg/ml, and 7.5 mg/ml. An initial hit for Fic-1 (258-508) E274G was observed in Protein Complex F9 (1 M AmSO4, 0.1 M MES pH6.5) at protein concentrations 1.8 mg/ml (Crystal appeared at Day 8, and stopped growing at about Day 13) and 3.7 mg/ml (Day 2, and Day 8), although the crystals were small and abundant (Fig. 25A). At 7.5 mg/ml, the drop was clear at setup, but precipitated after a day.

2.2.4 Optimization of Fic-1 crystals proved hard.

An additive screen was performed based on Fic-1 (145-508) E247G Protein Complex F9 hit. From the screen with protein concentrations of 2.7 mg/ml, 3.7 mg/ml, and 5.7 mg/ml, 3 hits appeared. Dozens of crystals started to grow at Day 3 in G10 (3% ethanol), while drops were clear at the two lower protein concentrations. Fewer and bigger crystals formed from Day 4 in G11 (3% 2-propanol) at 5.7 mg/ml (Fig. 25B); lower concentrations yielded no crystals. In H1 (30 mM Gly-Gly-Gly) roundish small crystals formed at protein concentrations 2.7 and 3.7 mg/ml.

In a replicative screen, XG100 was set up at concentrations of 4 and 6.3 mg/ml. Only G10 reproduced the crystal at 6.3 mg/ml. In G11 and H1, there were only irregular precipitates.

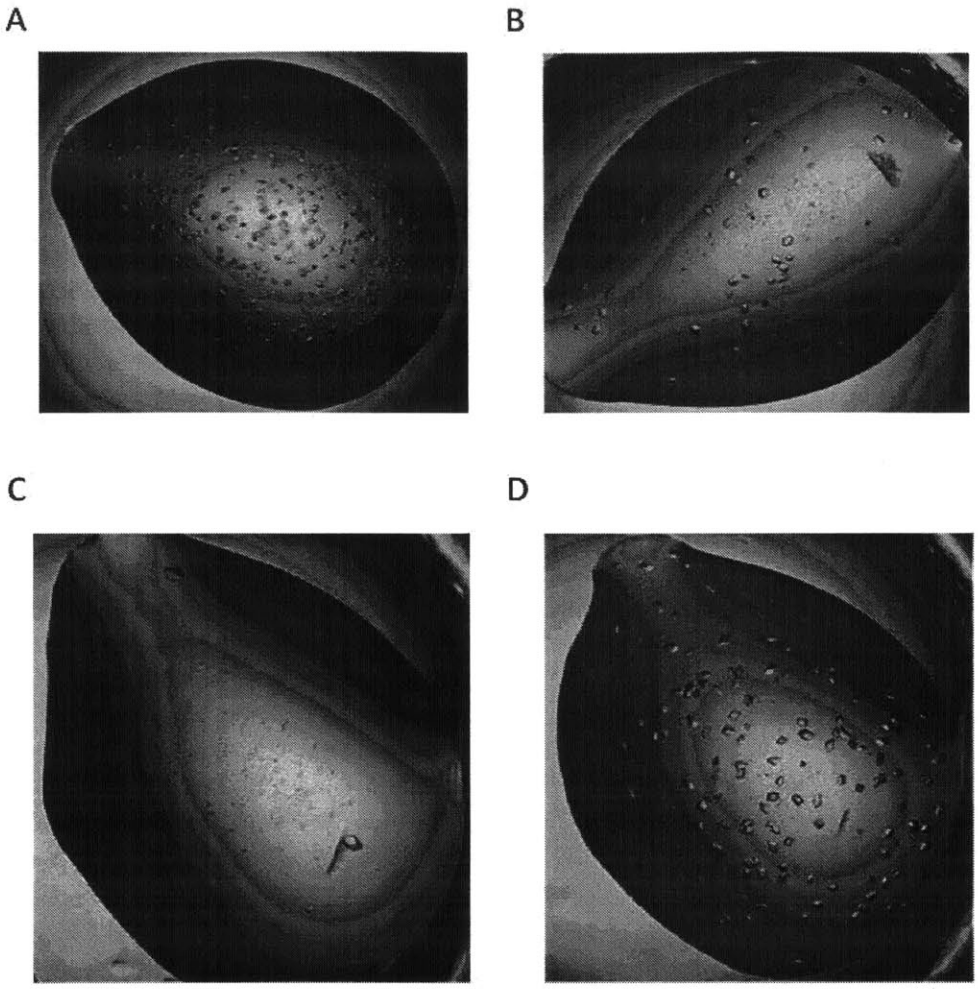

Fig. 25. Crystal hits of XG100 obtained in various screens. (A) XG100 crystal hit in Protein Complex F9 (1 M AmSO4, 0.1 M MES pH6.5) with protein concentration of 3.7 mg/ml at 3 weeks. (B) XG100 crystal hit in Additive screen G11 (3% 2-propanol) based on Protein Complex F9 with protein concentration of 5.7 mg/ml at 3 weeks. (C and D) XG100 crystal hits in grid screen H5 (0.9 M AmSO4, 0.1 M MES 6.7) based on Protein Complex F9 with protein concentration of 5.1 mg/ml and 6.3 mg/ml.

A grid screen was conducted to refine the F9 condition using Formulatrix robot. In one dimension, the ammonium sulfate concentration ranged from 0.7 M to 1.25 M, and in the other, the pH of 0.1 M MES buffer ranged from 5.5 to 6.7. There was a hit in H5 (0.9 M AmSO4, 0.1 M MES 6.7) at protein concentrations 5.1 mg/ml (several crystals of 20 μm) (Fig. 25C) and 6.3 mg/ml (hundreds of crystals of 10 μm) (Fig. 25D). However, it was unusual to obtain only two hits from a 96-condition grid screen.

A 24-well hanging-drop screen of Fic-1 (145-508) E247G was set up based on the additive screen and the grid screen. In one dimension, AmSO4 concentration varied from 0.7 M to 1.2 M, and in the other dimension, pH of 0.1 M MES varied from 6.1 to 6.7. There were hits in conditions 0.9 M AmSO4, 0.1 M MES pH6.3, and 0.8 M AmSO4, 0.1 M MES pH6.5. In another plate with addition of 3% 2-propanol, there were hits in conditions 1 M AmSO4, 0.1 M MES pH6.3; 0.9 M AmSO4, 0.1 M MES pH6.5; and 0.9 M AmSO4, 0.1 M MES pH6.7. Among all the hits, 0.9 M 0.1 M MES pH6.7 with 3% 2-propanol produced the best crystal. But still, the drop contains hundreds of small crystals.

Finer gradients of each component in the hit conditions were tried to optimize the crystal, and other crystallization agents and techniques as well.

Reservoir composition

a. I varied 2-propanol and ethanol concentration from 1.2% to 5.1% and tried adding 0.03 M Gly-Gly-Gly.
b. I varied ammonium sulfate concentration from 0.5 M to 1.3 M.
c. I added PEG200 and glycerol in amount from 0.05% to 4.8%, and tried adding 0.5% ethylene glycol (PEG200, glycerol, and ethylene glycol inhibit nucleation.)
d. I replaced 0.1 M MES with 0.1 M HEPES from pH6.7 to pH7.3. (Because buffer capacity for MES goes as high as pH6.7.)
e. I varied MES concentration from 0.05 M to 0.2 M.
f. I varied MES pH from 6.1 to 6.7 and refined pH between 6.5 and 6.7.

Drop composition

g. I tried adding 10-time excess of AMPPNP to XG100.
h. I tried varying reservoir-protein composition from 1.5:0.5 to 0.8:1.2.
i. I tried increasing reservoir volume to 1.5 ul.
j. I tried increasing both reservoir and protein to 2 ul.

Seeding

k. I prepared micro-seeds using seed beads and seeded the drops from after 24h to after 4d using serial dilution of seeds of up to 100,000-fold dilution. I seeded with either a cat whisker or by pipetting.
l. I tried macro-seeding by fishing crystals or pipetting crystal-containing drops and transferring them into fresh drops from after 24h to after 4 days. I cleaned crystals by dipping them in a cleaning solution.

Other crystallization methods

m. I performed microbatch experiment under silicon oil varying pH between 6.5 and 6.7.
n. I also tried sitting-drop method varying reservoir-protein composition from 1.5:0.5 to 0.8:1.2.

Unfortunately, despite all these efforts, I was not able to produce bigger crystals. The challenge appears to be reducing nucleation, without completely abolishing it. Also, reproducibility is a problem, which possibly points to an uncontrolled factor that might be important in this protein preparation.

2.2.5 Current Fic-1 crystals produced data with 7 Å resolution at best.

Fic-1 crystals were tested for diffraction twice.

For the first data collection, the crystals were fished with MicroMeshes (MiTeGen). Unfortunately, I was not able to spot any crystal on the frosty MicroMesh, hindering useful data collection.

For the second data collection, crystals were fished with 50 μm loops despite the smallness of the crystals. I was able to fish single crystals into the loops and they froze well with the cryoprotectant (1.1 M AmSO4, 0.1 M MES pH6.5, 3% 2-propanol, and 20% glycerol). Through adjusting beam size and intensity and other parameters, I was able to collect diffraction data from some crystals up to 7 Å resolution.

3. Materials and methods

Plasmids, protein expression, and purification

Recombinant proteins were expressed in LOBSTR *Escherichia coli*. All LINC Complex proteins were expressed from a modified ampicillin resistant pETDuet-1 (EMD Millipore, Billerica, MA) vector. All Fic-1 proteins were expressed from a modified kanamycin resistant pETDuet-1 (EMD Millipore, Billerica, MA) vector.

LINC complex proteins were expressed in LOBSTR(DE3) RIL strains (Kerafast, Boston, MA) (Andersen et al., 2013). Bacterial cultures were grown at 37°C to an optical density (OD600) of 0.6, shifted to 18c for 20 min, and induced ονernight at 18°C with 0.2 mM IPTG. Cells were resuspended in lysis buffer A (50 mM potassium phosphate pH 8.0, 500 mM NaCl, 40 mM imidazole, 5 mM beta-ME for single-protein expression) and lysed. The lysate was supplemented with 1 U/ml Benzonase (Sigma-Aldrich, St. Louis, MO) and 1 mM PMSF, cleared by centrifugation, and loaded onto a Ni-affinity resin. After washing with lysis buffer, bound protein was eluted with elution buffer (10 mM Tris pH 8 at 4°C, 150 mM sodium chloride, 250 mM imidazole, 1 mM KCl, 5 mM beta-ME for single-protein expression). The protein is further purified by size exclusion chromatography on a Superdex S200 or S75 column (GE Healthcare) equilibrated in buffer (10 mM Tris/HCl pH 8 at 4°C, 150 mM NaCl, 1 mM KCl).

Fic-1 proteins were processed in the same way except that no KCl was added to either elution buffer or size exclusion buffer.

Protein crystallization

Purified proteins were concentrated to saturation prior to crystallization. Fic-1 (134-508) E274G crystalized in 1 M AmSO4, 0.1 M MES pH6.5 by the hanging drop vapor diffusion method in 2 µl drops at 18°C. Crystallization conditions were optimized and best crystals grew within a week with dimensions of 20 µm. Prior to X-ray data collection, crystals were cryoprotected in the reservoir solution of 1.1 M AmSO4, 0.1 M MES pH6.5, 3% 2-propanol, and 20% glycerol. Data were collected at beamlines 24ID-C/-E at Argonne National Laboratory.

Limited proteolysis

2 µl of trypsin, chymotrypsin, and V8 protease in 3-fold serial dilution was added to 22 µl 1 mg/ml Fic-1 (65-508), and the mixture was incubated at room temperature for 30 min. To quench the reaction, 6 µl 5x SDS loading buffer was added and the mixture was incubated at 95°C for 3 min. The samples were analyzed by SDS-PAGE, and the protease concentration that yields stable fragments was identified.

Proteolysis was repeated in larger scale by adding 50 µl of protease of optimal concentration to 500 µl sample of protein. After 30-minute incubation at room temperature, 10 µl of 100 mM PMSF was added to quench the reaction. 5 µl sample was sent for MALDI-TOF mass spectrometry. 500 µl sample was analyzed by gel filtration on Superdex 75/200 10 300.

4. Acknowledgement

I would like to take this opportunity to thank my supervisor, Prof. Thomas Schwartz, who has been my consistently reliable source of guidance and support in research and life. I would also like to thank Dr. Brian Sosa, who readily shared constructs and expert opinions with me in LINC complex project. Beside being a fun bay-mate, Kevin Knockenhauer has taught me a great many crystallization techniques during my rotation and after I joined the lab. I would like to thank Dr. Matthias Truttmann (Ploegh Lab) for bringing the project of Fic-1 to me and for being a wonderful collaborator. Last but not least, my thanks go to the whole Schwartz Lab, including Dr. Esra Demircioglu, Kotaro Kelley, Victor Cruz, Dr. Greg Kabachinski, Dr. Kasper Andersen, Dr. James Chen, and Dr. Phat Dip, for their inexhaustible advice and for making doing research here exciting and fun.

1. Tapley EC and Starr DA. (2013). Connecting the nucleus to the cytoskeleton by SUN-KASH bridges across the nuclear envelope. Curr Opin Cell Biol. *25*, 57-62.
2. Sosa BA, Rothballer A, Kutay U, and Schwartz TU. (2012). LINC complexes form by binding of three KASH peptides to domain interfaces of trimeric SUN proteins. Cell. *149*, 1035-47.
3. Sosa BA, Kutay U, and Schwartz TU. (2013). Structural insights into LINC complexes. Curr Opin Struct Biol. *23*, 285-91.
4. Rothballer A and Kutay U. (2013). The diverse functional LINCs of the nuclear envelope to the cytoskeleton and chromatin. Chromosoma. *122*, 415-29.
5. Luxton GW and Starr DA. (2014). KASHing up with the nucleus: novel functions of KASH proteins at the cytoplasmic surface of the nucleus. Curr Opin Cell Biol. *28*, 69-75.
6. Pedelacq JD, Cabantous S, Tran T, Terwilliger TC, and Waldo GS. (2006). Engineering and characterization of a superfolder green fluorescent protein. Nat Biotechnol. *24*, 79-88.
7. Garcia-Pino A, Zenkin N, and Loris R. (2014). The many faces of Fic: structural and functional aspects of Fic enzymes. Trends Biochem Sci. *39*, 121-9.
8. Worby CA, Mattoo S, Kruger RP, Corbeil LB, Koller A, Mendez JC, Zekarias B, Lazar C, and Dixon JE. (2009). The fic domain: regulation of cell signaling by adenylation. Mol Cell. *34*, 93-103.
9. Kinch LN, Yarbrough ML, Orth K, and Grishin NV. (2009). Fido, a novel AMPylation domain common to fic, doc, and AvrB. PLoS One. *4*, e5818.
10. Rahman M, Ham H, Liu X, Sugiura Y, Orth K, and Kramer H. (2012). Visual neurotransmission in Drosophila requires expression of Fic in glial capitate projections. Nat Neurosci. *15*, 871-5.
11. Campanacci V, Mukherjee S, Roy CR, and Cherfils J. (2013). Structure of the Legionella effector AnkX reveals the mechanism of phosphocholine transfer by the FIC domain. EMBO J. *32*, 1469-77.
12. Xiao J, Worby CA, Mattoo S, Sankaran B, and Dixon JE. (2010). Structural Basis of Fic-mediated adenylylation. Nat Struct Mol Biol. *17*, 1004-10.
13. Palanivelu DV, Goepfert A, Meury M, Guye P, Dehio C, and Schirmer T. (2011). Fic domain-catalyzed adenylylation: insight provided by the structural analysis of the type IV secretion system effector BepA. Protein Sci. *20*, 492-9.
14. Das D, Krishna SS, McMullan D, Miller MD, Xu Q, Abdubek P, Acosta C, Astakhova T, Axelrod HL, Burra P, Carlton D, Chiu HJ, Clayton T, Deller MC, Duan L, Elias Y, Elsliger MA, Ernst D, Feuerhelm J, Grzechnik A, Grzechnik SK, Hale J, Han GW, Jaroszewski L, Jin KK, Klock HE, Knuth MW, Kozbial P, Kumar A, Marciano D, Morse AT, Murphy KD, Nigoghossian E, Okach L, Oommachen S, Paulsen J, Reyes R, Rife CL, Sefcovic N, Tien H, Trame CB, Trout CV, van den Bedem H, Weekes D, White A, Hodgson KO, Wooley J, Deacon AM, Godzik A, Lesley SA, and Wilson IA. (2009). Crystal structure of the Fic (Filamentation

induced by cAMP) family protein SO4266 (gi|24375750) from Shewanella oneidensis MR-1 at 1.6 A resolution. Proteins. *75*, 264-71.
15. Castro-Roa D, Garcia-Pino A, De Gieter S, van Nuland NA, Loris R, and Zenkin N. (2013). The Fic protein Doc uses an inverted substrate to phosphorylate and inactivate EF-Tu. Nat Chem Biol. *9*, 811-7.
16. Engel P, Goepfert A, Stanger FV, Harms A, Schmidt A, Schirmer T, Dehio C. (2012). Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins. Nature. *482*, 107-10.
17. Goepfert A, Stanger FV, Dehio C, and Schirmer T. (2013). Conserved inhibitory mechanism and competent ATP binding mode for adenyleltransferases with Fic fold. PLoS One. *8*, e64901.
18. Bunney TD, Cole AR, Broncel M, Esposito D, Tate EW, and Katan M. (2014). Crystal structure of the human, FIC-domain containing protein HYPE and implications for its functions. Structure. *2126*, 334-7.
19. Andersen KR, Leksa NC, and Schwartz TU. (2013). Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. Proteins. 81, 1857-61.