# The Molecular Basis of LINC Complex Formation

By

Brian A. Sosa
B.S. University of Puerto Rico at Rio Piedras
San Juan, PR 2008

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

DOCTOR OF PHILOSOPHY

AT THE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

DECEMBER 2013
[FEBRUARY 2014]





Signature of Author.....
Brian A. Sosa
Department of Biology
December 4, 2010

Certified by.....
Thomas U. Schwartz
Professor of Biology
Thesis Supervisor

Accepted by.....
Stephen P. Bell
Professor of Biology
Co-chair, Biology Graduate Committee

# The Molecular Basis of LINC Complex Formation

By

Brian A. Sosa

Submitted to the Department of Biology in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology

## Abstract

The nucleus is the hallmark of the eukaryotic cell. It contains most of the genetic material and it separates the processes of replication and transcription from that of translation. Communication between the nucleus and the cytoplasm occurs mostly through openings in the nuclear envelope composed of nuclear pore complexes. These massive assemblies allow for regulated transport of macromolecules across the barrier that is the nuclear envelope. However, another means of communication between the nucleus and the cytoskeleton has been characterized: linkers of the nucleoskeleton and cytoskeleton, or LINC, complexes mechanically connect the nucleus with its surroundings allowing for nuclear anchorage and nuclear movement during development and chromosome movement during meiosis. At the heart of LINC complexes are inner nuclear membrane resident SUN proteins and outer nuclear membrane resident KASH proteins. In this thesis we structurally characterize the human SUN2-KASH1 and KASH2 complexes and provide the molecular basis for their interaction. The solved structures suggest plausible models for high-order of LINC complex assembly as well as LINC complex mediated spacing between the inner and outer nuclear membrane. Questions about how these complexes are regulated also arise from the structure. TorsinA is an AAA+ ATPase suggested to play a role in LINC complex regulation. Analysis of Torsin's binding partners LAP1 and LULL1 show that they are catalytically inactive AAA+ ATPases. We characterize the complex and show by EM that they form ring akin to other AAA+ ATPases. With these studies we provide the first structural analysis of LINC complexes and Torsin ATPases and also provide biochemical tools for the study of LINC complex regulation.

Thesis Supervisor: Thomas U. Schwartz

**ACKNOWLEDGEMENTS**

These past 5 years have been some of the most exciting in my life. The work put into this thesis not only has it made me a better scientist, but person as well. I've learned that is not only about book knowledge, but also about dedication, sacrifice, excitement, collaboration, enthusiasm and passion. It was about many long nights and early mornings trying to get that one piece of data that moves us all forward. It is also about many moments in which I found myself daydreaming about models and electron-densities. Needless to say that all this work and experiences were lived with some very important people that have made my stay at MIT invaluable.

To my early lab mates and friends, especially James and Teresa Partridge, Nina Leksa, Alex Ulrich and James Chen I thank you for welcoming me into the group, for the all the scientific advice and helping me find my niche in this institution. Specially, James Chen, who spend many hours in a dark room looking for that "ONE-GRID" that was going to make the difference. To the current and next Schwartz lab generation, I thank you for helping me in these difficult last few years and keeping the cool environment in which I started. I would also like to thank my committee members, Bob Sauer and Steve Bell for invaluable advice and encouraging comments.

To my adviser Thomas, I am profoundly grateful of you taking a chance with me and teaching me what science really is and how it should be done. You have been an incredible source of knowledge whose personal scientific experience is a perfect model of what I would like mine to be.

To my family in Puerto Rico, thanks for not letting the distance make us into strangers and for keeping me in touch with every single text message that make me laugh to this day. ¡Los adoro a todos! To my New England family, I am eternally grateful for your love and look forward to many years of paying you back. To Ray-Ray, you are a good boy! And to Lilly, I will thank you for the rest of my life with all the love I could possibly give, but as I've told you before, you have single-handedly made my life the happiest it could be. I love you.

## Table of Contents

# Chapter One

## Structural Insights into LINC Complexes

This chapter was adapted from Sosa, B. A., Kutay, U. & Schwartz, T. U. Structural insights into LINC complexes. *Current Opinion in Structural Biology* 23, 285–291 (2013).

**SUMMARY**

Communication between nucleus and cytoplasm extends past molecular exchange and critically includes mechanical wiring. Cytoskeleton and nucleoskeleton are connected via molecular tethers that span the nuclear envelope. SUN-domain proteins spanning the inner nuclear membrane and KASH-peptide bearing proteins residing in the outer nuclear membrane directly bind and constitute the core of the LINC complex. These connections appear critical for a growing number of biological processes and aberrations are implicated in a host of diverse diseases, including muscular dystrophies, cardiomyopathies, and premature aging. We discuss recent developments in this vibrant research area, particularly in context of first structural insights into LINC complexes. Furthermore, we begin to characterize the AAA+ ATPase Tor1A, a possible regulatory protein for LINC complex formation or disassembly.

## INTRODUCTION

The nuclear envelope (NE) is a double-lipid bilayer that separates the nucleus from the cytoplasm. This way, transcription and translation are spatially separated in eukaryotes, enabling sophisticated regulatory mechanisms for gene expression. The NE is an extension of the endoplasmic reticulum (ER) and consists of an outer nuclear membrane (ONM) and an inner nuclear membrane (INM), evenly separated by the perinuclear space (PNS) of ~50 nm width. ONM and INM are fused at circular openings, occupied by nuclear pore complexes (NPCs). Macromolecular trafficking between the nucleus and the cytoplasm occurs mainly through NPCs (Brohawn et al., 2009; Gorlich and Kutay, 1999), although recent findings suggest a vesicular transport mechanism across the PNS akin to nuclear egress by herpes viruses may be used for exceptionally large nuclear export cargo (Rose and Schlieker, 2012; Speese et al., 2012).

Mechanical communication has long been recognized for cells interacting with their surrounding (Ingber, 1997). That the nucleus is also mechanically tethered to its environment has only been uncovered more recently (Maniotis et al., 1997). Since that initial discovery, interest in the subject has grown rapidly. Mechanical coupling of the nucleus to the cytoskeleton can potentially serve many purposes. First, the position of the nucleus within a cell needs to be maintained in many cell types, notably in neurons and muscle cells, suggesting that nuclei require a mechanism to be pulled into their desired place (Tapley and Starr, 2013). Second, physical connections across the NE are attractive candidates for mediating mechanotransduction, a very fast signaling mechanism that results in transcription programs triggered by extracellular stimuli (Lombardi et al., 2011; Wang et al., 2009). Third, these nucleocytoplasmic linkages can also be used to determine the position of specific nuclear structures, like the ends of paired chromosomes during meiosis (Chikashige et al., 2006; Sato et al., 2009). Underscoring the general importance of the nucleoskeleton and its link to the cytoskeleton, a growing number of diverse genetic disorders, including

neurological, muscular, and premature aging have been linked to mutations in its constituents (Gruenbaum et al., 2005; Méndez-López and Worman, 2012; Worman and Bonne, 2007).

Physical interactions across the nuclear envelope are mediated via **LI**nkers of **N**ucleoskeleton and **C**ytoskeleton (LINC) complexes (Crisp et al., 2006; Starr et al., 2010). The center of LINC complexes is the interaction of INM-resident SUN (**S**ad1 and **UN**C-84) proteins with ONM-resident KASH (**K**larsicht, **A**NC-1 and **S**YNE/Nesprin-1 and -2 **H**omology) proteins within the PNS. The SUN-KASH interaction complex has recently been solved by X-ray crystallography, providing a rich basis for detailed studies on LINC function (Sosa et al., 2012).

**SUN proteins**

SUN proteins are type II membrane proteins conserved across all eukaryotes and typically found in the INM. At the N terminus they contain a variable nucleoplasmic region, followed by a transmembrane helix connecting into a predicted coiled-coil segment localized to the PNS. The most recognizable feature of SUN proteins is a stretch of ~175 amino acids, usually at the very C terminus, termed 'SUN domain' based on the homology between **S**ad1 from *Schizosaccharomyces pombe* and **UN**C-84 from *Caenorhabditis elegans* (Malone et al., 1999). With increasing complexity of the organism, the number of SUN proteins also increases. Although single cell organisms apparently carry only one SUN domain protein, nematodes and flies contain two genes for SUN proteins, and the mammalian genome encodes at least five distinct members of the SUN protein family, Sun1-5 (Starr et al., 2010). The expression of the individual SUN proteins depends on the cell type, suggesting cell type-specific adaption of LINC complexes to meet distinct cellular and physiological requirements. For example, in *C. elegans*, UNC-84 is expressed in most cells, whereas SUN-1/Matefin expression is restricted to germ cells. Notably, the *C. elegans* SUN proteins show no overlapping activity, indicating that their

respective LINC complexes occupy distinct functions (Fridkin et al., 2004). As in *C. elegans*, the Drosophila SUN protein Klaroid is present in almost every cell type, whereas SUN4/Spag4 appears to be strictly confined to the male germ line (Kracklauer et al., 2007; Technau et al., 2008). Similarly, the two major mammalian SUN proteins, Sun1 and Sun2, are widely expressed in different cell types (Crisp et al., 2006; Padmakumar, 2005). In contrast, the three additional genes coding for Sun3, Sun4, and Sun5, respectively, all show a much more restricted, testis-specific expression pattern (Göb et al., 2010). In addition to the SUN domains at the C terminus, there are also SUN-like proteins (Field et al., 2012; Shimada et al., 2010) with the SUN-like domain in the center of the protein. To what extent SUN- and SUN-like proteins are functionally related is currently a matter of speculation.

One intriguing aspect of SUN proteins is their apparent exclusive localization to the INM. Therefore, a number of studies have focused on understanding the molecular parameters that determine transport to the INM (Tapley et al., 2011; Turgay et al., 2010), as part of an effort to understand INM targeting in general. It appears that multiple factors contribute to targeting and it is not yet clear whether they are universal or whether instead different INM proteins employ distinct mechanisms.

**KASH proteins**

Thus far, KASH proteins are exclusively found at the ONM and share several common features. They are tail-anchored, single-span transmembrane proteins with a short luminal C terminus. The transmembrane helix together with the 8-30 residues of the luminal tail are well conserved and form the KASH 'domain', recognizable by primary sequence analysis (Figure 1) (Starr and Han, 2002). The most striking feature of the luminal KASH peptide is the C-terminal 'PPPX' motif, where X is always the very terminal residue.

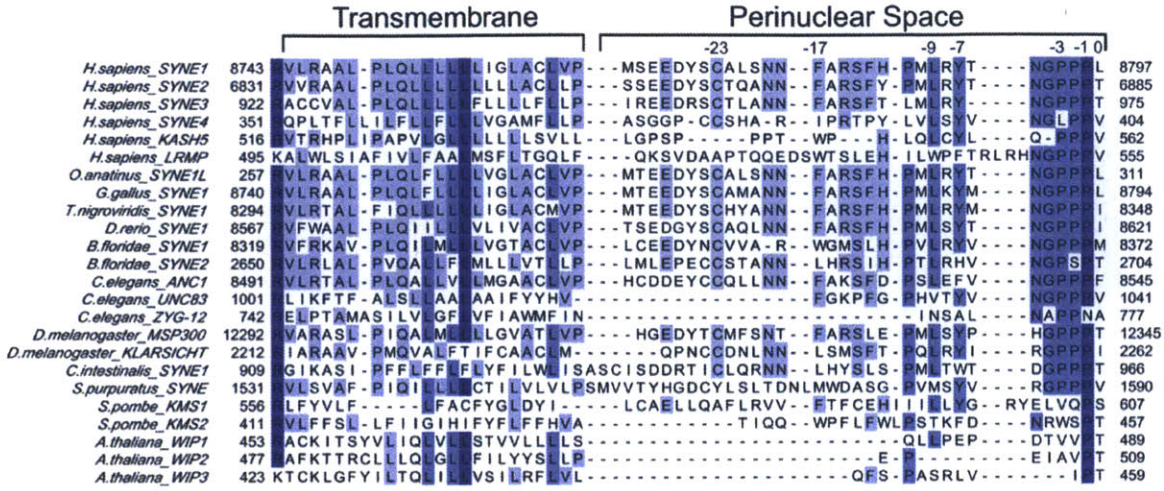


**Figure 1. Multiple sequence alignment of KASH peptides from highly diverged eukaryotes, including vertebrates, various metazoans, yeast, and plants.** The numbers above the sequence mark residues in human Nesprin2/SYNE2 important for SUN binding according to Sosa et al., 2012.

Much like SUN proteins, complex organisms tend to have more KASH proteins than simpler ones. To date, we know six KASH proteins in humans, four of which are called Nesprins (for nuclear envelope spectrin repeat). Nesprin 1 and 2 in mammals are exceedingly large, 0.8 – 1 MDa actin-binding proteins comprising multiple spectrin repeats that serve in nuclear tethering. These proteins generate extensive fibers that emanate deep into the cytoplasm. Nesprin-3, a shorter molecule, binds to plectin, which in turn links to the actin and/or intermediate filaments. Nesprin-4 is restricted to a few cell types and connects to microtubules via the motor protein kinesin. It has recently been shown that C-terminal, KASH-free truncations of Nesprin-4 result in hearing loss (Horn et al., 2013) KASH5, which binds SUN1, was reported recently to be a germ cell-specific protein involved in meiotic homologue pairing (Morimoto et al., 2012). The ONM-resident, lymphoid-restricted membrane protein (LRMP) is the latest characterized putative KASH proteins (Lindeman and Pelegri, 2012). Although its cognate SUN protein is not yet known, LRMP is necessary for pronuclear congression in the fertilized egg. Thus, it fulfills a mechanical task involving the positioning of nuclei and it has the correct topology and KASH peptide sequence, implying that it is another member of the growing KASH family.

## SUN-KASH Complex

The principal function of the LINC complex is to tether nucleo- and cytoskeleton mechanically, which presumably requires a very strong and stable interaction between SUN and KASH. The crystal structure of the LINC complex published in 2012 illustrates this property (Sosa et al., 2012). The SUN domain itself folds into a beta-sandwich structure, a fairly common fold in eukaryotes. The basic structure is decorated with several SUN-specific features (Figure 2). First, N-terminal to the beta-sandwich, a helical extension forms a triple-stranded coiled coil with adjacent SUN domains, effectively generating a SUN homotrimer. Even though the SUN domains also directly touch neighboring SUN domains, their interactions are too weak to stabilize a trimer without the coiled-coil extension. Second, a ~20 amino-acid extension, the KASH-lid, emanates from the central beta-sandwich and is critical for KASH binding. Third, a conserved ~10 residue loop structure is positioned by a disulfide bond between two highly conserved cysteines. This loop binds a metal ion, and is also involved in KASH binding.

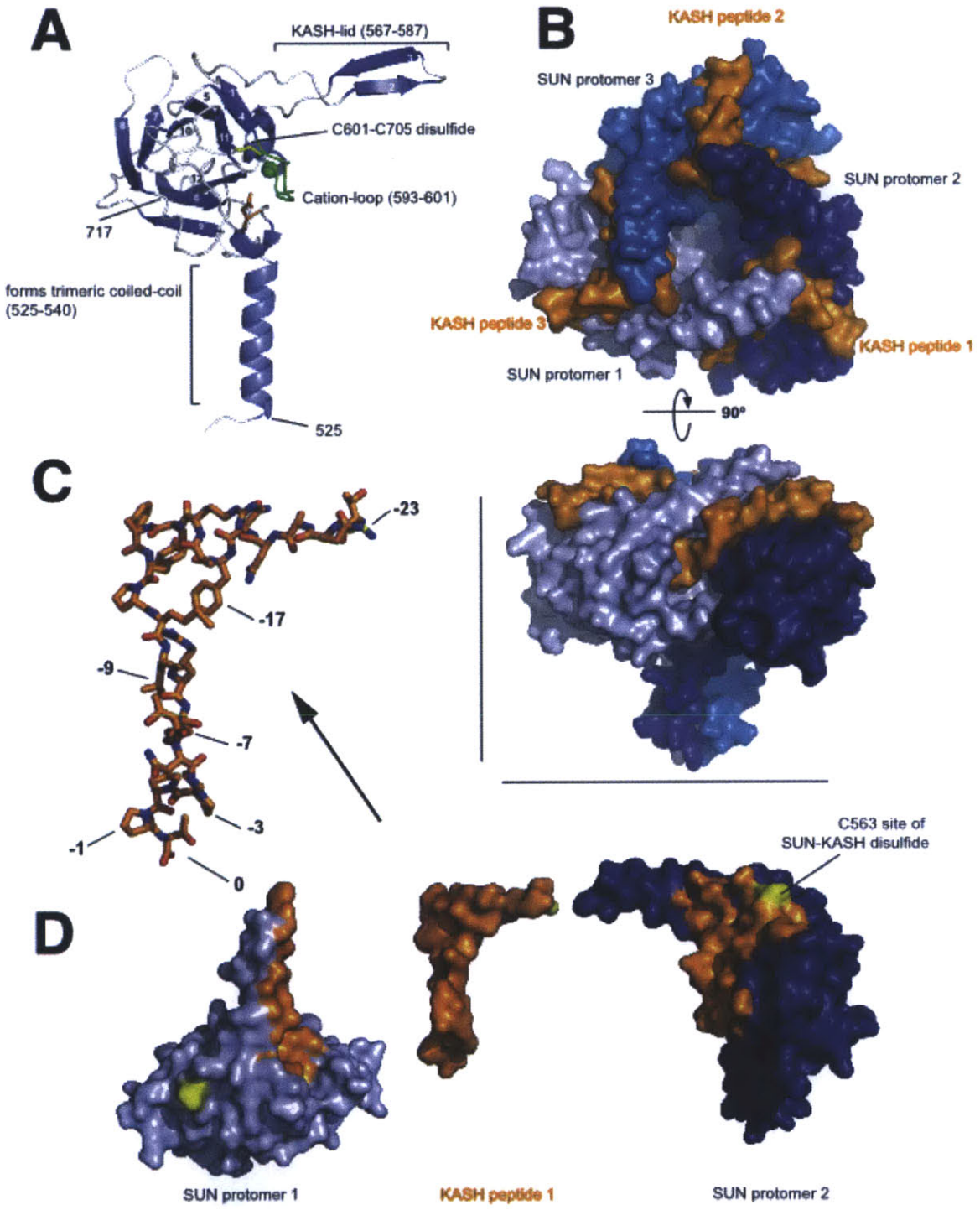


**Figure 2. Structural overview of the SUN-KASH complex.** (A) Overview of a human SUN2 protomer isolated from its Nesprin-2 binding partners in the trimeric SUN-KASH complex (Sosa et al., 2012). The protein is organized around a compact β-sandwich core, decorated with features important for function (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.

The SUN homotrimer serves as a tailored platform to interact with three KASH peptides (Figure 2B). The peptides are individually bound in three deep grooves, each created by two neighboring SUN protomers. Trimerization of SUN is therefore a prerequisite for KASH binding. The very C terminus of the KASH peptide, amino acid position 0, is buried in a pocket on SUN protomer 1 and the carboxylate makes numerous contacts, explaining the strict conservation of the peptide length (Figure 2C). Adding just one residue abolishes binding (Sosa et al., 2012).Positions -1 to -3 are typically trans-prolines and are bound in shape-dependent, van-der-Waals manner. After a few more exposed positions, residues -7 to -10 form a short beta-strand that connects the KASH-lid beta-hairpin of protomer 1 with the SUN core domain of protomer 2. Positions -7 and -9 are conserved as large hydrophobic amino acids, and are buried in a cleft between the SUN protomers. Following residue -11 the KASH peptide sharply kinks and lines up on the surface of SUN protomer 2. Except for the buried hydrophobic residue at position -17 conservation is weak in this region, consistent with data indicating that the last 11-14 residues are sufficient for stable SUN2-KASH2 binding (Sosa et al., 2012). Importantly though, residue -23 encodes a conserved cysteine that crosslinks with the conserved SUN2-cysteine 563, drastically stabilizing the interaction further. In summary, SUN trimerization and the formation of large interfaces between three KASH peptides, each clasped between two neighboring SUN protomers and making many stabilizing interactions, ensure very stable binding.

Although the SUN-KASH interaction seen in the crystal structure has no close precedent in other peptide-protein interactions, some beta-sandwich proteins that are distantly related to SUN bind small molecules, often lectins, in a pocket that overlaps the 'PPPX' binding site. These F-lectins also contain a metal binding

loop, and they also use an intramolecular disulfide bond to stabilize this loop, but they typically do not trimerize. Thus SUN proteins and F-lectins may have evolved from a common ancestor that already had a small molecule binding moiety, which diverged into the different binding interfaces seen in the extant domains.

**Conservation of the SUN-KASH interaction**

The existing SUN2-KASH1/-2 complex structures suggest that the principal arrangement of three KASH peptides binding a SUN trimer is universally conserved. All SUN domains characterized to date are immediately preceded by predicted coiled-coil segments, suggesting that the characteristic cloverleaf-like, trimeric SUN arrangement is also conserved. In addition, the binding mode whereby two adjacent SUN domains form an elaborate, shared binding site also strongly supports the notion that the general SUN-KASH heterohexameric arrangement is conserved. The KASH peptides in vertebrates are similar enough to expect them all to bind SUN in similar manner. In nematodes, yeast and plants, however, the sequences are quite diverged (Figure 1), and experimental data is therefore needed to make definitive statements about universally conserved, common features.

**The coiled-coil segment of SUN proteins**

In all SUN proteins the C-terminal SUN domain is preceded by a predicted helical region that spans the remainder of the perinuclear domain. Detailed analyses indicate that these helices can form coiled-coils. As the independently determined apo- and KASH-bound SUN2 structures now show (Sosa et al., 2012; Wang et al., 2012; Zhou et al., 2012), the region immediately adjacent to the SUN domain forms a non-canonical, right-handed trimeric coiled-coil arrangement with undecan rather than heptad repeat character (Gruber and Lupas, 2003). Since coiled-coil prediction methods are trained to detect

canonical structures, it is plausible that the entire helical domain forms an extended, continuous trimeric coiled-coil. Whether it is right-handed throughout or flips handedness is an open question. The reversal of coiled-coil handedness has been observed previously, and the energetic barrier for the conversion is rather low (Alvarez et al., 2010). The length of the coiled-coil segment can be regarded as a ruler that is either adapted to or determines the width of the PNS. The coiled-coil segments of the ubiquitous SUN1/2 proteins are almost equal in length and would generate an ~45 nm rod (Figure 3). The testis-specific SUN3, SUN4, and SUN5 proteins, however, have shorter predicted coiled-coils. These SUNs are localized in specific NE regions and the shorter intermembrane distance forced by these LINC complexes may have functional importance (Göb et al., 2010; Kracklauer et al., 2013).

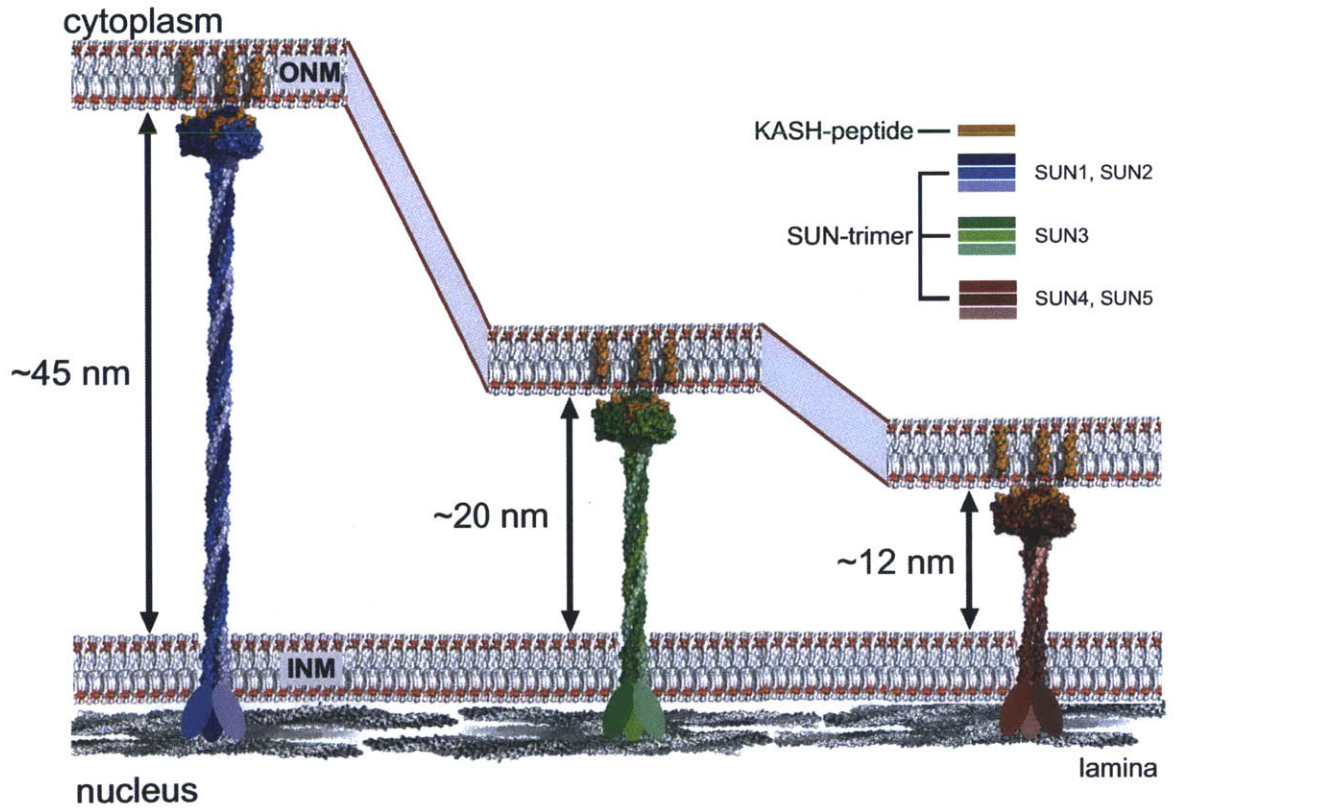


**Figure 3. LINC complex mediated nuclear envelope spacing model.** Various SUN proteins exhibit predicted perinuclear α-helical coiled-coil domains of various length. If these elements are modeled as trimeric coiled-coils, currently the most likely scenario, spacing between INM and ONM would vary dependent on the employed SUN protein.

## Higher-order arrays of LINC complexes

To move the entire nucleus through the cell, LINC complexes must bear substantial mechanical load. Having a connection involving three KASH peptides interacting with a SUN trimer is one way of strengthening the LINC complex. Adding a disulfide bridge covalently linking SUN and KASH is another, but are these interactions sufficient? In fibroblasts, nuclei move away from a wound edge by harnessing the retrograde flow of actin. Here, SUN-KASH bridges in the NE arrange linearly, creating transmembrane actin-associated nuclear, or TAN, lines parallel to the actin fibers to which they are presumably attached (Borrego-Pinto et al., 2012; Luxton and Gundersen, 2011; Luxton et al., 2010). It is attractive to think that this 2D arrangement strengthens actin tethering, but this idea awaits experimental verification. In meiotic cells, SUN proteins cluster to bring telomeric chromosome ends together during bouquet formation (Chikashige et al., 2006; Ding et al., 2007; Penkner et al., 2009; Schmitt et al., 2007), again suggesting that higher order LINC clusters are functionally important. Spectrin repeats in Nesprins, as well as additional coiled-coil segments in Nesprin2, are strong candidates for linking individual hetero-hexameric SUN-KASH complexes into larger 2D arrays (Sosa et al., 2012). The molecular details of such networks and their regulation are exciting topics for future studies.

## Regulation of LINC complex formation

So far, LINC complexes are exclusively found in the NE, with SUN proteins crossing the INM and KASH proteins crossing the ONM. SUNs are expected to be inserted cotranslationally into the ER via the Sec61 channel (Rapoport, 2007), while KASH proteins are tail anchored and might be integrated into the ER via the action of the GET complex (Denic, 2012). Because both SUNs and KASH proteins are inserted into the ER it is intriguing to consider what prevents premature formation of LINC complexes, and how the cell ensures exclusive formation of SUN-KASH bridges in the PNS. One possibility is that at least one

partner is kept in a binding-incompetent state before it reaches its final destination. Perhaps SUN only trimerizes after reaching the INM, and would therefore only gain KASH-binding competence at the target membrane. Another option is that one of the two binding partners is bound to a chaperone before engaging with the proper partner at the NE. The conserved prolines in KASH proteins also invite to speculations about a regulatory role. Maintaining one or more in the *cis*-conformation (they are all *trans* in the bound peptide) would be another elegant way of preventing LINC formation at the wrong site.

**Laminopathies and LINCs**

Nuclear envelopathies are a subset of genetic diseases caused by mutations in NE proteins. Amongst the most studied are laminopathies, which are caused by mutations in genes encoding proteins of the nuclear lamina. Over 180 mutations have been described for the *LMNA* gene leading to at least 13 known diseases (Capell and Collins, 2006; Worman and Bonne, 2007). The *LMNA* gene encodes for lamin-A and -C. Imaging of cells carrying mutations in the lamina shows severe NE structural defects suggesting that nuclear morphology alterations might lead to disease (Capell and Collins, 2006).

As mentioned above, SUN proteins are located in the INM where they interact with other INM and nuclear proteins. In mammals, SUN1 interacts with lamin-A/C. A fascinating study provided some clues as to how SUN1 plays a role in laminopathies (Chen et al., 2012; Suh et al., 2012). Several mouse models of different laminopathies exist that illustrate well disease progression. In this study, several of these mouse models were further modified to also carry a *SUN1* deletion effectively creating a *SUN1 LMNA* double knocked-out ($Lmna^{-/-}$ $Sun1^{-/-}$). Unexpectedly, these mice survived longer than LMNA single knockout ($Lmna^{-/-}$). Analysis of $Lmna^{-/-}$ cells showed an overexpression of SUN1 and its accumulation in the golgi after the NE was saturated. Interestingly, nuclear morphology and proliferation defects were rescued by RNAi mediated

knockdown of SUN1. This study illustrates the role of LINC complex in disease and distinguishes SUN1 as an attractive therapeutic target for the treatment of laminopathies.

**Torsin ATPases**

Torsin proteins belong to the ATPases Associated with a variety of cellular Activities (AAA+ ATPases) superfamily. AAA+ ATPases use ATP hydrolysis to undergo conformational changes and to exert force onto a substrate (Erzberger and Berger, 2006; Wendler et al., 2012). These proteins are involved in many cellular processes including protein unfolding, refolding, degradation, chromatin modification, DNA replication, vesicle fusion, etc (Erzberger and Berger, 2006; Narlikar et al., 2013; Scott et al., 2005; Stinson et al., 2013). Despite being so functionally diverse, members of this superfamily have a set of canonical motifs (Erzberger and Berger, 2006; Wendler et al., 2012). On the primary sequence level, they all contain Walker-A and -B motifs. The Walker-A motif, also known as the P-loop, consists of the consensus sequence of GxxGxGKT/S (where x is any amino acid) and its involved in binding to ATP. The Walker-B motif, consensus sequence hhhhDE (where h is any hydrophobic amino acid) is involved in the hydrolysis step. Together the Walker-A and Walker-B motifs coordinate the β and γ phosphates of ATP and the water activating magnesium ion. Importantly, mutating the lysine residue of the Walker-A motif renders the protein incapable of binding ATP. Mutating the glutamate residue of the Walker-B motif to a glutamine (E/Q or "substrate-trap" mutant) renders ATPases incapable of hydrolyzing ATP, yet capable of engaging with substrates, a very important biochemical tool for the study of AAA+ ATPases (Goodchild and Dauer, 2005; Jokhi et al., 2013; Martin et al., 2008; Zhu et al., 2010). Other motifs that are also well conserved and play critical roles in the function of AAA+ ATPases include the sensor-I motif, sensor-II and arginine-finger (Wendler et al., 2012). The sensor-I motif is composed of a polar residue that helps coordinate the attacking water for hydrolysis. The sensor-II motif has the consensus sequence GAR and plays a role in ATP

turnover. The arginine-finger, works by activating an adjacent ATPase subunit, with a mechanism akin to GTPase activating proteins. Torsin ATPases also contain a signal peptide in their N-terminus, which targets them to the endoplasmic reticulum (Vander Heyden et al., 2011; 2009). Furthermore, they contain some fundamental variations in two of the motifs described above. First is a single residue variation in the Walker-A: GxxGxGK**T/S** to GxxGxGK**N**. The second consists of a G**AR** to G**CK** change in the sensor-II motif. The implications of this variation are discussed further below.

AAA+ ATPases have the same domain organization composed of a variable N domain, a large domain, also called the AAA+ core, and a small domain. The AAA+ core adopts a compact αβα fold containing the Walker-A, Walker-B and sensor-I ATP binding motifs. The small domain is located C terminal to the AAA+ core. It adopts a compact helical bundle that contains the sensor-II motif. The high structural homology between AAA+ ATPases and the conservation of the motifs makes the structural prediction an attractive approach to biochemically and structurally understand Torsin ATPases. Structural models of the Torsin homologue OOC-5 in *Caenorhabditis elegans* have provided some clues for the role of the sensor-II motif in ATPase activity (Zhu et al., 2008). In the model the sensor-II motif cysteine is located adjacent to another conserved cysteine in the small domain and therefore predicted to form an intramolecular disulfide bond. Furthermore, the authors go on to show that the redox state of OOC-5 influences its ability to bind ATP and ADP and cause local conformational changes. Moreover, cysteine to serine mutations on the small domain severely affect embryo hatch rate when compared to wild-type, indicating that these cysteines are essential. Further analysis of human Tor1A shows that mutating these cysteines hinders its ability to bind LAP1 and LULL1 similar to the effect seen with the ΔE mutation (Zhu et al., 2010).

## Dystonia and TorsinA

Dystonia is a neurological movement disorder that affects a minimum of 300,000 people in North America alone. The most common type of dystonia is the early-onset generalized dystonia or DYT1 dystonia, which is inherited in autosomal-dominant fashion with reduced penetrance (Granata et al., 2010; Warner et al., 2010). The age of onset is typically between 5 and 28 years of age, which corresponds to a period of motor learning and synaptic plasticity in the basal ganglia (Granata et al., 2010). Most cases of DYT1 dystonia are caused by an in-frame three base pair deletion (ΔGAG) leading to a single glutamate deletion (ΔE) in the small domain of Tor1A (Ozelius et al., 1997a; 1997b). Interestingly, this disease shows no neuropathology or neurodegeneration suggesting that circuitry defects in neurons are the problem rather than cell death. Tor1A is expressed ubiquitously, yet the ΔE mutation seems to only affect neurons, most likely due to compensation by one of at least three Tor1A homologues, Tor1B, Tor2A and Tor3A (Jungwirth et al., 2010). In particular Tor1B is 67% identical and has been suggested to compensate for Tor1A in non-neuronal tissue.

## Tor1A-LAP1 and Tor1A-LULL1 complexes

A specific cellular function for Tor1A remains enigmatic, however, several binding partners have been characterized providing some clues about its role in the cell. The two main binding partners for Tor1A were discovered using FRAP (fluorescence recovery after photobleaching) experiments in which the dynamics of GFP labeled Tor1A-wt, Tor1A-E171Q (substrate-trap) and Tor1A-ΔE (disease mutant) were analyzed (Goodchild and Dauer, 2005). In the ER, the rate of recovery for Tor1A-wt, Tor1A-E/Q and Tor1A-ΔE were comparable, however the rate of recovery in the NE was significantly slower for Tor1A-E/Q and Tor1A-ΔE when compared to Tor1A-wt. This implied that Tor1A-E/Q and Tor1A-ΔE were engaging with a substrate in the NE. This led to the characterization of Lamina-Associated Polypeptide-1 (LAP1) as a Tor1A binding partner. Further sequence

analysis uncovered another ER resident protein capable of binding Tor1A that shared high sequence conservation with the luminal domain of LAP1, termed Luminal domain-Like LAP1 or LULL1. LAP1 localizes Tor1A to the NE whereas LULL1 targets Tor1A to both the ER and NE.

Specific functions of these two proteins remain unclear. However, LAP1 also interacts with lamina, chromosomes and the INM protein emerin and is phosphorylated during interphase and mitosis (Goodchild and Dauer, 2005; Shin et al., 2013). Additionally, LAP1 was recently shown to bind and be dephosphorylated by Protein Phosphatase 1 (PP1), a step suggested to be important for NE assembly (Santos et al., 2013).

Since LAP1 and LULL1 bind more tightly to ATP-bound Tor1A it was assumed that they were substrates (Goodchild and Dauer, 2005). However, a recent study suggests that LAP1 and LULL1 are activators of Torsin ATPases rather than substrates (Zhao et al., 2013). The authors show Tor1A, Tor1B and Tor3A are slow ATPases by themselves. Nevertheless, upon the addition of LAP1 or LULL1, ATPase activity increased significantly, implying that LAP1 and LULL1 work as cofactors required for Torsin ATPase activity. As expected, Torsin ATPase activity was abolished by the E171Q and the ΔE mutations. This report also shows that Tor1A-LAP1/LULL1 complexes exhibit a slow ATP hydrolysis rate suggesting a "switch-like" ATPase activity.

**Torsin ATPases in the Endoplasmic Reticulum and Nuclear Envelope**

Torsin ATPases have not been attributed a specific function. However, growing evidence suggest a role on direct or indirect ER/NE membrane architecture and remodeling. Human Tor1A and Tor1B contain a hydrophobic stretch of 20 amino acids in their N-terminus thought to be a TM domain. By epitope tagging Tor1A between the N-terminal signal peptide and the hydrophobic region and then either permeabilizing all membranes with Triton-X-100 or selectively

permeabilizing the plasma membrane with digitonin, it was shown that the tag was only accessible to antibodies upon full membrane permeabilization suggesting that the hydrophobic region of Torsins do not traverse the lipid bilayer (Vander Heyden et al., 2011). It was also shown that this hydrophobic region is sufficient to direct membrane interaction of unrelated soluble proteins. Tor1A also displayed a preference to flat ER sheets and NE, suggesting that this helix serves as ER/NE retention signal.

Molecular transport between the cytoplasm and nucleus has long been attributed to only the NPC (Brohawn et al., 2009). However, emerging evidence suggest that transport of molecules too large to traverse across the NPC can occur via budding through the NE. Two examples of this process have been characterized. The first is herpes virus capsid nuclear egress (Maric et al., 2011; Mettenleiter et al., 2006). Herpes virus capsid pre-assembly happens inside the nucleus. After its formation, the pro-capsid interacts with the INM, budding into the perinuclear space, becoming membrane enclosed. This vesicle fuses with the ONM allowing for Herpes virus capsid nuclear exit. Interestingly, upon overexpression of Tor1A, capsids were able to enter the perinuclear space, however, the resulting vesicles are unable to escape and accumulated in the ER/NE lumen (Maric et al., 2011). Another example of budding through the NE is that of mega Ribonucleoprotein (RNP) particles transport. These particles are much too large to traverse the NPC and their formation had long been debated. A recent study showed that megaRNPs form inside the nucleus and co-localize with the A-type lamin LamC (Speese et al., 2012). This is followed by envelopment of the particle into the perinuclear space and exit through the ONM. Torsin ΔE and E/Q mutants exhibited an accumulation of vesicles inside the NE (Jokhi et al., 2013). Furthermore, Torsin E/Q mutants accumulated in the base of the vesicles in the INM leading to the hypothesis that Torsin ATPases might play a role in membrane vesicle scission and that this process might be affected in dystonia.

Importantly, Tor1A and Tor1B are able to bind the KASH peptides of Nesprin-1, Nesprin-2 and Nesprin-3 (Nery et al., 2008). Upon overexpression, YFP tagged Nesprin-3 co-inmunoprecipitates with Tor1A variants with Tor1A-ΔE showing the strongest interaction. In the absence of Tor1A, YFP-Nesprin-3 localized primarily to the ER. Additionally, YFP-Nesprin-3 was also enriched in the ER in cells from DYT patients. Further, wound healing assays with cells lacking Tor1A showed a delay in migration and nuclear polarization when compared to wild-type cells. A separate study shows that upon LULL1 overexpression, Tor1A concentrates in the NE leading to the displacement of LINC complexes (Vander Heyden et al., 2009). Together these studies suggest that Tor1A and Tor1B play an important role in the regulation of LINC complex formation by promoting assembly or disassembly of the complex.

## REFERENCES


Alvarez, B.H., Gruber, M., Ursinus, A., Dunin-Horkawicz, S., Lupas, A.N., and Zeth, K. (2010). A transition from strong right-handed to canonical left-handed supercoiling in a conserved coiled-coil segment of trimeric autotransporter adhesins. J. Struct. Biol. *170*, 236–245.

Borrego-Pinto, J., Borrego-Pinto, J., Jegou, T., Jegou, T., Osorio, D.S., Osorio, D.S., Auradé, F., Auradé, F., Gorjánácz, M., Gorjánácz, M., et al. (2012). Samp1 is a component of TAN lines and is required for nuclear movement. J Cell Sci.

Brohawn, S.G., Brohawn, S.G., Partridge, J.R., Partridge, J.R., Whittle, J.R.R., Whittle, J.R.R., Schwartz, T.U., and Schwartz, T.U. (2009). The nuclear pore complex has entered the atomic age. Structure *17*, 1156–1168.

Capell, B.C., and Collins, F.S. (2006). Human laminopathies: nuclei gone genetically awry. Nat Rev Genet *7*, 940–952.

Chen, C.-Y., Chen, C.-Y., Chi, Y.-H., Chi, Y.-H., Mutalif, R.A., Mutalif, R.A., Starost, M.F., Starost, M.F., Myers, T.G., Myers, T.G., et al. (2012). Accumulation of the inner nuclear envelope protein sun1 is pathogenic in progeric and dystrophic laminopathies. Cell *149*, 565–577.

Chikashige, Y., Chikashige, Y., Tsutsumi, C., Tsutsumi, C., Yamane, M., Yamane, M., Okamasa, K., Okamasa, K., Haraguchi, T., Haraguchi, T., et al. (2006). Meiotic proteins bqt1 and bqt2 tether telomeres to form the bouquet arrangement of chromosomes. Cell *125*, 59–69.

Crisp, M., Crisp, M., Liu, Q., Liu, Q., Roux, K., Roux, K., Rattner, J.B., Rattner, J.B., Shanahan, C., Shanahan, C., et al. (2006). Coupling of the nucleus and cytoplasm: role of the LINC complex. J Cell Biol *172*, 41–53.

Denic, V. (2012). A portrait of the GET pathway as a surprisingly complicated young man. Trends in Biochemical Sciences *37*, 411–417.

Ding, X., Ding, X., Xu, R., Xu, R., Yu, J., Yu, J., Xu, T., Xu, T., Zhuang, Y., Zhuang, Y., et al. (2007). SUN1 is required for telomere attachment to nuclear envelope and gametogenesis in mice. Developmental Cell *12*, 863–872.

Erzberger, J.P., and Berger, J.M. (2006). Evolutionary relationships and structural mechanisms of AAA+ proteins. Annu Rev Biophys Biomol Struct *35*, 93–114.

Field, M.C., Horn, D., Alsford, S., Koreny, L., and Rout, M.P. (2012). Telomeres, tethers and trypanosomes. Nucleus (Austin, Tex) *3*, 478–486.

Fridkin, A., Mills, E., Margalit, A., Neufeld, E., Lee, K.K., Feinstein, N., Cohen, M., Wilson, K.L., and Gruenbaum, Y. (2004). Matefin, a Caenorhabditis elegans germ line-specific SUN-domain nuclear membrane protein, is essential for early embryonic and germ cell development. Proc Natl Acad Sci USA *101*, 6987–6992.

Goodchild, R.E., and Goodchild, R.E. (2005). The AAA+ protein torsinA interacts with a conserved domain present in LAP1 and a novel ER protein. J Cell Biol *168*, 855–862.

Gorlich, D., and Kutay, U. (1999). Transport between the cell nucleus and the cytoplasm. Annu. Rev. Cell Dev. Biol. *15*, 607–660.

Göb, E., Göb, E., Schmitt, J., Schmitt, J., Benavente, R., Benavente, R., Alsheimer, M., and Alsheimer, M. (2010). Mammalian sperm head formation involves different polarization of two novel LINC complexes. PLoS ONE *5*, e12072.

Granata, A., Granata, A., Warner, T.T., and Warner, T.T. (2010). The role of torsinA in dystonia. European Journal of Neurology *17*, 81–87.

Gruber, M., and Lupas, A.N. (2003). Historical review: another 50th anniversary--new periodicities in coiled coils. Trends in Biochemical Sciences *28*, 679–685.

Gruenbaum, Y., Margalit, A., Goldman, R.D., Shumaker, D.K., and Wilson, K.L.

(2005). The nuclear lamina comes of age. Nat. Rev. Mol. Cell Biol. *6*, 21–31.

Horn, H.F., Brownstein, Z., Lenz, D.R., Shivatzki, S., Dror, A.A., Dagan-Rosenfeld, O., Friedman, L.M., Roux, K.J., Kozlov, S., Jeang, K.-T., et al. (2013). The LINC complex is essential for hearing. J. Clin. Invest. *123*, 740–750.

Ingber, D.E. (1997). Tensegrity: the architectural basis of cellular mechanotransduction. Annu. Rev. Physiol. *59*, 575–599.

Jokhi, V., Ashley, J., Nunnari, J., Noma, A., Ito, N., Wakabayashi-Ito, N., Moore, M.J., and Budnik, V. (2013). Torsin mediates primary envelopment of large ribonucleoprotein granules at the nuclear envelope. Cell Reports *3*, 988–995.

Jungwirth, M., Dear, M.L., Brown, P., Holbrook, K., and Goodchild, R. (2010). Relative tissue expression of homologous torsinB correlates with the neuronal specific importance of DYT1 dystonia-associated torsinA. Hum Mol Genet *19*, 888–900.

Kracklauer, M.P., Kracklauer, M.P., Banks, S.M.L., Banks, S.M.L., Xie, X., Xie, X., Wu, Y., Wu, Y., Fischer, J.A., and Fischer, J.A. (2007). Drosophila klaroid encodes a SUN domain protein required for Klarsicht localization to the nuclear envelope and nuclear migration in the eye. Fly (Austin) *1*, 75–85.

Kracklauer, M.P., Link, J., and Alsheimer, M. (2013). Chapter Five - LINCing the Nuclear Envelope to Gametogenesis. In Gametogenesis, P.M. Wassarman, ed. (Academic Press), pp. 127–157.

Lindeman, R.E., and Pelegri, F. (2012). Localized products of futile cycle/lrmp promote centrosome-nucleus attachment in the zebrafish zygote. Curr. Biol. *22*, 843–851.

Lombardi, M.L., Lombardi, M.L., Lammerding, J., and Lammerding, J. (2011). Keeping the LINC: the importance of nucleocytoskeletal coupling in intracellular force transmission and cellular function. Biochem. Soc. Trans. *39*, 1729–1734.

Luxton, G.W.G., and Gundersen, G.G. (2011). Orientation and function of the nuclear-centrosomal axis during cell migration. Curr. Opin. Cell Biol. *23*, 579–588.

Luxton, G.W.G., Luxton, G.W.G., Gomes, E.R., Gomes, E.R., Folker, E.S., Folker, E.S., Vintinner, E., Vintinner, E., Gundersen, G.G., and Gundersen, G.G. (2010). Linear arrays of nuclear envelope proteins harness retrograde actin flow for nuclear movement. Science *329*, 956–959.

Malone, C.J., Fixsen, W.D., Horvitz, H.R., and Han, M. (1999). UNC-84 localizes to the nuclear envelope and is required for nuclear migration and anchoring during C. elegans development. Development *126*, 3171–3181.

Maniotis, A.J., Chen, C.S., and Ingber, D.E. (1997). Demonstration of mechanical connections between integrins, cytoskeletal filaments, and nucleoplasm that stabilize nuclear structure. Proc Natl Acad Sci USA *94*, 849–854.

Maric, M., Shao, J., Ryan, R.J., Wong, C.S., Gonzalez-Alegre, P., and Roller, R.J. (2011). A Functional Role for TorsinA in Herpes Simplex Virus 1 Nuclear Egress. Journal of Virology *85*, 9667–9679.

Martin, A., Baker, T.A., and Sauer, R.T. (2008). Protein unfolding by a AAA+ protease is dependent on ATP-hydrolysis rates and substrate energy landscapes. Nat Struct Mol Biol *15*, 139–145.

Mettenleiter, T.C., Klupp, B.G., and Granzow, H. (2006). Herpesvirus assembly: a tale of two membranes. Host Microbe Interactions: Fungi/Host Microbe Interactions: Parasites/Host Microbe Interactions: Viruses *9*, 423–429.

Méndez-López, I., and Worman, H.J. (2012). Inner nuclear membrane proteins: impact on human disease. Chromosoma *121*, 153–167.

Morimoto, A., Shibuya, H., Zhu, X., Kim, J., Ishiguro, K.-I., Han, M., and Watanabe, Y. (2012). A conserved KASH domain protein associates with telomeres, SUN1, and dynactin during mammalian meiosis. J Cell Biol *198*, 165–

172.

Narlikar, G.J., Sundaramoorthy, R., and Owen-Hughes, T. (2013). Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. Cell *154*, 490–503.

Nery, F.C., Nery, F.C., Zeng, J., Zeng, J., Niland, B.P., Niland, B.P., Hewett, J., Hewett, J., Farley, J., Farley, J., et al. (2008). TorsinA binds the KASH domain of nesprins and participates in linkage between nuclear envelope and cytoskeleton. J Cell Sci *121*, 3476–3486.

Ozelius, L.J., Hewett, J., Kramer, P., Bressman, S.B., Shalish, C., de Leon, D., Rutter, M., Risch, N., Brin, M.F., Markova, E.D., et al. (1997a). Fine localization of the torsion dystonia gene (DYT1) on human chromosome 9q34: YAC map and linkage disequilibrium. Genome Res. *7*, 483–494.

Ozelius, L.J., Hewett, J.W., Page, C.E., Bressman, S.B., Kramer, P.L., Shalish, C., de Leon, D., Brin, M.F., Raymond, D., Corey, D.P., et al. (1997b). The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein. Nat Genet *17*, 40–48.

Padmakumar, V.C. (2005). The inner nuclear membrane protein Sun1 mediates the anchorage of Nesprin-2 to the nuclear envelope. J Cell Sci *118*, 3419–3430.

Penkner, A.M., Penkner, A.M., Fridkin, A., Fridkin, A., Gloggnitzer, J., Gloggnitzer, J., Baudrimont, A., Baudrimont, A., Machacek, T., Machacek, T., et al. (2009). Meiotic chromosome homology search involves modifications of the nuclear envelope protein Matefin/SUN-1. Cell *139*, 920–933.

Rapoport, T.A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature *450*, 663–669.

Rose, A., and Schlieker, C. (2012). Alternative nuclear transport for cellular protein quality control. Trends in Cell Biology *22*, 509–514.

Santos, M., Rebelo, S., Van Kleeff, P.J.M., Kim, C.E., Dauer, W.T., Fardilha, M., da Cruz e Silva, O.A., and da Cruz e Silva, E.F. (2013). The Nuclear Envelope Protein, LAP1B, Is a Novel Protein Phosphatase 1 Substrate. PLoS ONE *8*, e76788.

Sato, A., Sato, A., Isaac, B., Isaac, B., Phillips, C.M., Phillips, C.M., Rillo, R., Rillo, R., Carlton, P.M., Carlton, P.M., et al. (2009). Cytoskeletal forces span the nuclear envelope to coordinate meiotic chromosome pairing and synapsis. Cell *139*, 907–919.

Schmitt, J., Schmitt, J., Benavente, R., Benavente, R., Hodzic, D., Hodzic, D., Höög, C., Höög, C., Stewart, C.L., Stewart, C.L., et al. (2007). Transmembrane protein Sun2 is involved in tethering mammalian meiotic telomeres to the nuclear envelope. Proc Natl Acad Sci USA *104*, 7426–7431.

Scott, A., Scott, A., Chung, H.-Y., Chung, H.-Y., Gonciarz-Swiatek, M., Gonciarz-Swiatek, M., Hill, G.C., Hill, G.C., Whitby, F.G., Whitby, F.G., et al. (2005). Structural and mechanistic studies of VPS4 proteins. Nat. Rev. Mol. Cell Biol. *24*, 3658–3669.

Shimada, N., Shimada, N., Inouye, K., Inouye, K., Sawai, S., Sawai, S., Kawata, T., and Kawata, T. (2010). SunB, a novel Sad1 and UNC-84 domain-containing protein required for development of Dictyostelium discoideum. Development, Growth & Differentiation *52*, 577–590.

Shin, J.-Y., Méndez-López, I., Wang, Y., Hays, A.P., Tanji, K., Lefkowitch, J.H., Schulze, P.C., Worman, H.J., and Dauer, W.T. (2013). Lamina-associated polypeptide-1 interacts with the muscular dystrophy protein emerin and is essential for skeletal muscle maintenance. Developmental Cell *26*, 591–603.

Sosa, B.A., Rothballer, A., Kutay, U., and Schwartz, T.U. (2012). LINC Complexes Form by Binding of Three KASH Peptides to Domain Interfaces of Trimeric SUN Proteins. Cell *149*, 1035–1047.

Speese, S.D., Speese, S.D., Ashley, J., Ashley, J., Jokhi, V., Jokhi, V., Nunnari, J., Nunnari, J., Barria, R., Barria, R., et al. (2012). Nuclear envelope budding enables large ribonucleoprotein particle export during synaptic Wnt signaling. Cell *149*, 832–846.

Starr, D.A., and Han, M. (2002). Role of ANC-1 in tethering nuclei to the actin cytoskeleton. Science *298*, 406–409.

Starr, D.A., Starr, D.A., Fridolfsson, H.N., and Fridolfsson, H.N. (2010). Interactions between nuclei and the cytoskeleton are mediated by SUN-KASH nuclear-envelope bridges. Annu. Rev. Cell Dev. Biol. *26*, 421–444.

Stinson, B.M., Nager, A.R., Glynn, S.E., Schmitz, K.R., Baker, T.A., and Sauer, R.T. (2013). Nucleotide Binding and Conformational Switching in the Hexameric Ring of a AAA+ Machine. Cell *153*, 628–639.

Suh, Y., Suh, Y., Kennedy, B.K., and Kennedy, B.K. (2012). Dialing Down SUN1 for Laminopathies. Cell *149*, 509–510.

Tapley, E.C., and Starr, D.A. (2013). Connecting the nucleus to the cytoskeleton by SUN–KASH bridges across the nuclear envelope. Curr. Opin. Cell Biol. *25*, 57–62.

Tapley, E.C., Ly, N., and Starr, D.A. (2011). Multiple mechanisms actively target the SUN protein UNC-84 to the inner nuclear membrane. Mol Biol Cell.

Technau, M., Technau, M., Roth, S., and Roth, S. (2008). The Drosophila KASH domain proteins Msp-300 and Klarsicht and the SUN domain protein Klaroid have no essential function during oogenesis. Fly (Austin) *2*, 82–91.

Turgay, Y., Turgay, Y., Ungricht, R., Ungricht, R., Rothballer, A., Rothballer, A., Kiss, A., Kiss, A., Csucs, G., Csucs, G., et al. (2010). A classical NLS and the SUN domain contribute to the targeting of SUN2 to the inner nuclear membrane. Embo J. *29*, 2262–2275.

Vander Heyden, A.B., Naismith, T.V., Snapp, E.L., and Hanson, P.I. (2011). Static retention of the lumenal monotopic membrane protein torsinA in the endoplasmic reticulum. Embo J. *30*, 3217–3231.

Vander Heyden, A.B., Vander Heyden, A.B., Naismith, T.V., Naismith, T.V., Snapp, E.L., Snapp, E.L., Hodzic, D., Hodzic, D., Hanson, P.I., and Hanson, P.I. (2009). LULL1 retargets TorsinA to the nuclear envelope revealing an activity that is impaired by the DYT1 dystonia mutation. Mol Biol Cell *20*, 2661–2672.

Wang, N., Tytell, J.D., and Ingber, D.E. (2009). Mechanotransduction at a distance: mechanically coupling the extracellular matrix with the nucleus. Nat. Rev. Mol. Cell Biol. *10*, 75–82.

Wang, W., Wang, W., Shi, Z., Shi, Z., Jiao, S., Jiao, S., Chen, C., Chen, C., Wang, H., Wang, H., et al. (2012). Structural insights into SUN-KASH complexes across the nuclear envelope. Cell Res.

Warner, T.T., Granata, A., and Schiavo, G. (2010). TorsinA and DYT1 dystonia: a synaptopathy? Biochem. Soc. Trans. *38*, 452.

Wendler, P., Wendler, P., Ciniawsky, S., Ciniawsky, S., Kock, M., Kock, M., Kube, S., and Kube, S. (2012). Structure and function of the AAA+ nucleotide binding pocket. Biochim Biophys Acta *1823*, 2–14.

Worman, H.J., and Bonne, G. (2007). "Laminopathies": a wide spectrum of human diseases. Experimental Cell Research *313*, 2121–2133.

Zhao, C., Brown, R.S.H., Chase, A.R., Eisele, M.R., and Schlieker, C. (2013). Regulation of Torsin ATPases by LAP1 and LULL1. Proc Natl Acad Sci USA *110*, E1545–E1554.

Zhou, Z., Zhou, Z., Du, X., Du, X., Cai, Z., Cai, Z., Song, X., Song, X., Zhang, H., Zhang, H., et al. (2012). Structure of Sad1-UNC84 Homology (SUN) Domain Defines Features of Molecular Bridge in Nuclear Envelope. Journal of Biological Chemistry *287*, 5317–5326.

Zhu, L., Millen, L., Mendoza, J.L., and Thomas, P.J. (2010). A unique redox-sensing sensor II motif in TorsinA plays a critical role in nucleotide and partner binding. Journal of Biological Chemistry *285*, 37271–37280.

Zhu, L., Zhu, L., Wrabl, J.O., Wrabl, J.O., Hayashi, A.P., Hayashi, A.P., Rose, L.S., Rose, L.S., Thomas, P.J., and Thomas, P.J. (2008). The torsin-family AAA+ protein OOC-5 contains a critical disulfide adjacent to Sensor-II that couples redox state to nucleotide binding. Mol Biol Cell *19*, 3599–3612.

## Chapter Two

# LINC Complexes Form by Binding of Three KASH Peptides to Domain Interfaces of Trimeric SUN Proteins

This chapter was previously published as *Sosa, B. A., Rothballer, A., Kutay, U. & Schwartz, T. U. LINC Complexes Form by Binding of Three KASH Peptides to Domain Interfaces of Trimeric SUN Proteins. *Cell* **149,** 1035–1047 (2012).

*Sosa, B.A. performed the experiments for table 1 and figures 3, 4, 5, 7, 8, 10A, 10B and 11.

**SUMMARY**

Linker of Nucleoskeleton and Cytoskeleton (LINC) complexes span the nuclear envelope and are composed of KASH and SUN proteins residing in the outer and inner nuclear membrane, respectively. LINC formation relies on direct binding of KASH and SUN in the perinuclear space. Thereby, molecular tethers are formed that can transmit forces for chromosome movements, nuclear migration and anchorage. We present crystal structures of the human SUN2-KASH1/2 complex, the core of the LINC complex. The SUN2 domain is rigidly attached to a trimeric coiled-coil that prepositions it to bind three KASH peptides. The peptides bind in three deep and expansive grooves formed between adjacent SUN domains, effectively acting as molecular glue. In addition, a disulfide between conserved cysteines on SUN and KASH covalently links both proteins. The structure provides the basis of LINC complex formation and suggests a model for how LINC complexes might arrange into higher-order clusters to enhance force-coupling.

## INTRODUCTION

The nuclear envelope (NE), a double lipid bilayer, forms the boundary of the nuclear compartment and serves as a physical barrier between chromatin in the nuclear interior and the cytoplasm. The integrity and physical properties of the NE play pivotal roles in cellular homeostasis, signaling and organization of intracellular architecture (Mekhail and Moazed, 2010; Stewart et al., 2007). The exchange of material between nucleus and cytoplasm is guaranteed by passageways through the NE that are formed by nuclear pore complexes (Onischenko and Weis, 2011).

Communication between nucleus and nuclear exterior is not limited to transport of macromolecules across the NE but also involves connections of the NE and the cytoskeleton. These physical linkages are instrumental for plasticity of cellular organization and are required for processes such as nuclear migration and anchorage, meiotic chromosome movements, centrosome positioning, and global organization of the cytoskeleton (reviewed in Burke and Roux, 2009; Starr and Fridolfsson, 2010; Tzur et al., 2006; Worman and Gundersen, 2006). The molecular tether between NE and cytoskeletal elements is built by protein complexes referred to as LINC (Linker of Nucleoskeleton and Cytoskeleton) (Crisp et al., 2006). LINC complexes are conserved from yeast to human. They form bridges across the perinuclear space by pairing of KASH (Klarsicht, ANC-1, and Syne Homology) and SUN (Sad1, UNC-84) proteins that localize to outer (ONM) and inner nuclear membrane (INM), respectively (Razafsky and Hodzic, 2009; Starr and Fischer, 2005).

KASH proteins are tail-anchored membrane proteins. Their large cytoplasmic domains extend into the cytoplasm, where they interact with actin filaments, microtubules, intermediate filaments or centrosomes. At their C termini, KASH proteins expose a short peptide of around 30 residues into the perinuclear space. This luminal peptide together with the preceding transmembrane segment

provides the characteristic signature of KASH proteins referred to as KASH domain (Starr and Han, 2002). The luminal KASH peptides of many family members are similar. Their conserved features are the presence of aromatic/hydrophobic residues at defined positions and proline(s) close to their C-terminal ends. Studies in various organisms have shown that KASH domains are necessary and sufficient for localization of KASH proteins to the ONM (reviewed in Starr and Fridolfsson, 2010). This is achieved by a direct interaction of the KASH peptide with members of the SUN protein family residing in the INM (Ostlund et al., 2009; Padmakumar et al., 2005).

Vertebrates possess at least 4 different KASH proteins (SYNE/Nesprin-1 to -4) that contain spectrin repeats in their cytoplasmic domains (Apel et al., 2000; Zhang et al., 2001). Nesprin-1 and -2 (SYNE1, 2) exist in numerous splice isoforms. The giant isoforms of Nesprin-1 and -2 interact with actin via their N-terminal actin-binding domain, whereas other isoforms can associate with microtubule motors (Padmakumar et al., 2004; Yu et al., 2011; Zhang et al., 2009; Zhen et al., 2002). Nesprin-3 and -4 bind to plectin and kinesin-1, respectively, thereby connecting the nucleus to intermediate filaments and microtubules (Roux et al., 2009; Wilhelmsen et al., 2005). Reflecting the importance of NE-cytoskeletal interactions for human health, mutations in Nesprins have been linked to the pathogenesis of Emery-Dreifuss muscular dystrophy (Puckelwartz et al., 2009; Zhang et al., 2007a).

SUN proteins form the inner half of the LINC complex. They are anchored in the INM by at least one transmembrane segment, exposing their N termini to the nucleoplasm. These nucleoplasmic domains harbor signals that aid targeting of SUN proteins to the INM and mediate association with nuclear binding partners such as the nuclear lamina (Fridkin et al., 2004; Haque et al., 2006; Lee et al., 2002). The C-terminal segments of SUN proteins reside in the perinuclear space and consist of coiled-coil regions followed by the conserved SUN domain, which is required for KASH binding. Among the five characterized mammalian SUN

proteins, only SUN1 and SUN2 are widely expressed (Crisp et al., 2006; Wang et al., 2006) and can bind to all four Nesprins (Ketema et al., 2007; Roux et al., 2009; Stewart-Hutchinson et al., 2008). A double knockout of SUN1 and SUN2 in mice is lethal (Lei et al., 2009), similar to Nesprin-1/2 deletion (Zhang et al., 2007b). Through their cellular function in chromatin organization and nuclear migration and anchorage, SUN1/2 and Nesprin-1/2 play pivotal roles in a variety of developmental processes in vertebrates, ranging from gametogenesis to the development of muscles, brain and the retina (Ding et al., 2007; Lei et al., 2009; Puckelwartz et al., 2009; Yu et al., 2011; Zhang et al., 2010; Zhang et al., 2009). How SUN and KASH interact to provide mechanical coupling between structural components of the nucleus and the cytoskeleton has remained elusive. Here, we present crystal structures of the SUN domain of human SUN2 in its free and KASH-bound conformations. These structures reveal a trimeric organization of SUN domains intimately bound to three KASH peptides at the SUN protomer interfaces. Biochemical and cell biological studies support and further qualify our structural findings.

## RESULTS

### Mapping the KASH-binding competent region of human SUN2

To define the minimal region in human SUN2 required for binding of the luminal part of KASH domains, we used recombinant KASH fusion proteins composed of an N-terminal protein A tag (zz) and the C-terminal 29 residues of human Nesprin-1 or -2 (Figure 1A, zz-KASH1, zz-KASH2) in pulldown experiments. First, we tested whether the KASH fusion proteins could interact with endogenous SUN proteins present in HeLa cell detergent extract. Both KASH1 and KASH2 retrieved SUN1 and SUN2 from the extract as verified by immunoblotting (Figure 1B). Binding of SUN proteins to KASH1/2 was specific, as no binding of SUN proteins was observed to either zz alone or zz-KASH2Δ, in which the four C-terminal amino acids of KASH that are required for interaction with SUN proteins are missing (Padmakumar et al., 2005).

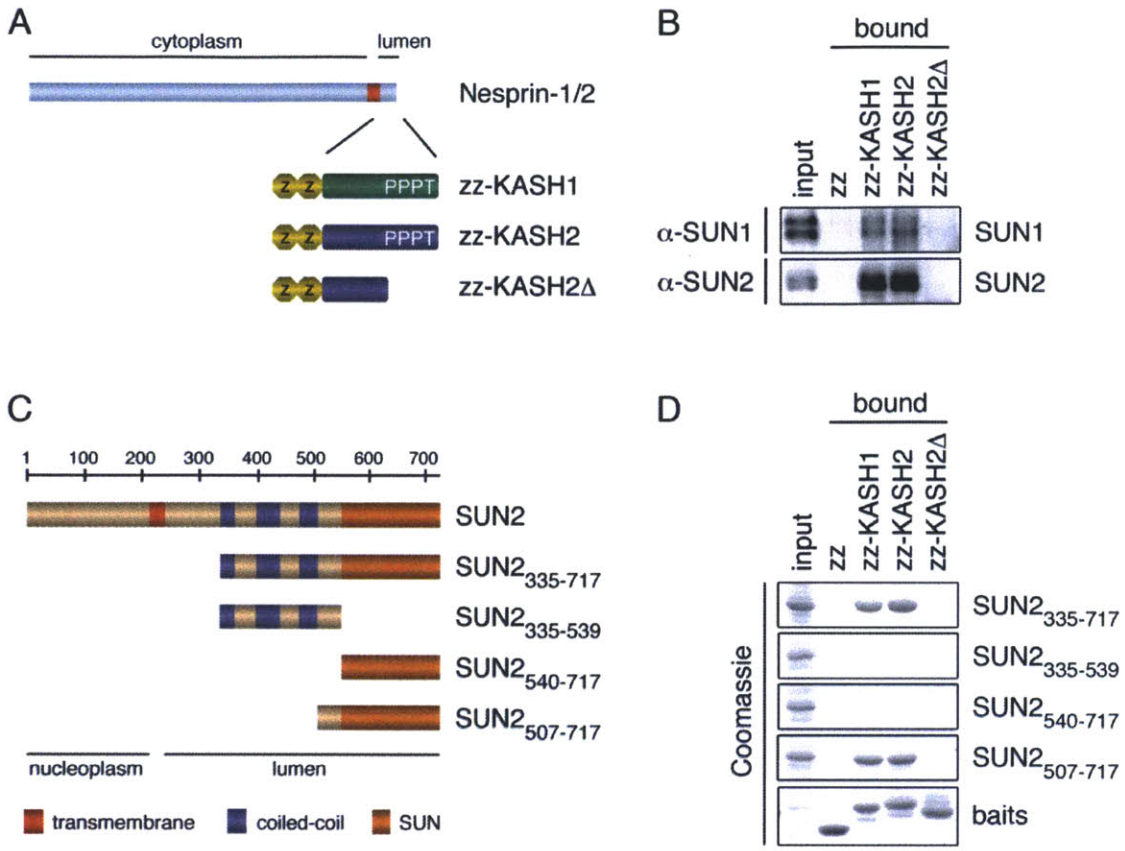


**Figure 1. A Minimal Coiled-coil and the SUN Domain are Required for KASH Binding.** (A) Schematic representation of zz-KASH fusion proteins. The luminal peptides of Nesprin-1 and Nesprin-2 (29 aa) were fused to an N-terminal protein A tag (zz-KASH1 and zz-KASH2). zz-KASH2Δ lacks the C-terminal amino acids PPPT. (B) SUN1 and SUN2 from HeLa cells bind to zz-KASH peptides in vitro. HeLa cells were lysed in RIPA buffer and the extract was incubated with immobilized zz-KASH fusion proteins. Bound proteins were eluted and analyzed by immunoblotting using α-SUN1 and α-SUN2 antibodies. Loads in input and pulldown lanes correspond to 2.5% and 20% of the total, respectively. (C) Schematic representation of SUN2 constructs used in pulldown experiments in (D). Coiled-coil regions were predicted using Parcoil2 (McDonnell et al., 2006). (D) KASH binding of recombinant SUN2 fragments. His-tagged SUN2 fragments were expressed in *E. coli*, purified and added at 0.25 μg/μl to *E. coli* lysate supplemented with RIPA detergents. Mixtures were incubated with immobilized zz-KASH fusion proteins. Bound proteins were eluted and analyzed by SDS-PAGE and Coomassie staining. Loads in the input and pulldown lanes correspond to 1.25% and 10% of the total, respectively.

Next, we tested direct KASH binding of recombinant SUN2 protein fragments comprising different regions of its luminal domain (Figure 1C,D). In agreement

with published data (Ostlund et al., 2009), we observed that a protein fragment comprising both the predicted coiled-coil regions and the C-terminal SUN domain of SUN2 ($SUN2_{335-717}$) efficiently bound to both KASH peptides. However, neither the isolated coiled-coil region ($SUN2_{335-539}$) nor the SUN domain alone ($SUN2_{540-717}$) was sufficient for KASH interaction. Further deletion analysis revealed that the coiled-coil region preceding the SUN domain could be significantly shortened without loosing KASH interaction. A fragment comprising an N-terminally extended SUN domain ($SUN2_{507-717}$) was found sufficient for efficient binding to both KASH peptides. When we introduced further truncations (Figure 2), we observed that binding to KASH decreased ($SUN2_{514-717}$) and finally became barely detectable ($SUN2_{521-717}$). As the protein segment 507-539 is not part of the conserved SUN domain and is just adjacent to the last predicted coiled-coil of SUN2, we reasoned that the contribution of residues 507-539 to KASH binding may be attributed to a potential role in organization of the SUN domain into a higher order structure. To test this assumption, we fused an unrelated, well-characterized coiled-coil region derived from the yeast transcription factor Gcn4 (Ciani et al., 2010) in front of the three extended SUN domain fragments ($SUN2_{521-717}$, $SUN2_{514-717}$, and $SUN2_{507-717}$). Strikingly, addition of this unrelated coiled-coil (ucc) restored the KASH binding capacity of $SUN2_{521-717}$, demonstrating that residues 521-717 of SUN2 are sufficient for interaction with KASH peptides and suggesting that coiled-coil regions in front of the SUN domain contribute to the higher order organization of SUN domains into a KASH-binding competent state.

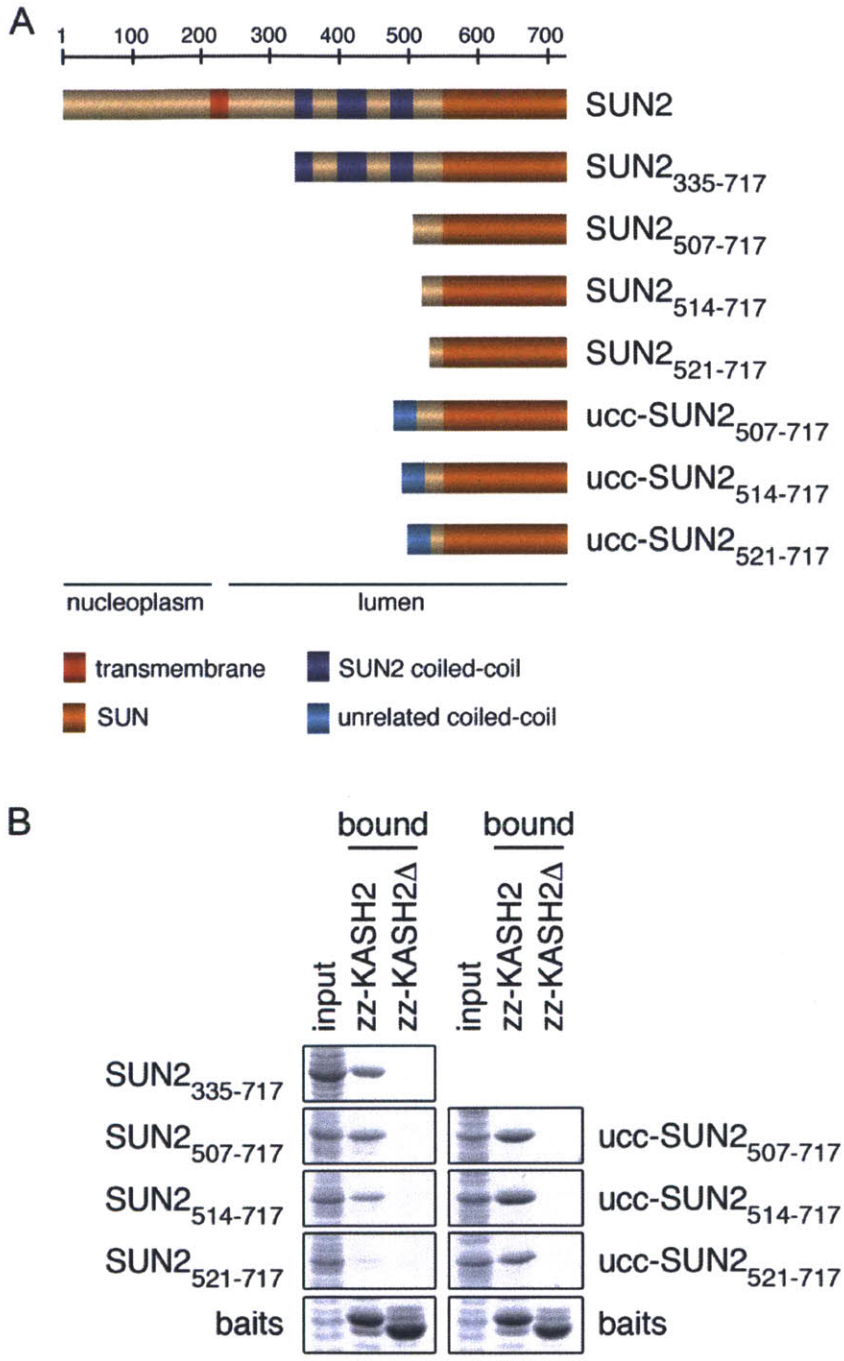


**Figure 2. Addition of an Unrelated Coiled-coil to the SUN Domain Supports KASH Binding.** (A) Schematic representation of SUN2 fragments and their derivatives. A luminal fragment of SUN2 was truncated stepwise to define the minimal requirements for KASH interaction. Note that also an unrelated coiled-coil (ucc) corresponding to a trimeric version of the GCN4 coiled-coil (Ciani et al., 2010) was used to replace the predicted coiled-coil region of SUN2 preceding the SUN domain. (B) KASH binding of SUN2 and ucc-SUN2 constructs. The indicated purified His-tagged SUN2 derivatives were added to *E. coli* lysate to the same concentration (0.8 μM for trimer) and incubated with immobilized zz-KASH2 or zz-KASH2Δ as in Figure 1D. Bound proteins were retrieved and analyzed by SDS-PAGE followed by Coomassie staining. Loads in input and pulldown lanes correspond to 2.5% and 20% of the total, respectively.

**Structure determination of apo-SUN2, SUN2-KASH1, and SUN2-KASH2**

After determining the minimal SUN domain of human SUN2, we set up crystallization screens for apo-$SUN2_{522-717}$, $SUN2_{522-717}$-KASH1 (Nesprin-1, residues 8769-8797) and $SUN2_{522-717}$-KASH2 (Nesprin-2, residues 6857-6885). We first obtained crystals of apo-SUN2, diffracting to 2.2 Å (Table 1). We used selenomethionine-derivatized protein to solve the structure with the single-anomalous dispersion (SAD) method (for details see Experimental Procedures). The refined apo-SUN2 structure ($R/R_{free}$ 18.9%/23.3%) was subsequently used to solve the structures of the SUN2-KASH1 and SUN2-KASH2 complexes by molecular replacement.

| Table 1. Data Collection and Refinement Statistics | | | | |
|---|---|---|---|---|
| | **hSUN2** | **hSUN2-L546M** | **hSUN2:KASH1** | **hSUN2:KASH2** |
| | | **Data collection** | | |
| Data set | Native | Selenomethionine | Native | Native |
| Space group | R32 | R32 | R32 | R32 |
| *a*, *b*, *c* (Å) | 79.6, 79.6, 199.1 | 79.7, 79.7, 200.2 | 79.6, 79.6, 256.4 | 79.3, 79.3, 260.0 |
| Wavelength (Å) | 0.9795 | 0.9792 | 0.9792 | 0.9792 |
| Resolution range (Å) | 56.7-2.2 (2.3-2.2) | 56.8-2.4 (2.5-2.4) | 85.5-2.3 (2.4-2.3) | 43.3-2.7 (2.8-2.7) |
| Total reflections | 72,988 | 56,714 | 49,146 | 52,267 |
| Unique reflections | 12,305 | 9,510 | 13,944 | 8,912 |
| Completeness (%) | 99.8 (99.9) | 99.6 (99.8) | 99.6 (97.7) | 99.9 (99.8) |
| Redundancy | 5.9 (6.1) | 3.1 (3.1) | 3.5 (3.4) | 5.9 (6.1) |
| $R_{merge}$(%) | 4.8 (57.0) | 4.6 (65.0) | 7.6 (70.7) | 5.5 (69.5) |
| $R_{r.i.m.}$(%) | 5.3 (62.3) | 6.6 (78.8) | 9.0 (83.3) | 6.1 (76.1) |
| $R_{p.i.m.}$(%) | 2.1 (24.8) | 2.7 (31.8) | 4.7 (43.4) | 2.5 (30.7) |
| $I/\sigma$ | 20.0 (3.5) | 22.2 (2.5) | 24.9 (1.5) | 21.0 (2.7) |
| Wilson B factor (Å$^2$) | 53.0 | 57.9 | 45.4 | 66.5 |
| | | **Refinement** | | |
| Resolution range (Å) | 56.7-2.2 | | 45.0-2.3 | 36.1-2.7 |
| $R_{work}$ (%) | 18.9 | | 19.6 | 18.4 |
| $R_{free}$ (%) | 23.3 | | 21.1 | 23.9 |
| | | **Number of Reflections** | | |
| Total | 12,304 | | 13,930 | 8,909 |
| $R_{free}$ reflections | 1,250 | | 1,393 | 891 |
| | | **Number of Atoms** | | |
| Protein | 1,585 | | 1,762 | 1,751 |
| Water | 101 | | 129 | 41 |
| n-Decyl-ß-maltoside | | | 25 | |
| | | **B Factors (Å$^2$)** | | |
| SUN2 | 45.8 | | 44.8 | 65.0 |
| KASH | | | 59.8 | 93.2 |
| Water | 48.2 | | 50.4 | 61.6 |
| n-Decyl-ß-maltoside | | | 71.7 | |
| | | **R.m.s. Deviations** | | |
| Bond length (Å) | 0.009 | | 0.007 | 0.009 |
| Bond angles (°) | 1.170 | | 1.023 | 1.141 |
| | | **Ramachandran Plot** | | |
| Favored (%) | 94.5 | | 94.5 | 93.1 |
| Allowed (%) | 5.5 | | 5.5 | 6.4 |
| Outliers (%) | 0 | | 0 | 0.5 |

The highest resolution shell is in parenthesis. $R_{merge}$ is the merging R factor. $R_{r.i.m.}$ is the redundancy independent merging R factor. $R_{p.i.m.}$ is the precision-indicating merging R factor. For definitions, see Weiss (2001).

**Structure of the SUN2-KASH2 complex**

The principal SUN domain of SUN2 folds into a compact β-sandwich (Figure 3A,B). Two sets of antiparallel β-sheets composed of 3 and 5 strands form this basic structure. The β-sandwich is decorated by two main additions, which are of major functional importance. At the N terminus, we find a helical extension (residues 525-540) involved in the formation of a coiled-coil. Between residues 567 and 587, there is a 20-residue extension that folds into a protruding antiparallel β-sheet, which we term the KASH-lid. As a third addition, residues 593-601 form a well-defined loop, which surrounds and coordinates a bound cation via five backbone carbonyls, thus we call this detail the cation-loop. Judged by the coordination sphere, the bond distances (Harding et al., 2006), and the temperature factor, we suggest that the bound cation is a potassium ion in our crystal. However, due to the possible malleability of the loop, the evolutionary relationship to other proteins (see below), and the ionic environment in the perinuclear space, a calcium ion might be coordinated *in vivo*. The arrangement of the cation-loop is further defined by an intrachain disulfide bond formed between conserved cysteines at positions 601 and 705.

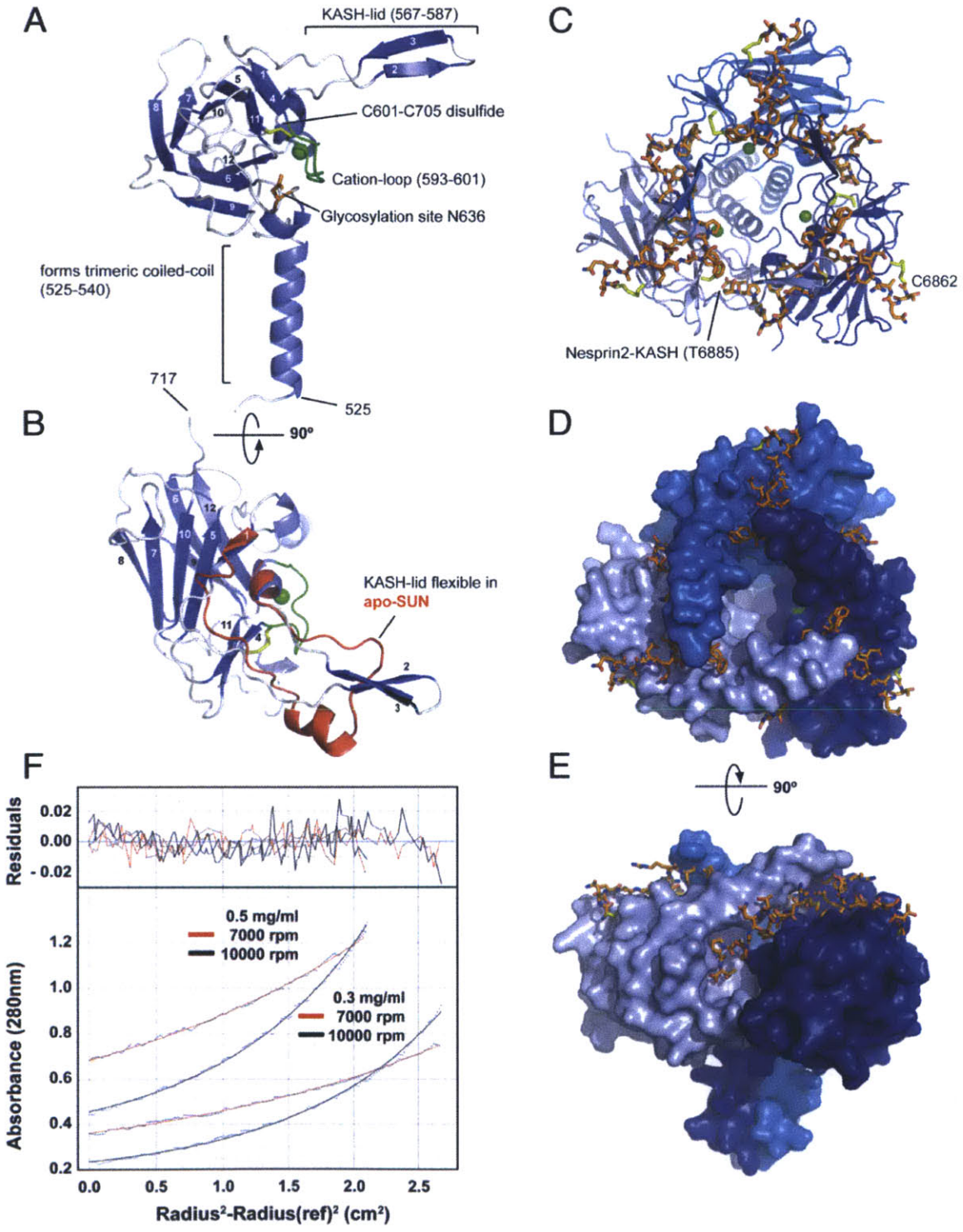


**Figure 3. Structural Analysis of the SUN-KASH Interaction.** (A) Overview of a $SUN2_{522-717}$ protomer isolated from its binding partners in the trimeric SUN-KASH complex. The protein is organized around a compact β-sandwich core, decorated with features important for function (labeled). Bound cation in green. (B) Top view of the $SUN2_{522-717}$ protomer (facing the outer nuclear membrane), rotated by 90° around the horizontal axis relative to the view in (A). The apo-protomer of $SUN2_{522-717}$ is superimposed in red. Only the

region with significant change is shown, coinciding largely with the KASH-lid. (C) Top view of the trimeric $SUN2_{522-717}$-KASH2 complex. The trimerizing SUN2 domains in cartoon representation and colored in shades of blue, the KASH2 peptide in orange. The peptide is covalently bound via a disulfide bridge between KASH2-C6862 and SUN2-C563. (D) Same view as (C), but $SUN2_{522-717}$ in surface representation, illustrating how each bound KASH2 peptide is clamped between two SUN2 protomers. (E) Side view of the SUN2-KASH2 complex, illustrating the deep binding pocket on SUN2 into which the terminal four residues of the KASH peptide bind. (F) Sedimentation equilibrium ultracentrifugation analysis of apo-$SUN2_{335-717}$. The experiment was performed at two protein concentrations and two centrifugation speeds. Data was fitted for a single species. Fitted curves overlaid over primary data (dots) in red and black for the two speeds. Residuals in the upper panel. The mass was determined to be 131.3 ± 6.1 kDa, the calculated mass is 138 kDa for the trimer.

SUN2 is organized around the threefold crystallographic axis to form a stable trimer (Figure 3C,D,E), consistent with a recent crystallographic study of apo-SUN2 (Zhou et al., 2012). The contacts between the SUN2 protomers are substantial. First, the N-terminal helical extension forms an unusual, right-handed supercoiled three-helix bundle around a neatly packed hydrophobic core. Second, each SUN protomer interacts with its two neighbors via two independent binding surfaces. As a result, 1766 $Å^2$ of surface exposed area is buried by SUN-SUN interactions per interface.

The SUN2 trimer harbors three independent binding sites for KASH2 peptides such that the overall complex is hexameric. The 23 visible residues of an individual KASH2 peptide bind in an extended groove meandering over the surface of the SUN2 trimer. The four C-terminal residues are buried in a deep pocket on SUN2 protomer 1. The next 11 residues are clamped between the KASH-lid of protomer 1 and the β-sandwich core of protomer 2. The eight N-terminal residues are bound by protomer 2 alone. C6862 is the last interacting residue and forms a disulfide bridge with C563 of SUN2. The SUN-KASH interaction is massive and altogether 1520 $Å^2$ are additionally buried on each SUN2 protomer surface upon KASH binding, presumably generating high affinity for KASH and further stabilizing the trimer.

It has been suggested previously that SUN proteins form higher order oligomers, assuming a dimeric organization as the basic unit (Lu et al., 2008). Oligomerization was also observed during gelfiltration of our samples (data not shown). To confirm that the trimeric form is the physiologically relevant oligomerization state of human SUN2, we studied its behavior in solution by analytical ultracentrifugation (Figure 3F). We used $SUN2_{335-717}$ that comprises almost the entire luminal domain of SUN2, including both the predicted coiled-coil regions and the SUN domain. Data obtained by sedimentation equilibrium analysis using two different centrifugal speeds and two protein concentrations was globally fitted to a mass of 131.3 ± 6.1 kDa, in good agreement with the calculated mass of 138 kDa of a $SUN2_{335-717}$ homotrimer. We conclude that the luminal domain of SUN2 is a homotrimer in solution.

**Details of the SUN2-KASH2 interaction**

A hallmark feature of all KASH peptides identified to date is their strict location at the very C terminus of a protein. Thus, in the following discussion, we count the KASH peptide residues inversely, starting at the C terminus with 0 for the last residue followed by negative numbers (Figure 4). Residues 0 to -3 line a pocket crafted into the SUN2 surface of protomer 1 that is perfectly complimentary to the shape of the peptide. The terminal carboxyl group points into this pocket, indicating that the positioning of C-terminal end of KASH should be incompatible with extending the peptide by even one residue. The carboxyl group is specifically coordinated by the hydroxyl groups of S641, Y703, and Y707. H628, the C601-C705 disulfide, and a part of the cation-loop form the backside of the pocket. The three prolines in the -1, -2, and -3 positions are all in the trans-conformation and in close van-der-Waals contact with SUN2. Residues in positions -4, -5, and -6 are more solvent exposed and their side chains are not interacting strongly with SUN2. Tyr at -7 is again deeply buried in a slot lined by SUN2 protomer 1 and 2, explaining its strong conservation. The next 3 KASH-residues are sandwiched between the KASH-lid of protomer 1 and the core of

protomer 2. They hydrogen-bond to the KASH-lid, in turn forming a third β-strand. Like Tyr at -7, Leu at -9 is also deeply buried in a cleft between SUN2 protomers 1 and 2, and is again highly conserved. Pro at -11 ends the β-strand. Between positions -11 and -12, the peptide kinks by about 90° to continue to the N-terminal Cys at -23. After the kink, the peptide exclusively interacts with SUN2 protomer 2. Other than bridging the distance to Cys at -23, the most noticeable interaction is burial of Phe at -17 into a hydrophobic pocket. Similar to positions -7 and -9, the large hydrophobic residue at -17 is also conserved. Cys at -23 is perfectly positioned to engage in a disulfide bond with SUN2-C563. The conservation of SUN2-C563 (Figure 5), and the presence of the SUN-KASH disulfide bond in living cells (Figure 6) suggest that it is physiologically important. Altogether 24 hydrogen bonds, 11 between SUN2 protomer 1 and KASH2, and 13 between protomer 2 and KASH2, tie the peptide strongly to the SUN2 trimer base. In comparison, merely three hydrogen bonds are formed between the protomer interfaces.

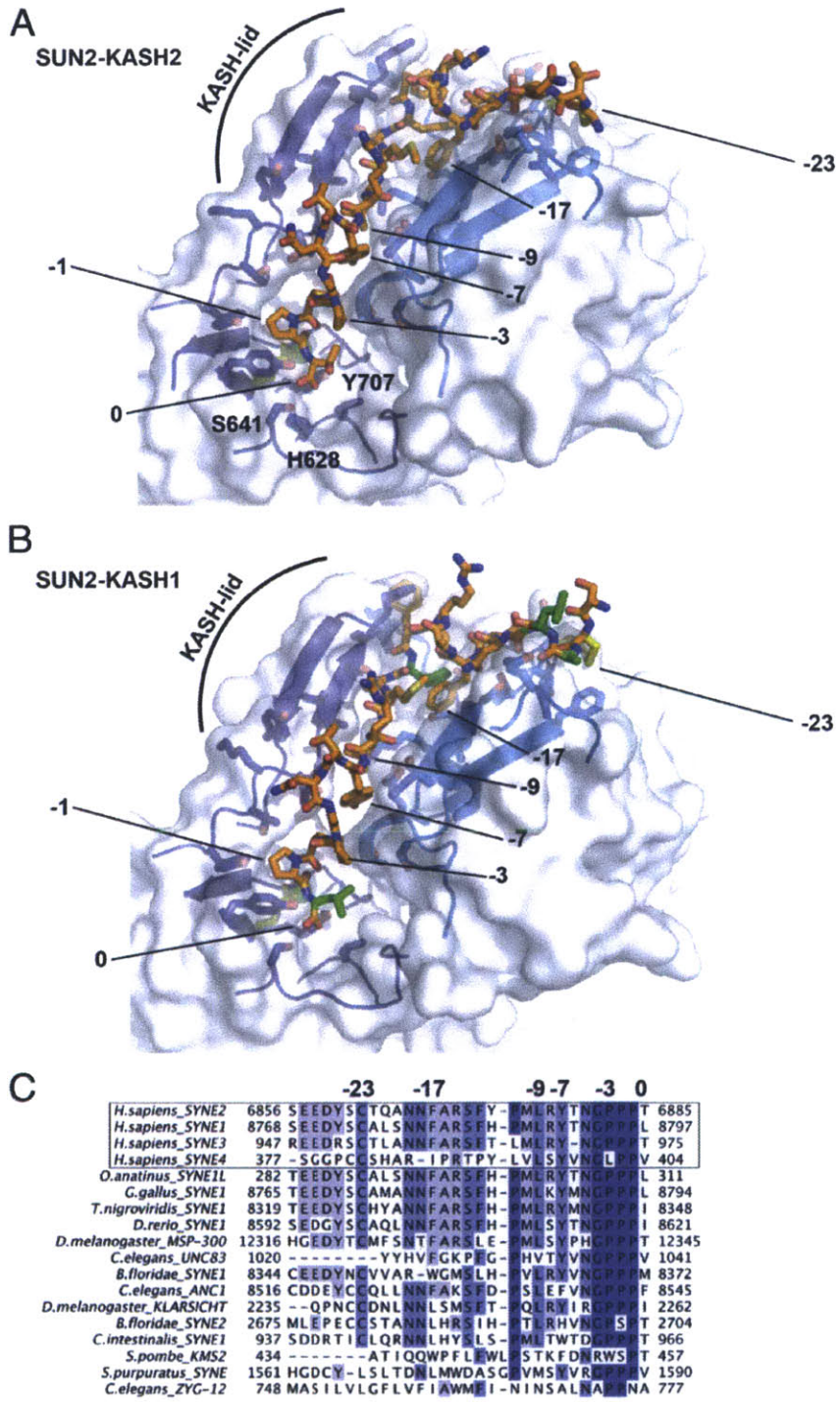


**Figure 4. Details of the SUN-KASH Interaction.** (A) Close-up view of the SUN2-KASH2 interaction. Two neighboring SUN protomers are shown in two shades of blue, with the KASH2 peptides in between in orange. Surface of the SUN2 binding area is half-transparent. KASH residues crucial for interaction are numbered. '0' denotes the C-terminal residue of the peptide. Pocket residues that abolish KASH-binding if mutated are labeled. (B) Same view of the SUN2-KASH1 interaction. Residues that differ between KASH1 and KASH2 are colored in green. (C) Multiple sequence alignment of the four identified human KASH

proteins, followed by a list of KASH peptides from highly diverged eukaryotes. The numbers match residues important for SUN binding.

Notably, SUN2 is glycosylated at N636 (Stewart-Hutchinson et al., 2008). N636 lies on the SUN2 surface (Figure 3A) and does not contact the KASH peptide, consistent with the previous finding that N-glycosylation is dispensable for KASH binding.

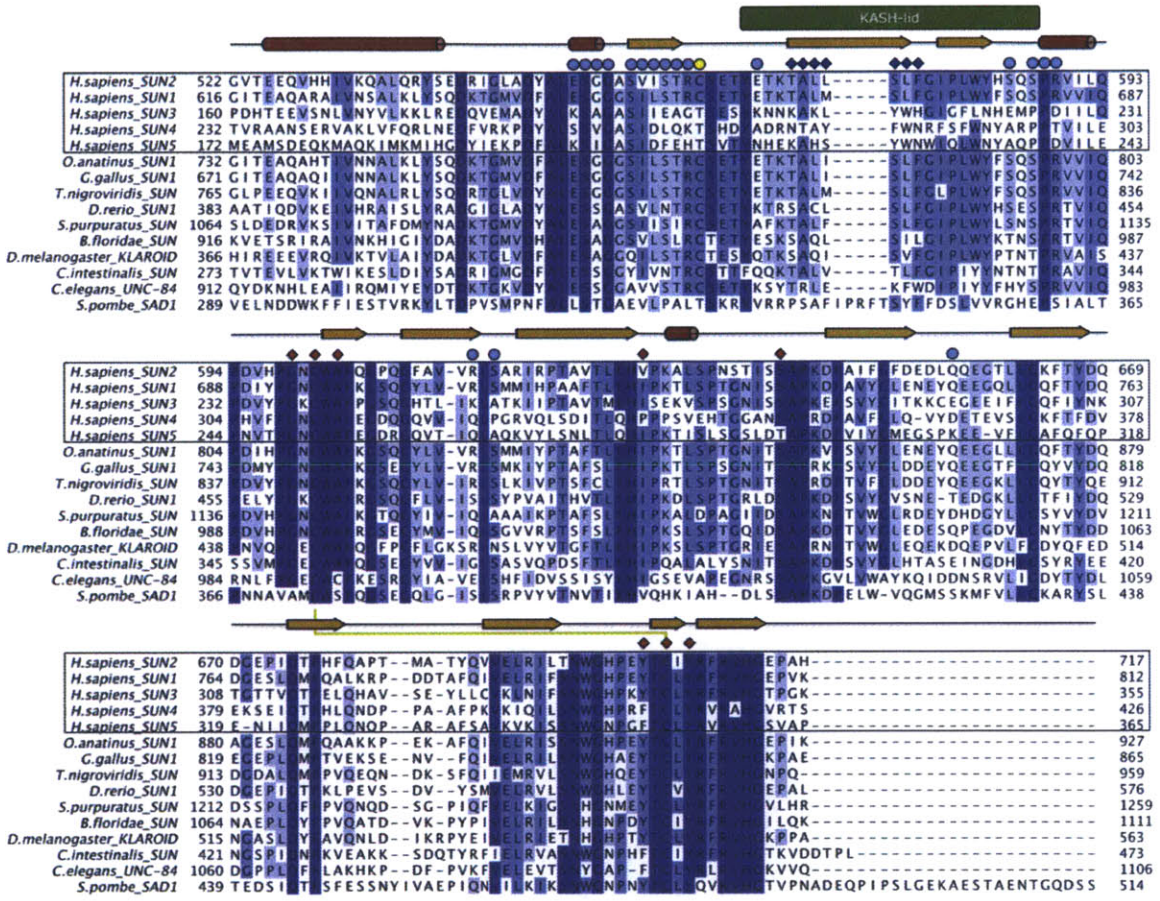


**Figure 5. Multiple Sequence Alignment of SUN Domains**. SUN domains from human SUN proteins and from highly diverged eukaryotes are shown. Secondary structure elements are shown above the sequence. The intramolecular disulfide bond is indicated by a yellow line. KASH contact sites are labeled. Red diamonds indicate pocket residues on SUN protomer 1, blue diamonds KASH-lid contacts. Colored circles indicate KASH binding sites on SUN protomer 2. The covalent link between SUN and KASH is indicated by a yellow dot.

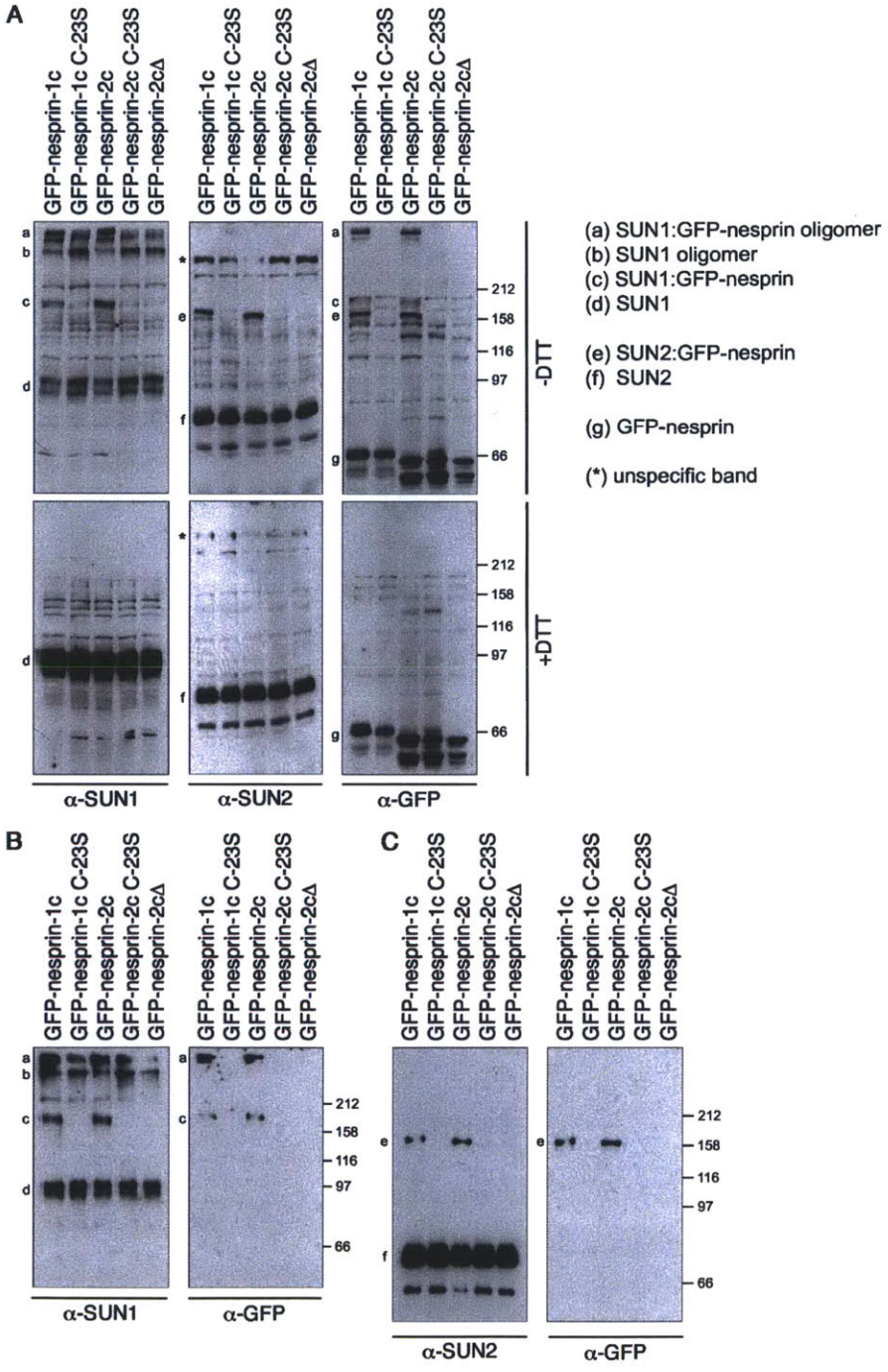


**Figure 6. The SUN-KASH Disulfide Bridge Exist in vivo.** (A) HeLa cells transiently expressing GFP-tagged C-terminal fragments of Nesprins (GFP-Nesprin-1c, GFP-Nesprin-2c) or C-23S mutants were analyzed by SDS-PAGE under reducing and non-reducing conditions followed by Western blotting using

anti-SUN1, anti-SUN2 or anti-GFP antibodies. In the non-reducing gel, disulfide-bridged SUN1 or SUN2 and GFP-Nesprins form the indicated high molecular weight complexes (a,c,e), which are destroyed under reducing conditions. (B, C) Extracts of cells (A) were subjected to immunoprecipitation with anti-SUN1 (B) or anti-SUN2 (C) antibodies and isolated proteins analyzed by non-reducing SDS-PAGE followed by Western blotting using the indicated antibodies. Note that disulfide bridged SUN-Nesprin wildtype complexes are precipitated. Note that the mutant Nesprins are not co-precipitated with SUN proteins under these conditions due to alkylation by NEM.

## Comparison of SUN2-KASH2 with apo-SUN2

Apo-SUN2 forms a structure very similar to KASH-bound SUN2 (Figure 3B). Both structures superpose with an rmsd of 0.8 Å over 180 Cα positions. The one major difference between apo-SUN2 and KASH-bound SUN2 is the conformation of the KASH-lid. While ordered in the bound form, the KASH-lid is poorly ordered and lacks the β-hairpin in the apo-form. SUN2 trimers tightly pack sideways into a two-dimensional array in the crystal, with the KASH-lids of neighboring trimers caking the 2D-arrays into a 3D-lattice. Under physiological conditions, the KASH-lid most likely adopts a random conformation that becomes ordered upon KASH-binding. In the apo-form published recently (Zhou et al., 2012), the KASH-lid adopts yet another varied conformation, further supporting the flexibility of this region in the absence of KASH peptide.

## Comparison of SUN2-KASH2 with SUN2-KASH1

We also solved the structure of the SUN2-KASH1 complex. Since the KASH peptide is not involved in crystal-packing contacts, we anticipated that both complexes would crystallize in the same crystal form, which they did. Overall, both complexes look very similar, and the KASH peptides are bound in the same binding groove. Compared to KASH2, KASH1 differs in 5 positions within the 24 SUN2-binding KASH residues (Figure 4B). For position 0, the Thr to Leu substitution is subtle enough to not change the interaction with the SUN2 binding pocket. The four other changes in positions -12, -20, -21, and -22, respectively, can be accommodated easily by SUN2. There are hardly restrictions on size and

charge of the KASH side chains in these positions, well reflected by their poor conservation (Figure 7).

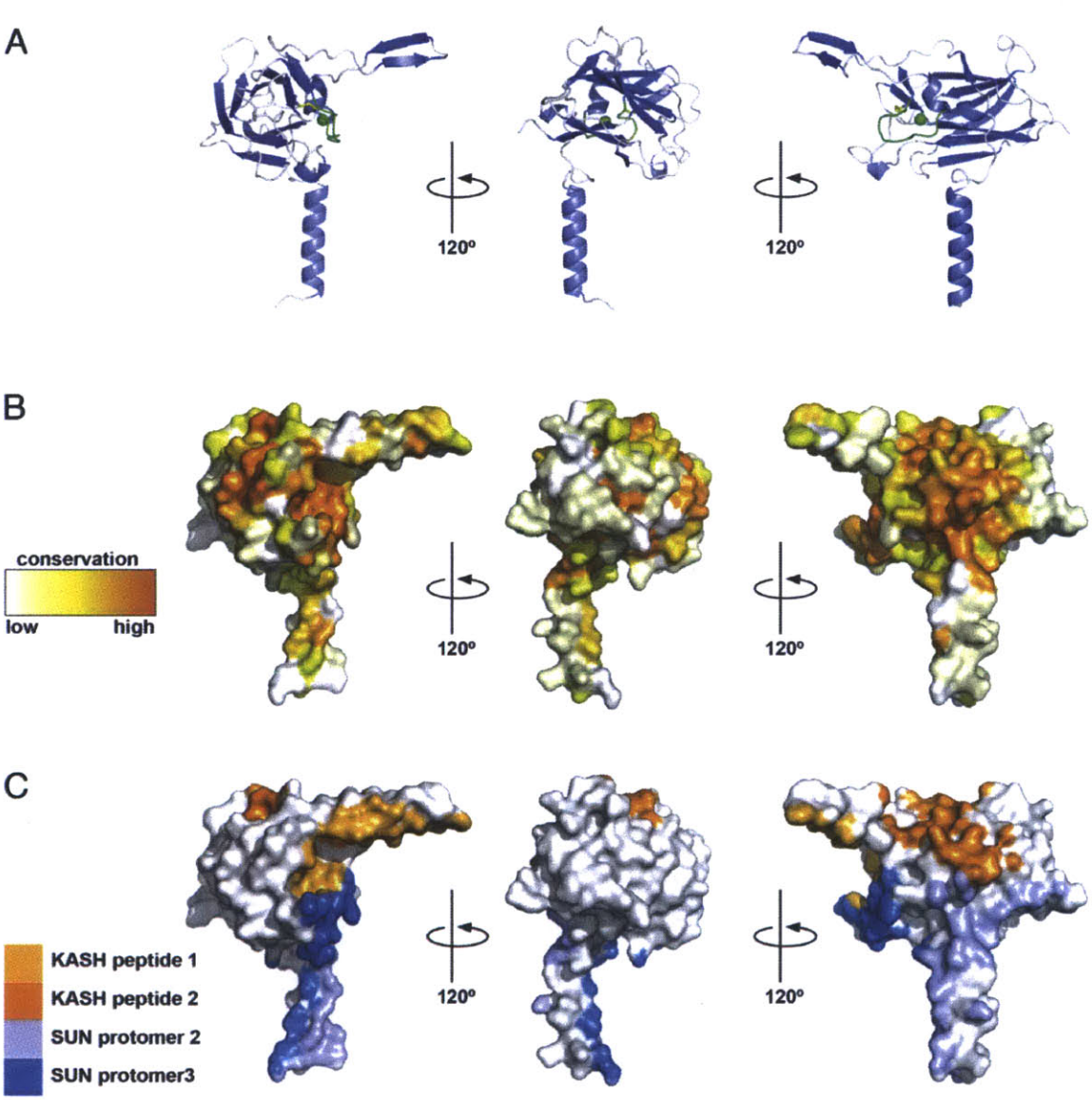


**Figure 7. Surface Analysis of the SUN protomer.** The SUN protomer in three different orientations, related to one another by 120° rotations around the vertical axis. (A) Cartoon representation, coloring as in Figure 2A. (B) Surface representation, gradient-colored to illustrate the conservation. (C) Surface representation, colored to show the different binding interfaces to the two neighboring SUN protomers (shades of blue), and the two KASH peptides (shades of orange).

## SUN2 Conservation

We generated an alignment of maximally diverged eukaryotic SUN domains to analyze their sequence conservation (Figure 5). The SUN domain is relatively well conserved, despite that the basic β-sandwich fold itself can be adopted by highly divergent sequences. The conservation can be explained by the extensive surface areas that are involved in protein-protein interactions (Figure 7). Examining the conservation profile of surface residues of a SUN2 protomer, one can appreciate that conserved regions match with both KASH binding and homotrimerization interfaces. This analysis strongly supports that the binding mode observed in the crystal is physiologically relevant.

## Comparison of SUN2 with F-lectins

A VAST search for structural homologs of SUN2 reveals similarity to a number of functionally diverse β-sandwich proteins implicated e.g. in DNA repair, intraflagellar transport, and various catalytic activities. Most intriguing among the close homologs are F-lectins involved in fucose binding. A comparison of the KASH-bound SUN domain to eel agglutinin bound to fucose (Bianchet et al., 2002) demonstrates that not only is the basic β-sandwich fold present, but also the additional cation-loop and the stabilizing disulfide bridge are maintained (Figure 8). In both proteins, these elements are involved in substrate binding. In the agglutinin structure, fucose is bound close to the position where the pocket in SUN2 is found, which binds the terminal PPPX motif of KASH. This similarity suggests that both proteins are ancestrally connected and diverged into today's function by crafting the peptide/sugar interacting moiety (Mans et al., 2004).

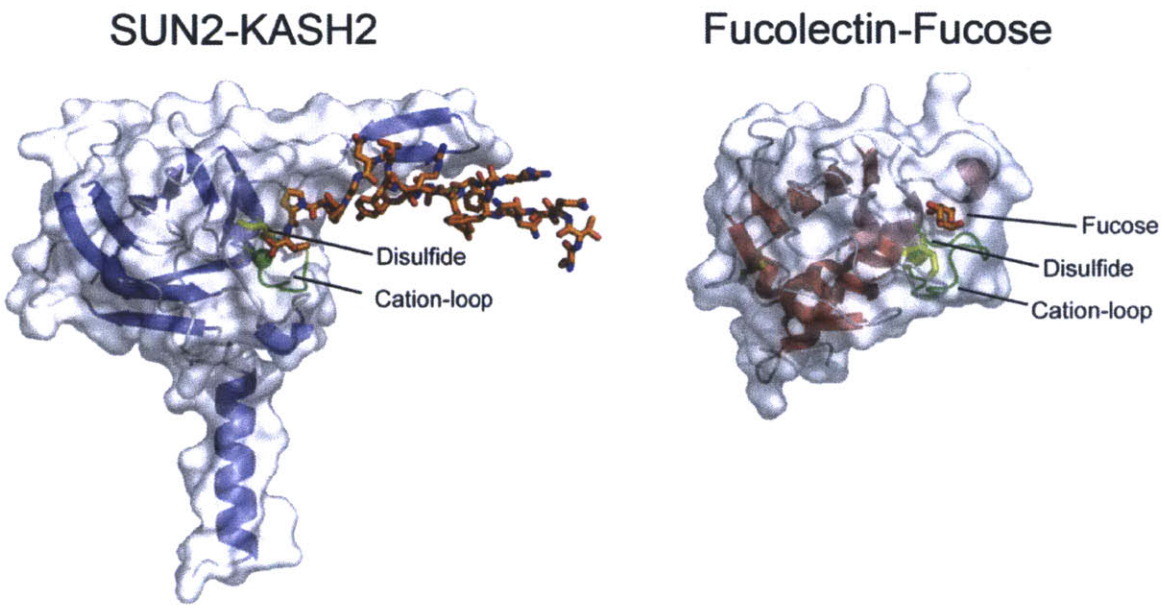


**Figure 8. Structural Homology with Lectins.** On the left, a SUN2 protomer with bound KASH2 peptide. On the right, Fucolectin from European eel bound to fucose. Both proteins are shown in the same orientation. In both proteins, a metal cation organizes a loop that interacts with KASH and fucose, respectively. The loop is further held in place by an internal disulfide bond.

### Mutational Analysis of SUN-KASH interaction

To support the conclusions derived from the structural model of SUN-KASH, and to determine the contribution of features in SUN and KASH required for complex formation, we performed mutational analysis of the KASH peptide and the SUN domain.

First, we analyzed which region of KASH2 is required for interaction with SUN in pulldown experiments. Truncation of the N-terminal 11 or 15 residues of the KASH2 29-mer yielding a KASH 18-mer and a 14-mer, respectively, did only slightly diminish interaction of the peptides with the luminal domain of SUN2 ($SUN2_{335-717}$) (Figure 9A,B). As these deletions removed the conserved cysteine at position -23, the disulfide bond generated between this residue and C563 of SUN2 appears dispensable for binding. Still, it is possible that the disulfide might influence functionality of the LINC complex in another way (see discussion).

When the KASH peptide was shortened by additional 4 residues, the resulting 10-mer was incompetent for interaction with SUN2 in our assay, indicating that a short PPPX containing peptide is not sufficient to promote binding. Clearly, this effect is not explained solely by reducing the length of the peptide, as an 18-mer in which residues between positions -9 and -17 were substituted for alanines also failed to recruit $SUN2_{335-717}$.

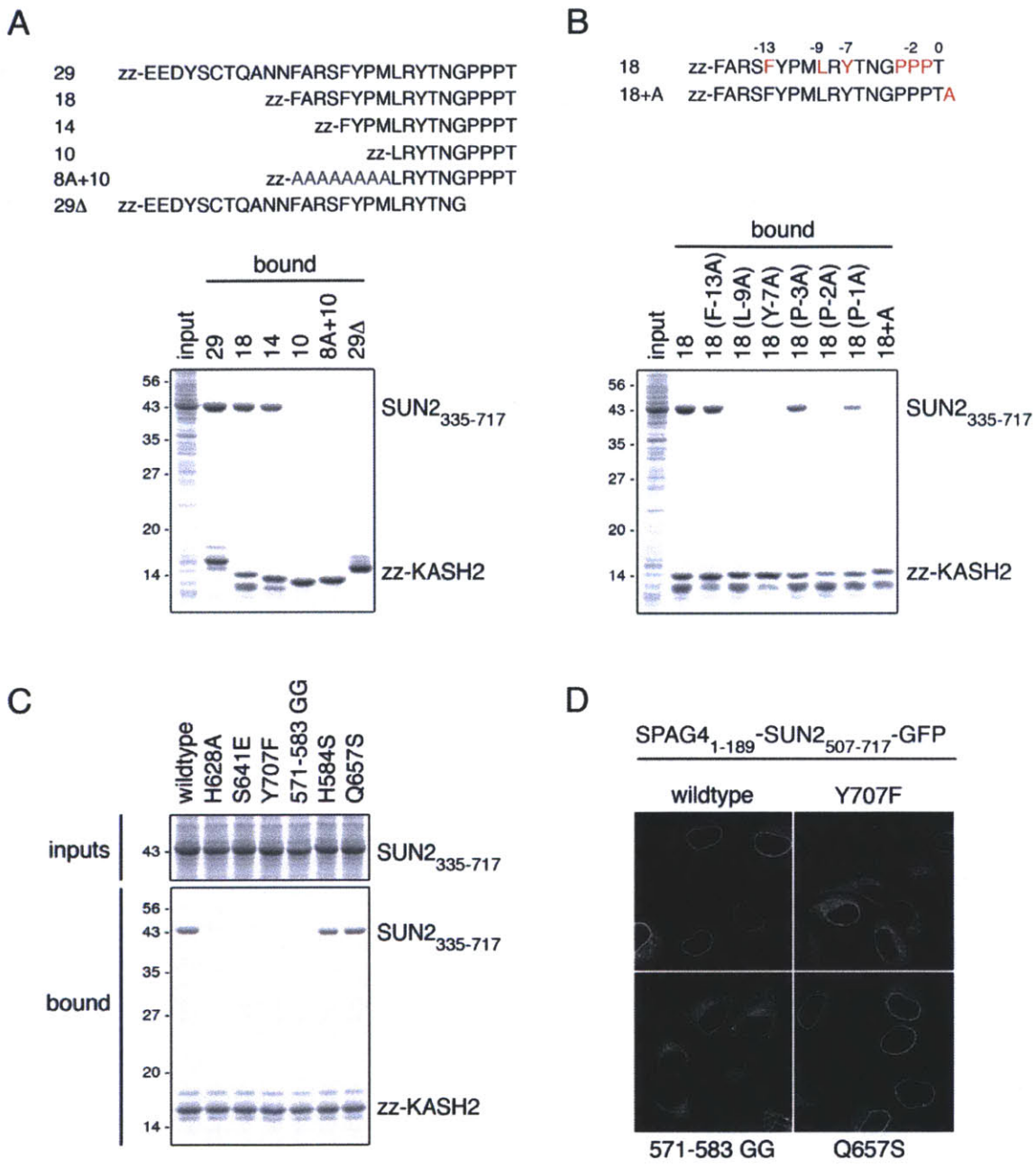


**Figure 9. KASH Interaction Depends on Both the Terminal Binding Pocket and the β-Hairpin of the SUN Domain**. (A) The C-terminal 14 aa of KASH2 are sufficient for interaction with the SUN domain. The depicted zz-KASH2 derivatives were bound to beads and incubated with *E. coli* lysate of cells expressing His-SUN2$_{335-717}$. Bound proteins were eluted and analyzed by SDS-PAGE and Coomassie staining. Loads in the input and pulldown lanes correspond to 1.25% and 10% of the total, respectively. (B) Contribution of conserved residues within the C-terminal 14 aa of KASH2 to SUN binding. Binding of His-SUN2$_{335-717}$ to the depicted zz-KASH2 derivatives was analyzed as in (A). Note that the extension of the KASH peptide by one residue abolishes SUN interaction. (C) KASH binding of SUN domain mutants. Wildtype His-SUN2$_{335-717}$ or mutant derivatives were added to *E. coli* lysate (1.6 μM of trimer) and incubated with immobilized zz-KASH2

as in Figure 1D. Bound proteins were eluted and analyzed as in (A). (D) SUN domain mutants deficient in KASH interaction fail to mediate NE targeting of a SUN domain-dependent reporter construct in vivo. The localization of $SPAG4_{1-189}$-$SUN2_{507-717}$-GFP wildtype or indicated SUN domain mutants was analyzed after transfection of HeLa cells. Cells were analyzed by confocal fluorescence microscopy.

To define the contribution of specific residues in KASH for SUN binding, we mutated conserved amino acids residing within the last 14 residues to alanines in the context of the KASH2 18-mer (Figure 9B). Residues at -7 and -9 both fit snuggly in deep hydrophobic pockets that are created between the KASH-lid of SUN protomer 1 and core of protomer 2. Consistently, mutating either position into alanine abolishes SUN-KASH interaction. Ala substitution of Phe at -13, however, only slightly reduces KASH binding. This suggests that the groove formed upon KASH-binding between two SUN protomers, which accommodates the KASH β-strand comprising residues -11 to -7, is essential for KASH interaction.

The crystal structure suggests that the three C-terminal prolines at -1 to -3 of KASH are important for positioning the C terminus of the peptide into the surface binding pocket. Pro at -2 plays a central role in this, as its substitution to Ala abolished SUN binding. Mutation of Pro at either -1 or -3 reduced SUN interaction but did not prevent it, suggesting that the central proline might be key to orient the KASH2 C terminus to dip into its binding pocket on SUN.

If the correct positioning of KASH's C terminus into its binding pocket on SUN is required for SUN-KASH interaction, then extending KASH at the C terminus should not be tolerable. We tested this prediction by adding one alanine to the C terminus of KASH2 and observed that binding of this KASH mutant (18+A) was indeed prevented. Consistently, it was previously shown that extending KASH C-terminally by about 10 residues abolishes SUN binding (Stewart-Hutchinson et al., 2008). Taken together, the mutational analysis suggests that key

determinants for SUN-KASH binding reside in the last 14 residues of KASH, including both contributions of the KASH β-strand and the terminal PPPX motif.

The structural data and mutational analysis of KASH indicated that both the KASH binding pocket and the KASH-lid on SUN are expected to be pivotal for complex formation. To test the significance of both elements on KASH binding, we introduced mutations into the SUN domain (using $SUN2_{335-717}$). Mutation of H628 at the back site of the KASH binding pocket (H628A) as well as mutations of residues involved in coordination of KASH's C terminus (S641E, Y707F) impaired KASH interaction (Figure 9C). Likewise, deletion of the KASH-lid, which was achieved by replacing residues 571 to 583 by two glycines (571-583 GG), strongly diminished binding to KASH. Two unrelated mutations on the surface of the SUN domain, H584S and Q657S, had no effect on SUN-KASH complex formation. Thus, both KASH binding pocket and KASH-lid are instrumental for SUN-KASH interaction.

Next, we tested the importance of SUN2 trimerization and of the cation loop for KASH interaction. We disrupted the relative orientation of the three SUN domains by distorting the short trimeric coiled-coil segment 525-539, which is not in direct contact with KASH. L536D interferes with the hydrophobic core of the coiled-coil, while ΔR538 changes the phase of the helix. Both mutations expectedly disrupt SUN-KASH interaction (Figure 10), strongly indicating that the specific trimeric orientation of the SUN domains is necessary for KASH binding. We further examined a D542N mutation that was previously shown to exert a nuclear migration defect in *C. elegans* (Malone et al., 1999). D542 forms a salt bridge with R707 to anchor the coiled-coil to the SUN domain. D542N also disrupts KASH binding, but the protein fold seems to be compromised since protein solubility was significantly reduced. The importance of the cation-loop was probed using a ΔN600 mutant, which shortens the loop. As expected, the mutant does not bind KASH (Figure 10).

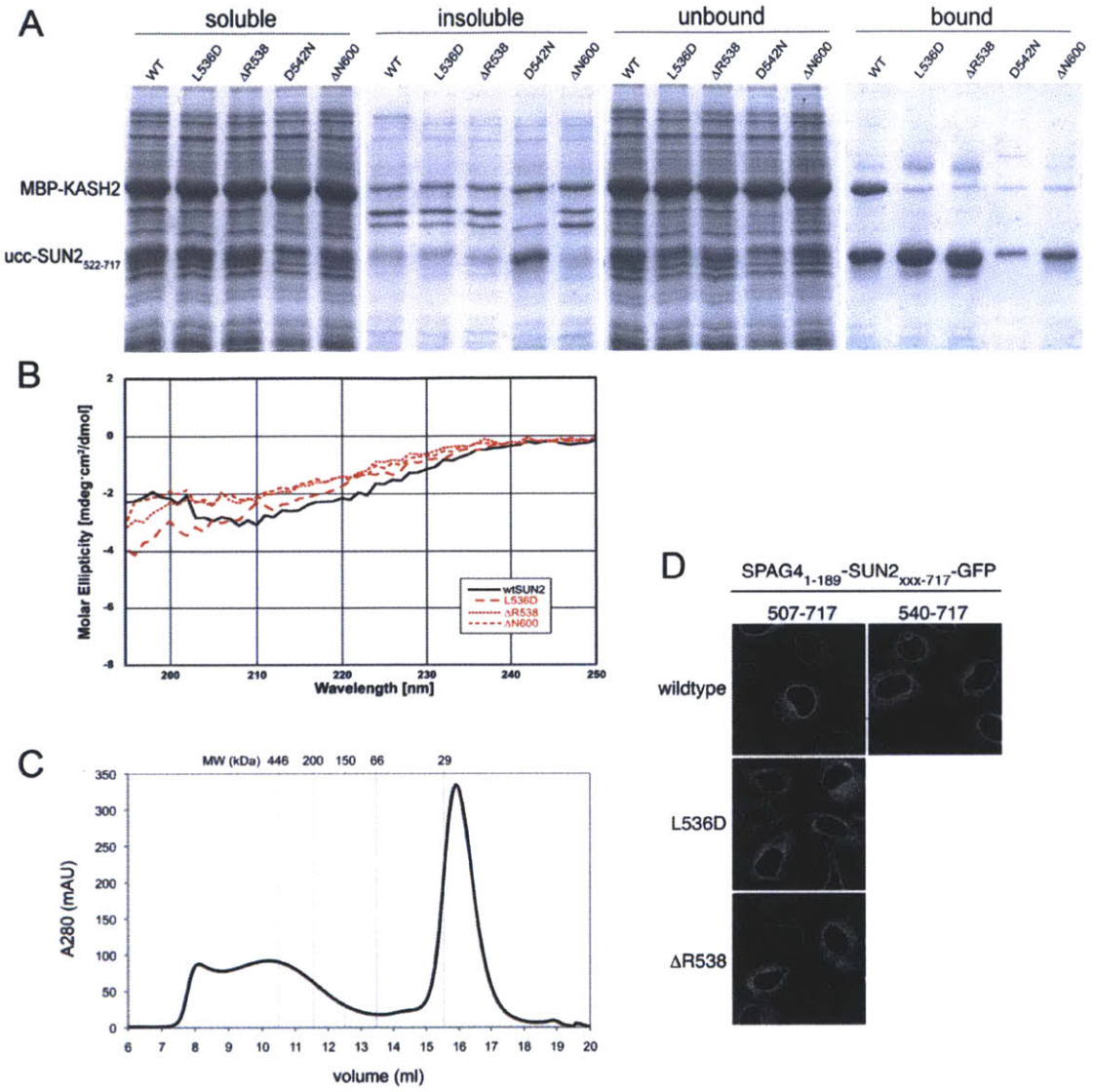


**Figure 10. Importance of SUN Trimerization and the Cation-loop for KASH-Binding.** (A) Point mutations were introduced in the SUN domain of human SUN2 in the backbone of His-tagged ucc-SUN2$_{522-717}$ to either distort trimeric organization (L536D, ΔR538, D542N) or the structure of the cation-loop (ΔN600). Wildtype ucc-SUN2$_{522-717}$ or the indicated mutants were co-expressed with MBP-KASH2 in E. coli. Interaction between MBP-KASH2 and the SUN domain variants was analyzed by a Ni-affinity pull-down assay of His-tagged SUN2. SDS-PAGE analysis of relevant fractions is shown. All mutants impaired binding of MBP-KASH2. Note that the D542N mutant of SUN2 appears to be compromised in folding, since it is largely insoluble. (B) Analysis of wildtype SUN2$_{522-717}$ and the indicated mutants by CD spectroscopy. All proteins were purified and the unrelated coiled-coil removed by protease cleavage prior to CD analysis. The spectra do not reveal any gross variations and are typical for β-sandwich proteins, indicating that the point mutations do not affect the core fold of the SUN2 domain. (C) The isolated SUN domain of SUN2 (SUN2$_{540-717}$) is a monomer. His-tagged SUN2$_{540\,717}$ (calculated MW 23 kDa) was expressed in E. coli, purified by Ni-affinity chromatography and analyzed by gelfiltration on a Superdex S200 HR10/300 column using Tris/HCl

pH 7.5, 200 mM NaCl as running buffer. Calibration of the column relative to a molecular weight standard is indicated on top. (D) Mutations disrupting the correct trimeric arrangement of SUN domains, which are deficient in KASH interaction (A), impair NE targeting of a SUN domain-dependent reporter construct in vivo. Wildtype $SPAG4_{1-189}$-$SUN2_{507-717}$-GFP, two point mutant derivatives (L536D, ΔR538) or $SPAG4_{1-189}$-$SUN2_{540-717}$-GFP, which lacks the trimeric coiled-coil region preceding the SUN domain, were expressed in HeLa cells. Cells were analyzed by confocal fluorescence microscopy.

Finally, we verified that the identified features of SUN2 are important for functionality of the SUN domain in living cells. Based on our previous report that the SUN domain serves as one of several determinants for localization of SUN2 to the NE, we employed NE targeting of a SUN domain-dependent reporter protein as readout (Turgay et al., 2010). The reporter consists of the N-terminal, membrane-bound domain of the SUN family member SPAG4 ($SPAG4_{1-189}$), followed by $SUN2_{507-717}$ comprising the SUN domain and part of its preceding trimeric coiled-coil, and a C-terminal GFP (Figure 9D). During biosynthesis, the reporter protein is inserted into the ER membrane by help of the SPAG4 transmembrane domain and then targeted to the INM using the SUN domain as localization signal (Turgay et al., 2010). Consistent with our in vitro analysis (Figure 9C), mutation of either the KASH binding pocket on SUN2 (Y707F) or deletion of the KASH-lid (571-583 GG) compromised NE targeting, resulting in an accumulation of the $SPAG4_{1-189}$-$SUN2_{507-717}$-GFP reporter in the ER. The same effect was observed for mutations L536D and ΔR538, which were designed to change the relative orientation of the three SUN domains (Figure 10). Notably, also deletion of the coiled-coil region in front of the SUN domain ($SPAG4_{1-189}$-$SUN2_{540-717}$-GFP) impaired NE accumulation, consistent with the inability of this SUN2 fragment, which is monomeric in solution, to bind KASH in vitro (Figures 1, S5). Collectively, these data demonstrate that the respective mutations interfere with the functionality of the SUN domain in its physiological environment, verifying the conclusions drawn from both the structure and the biochemical analysis, and further indicating that the function of the SUN domain in NE targeting relies on KASH interaction.

## DISCUSSION

One of the key findings of this study is that three KASH peptides bind to a trimeric arrangement of SUN domains. We hypothesize that the trimeric arrangement is universally conserved for all SUN domains, since it is the prerequisite for the mode of KASH binding as we observe it in our structures. All known SUN homologs from divergent species contain predicted coiled-coil segments N-terminal to the SUN domain, supporting our hypothesis. It is curious how well the position of the trimeric coiled-coil, relative to the SUN domain, is conserved. The coiled-coil rigidly emanates from the β-sandwich domain and is held in position by a highly conserved salt bridge between D542 at the base of the coil and the R708/710 pair in strand 12 (Figure 5). Why is the SUN domain not a stable trimer itself, without the coiled-coil? One possibility is that in the observed configuration, it might be possible to regulate the SUN-KASH bridge formation by pivoting the SUN domain away from the coiled-coil thereby efficiently inhibiting interaction.

The SUN2-KASH1 and SUN2-KASH2 complexes excellently explain conservation of the KASH peptide as defined in the Pfam database. KASH peptide positions -1 to -11 are generally the most conserved, with a particularly high level of identity at positions -1, -2, -3, -4, -7, -9, and -11. Further upstream, the hydrophobic residue at -17 and Cys at -23 are strongly conserved (Figure 4C). Accordingly, the minimal KASH peptide displaying SUN interaction in our assay contains 14 residues (Figure 9). Thus, we expect the largest variations within the part of the KASH peptide that exclusively interacts with SUN protomer 2, positions -12 to -23, which shows the strongest sequence variations and is less ordered in our structures. Some of the characterized metazoan KASH peptides are shorter than others, for example the luminal domains of *C. elegans* UNC83 and ZYG-12 (Figure 4C). It will be interesting to see structural details of how those sequences are recognized. We hypothesize that the principal interaction mode will be the same, i.e. peptide-in-pocket interaction of the four C-

terminal residues, and sandwiching of the preceding extended stretch, positions -5 to -11, between two SUN protomers.

In a recently published apo-SUN structure, the authors attempted to model the KASH-binding site and proposed that each SUN protomer would bind a single KASH peptide (Zhou et al., 2012). It was argued that binding of KASH2 to SUN2 would resemble the interaction of a CD40 peptide with tumor necrosis receptor associated factor 2 (TRAF2). However, although TRAF2 is a trimerized β-sandwich like SUN2, the surface area in question is not conserved in SUN2 proteins and remote from the three KASH binding sites revealed by our structure. Consistent with our data, Zhou et al. find that modifying their proposed surface area through A647E and F610A mutations on SUN2 does in fact not significantly affect KASH binding. In contrast, changing the well conserved Gly608 to Asp, disrupts KASH binding, however, Gly608 is a buried residue and its mutation expected to affect the β-sandwich fold rather than the local surface area. Finally, the model proposed by Zhou et al. does not explain why SUN trimerization is necessary for KASH binding, which we show is critical for LINC complex formation.

In contrast to metazoa, putative SUN-KASH interactions in yeasts are more difficult to predict. The Sad1-Kms2 interaction in *S. pombe* might still follow the same scheme, although the difference in the KASH-lid of Sad1 might explain the substantial sequence variation in the KASH peptide of Kms2 (Figure 5). The candidate SUN-KASH bridge in *S. cerevisiae* between Mps3 and Mps2 or Csm4 appears to be very different. In Mps3, the intramolecular disulfide as well as the salt bridge stabilizing the coiled-coil position are missing, and the KASH-lid is hardly recognizable by sequence.

One major function of LINC complexes is to provide mechanical coupling of the nucleus to the cytoskeleton, allowing for force transmission across the NE (Lombardi and Lammerding, 2011). The large KASH-peptide binding groove

between neighboring SUN domains appears to be the foundation for strong SUN-KASH interactions. Yet, our structural model bears several additional implications for how the structural organization of LINC complexes contributes to the establishment of a force-resistant coupling device. First, the presence of three KASH peptide binding sites on each SUN trimer entails a contribution of binding avidity to the strength of SUN-KASH interaction. Second, a disulfide bond covalently links SUN and KASH, thus explaining how this bridge can resist the substantial forces that nuclear migration and chromosome rearrangements might require. And third, the trimeric arrangement of SUN proteins is ideally suited to allow for organization of higher-order SUN-KASH complexes, which would obviously potentiate possible forces that LINC complexes can transmit.

The disulfide bond between SUN and KASH raises the intriguing question of how the LINC complex can get separated after it has been formed. The perinuclear space, as an extension of the ER, is rife with protein disulfide isomerases (PDIs) that help remodel disulfide bonds. Therefore PDIs are prime candidates for breaking the SUN-KASH disulfide bond. Almost all of the 20 PDI families detected in humans are ER-resident proteins, but their specific substrates are poorly defined (Appenzeller-Herzog and Ellgaard, 2008). It is an interesting possibility to speculate whether one of these enzymes has a regulatory role in LINC complex remodeling.

It is presently unclear how KASH proteins are organized and whether each SUN trimer interacts with KASH peptides originating from individual KASH protein monomers or from KASH proteins present in form of dimers, trimers or even higher-order organizational units. For some KASH proteins, there already exists evidence that they are organized by multimerization (Ketema et al., 2007; Mans et al., 2004; Mislow et al., 2002). Using Cys at -23 as marker, the three KASH peptides are ~50 Å apart on the ONM-facing, bottom surface of the SUN-trimer in our structures. The distance between the N-terminal end of the KASH peptide and the transmembrane helix is only ~7 residues, suggesting that the TM helices

of an oligomeric KASH protein cannot interact with each other in the membrane if all peptides of a KASH oligomer are bound to one SUN trimer. However, it is striking that also the TM helices of KASH proteins are quite well conserved in sequence, suggesting that they might engage in protein-protein interactions. It remains to be seen how these transmembrane domains are organized and whether their conservation is explained by self-association or by interaction with other partners. If they self-associate, we predict that they cannot interact with the same SUN trimer and thus enable higher-order complexes. The association of KASH peptides from oligomeric Nesprins with neighboring SUN trimers would trigger two-dimensional clusters of SUN-KASH bridges in the NE. Alternatively, clustering of SUN proteins has recently been discussed to originate from the perinuclear coiled-coil domains forming a hybrid of N-terminal dimeric and C-terminal trimeric coils (Zhou et al., 2012). In this hypothesis, one of the three hypothetical N-terminal coils would have to interact with a neighboring SUN trimer to form a coiled-coil, thus would trigger clustering. We note that this hypothesis is in conflict with our solution data that shows that $SUN_{335-717}$, which includes the predicted dimeric coiled-coil segment, is trimeric and does not from higher order assemblies (Figure 3).

Arrays of LINC complexes have been described in at least two cases. First, TAN lines, linear, actin-associated arrays of SUN2 and Nesprin-2G, reorient the nucleus in wounded cell monolayers (Luxton et al., 2010). Second, SUN-domain clustering is also observed at the attachment sites of meiotic telomeres in diverse organisms (Chikashige et al., 2006; Ding et al., 2007; Penkner et al., 2007; Schmitt et al., 2007). Clearly, a next step in illuminating LINC complex formation must unravel the organizational principle of KASH proteins.

It has been previously observed that depletion of SUN proteins from HeLa cells leads to loss of the regular, 40 - 50 nm spacing between the nuclear membranes, accompanied by irregular expansions of the perinuclear space (Crisp et al., 2006). This suggests that SUN proteins function to maintain the defined proximity

of ONM and INM. In fact, the luminal coiled-coil domains of both SUN2 and SUN1 are ideally suited to define the distance of both membranes. If one assumes that the luminal part of SUN2 preceding the SUN domain is composed of a trimeric coiled-coil with the potential of forming a rigid rod, it would extend over a length of ~ 45 nm when compared to experimentally determined trimeric coiled-coils (Testa et al., 2009). Together with the extension of the SUN domain trimer along its longitudinal axis, this results in a predicted length of the entire luminal domain of SUN2 of ~ 48 nm (Figure 11). Notably, the luminal domains of SUN1 and SUN2 share the same organizational principles and are of similar length (477 and 482 aa). Thus, the predicted length of these luminal domains correlates well with the observed spacing of both nuclear membranes. Similarly, it has been suggested that the coiled-coil protein CLIMP63 serves a luminal spacer defining the width of ER sheets (Shibata et al., 2010).

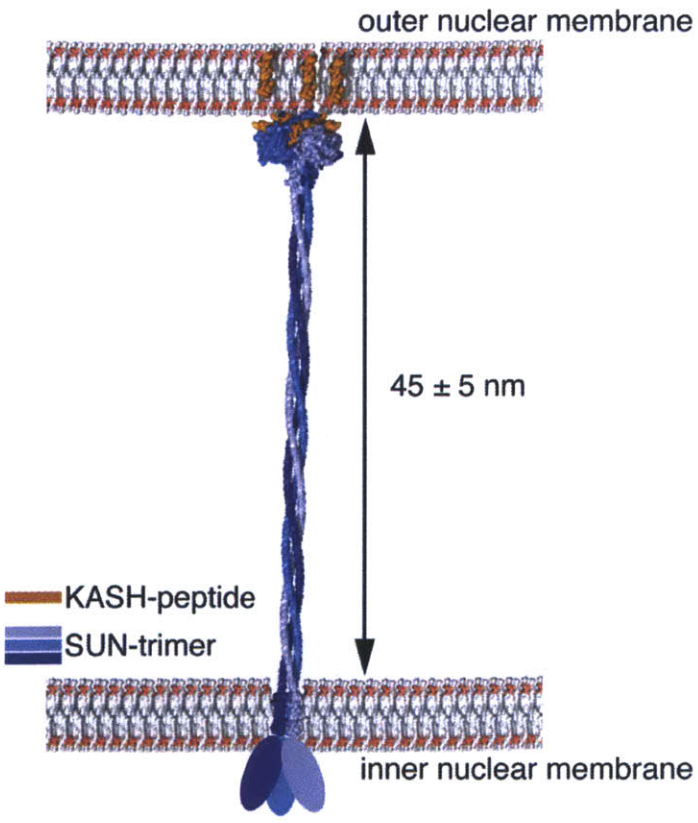


**Figure 11. Model for the LINC-Complex Bridging the Nuclear Envelope.** The LINC complex is displayed to scale within the perinuclear space between ONM and INM. The trimeric SUN2 is modeled based on our experimental structure, with a theoretically calculated trimeric coiled-coil extending N-terminally to the INM

(arteni.cs.dartmouth.edu/cccp/). The N-terminal, nucleoplasmic domain is indicated as an oval sphere of approximate size assuming a globular shape. Atomic coordinates for the lipid bilayers and the transmembrane segments are from public sources.

The tight association of SUN and KASH raises the issue of whether and how these linkages can be taken apart. One attractive possibility is that the putative AAA+ ATPase torsinA could be involved in remodeling or disassembly of SUN-KASH bridges. This hypothesis is strengthened by the observation that torsinA harbors a redox sensor motif (Zhu et al., 2010), which, we speculate, might be necessary for breaking disulfide-bonded SUN-KASH pairs. TorsinA is found in the ER lumen, and, interestingly, a torsinA mutant bearing a single glutamate deletion accumulates in the perinuclear space (Goodchild and Dauer, 2004) and has reported affinity for KASH peptides (Nery et al., 2008). The same mutation in torsinA causes the majority of torsinA-associated cases of early-onset dystonia, a severe movement disorder (Ozelius et al., 1997). Although there is no data yet that the development of dystonia can be attributed to a potential function of torsinA in remodeling SUN-KASH bridges, there is increasing evidence that functionality of the LINC complex is important for human health. It might not be a coincidence that mutations in Nesprins have recently been linked to another movement disorder, Emery-Dreifuss muscular dystrophy (EDMD) (Puckelwartz et al., 2009; Zhang et al., 2007a). Most commonly, EDMD is caused by mutations in the emerin (EMD) gene, and less frequently by mutations in LMNA, the gene encoding for A-type lamins. Notably, lamins and the INM protein emerin are associated with LINC complexes (Fridkin et al., 2009). In addition, mutations in LUMA (TMEM43), another SUN2 interacting protein, have recently been identified in EDMD-related myopathy patients (Liang et al., 2011). Thus, the LINC complex and associated proteins emerge at the root of several inherited diseases associated with impaired locomotion.

In summary, our structural and biochemical characterization of the SUN-KASH complex has revealed first insights into the organization of this important link.

The structure can now help to formulate specific, testable hypotheses, which will undoubtedly enrich this field in the short and long term.

## EXPERIMENTAL PROCEDURES

### DNA Constructs

DNA fragments coding for the C-terminal 29 aa of Nesprin-1 and Nesprin-2 (KASH1 and KASH2), or a deletion variant of KASH2 lacking the C-terminal 4 aa (KASH2Δ), were amplified from HeLa cell cDNA by PCR. Coding sequences of all other KASH variants were cloned using oligonucleotides. All KASH constructs were cloned into pQE60-2z (Kutay et al., 1997). SUN2 fragments were amplified by PCR using pEGFP-N3-SUN2 (Turgay et al., 2010) as template. PCR fragments were cloned into pETDuet-1 (Novagen) to yield an N-terminally 6xHis-tagged fusion protein. Mutations were introduced into $SUN2_{335-717}$ and $SPAG4_{1-189}$-$SUN2_{507-717}$-GFP (Turgay et al., 2010) by QuikChange site-directed mutagenesis (Stratagene).

For crystallographic purposes, $SUN2_{522-717}$ was cloned similarly into pETDuet-1, but in addition a short coiled-coil fragment of GCN4 and a cleavage site for human rhinovirus 3C protease were inserted following the 6xHis-tag (Ciani et al., 2010). To co-express SUN2-KASH complexes, human Nesprin-1 (KASH1, residues 8769-8797) and Nesprin-2 (KASH2, residues 6857-6885) were cloned into the second cassette of the 6xHis-tri-GCN4 tagged SUN2 as 3C-cleavable maltose-binding-protein (MBP) fusions.

### Recombinant Protein Expression and Purification

For protein binding assays, zz-KASH constructs were expressed in *E. coli* BLR(pRep4) for 4 hr at 37°C after induction with 0.5 mM IPTG. His-SUN2 constructs were expressed in *E. coli* BL21 Rosetta for 16 hr at 18°C after induction with 0.5 mM IPTG for $SUN2_{335-717}$ and $SUN2_{335-539}$, or with 0.1 mM IPTG for all other constructs. Cells were lysed by sonication in 50 mM Tris/HCl, pH 7.5, 200 mM NaCl. Lysates were cleared by ultracentrifugation (1 h,

50 000 rpm, Ti70, Beckman). SUN2 constructs were purified over Ni-NTA agarose (Qiagen), eluted in 50 mM Tris/HCl pH 7.5, 200 mM NaCl, 400 mM imidazole, and precipitated with 30% saturated $(NH_4)_2SO_4$ solution. Precipitates were dissolved in 50 mM Tris/HCl, pH 7.5, 200 mM NaCl, and residual ammonium sulfate was removed using illustra MicroSpin G-50 columns (GE Healthcare).

For crystallography, SUN2-KASH1 and SUN2-KASH2 were expressed in *E. coli* BL21(DE3) RIL strains (Stratagene). The bacterial cultures were grown at 30°C to an optical density (OD) of 0.6. Then, the culture was shifted to 18°C for 30 min and induced overnight with 0.2 mM IPTG. Apo-$SUN2_{522-717}$ expressing cells were resuspended in lysis buffer (50 mM potassium phosphate pH 8.0, 400 mM NaCl, 3 mM β-ME, 40 mM imidazole) and lysed. The lysate was supplemented with 1U/ml Benzonase (Sigma) 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/HCl, 150 mM NaCl, 3 mM β-ME, 250 mM imidazole, pH 8.0). The eluted protein was purified by size exclusion chromatography on a Superdex S200 column (GE Healthcare) equilibrated in crystallization buffer (10 mM Tris/HCl pH 8.0, 150 mM NaCl, 0.1 mM EDTA and 1 mM DTT). After cleaving with 3C protease the protein was separated from the fusion tag by a second size exclusion step under identical conditions. SUN2-KASH1 and SUN2-KASH2 were purified using the protocol developed for apo-SUN2, except that reducing and chelating agents were omitted and elution as well as crystallization buffers contained 1 mM $CaCl_2$.

To allow selenium phasing, a second methionine was introduced into the SUN2522-717 sequence (L546M) using PCR mutagenesis. The seleno-methionine (SeMet)-derived protein was obtained using methionine-pathway inhibition as described before (Brohawn et al., 2008).

For the SUN-KASH pull-down experiment shown in Figure 8, the 6xHis-SUN2-MBP-KASH2 pairs were co-expressed in *E. coli* and Ni-purified under the conditions described for the preparation of the crystallographic samples.

**In Vitro Binding Experiments**

Per binding reaction, 15 µl of IgG sepharose were saturated with zz-KASH derivatives by incubation in cleared lysate of *E. coli* expressing the respective construct. Beads were washed and added to binding reactions.

For binding of endogenous SUN proteins, 2 x $10^6$ HeLa cells were lysed in 200 µl RIPA buffer (50 mM Tris/HCl pH 7.5, 200 mM NaCl, 0.5% NP40, 0.1% DOC, 0.025% SDS) containing protease inhibitors (1 mM PMSF, 10 µg/ml aprotinin, 10 µg/ml leupeptin, 1 µg/ml pepstatin A).

For binding of recombinant SUN2 constructs, purified proteins were added to 200 µl 'empty' *E. coli* lysate supplemented with RIPA detergents (final buffer composition: 50 mM Tris/HCl, pH 7.5, 200 mM NaCl, 0.5% NP40, 0.1% DOC, 0.025% SDS). Alternatively (Figure 11), cleared lysate of *E. coli* expressing recombinant $SUN2_{335-717}$ was used directly.

Samples were incubated for 4 hr at 4°C in an overhead shaker. Beads were washed and bound proteins were eluted with SDS sample buffer.

**Protein Crystallization**

Purified $SUN2_{522-717}$ and SeMet-derivatized SUN2(L546M) were concentrated to 15 mg/ml prior to crystallization. The apo-proteins crystallized in 16% (w/v) polyethylene glycol (PEG) 3350 and 200 mM potassium thiocyanate by the hanging drop vapor diffusion method in 2 µl drops at 18°C. Rhombohedral crystals grew within 2-5 days with dimensions of 15 $\mu m^3$ x 15 µm x 15 µm. The

co-purified SUN2-KASH1 complex was concentrated to 13 mg/ml and similarly-shaped crystals were obtained from 100 mM HEPES pH 7.4, 7% PEG 4000, 10% 1,6-hexandediol and 0.25% n-decyl-β-D-maltoside (DM). For the SUN2-KASH2 crystals, the complex was concentrated to 10 mg/ml and was set up in 100 mM HEPES pH 7.5, 200 mM ammonium acetate, 25% 2-propanol and 0.3% DM. All crystals were cryoprotected in the reservoir solution supplemented with 18% (v/v) glycerol. Data were collected at beamlines 24ID-C/-E at Argonne National Laboratory.

**Structure Determination**

Data reduction was carried out using HKL2000 (Otwinowski and Minor, 1997). The structure of apo-SUN2 was solved using single anomalous dispersion (SAD) data from the SeMet derivatized SUN2 L546M mutant. The AutoSol procedure from the PHENIX suite was used for phasing and initial refinement (Adams et al., 2010). The final model was refined against native data extending to 2.2 Å. Model building was carried out with Coot (Emsley et al., 2010) and refinement was done with phenix.refine from the PHENIX suite. All residues were assigned including an additional Gly-Pro dipeptide at the N terminus, a remnant of the proteolytic cleavage site.

The SUN2-KASH1 and SUN2-KASH2 structures were subsequently solved by molecular replacement using apo-SUN2 as the search model. Model building was carried out with Coot and refinement was done with phenix.refine. Both complex models are complete for the $SUN2_{522-717}$ sequence. For the KASH1 and KASH2 peptides, the N-terminal residues 8769-8771 and 6857-6860, respectively, could not be positioned, presumably due to their flexibility in the crystal.

## Analytical Ultracentrifugation

Purified $SUN2_{335-717}$ was gel-filtered in 10 mM Tris/HCl pH 8.0, 150 mM NaCl and 1 mM $CaCl_2$ immediately prior to the experiments. Analytical ultracentrifugation experiments were carried out with an Optima XL-A centrifuge using an An60Ti rotor. Samples for sedimentation equilibrium (110 μl sample or 120 μl buffer) were loaded into Epon-charcoal 6 channel centerpieces, fit with quartz windows, and spun at 7000 and 10000 rpm, respectively. Two concentrations (0.3 and 0.5 mg/ml) were analyzed. Reaching of the sedimentation equilibrium was monitored with Winmatch. Absorbance data was collected at 280 nm at 5 replicates per 1 nm step. Sedimentation equilibrium data were globally fitted with Ultrascan II (http://ultrascan.uthscsa.edu) with a single ideal species model.

## Circular Dichroism

The experiments were carried out at 25°C using an Aviv Model 202 spectrometer. Spectra were recorded at 25 μM protein concentration in a cuvette with 0.1 mm optical pathlength. Spectra were recorded in 1 nm steps, averaged for 10 sec, and corrected for buffer baseline.

## Cell Culture, Transfection and Microscopy

HeLa cells were maintained in DMEM containing 10% FCS, 100 U/ml penicillin, and 100 μg/ml streptomycin at 37°C and 5% $CO_2$. Transient transfection was performed using X-tremeGENE transfection reagent (Roche). 24 hr after transfection, cells were fixed with 1% PFA for 10 min and then washed with PBS. Coverslips were mounted on VectaShield (VectorLabs). Confocal images were acquired with a Leica TCS-SP1 microscope using a HCX PI APO lbd.Bl. 40x, NA 1.25 oil immersion objective.

## Analysis of DTT-sensitive disulfide bridges

HeLa cells were transiently transfected with plasmid DNA encoding GFP-Nesprin-1c (aa 8444-8797 of Nesprin-1 isoform 1, NP_892006.3), GFP-Nesprin-2c (aa 6555-6885 of Nesprin-2 isoform 1, NP_055995.4), or their mutant versions. After 48 hr, cells were harvested and treated with 20 mM N-ethyl maleimide (NEM). Cells were lysed in SDS sample buffer either without (non-reducing) or with 50 mM DTT (reducing). Lysates were analyzed by SDS-PAGE followed by Western Blotting.

## Antibodies

α-SUN1 was raised in rabbits against a recombinant fragment ($SUN1_{520-626}$) of human SUN1. α-SUN2 has been described (Turgay et al., 2010).

## Accession Numbers

The coordinate file and structure factors for the SUN2-KASH1, SUN2-KASH2, and apo-SUN2 crystal structure were deposited in the Protein Data Bank under accession code 4DXR, 4DXS, and 4DXT, respectively.

## REFERENCES


Apel, E.D., Lewis, R.M., Grady, R.M., and Sanes, J.R. (2000). Syne-1, a dystrophin- and Klarsicht-related protein associated with synaptic nuclei at the neuromuscular junction. J. Biol. Chem. *275*, 31986-31995.

Appenzeller-Herzog, C., and Ellgaard, L. (2008). The human PDI family: versatility packed into a single fold. Biochim. Biophys. Acta *1783*, 535-548.

Bianchet, M.A., Odom, E.W., Vasta, G.R., and Amzel, L.M. (2002). A novel fucose recognition fold involved in innate immunity. Nat. Struct. Biol. *9*, 628-634.

Burke, B., and Roux, K.J. (2009). Nuclei take a position: managing nuclear location. Dev. Cell *17*, 587-597.

Chikashige, Y., Tsutsumi, C., Yamane, M., Okamasa, K., Haraguchi, T., and Hiraoka, Y. (2006). Meiotic proteins bqt1 and bqt2 tether telomeres to form the bouquet arrangement of chromosomes. Cell *125*, 59-69.

Ciani, B., Bjelic, S., Honnappa, S., Jawhari, H., Jaussi, R., Payapilly, A., Jowitt, T., Steinmetz, M.O., and Kammerer, R.A. (2010). Molecular basis of coiled-coil oligomerization-state specificity. PNAS *107*, 19850-19855.

Crisp, M., Liu, Q., Roux, K., Rattner, J.B., Shanahan, C., Burke, B., Stahl, P.D., and Hodzic, D. (2006). Coupling of the nucleus and cytoplasm: role of the LINC complex. J. Cell Biol. *172*, 41-53.

Ding, X., Xu, R., Yu, J., Xu, T., Zhuang, Y., and Han, M. (2007). SUN1 is required for telomere attachment to nuclear envelope and gametogenesis in mice. Dev. Cell *12*, 863-872.

Fridkin, A., Mills, E., Margalit, A., Neufeld, E., Lee, K.K., Feinstein, N., Cohen, M., Wilson, K.L., and Gruenbaum, Y. (2004). Matefin, a C. elegans germ line-specific SUN-domain nuclear membrane protein, is essential for early embryonic and germ cell development. PNAS *101*, 6987-6992.

Fridkin, A., Penkner, A., Jantsch, V., and Gruenbaum, Y. (2009). SUN-domain and KASH-domain proteins during development, meiosis and disease. CMLS *66*, 1518-1533.

Goodchild, R.E. and Dauer, W.T. (2004). Mislocalization to the nuclear envelope: an effect of the dystonia-causing torsinA mutation. PNAS *101*, 847-852.

Harding, M.M. (2006). Small revisions to predicted distances around metal sites in proteins. Acta Cryst. D *62,* 678-682.

Haque, F., Lloyd, D.J., Smallwood, D.T., Dent, C.L., Shanahan, C.M., Fry, A.M., Trembath, R.C., and Shackleton, S. (2006). SUN1 interacts with nuclear lamin A and cytoplasmic nesprins to provide a physical connection between the nuclear lamina and the cytoskeleton. Mol. Cell. Biol. *26*, 3738-3751.

Ketema, M., Wilhelmsen, K., Kuikman, I., Janssen, H., Hodzic, D., and Sonnenberg, A. (2007). Requirements for the localization of nesprin-3 at the NE and its interaction with plectin. J. Cell Sci. *120*, 3384-3394.

Lee, K.K., Starr, D., Cohen, M., Liu, J., Han, M., Wilson, K.L., and Gruenbaum, Y. (2002). Lamin-dependent localization of UNC-84, a protein required for nuclear migration in C. elegans. Mol. Biol. Cell *13*, 892-901.

Lei, K., Zhang, X., Ding, X., Guo, X., Chen, M., Zhu, B., Xu, T., Zhuang, Y., Xu, R., and Han, M. (2009). SUN1 and SUN2 play critical but partially redundant

roles in anchoring nuclei in skeletal muscle cells in mice. PNAS *106*, 10207-10212.

Liang, W.C., Mitsuhashi, H., Keduka, E., Nonaka, I., Noguchi, S., Nishino, I., and Hayashi, Y.K. (2011). TMEM43 mutations in Emery-Dreifuss muscular dystrophy-related myopathy. Ann. Neurol. *69*, 1005-1013.

Lombardi, M.L., and Lammerding, J. (2011). Keeping the LINC: the importance of nucleocytoskeletal coupling in intracellular force transmission and cellular function. Biochem. Soc. Trans. *39*, 1729-1734.

Lu, W., Gotzmann, J., Sironi, L., Jaeger, V.M., Schneider, M., Luke, Y., Uhlen, M., Szigyarto, C.A., Brachner, A., Ellenberg, J., *et al.* (2008). Sun1 forms immobile macromolecular assemblies at the NE. Biochim. Biophys. Acta *1783*, 2415-2426.

Luxton, G.W., Gomes, E.R., Folker, E.S., Vintinner, E., and Gundersen, G.G. (2010). Linear arrays of nuclear envelope proteins harness retrograde actin flow for nuclear movement. Science *329*, 956-959.

Malone, C.J., Fixsen, W.D., Horvitz, H.R., and Han, M. (1999). UNC-84 localizes to the nuclear envelope and is required for nuclear migration and anchoring during C. elegans evelopment. Development *126*, 3171-3181.

Mans, B.J., Anantharaman, V., Aravind, L., and Koonin, E.V. (2004). Comparative genomics, evolution and origins of the NE and NPC. Cell Cycle *3*, 1612-1637.

McDonnell, A.V., Jiang, T., Keating, A.E., and Berger, B. (2006). Paircoil2: improved prediction of coiled coils from sequence. Bioinformatics *22*, 356-358.

Mekhail, K., and Moazed, D. (2010). The nuclear envelope in genome organization, expression and stability. Nat. Rev. Mol. Cell Biol. *11*, 317-328.

Mislow, J.M., Holaska, J.M., Kim, M.S., Lee, K.K., Segura-Totten, M., Wilson, K.L., and McNally, E.M. (2002). Nesprin-1alpha self-associates and binds directly to emerin and lamin A in vitro. FEBS Lett. *525*, 135-140.

Nery, F.C., Zeng, J., Niland, B.P., Hewett, J., Farley, J., Irimia, D., Li, Y., Wiche, G., Sonnenberg, A., and Breakefield, X.O. (2008). TorsinA binds the KASH domain of nesprins and participtes in linkage between nuclear envelope and cytoskeleton. J. Cell Sci. *121*, 3476-3486.

Onischenko, E., and Weis, K. (2011). Nuclear pore complex-a coat specifically tailored for the nuclear envelope. Curr. Opin. Cell Biol. *23*, 293-301.

Ostlund, C., Folker, E.S., Choi, J.C., Gomes, E.R., Gundersen, G.G., and Worman, H.J. (2009). Dynamics and molecular interactions of linker of nucleoskeleton and cytoskeleton (LINC) complex proteins. J. Cell Sci. *122*, 4099-4108.

Ozelius, L.J., Hewett, J.W., Page, C.E., Bressman, S.B., Kramer, P.L., Shalish, C., de Leon, D., Brin, M.F., Raymond, D., Corey, D.P., *et al.* (1997). The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein. Nat. Genet. *17*, 40-48.

Padmakumar, V.C., Abraham, S., Braune, S., Noegel, A.A., Tunggal, B., Karakesisoglou, I., and Korenbaum, E. (2004). Enaptin, a giant actin-binding protein, is an element of the nuclear membrane and the actin cytoskeleton. Exp. Cell Res. *295*, 330-339.

Padmakumar, V.C., Libotte, T., Lu, W., Zaim, H., Abraham, S., Noegel, A.A., Gotzmann, J., Foisner, R., and Karakesisoglou, I. (2005). The INM protein Sun1 mediates the anchorage of Nesprin-2 to the nuclear envelope. J. Cell Sci. *118*, 3419-3430.

Penkner, A., Tang, L., Novatchkova, M., Ladurner, M., Fridkin, A., Gruenbaum, Y., Schweizer, D., Loidl, J., and Jantsch, V. (2007). The nuclear envelope protein Matefin/SUN-1 is required for homologous pairing in C. elegans meiosis. Dev.Cell *12*, 873-885.

Puckelwartz, M.J., Kessler, E., Zhang, Y., Hodzic, D., Randles, K.N., Morris, G., Earley, J.U., Hadhazy, M., Holaska, J.M., Mewborn, S.K., *et al.* (2009). Disruption of nesprin-1 produces an Emery Dreifuss muscular dystrophy-like phenotype in mice. Hum. Mol. Genet. *18*, 607-620.

Razafsky, D., and Hodzic, D. (2009). Bringing KASH under the SUN: the many faces of nucleo-cytoskeletal connections. J. Cell Biol. *186*, 461-472.

Roux, K.J., Crisp, M.L., Liu, Q., Kim, D., Kozlov, S., Stewart, C.L., and Burke, B. (2009). Nesprin 4 is an ONM protein that can induce kinesin-mediated cell polarization. PNAS *106*, 2194-2199.

Schmitt, J., Benavente, R., Hodzic, D., Hoog, C., Stewart, C.L., and Alsheimer, M. (2007). Transmembrane protein Sun2 is involved in tethering mammalian meiotic telomeres to the NE. PNAS *104*, 7426-7431.

Shibata, Y., Shemesh, T., Prinz, W.A., Palazzo, A.F., Kozlov, M.M., and Rapoport, T.A. (2010). Mechanisms determining the morphology of the peripheral ER. Cell *143*, 774-788.

Starr, D.A., and Fischer, J.A. (2005). KASH 'n Karry: the KASH domain family of cargo-specific cytoskeletal adaptor proteins. BioEssays *27*, 1136-1146.

Starr, D.A., and Fridolfsson, H.N. (2010). Interactions between nuclei and the cytoskeleton are mediated by SUN-KASH nuclear-envelope bridges. Annu. Rev. Cell Dev. Biol. *26*, 421-444.

Starr, D.A., and Han, M. (2002). Role of ANC-1 in tethering nuclei to the actin cytoskeleton. Science *298*, 406-409.

Stewart, C.L., Roux, K.J., and Burke, B. (2007). Blurring the boundary: the nuclear envelope extends its reach. Science *318*, 1408-1412.

Stewart-Hutchinson, P.J., Hale, C.M., Wirtz, D., and Hodzic, D. (2008). Structural requirements for the assembly of LINC complexes and their function in cellular mechanical stiffness. Exp. Cell Res. *314*, 1892-1905.

Testa, O.D., Moutevelis, E., and Woolfson, D.N. (2009). CC+: a relational database of coiled-coil structures. Nucleic Acids Res. *37*, D315-322.

Turgay, Y., Ungricht, R., Rothballer, A., Kiss, A., Csucs, G., Horvath, P., and Kutay, U. (2010). A classical NLS and the SUN domain contribute to the targeting of SUN2 to the INM. EMBO J. *29*, 2262-2275.

Tzur, Y.B., Wilson, K.L., and Gruenbaum, Y. (2006). SUN-domain proteins: 'Velcro' that links the nucleoskeleton to the cytoskeleton. Nat. Rev. Mol. Cell Biol. *7*, 782-788.

Wang, Q., Du, X., Cai, Z., and Greene, M.I. (2006). Characterization of the structures involved in localization of the SUN proteins to the NE and the centrosome. DNA Cell Biol. *25*, 554-562.

Wilhelmsen, K., Litjens, S.H., Kuikman, I., Tshimbalanga, N., Janssen, H., van den Bout, I., Raymond, K., and Sonnenberg, A. (2005). Nesprin-3, a novel ONM protein, associates with the cytoskeletal linker protein plectin. J. Cell Biol. *171*, 799-810.

Worman, H.J., and Gundersen, G.G. (2006). Here come the SUNs: a nucleocytoskeletal missing link. Trends Cell Biol. *16*, 67-69.

Yu, J., Lei, K., Zhou, M., Craft, C.M., Xu, G., Xu, T., Zhuang, Y., Xu, R., and Han, M. (2011). KASH protein Syne-2/Nesprin-2 and SUN proteins SUN1/2 mediate nuclear migration during mammalian retinal development. Hum. Mol. Genet. *20*, 1061-1073.

Zhang, J., Felder, A., Liu, Y., Guo, L.T., Lange, S., Dalton, N.D., Gu, Y., Peterson, K.L., Mizisin, A.P., Shelton, G.D., *et al.* (2010). Nesprin 1 is critical for nuclear positioning and anchorage. Hum. Mol. Genet. *19*, 329-341.

Zhang, Q., Bethmann, C., Worth, N.F., Davies, J.D., Wasner, C., Feuer, A., Ragnauth, C.D., Yi, Q., Mellad, J.A., Warren, D.T., *et al.* (2007a). Nesprin-1 and -2 are involved in the pathogenesis of Emery Dreifuss muscular dystrophy and are critical for NE integrity. Hum. Mol. Genet. *16*, 2816-2833.

Zhang, Q., Skepper, J.N., Yang, F., Davies, J.D., Hegyi, L., Roberts, R.G., Weissberg, P.L., Ellis, J.A., and Shanahan, C.M. (2001). Nesprins: a novel family of spectrin-repeat-containing proteins that localize to the nuclear membrane in multiple tissues. J. Cell Sci. *114*, 4485-4498.

Zhang, X., Lei, K., Yuan, X., Wu, X., Zhuang, Y., Xu, T., Xu, R., and Han, M. (2009). SUN1/2 and Syne/Nesprin-1/2 complexes connect centrosome to the

nucleus during neurogenesis and neuronal migration in mice. Neuron *64*, 173-187.

Zhang, X., Xu, R., Zhu, B., Yang, X., Ding, X., Duan, S., Xu, T., Zhuang, Y., and Han, M. (2007b). Syne-1 and Syne-2 play crucial roles in myonuclear anchorage and motor neuron innervation. Development *134*, 901-908.

Zhen, Y.Y., Libotte, T., Munck, M., Noegel, A.A., and Korenbaum, E. (2002). NUANCE, a giant protein connecting the nucleus and actin cytoskeleton. J. Cell Sci. *115*, 3207-3222.

Zhou, Z., Du, X., Cai, Z., Song, X., Zhang, H., Mizuno, T., Suzuki, E., Yee, M.R., Berezov, A., Murali, R., Wu, S.L., Karger, B.L., Greene M.I., and Wang, Q. (2012). Structure of the SUN domain defines features of a molecular bridge in the nuclear envelope. J. Biol. Chem., *287,* 5317-5326.

Zhu L., Millen, L., Mendoza, J.L., and Thomas P.J. (2010). A unique redox-sensing sensor motif in torsinA plays a critical role in nucleotide and partner binding. J. Biol. Chem. *285*, 37271-37280.

## Chapter Three

# LAP1 and LULL1 are Catalytically Inactive AAA+ ATPases that form Ring Structures with Tor1A

**SUMMARY**

Primary dystonia is an autosomal dominant movement disorder caused by a single glutamate deletion in the AAA+ ATPase Tor1A. Despite its well known role in the disease, little biochemical information is available for this protein. Here we characterize the complex of Tor1A with its binding partners LAP1 and LULL1. By bioinformatic analysis we show that LAP1 and LULL1 belong to a growing family of catalytically inactive AAA+ ATPases. We designed a purification scheme that allows for the production of large quantities of chromatographically pure Tor1A-LAP1/LULL1 complexes and define their stoichiometry using analytical ultracentrifugation. Negative stain single particle electron microscopy of copurified Tor1A/LULL1 shows that they form oligomeric rings similar to other AAA+ ATPases. We suggest that LAP1 and LULL1 are catalytically inactive AAA+ ATPases that are important at least for Tor1A complex assembly.

## INTRODUCTION

Dystonia is a movement disorder that affects 300,000 people in North America alone. This neurological disease is characterized by involuntary muscle contractions leading to repetitive, twisting movements and irregular postures (Granata et al., 2010; Hewett, 2000). In 1997, it was discovered that the most common and severe type of dystonia, the autosomal dominant early-onset generalized dystonia, was caused by an in-frame ΔGAG deletion in the DYT1 (TorsinA or Tor1A) gene leading to the deletion of a single glutamate residue within the protein (Figure 1) (Ozelius et al., 1997a; 1997b). DYT1 patients show no signs of neurodegeneration or neuropathology, suggesting that functional defects in neuronal tissue are the problem rather than cell death. Interestingly, despite being ubiquitously expressed, only neurons are affected by the disease, probably due to compensation by one of three known Tor1A homologues, Tor1B, Tor2A and Tor3A in non-neuronal tissue (Figure 1) (Hewett, 2000; Jungwirth et al., 2010; Zhu et al., 2010).

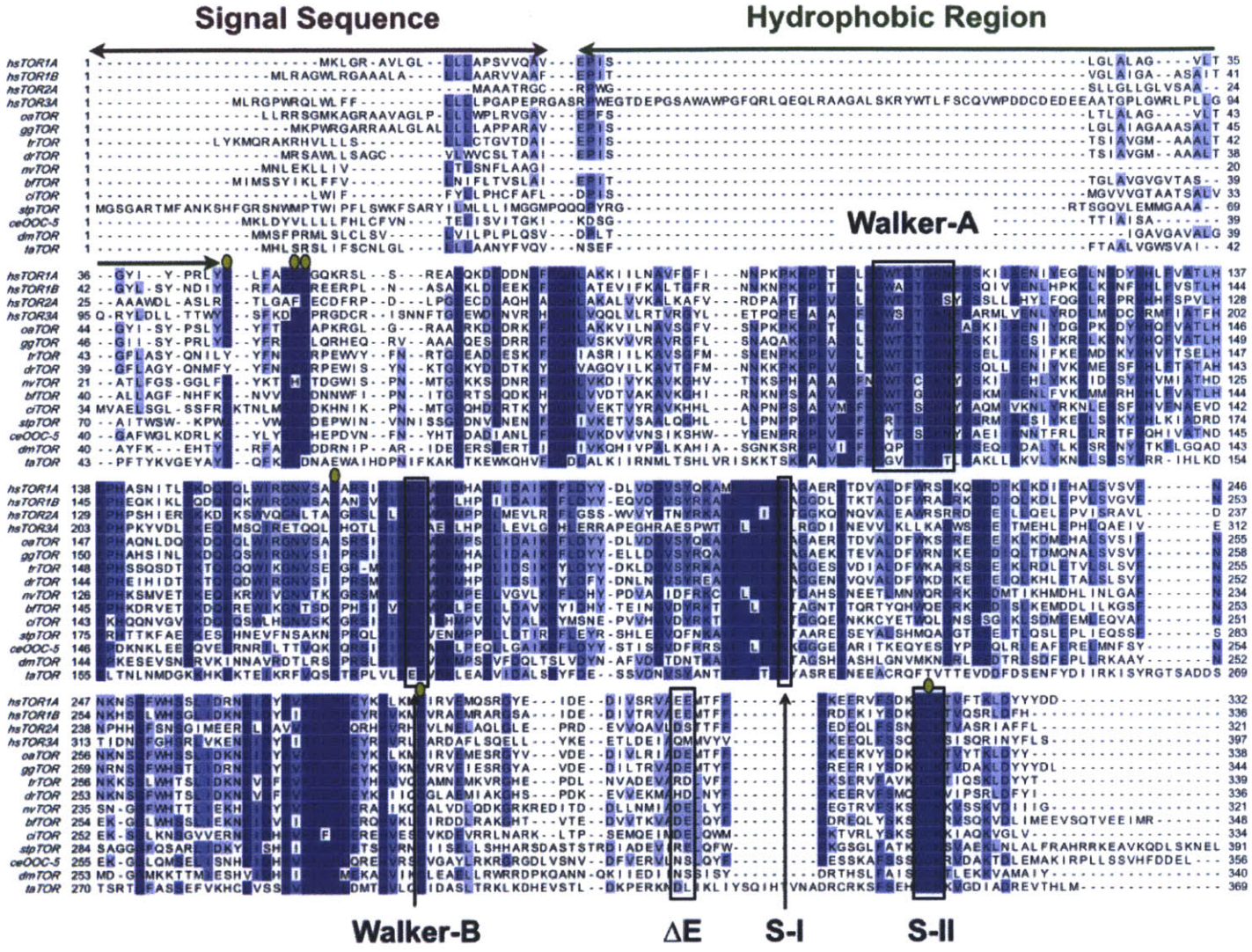


**Figure 1. Multiple sequence alignment of Torsin AAA+ ATPases.** Torsin AAA+ ATPases from human and highly diverged eukaryotes are shown. ER signal peptide, hydrophobic region, Walker-A, Walker-B, sensor-1 (S-I) and sensor-II (S-II) motifs, and the disease causing ΔE mutation are highlighted. Yellow dots indicate conserved cysteine residues.

Torsins are members of the ATPases Associated with a variety of cellular Activities (AAA+ ATPases) superfamily (Basham and Rose, 2001; Erzberger and Berger, 2006; Ozelius et al., 1997b; Warner et al., 2010; Zhao et al., 2013; Zhu et al., 2010). AAA+ ATPases use ATP hydrolysis to undergo conformational changes and to exert mechanical force onto a substrate. As their name implies, members of this family of proteins are involved in a vast array of processes in the cell including vesicle fusion and scission, protein unfolding, complex assembly or

disassembly, protein transport, nucleic acid remodeling, etc (Erzberger and Berger, 2006; Glynn et al., 2009; Narlikar et al., 2013; Wendler et al., 2012). AAA+ ATPases share the same overall fold, modulated by different appendages that allow performing different functions. On the primary sequence level, they all share several important motifs involved in the binding and hydrolysis of ATP (Erzberger and Berger, 2006). These canonical motifs are the Walker-A, Walker-B, sensor-I, sensor-II and arginine-finger. The Walker-A motif, consensus sequence of GxxGxGKT/S (where x is any amino acid), is involved in binding ATP, whereas, the Walker-B motif, consensus sequence of hhhhDE (where h is any hydrophobic residue), is involved in the hydrolysis step. The sensor-I, sensor-II (consensus sequence of GAR) and arginine-finger motifs are also involved in the hydrolysis step. Torsins however have slight variations in two motifs: Walker-A is mutated from GxxGxGK**T/S** to GxxGxGK**N** and sensor-II from G**AR** to G**CK** (Figure 2) (Zhu et al., 2010; 2008). Due to the endoplasmic reticulum (ER) localization of Torsins, this sensor-II variation has been suggested to form a disulfide bridge with another well conserved cysteine residue within the small domain of the AAA+ fold. Furthermore, this putative disulfide bridge has been implicated in the redox regulation of Torsin's ATPase activity (Zhu et al., 2010; 2008).

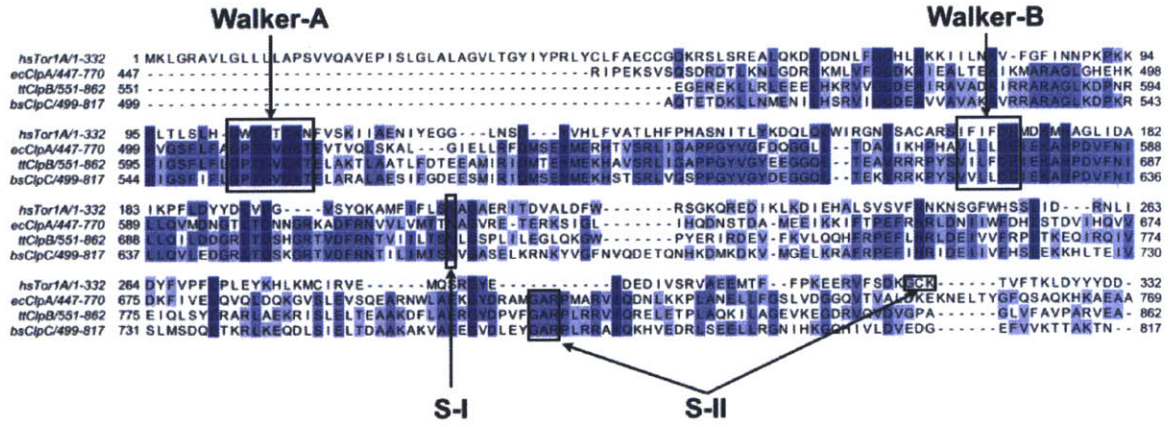


**Figure 2. Multiple sequence alignment of Tor1A with paralogous AAA+ ATPases.** Tor1A, ClpA, ClpB and ClpC are shown. Walker-A, Walker-B, sensor-1 (S-I) and sensor-II (S-II) motifs are highlighted.

As mentioned above, Torsins are ER resident proteins; hence they contain an N-terminal signal peptide that gets cleaved upon translation into the ER. After the signal peptide, Torsins contain a stretch of approximately 20 amino acid long hydrophobic region that has been shown to function in ER retention, possibly through membrane attachment (Vander Heyden et al., 2011). A conserved cysteine rich region whose function is unknown follows the hydrophobic region. Lastly, the AAA+ ATPase domain resides at the C terminus (Figure 1).

Despite extensive study, a specific function or substrates have not been attributed to Torsin ATPases. However, studies of the Torsin homologue OOC-5 in *Caenorhabditis elegans* have suggested that it is involved in nuclear rotation during oogenesis (Basham and Rose, 2001; 1999). It was also shown that mutating the sensor-II cysteines leads to a decrease in embryonic viability (Zhu et al., 2008). Mammalian Tor1A has been suggested to function in nuclear envelope (NE) organization, synaptic vesicle transport and turnover, secretory pathway, protein degradation, cytoskeletal organization and transport via NE budding (Granata et al., 2010; Jokhi et al., 2013; Vander Heyden et al., 2009; Warner et al., 2010). The specific role of Tor1A in any of these processes remains, however, largely unknown.

LInkers of Nucleoskeleton and Cytoskeleton (LINC) complexes connect the nucleus with the cytoskeleton via bridges that cross the NE (Starr et al., 2010). LINCs are composed of the inner nuclear membrane (INM) resident Sad1 UNC-84 homology (SUN) proteins and the outer nuclear membrane resident Klarsicht, ANC-1 and Syne/Nesprin Homology (KASH) motif containing proteins. The LINC complex allows for nuclear migration, nuclear anchorage and homologous chromosome pairing. The structures of the human SUN2-KASH1 and SUN2-KASH2 complexes have revealed a very tight hexameric assembly in which three SUN protomers interact with three KASH peptides (Sosa et al., 2012). Interestingly, a disulfide bridge is formed between the SUN protomers and KASH peptides effectively covalently linking the nucleus with the cytoskeleton. The

nature of this complex raises questions about its regulation throughout the cell cycle. The proposed redox regulated ATPase activity of Torsin makes it an interesting candidate for this role. Further, Tor1A was shown to bind the KASH peptides of Nesprin-1, -2 and -3 and upon overexpression, it displaces LINC complexes from the NE to the ER (Nery et al., 2008; Vander Heyden et al., 2009).

In the ER/NE lumen, Tor1A binds Lamina-Associated Polypeptide-1 (LAP1) and Luminal domain-Like LAP-1, (LULL1), as shown by multiple labs (Goodchild and Dauer, 2005; Zhao et al., 2013). Both, LAP1 and LULL1, are type-II integral membrane proteins with luminal domains that share 62% identity (Figure 3). LAP1, an INM protein, specifically localizes Tor1A to the NE whereas LULL1 targets Tor1A to both the ER and NE. LAP1 and LULL1 were originally thought to be Tor1A substrates, since they interact more tightly with ATP-bound Tor1A (Goodchild and Dauer, 2005). However, a more recent study on Tor1A-LAP1/LULL1 interaction suggests that LAP1/LULL1 are rather activating factors for ATP hydrolysis by Tor1A, which is otherwise an extremely slow enzyme (Zhao et al., 2013). This report also suggests that Tor1A-LAP1/LULL1 complexes exhibit a "switch-like" ATPase activity.

In the present study we set out to structurally characterize the Tor1A-LAP1/LULL1 complexes. We performed an extensive bioinformatic analysis of LAP1 and LULL1 and determine that they adopt an AAA+ fold, but are catalytically inactive. We designed novel purification strategies that allow for the recombinant production of large quantities of Tor1A-LAP1/LULL1 complexes using *Escherichia coli*. We show that the Tor1A-LAP1/LULL1 complexes have a 1:1 stoichiometry. Finally, using electron microscopy, we show that Tor1A-LAP1/LULL1 complexes form oligomeric rings, similar to other AAA+ ATPase systems. This study provides the first evidence of Tor1A-LAP1/LULL1 architecture and suggests models by which Tor1A gets targeted to either the ER and/or NE.

## RESULTS

### Bioinformatic analysis of Tor1A

Tor1A is a AAA+ ATPase, easily identifiable by the short sequence motifs that characterize the family (Figures 1 and 2). To model Tor1A, we performed an HHsearch on human Tor1A (residues 21-332) using HHpred (Soding, 2005; Soding et al., 2005; 2006). As expected, Tor1A shares high structural homology with several AAA+ ATPases, particularly with ClpA (pdb: 1r6b, chain X), ClpB (pdb: 4fcw, chain A) and ClpC (pdb: 3pxi, chain A) of the Clp family (Table 1) (Biter et al., 2012). With the profile alignment obtained from the HHpred search we proceeded to generate a structural model of Tor1A using MODELLER with 1qvr, chain A, 3pxi, chain A and 1r6b, chain X as templates (Figure 4). As expected, the model shows that Tor1A adopts a canonical AAA+ ATPase fold with the Walker-A and Walker-B motifs located in a pocket between the large and small domains. Glutamate 302/303, whose single deletion (ΔE) is responsible for primary dystonia, is located in the small domain helical bundle. The sensor-II motif is also located within the small domain. Cysteine 280 faces cysteine 219 of the sensor-II motif, suggesting disulfide bridge formation as previously determined (Zhu et al., 2008; 2010).

**Table 1. HHpred search results.**

| Tor1A | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Hit** | **Name** | **Prob** | **E-value** | **P-value** | **Score** | **SS** | **Cols** | **Query HMM** | **Template HMM** |
| 3pxi_A | ClpC | 99.9 | 4.90E-24 | 1.50E-28 | 200.8 | 17.5 | 229 | 5-278 | 476-719 (758) |
| 4fcw_A | ClpB | 99.9 | 7.40E-23 | 2.30E-27 | 177.8 | 16.3 | 256 | 5-279 | 2-282 (311) |
| 1r6b_X | ClpA | 99.9 | 1.40E-21 | 4.30E-26 | 189.2 | 18.4 | 248 | 11-278 | 449-718 (758) |
| **LAP1** | | | | | | | | | |
| **Hit** | **Name** | **Prob** | **E-value** | **P-value** | **Score** | **SS** | **Cols** | **Query HMM** | **Template HMM** |
| 3pxi_A | ClpC | 95.6 | 1.40E-01 | 4.40E-06 | 57.7 | 14.2 | 193 | 345-573 | 459-661 (758) |
| 4fcw_A | ClpB | 94.2 | 2.30E+00 | 7.10E-05 | 40.4 | 16.5 | 185 | 366-573 | 6-215 (311) |
| 1r6b_X | ClpA | 82.7 | 3.50E+01 | 1.10E-03 | 37.5 | 16.3 | 161 | 345-511 | 426-604 (758) |
| **LULL1** | | | | | | | | | |
| **Hit** | **Name** | **Prob** | **E-value** | **P-value** | **Score** | **SS** | **Cols** | **Query HMM** | **Template HMM** |
| 4fcw_A | ClpB | 92.9 | 2.50E+00 | 7.80E-05 | 39.3 | 14.4 | 186 | 250-459 | 3-215 (311) |
| 3pxi_A | ClpC | 92 | 2.90E+00 | 9.00E-05 | 46.3 | 15.4 | 174 | 260-449 | 487-673 (758) |
| 1r6b_X | ClpA | 71.4 | 1.10E+02 | 3.40E-03 | 32.8 | 15.9 | 158 | 231-397 | 426-604 (758) |

For definitions see Soding et al., 2005.

## Bioinformatic analysis of LAP1 and LULL1

To elucidate the function of the ER/NE luminal domains of LAP1 and LULL1 and how they interact with Tor1A, we performed a bioinformatic analysis to understand their structure. Sequence conservation analysis shows that they share highest homology within the last ~250 residues (62% identity) (Figure 3). Primary and secondary structure analysis reveals several features of these proteins. First, they both have one predicted transmembrane helix, which orients their respective C termini into the lumen of the ER/NE as previously shown (Goodchild and Dauer, 2005; Vander Heyden et al., 2009; Zhao et al., 2013). Second, both LAP1 and LULL1 have unstructured nucleo- and cytoplasmic regions, respectively. Third, their luminal regions appear folded and are composed of a mixture of alpha helices as well as beta strands.

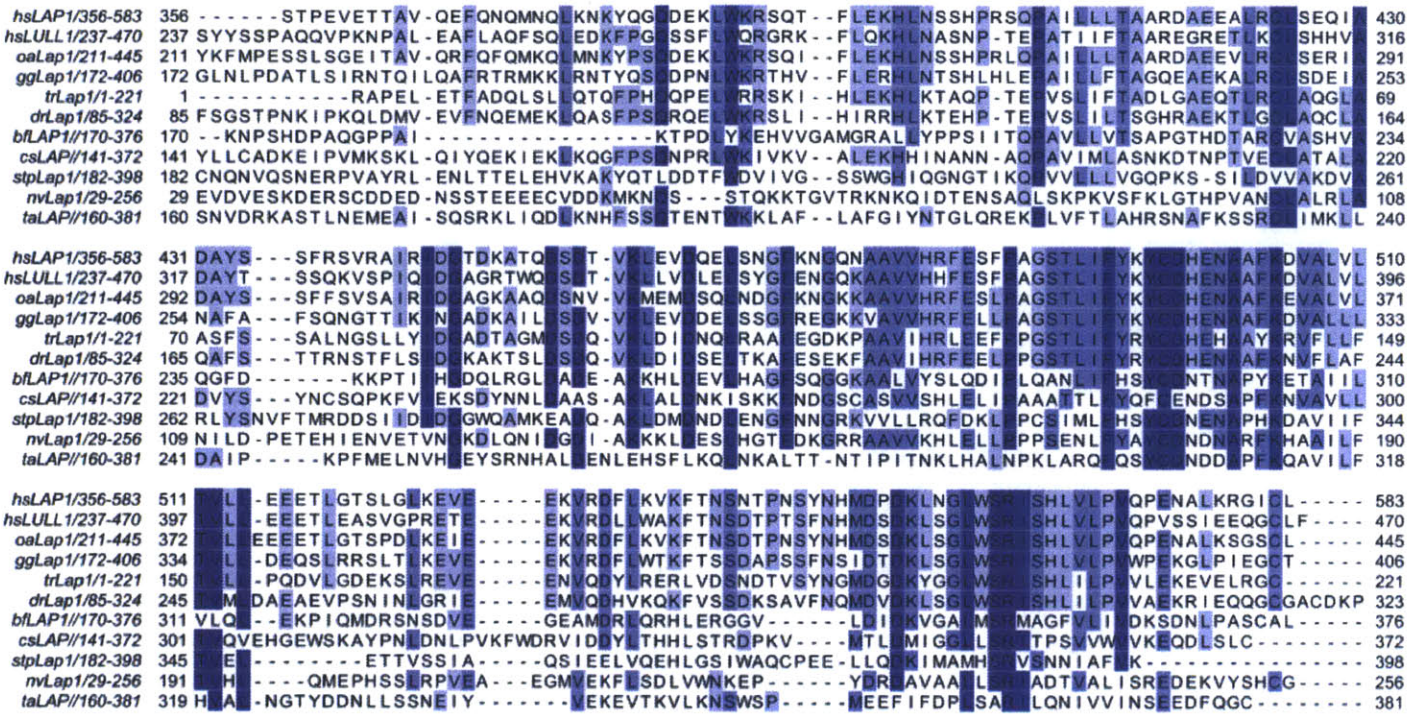


**Figure 3. Multiple sequence alignment of LAP1 and LULL1.** Luminal domains of human LAP1, LULL1 and highly diverged eukaryotic LAP1/LULL1 proteins are shown.

To obtain more detailed information about the structural content of the luminal regions of LAP1 and LULL1 we performed HHsearches. Interestingly, this analysis suggests that the luminal domains of LAP1 and LULL1 share high structural homology to AAA+ ATPases (Table 1). Similar to Tor1A, LAP1 and LULL1 both share homology with ClpA, ClpB and ClpC. Using the profile alignments obtained from the HHsearch, we generated structural models using MODELLER. The structural models obtained show a canonical AAA+ fold similar to the one obtained for Tor1A (Figure 4). However, neither LAP1 nor LULL1 are known to bind ATP. Primary sequence analysis shows that both are missing the defining features of ATPases, that is, they lack Walker-A, Walker-B, sensor-I and sensor-II motifs (Erzberger and Berger, 2006; Wendler et al., 2012). This observation was further corroborated by superposition of the models obtained with the known AAA+ ATPases structures of ClpA (pdb: 1r6b, chain X), ClpB (pdb: 4fcw, chain A) and ClpC (pdb: 3pxi chain A) (Figures 3 and 4).

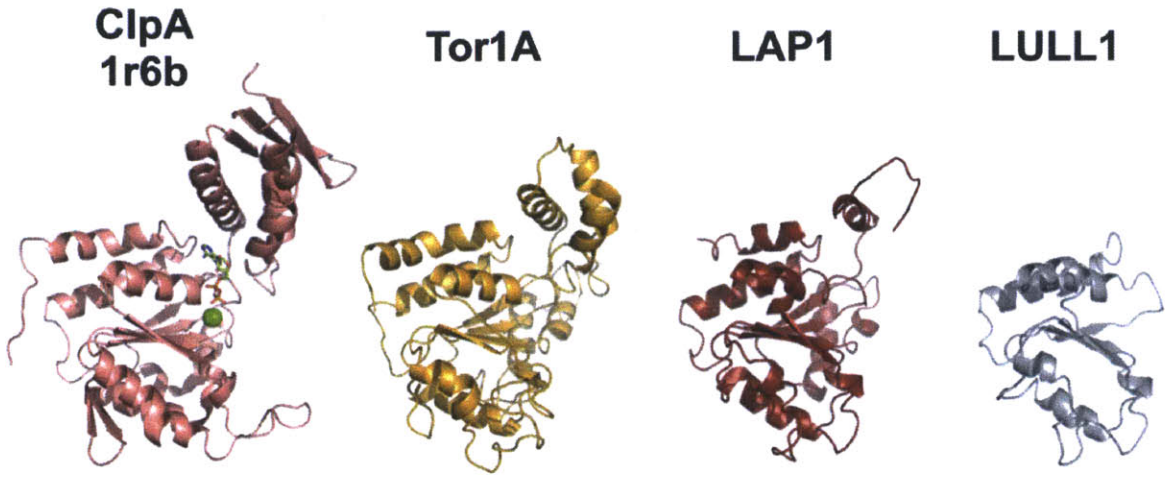


**Figure 4. Tor1A, LAP1 and LULL1 show high structural homology with members of the Clp family of AAA+ ATPases.** The crystal structure of ClpA (1r6b) and structural models of Tor1A, LAP1 and LULL1 are shown in the same orientation. For ClpA, bound ADP (sticks) and magnesium ion (sphere) are shown. Structural models were created using MODELLER and visualized using PyMOL.

## Purification of Tor1A-Lap1/LULL1 complexes

Purifying large quantities of Tor1A for in vitro studies has proven to be a difficult task. Some of the difficulties might arise from to the fact that Tor1A is an ER resident protein known to be glycosylated in two sites and likely to contain disulfide bonds that might be important for its stability and activity (Zhu et al., 2010). Additionally, Tor1A contains a hydrophobic region that might behave as a TM domain or as a monotypic membrane protein depending on the curvature of the membrane, which might provide difficulties when performing detergent extractions (Vander Heyden et al., 2011).

We set out to establish a purification protocol that would allow the production of large amounts of recombinantly expressed Tor1A in *E. coli*. First, we generated Tor1A constructs for expression by truncating Tor1A N- and C-terminally to residues 51-332, to remove the hydrophobic and cysteine-rich regions (Figure 1). However, expression of Tor1A by itself yielded insoluble and GroEL bound, thus presumably misfolded protein (Figure 5).

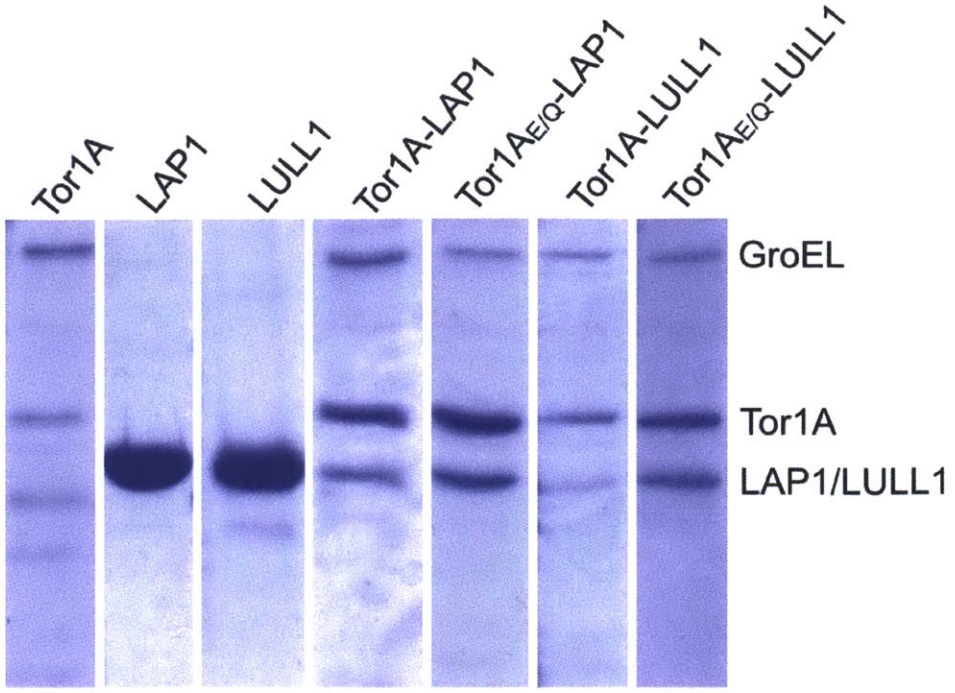


**Figure 5. Tor1A can be recombinantly purified when in complex with LAP1 or LULL1.** Single-step Ni-NTA purification of recombinantly expressed Tor1A, LAP1, and LULL1 under native conditions (analyzed by SDS-PAGE and Coomassie staining).. While LAP1 and LULL1 can be purified alone, Tor1A needs co-expression of either LAP1 or LULL1 for successful purification. The most common contaminant, GroEL is also shown.

After determining that LAP1 and LULL1 adopt catalytically inactive AAA+ ATPase folds, we reasoned that they might form stoichiometric complexes with Tor1A. Therefore, we generated *E. coli* strains that coexpress histidine tagged Tor1A together with the luminal domains of either LAP1 (residues 356-583) or LULL1 (residues 233-470). This strategy resulted in the expression of soluble Tor1A-LAP1/LULL1 complexes (Figure 5). Purification of Tor1A-LAP1/LULL1 followed a three-step regime, via Ni-affinity, cation-exchange, and size exclusion chromatography.

We compared the purification of wild-type versus E171Q "substrate-trap" mutant (herein Tor1A-E/Q) Tor1A. It has been shown that catalytically inactive, ATP-trapped Tor1A-E/Q mutant binds LAP1 and LULL1 stronger than wild-type (Goodchild and Dauer, 2005; Zhao et al., 2013; Zhu et al., 2010). We also observed this in our purifications during the Ni-affinity step, in which larger

amounts of complex eluted in the presence of the E171Q mutation than in the wild-type (Figure 5). Nonetheless, all complexes fell apart in later steps indicating rather moderate binding affinities. To solve the issue of complex stability we added ATP in the presence of $MgCl_2$ to the purification buffers. This significantly improved the stability of both Tor1A-LAP1/LULL1 complexes *in vitro*. Of note, purifications of the monomeric luminal domains of LAP1 and LULL1 yielded high amounts of pure proteins without the need of any additives in the purification (Figure 5).

## Characterization of Tor1A-LAP1 and Tor1A-LULL1 complexes

AAA+ ATPases oligomerize, usually into hexamers, in order to perform their function. ClpA, ClpB and ClpC, the closest structural homologs of Tor1A, assemble into hexameric double-rings (Biter et al., 2012; Lee et al., 2003; Schlee et al., 2001; Xia et al., 2004). This is because Clp family proteins contain two consecutive ATPase domains, termed D1 and D2. Upon oligomerization, these D1 and D2 domains assemble into two stacked hexameric rings, as shown by EM.

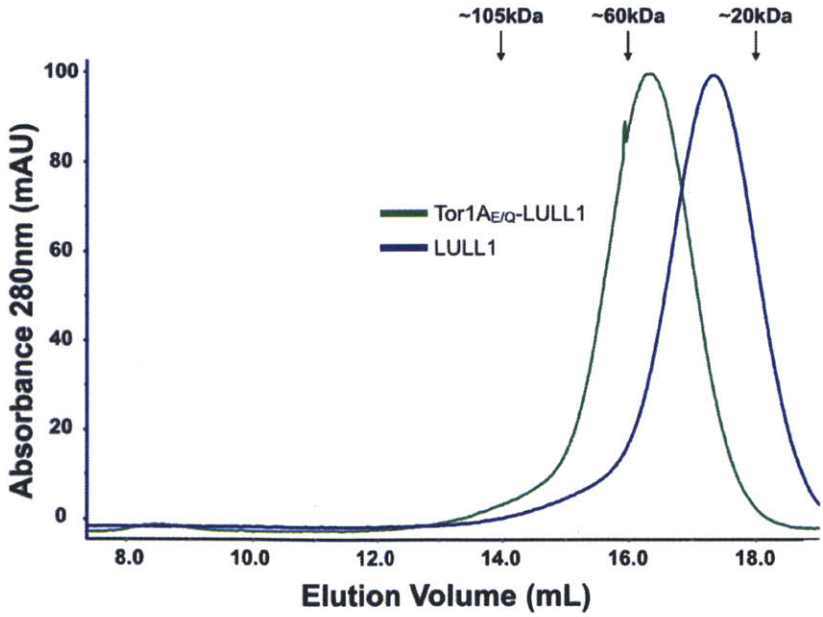


**Figure 6. Tor1A-LULL1 complex behaves as a dimer in solution**. Analytical size exclusion chromatography was performed on a Superdex 200 10/300 column. Tor1A-LULL1 complexes elute in a volume corresponding to a dimeric species whereas LULL1 behaves as a monomer in solution. Approximate calibration values are shown.

Tor1A, LAP1 and LULL1 all share structural homology to the Clp family of AAA+ ATPases (Table 1). Based on this we hypothesized that the interaction of Tor1A to LAP1 or LULL1 would result in either a single hetero-hexameric complex or into a double-stacked hexameric complex. To determine the oligomerization state, purified Tor1A-LULL1 complex was analyzed via size exclusion chromatography. The resulting chromatogram shows two chromatographically distinct species for the Tor1A-LULL1. A larger species accounts for ~10% of the sample and elutes at a volume that corresponds to a ~120 kDa molecule, whereas the remaining smaller species accounts elutes at a volume corresponding to a molecule of ~ 60 kDa. We observed that the larger species is not stable and dissociates into the smaller species over time (Figure 6). The more stable, smaller complex appears to be a 1:1 hetero-dimer. To corroborate the stoichiometry and molecular weight of the complex we performed analytical ultracentrifugation (AUC). Data obtained from sedimentation velocity analysis was fitted to a mass of 57.0 kDa, in good agreement with the predicted molecular weight of 59.4 kDa for a Tor1A-LULL1 1:1 complex (Figure 7).

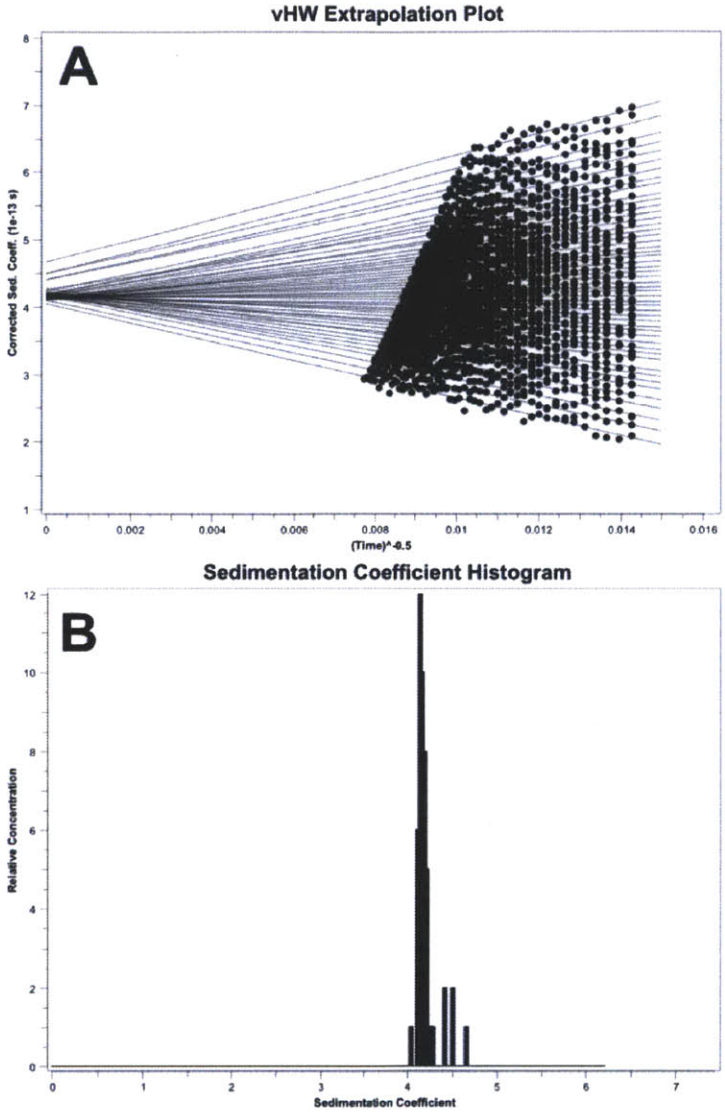


**Figure 7. Sedimentation velocity analysis of the Tor1A-LULL1 complex.** Enhanced van Holde-Weischet analysis of velocity data shows a sedimentation coefficient of 4.4754e-13 and an average molecular weight of 57.0 kDa in agreement with a 1:1 complex of Tor1A and LULL1.

## Structural Characterization of the Tor1A-LULL1 Complex

Determining a 1:1 stoichiometry of Tor1A to LULL1 (and LAP1) in combination with size exclusion chromatography does not allow us to find out whether the higher order assembly is a single- or double-stacked hexameric ring. There we set up a single particle electron microscopy (EM) experiment. Because of the instability of the oligomer, we sequentially setup EM grids for the different steps

of the purification. Samples obtained from the nickel elution, ion exchange chromatography and size exclusion chromatography were analyzed. Not surprisingly, the nickel elution sample, i.e. the first purification step, yielded the best particle density. Approximately 1000 particles were picked and averaged (Figure 8). The Tor1A-LULL1 complex forms a ring of ~120Å in diameter, similar to other AAA+ ATPases (Figure 8). Due to the preferred orientation of the ring on the grid, we lack a sufficient number of side views of the complex to unambiguously determine the number of stacked rings.

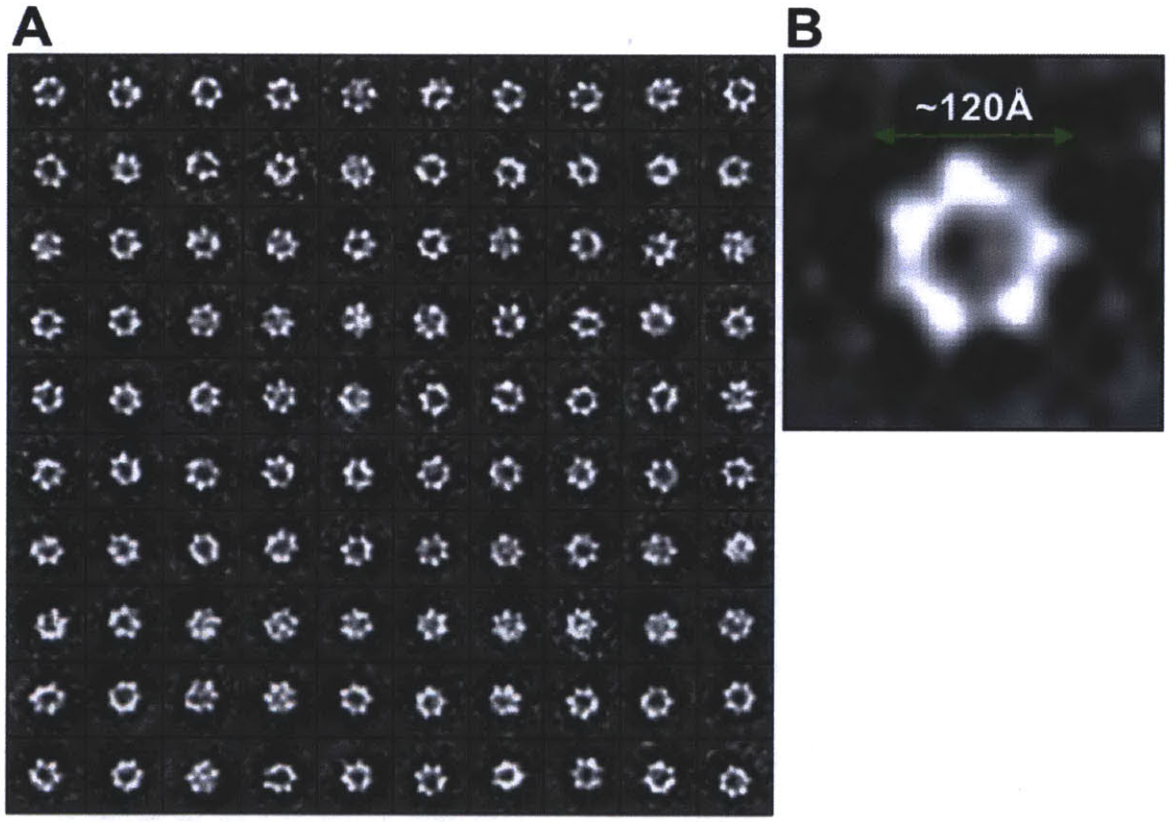


**Figure 8. Tor1A-LULL1 complexes form rings.** A dataset of 1,080 single particles was subjected to "direct classification" using PARTICLE (www.image-analysis.net/EM). Representative class images of the Tor1A-LULL1 complex (A) and a zoomed representative class image showing the approximate measured diameter of the rings.

## DISCUSSION

Despite its well characterized and critical role in the dystonia disease, a mechanism of Tor1A function has remained elusive (Granata et al., 2010). The lack of knowledge in part arises from the difficulty in producing the enzyme in high amount and purity. In this study we developed an approach to produce large amounts of chromatographically pure Tor1A, which will be critical in order to decipher, biochemically and structurally, how this ATPase works. Most likely, we will be able to extrapolate to the other Torsin-family members as well.

A key finding of this study is that the ER/NE luminal domains of LAP1 and LULL1 adopt an AAA+ fold, despite lacking the signature motifs, which renders them catalytically inactive. There are several examples of catalytically inactive AAA+ ATPases like the eukaryotic origin recognition complex (ORC) subunits Orc4 and Orc5, the hexamerization domains of p97 and N-ethylmaleimide-sensitive fusion protein (NSF), and the δ and δ′ subunits of the bacterial clamp loader (Brunger and DeLaBarre, 2003; Erzberger and Berger, 2006; Guenther et al., 1997a; 1997b; Jeruzalmi et al., 2001a; 2001b). These play important roles in regulating the activity of their complexes. An important difference between these and LAP1 and LULL1 is that Orc4, Orc5, p97, NSF and the clamp loader δ and δ′ subunits all contain somewhat degenerate Walker-A and Walker-B motifs. LAP1 and LULL1, however, lack Walker-A or –B motifs altogether. This suggests that LAP1 and LULL1 are not only inactive, but also lost the ability to bind nucleotide. It will be very interesting to investigate the evolutionary origin of LAP1 and LULL1 in order to understand whether they diverged from an AAA+ ATPase that lost its function but maintained its fold or converged from an unrelated protein that adopted this AAA+ fold.

The oligomerization state of Tor1A has been extensively debated in the field (Torres et al., 2004; Vander Heyden et al., 2009). As a member of the AAA+ ATPase superfamily that shares homology to the Clp family, it is widely assumed

that Tor1A should adopt a hexameric oligomerization state. Furthermore, several studies have suggested that Tor1A forms homo-dimers and homo-hexamers. These experiments consisted of overexpression of tagged Tor1A in human cells followed by *in vivo* crosslinking or blue native polyacrylamide gel electrophoresis (PAGE) and inmunoblotting. In the reported literature, the observed bands in those inmunoblots correspond to monomeric Tor1A as well as larger species, which might be the result of true oligomerization, or, alternatively, of binding to additional protein partner(s). In our experiments we did not observe self-oligomerization of purified Tor1A, LAP1 or LULL1 on their own. We show that a stable 1:1 heteromeric complex forms between Tor1A and LULL1 or LAP1, and that behaves as a dimer as shown by size exclusion chromatography and AUC. We can also show by negative stain EM that Tor1A-LULL1 can form hexameric rings, similar to other AAA+ ATPases. Additional data collection, ideally in thick vitreous ice to allow single-particle analysis, will resolve the important question of whether these rings are heteromeric single rings, or whether they are stacked. For simple steric reasons, a heteromeric rings is expected if the N-terminal hydrophobic extension of Tor1A indeed directly interacts with the membrane, as previous data suggests. In a double-ring model, the LAP1 or LULL1 ring has to be placed membrane-proximal, because of the transmembrane-segment directly preceding the AAA+ domain (Figure 9). The membrane-distal ring has to be composed of Tor1A. In that scenario, we would predict that the hydrophobic extension might not function in membrane-attachment. It is conceivable that the fragment instead is important for the stability of the double-ring by interacting with LAP1 or LULL1 through a largely hydrophobic interaction. It is also possible, that the hydrophobic extension in Tor1A simply interacts with the membrane when it is not bound to either LAP1 or LULL1.

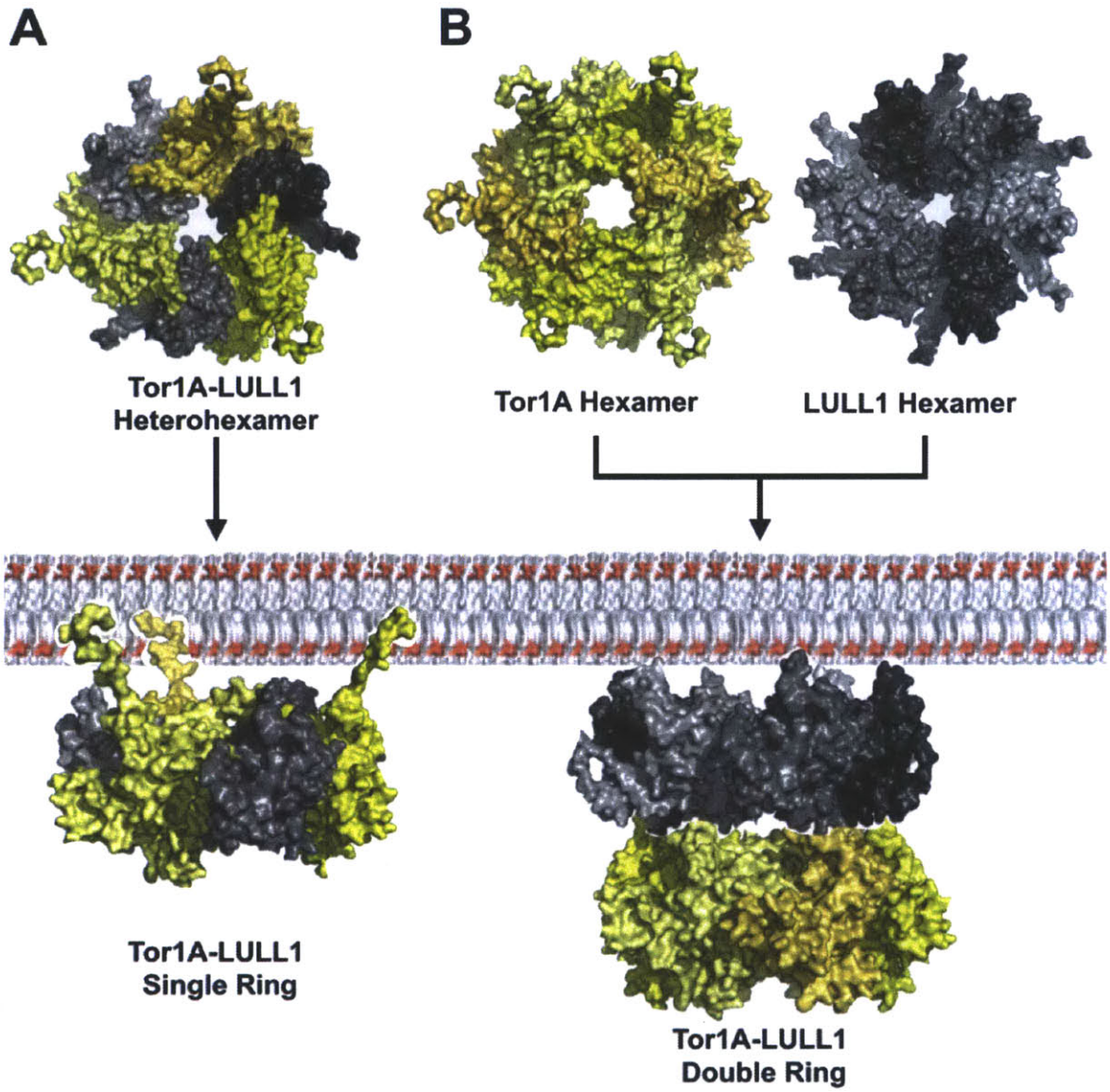


**Figure 9. Structural model of the Tor1A-LULL1 complex.** Based on our data, two different scenarios can be hypothesized for Tor1A-LULL complexes. Tor1A-LULL1 complexes arrange into heteroxeameric rings in which both proteins are capable to engage the membrane individually. Alternatively, based on the structural homology to Clp AAA+ ATPases, Tor1A can assemble into a single ring and LULL1 can assemble into another. This would lead into a double ring arrangement for the for the Tor1A-LULL1 complex.

The higher-order Tor1A-LULL1 complex that we observe in EM is not stable throughout the purification process. There are several possibilities to explain this behavior, if, indeed, the complex is stable *in vivo*, which may or may not be true. First, we have purified the complex recombinantly from E.coli. Since Tor1A and LULL1 are ER-resident proteins, one simple explanation is that reducing rather

than oxidizing conditions may prevent the formation of important disulfide bridges, either within or between the binding proteins. The strongly conserved cysteines raise a concern that purification from an oxidizing environment would be beneficial. Second, the hydrophobic region of Tor1A was absent from our constructs tested for complex formation. Third, Tor1A is glycosylated *in vivo*, which might stabilize the LULL1 interaction, but is absent in the recombinant protein. Fourth, the membrane itself could play a role in complex stability. Membrane-anchored LULL1 might oligomerize more readily due to avidity effects, imposed by the reduced entropic penalty of complex formation.

LAP1 and LULL1 share 62% identity in their luminal domains, while their cyto-/nucleoplasmic regions are not related. Intriguingly, they localize to different membranes: LAP1 is embedded in the INM while LULL1 is found throughout ER and NE (Goodchild and Dauer, 2005; Jungwirth et al., 2011; Naismith et al., 2009; Nery et al., 2008; Vander Heyden et al., 2009; 2011). The two proteins might be used to target Torsin ATPases to different membranes where they can act on different substrates. Alternatively, Torsins might engage different substrates depending on which binding partner is present. Further studies could help to answer these questions by determining structural differences between Tor1A-LAP1 and Tor1A-LULL1 complexes.

The role of Torsin ATPases still remains enigmatic. However, recent studies by several groups have shed light into Torsin function and mechanism. As previously mentioned, Tor1A has been implicated in the displacement of LINC complexes from the NE to the ER (Nery et al., 2008; Vander Heyden et al., 2009). The Schlieker group showed that Tor1A requires LAP1 or LULL1 for ATPase activity and that Tor1A is inactive by itself. Single-turnover experiments suggest that Tor1A-LAP1/LULL1 behaves as a "switch-like" ATPase (Zhao et al., 2013). In a study by the Budnik group Tor1A is implicated in non-NPC mediated transport across the NE via budding ('nuclear egress') suggesting that Tor1A might be involved in membrane remodeling akin to AAA+ ATPases like NSF, p97

and Vps4 (Jokhi et al., 2013). Our experiments that lead to the possibility to now recombinantly produce large quantities of Tor1A-LAP1 or Tor1A-LULL1 complexes enable structural studies, as well as rigorous *in vitro* analysis of the proposed enzymatic activities of Tor1A.

## EXPERIMENTAL PROCEDURES

### Bioinformatic Analysis

Multiple sequence alignments of Tor1A, LAP1 and LULL1 were performed using MUSCLE and were visualized using Jalview. The species nomenclature used for the alignments is as follows: *Homo sapiens* (hs), *Ornithorhynchus anatinus* (oa), *Gallus gallus* (gg), *Takifugu rubripes* (tr), *Danio rerio* (dr), *Branchiostoma floridae* (bf), *Strongylocentrotus purpuratus* (stp), *Ciona savignyi* (cs), *Ciona intestinalis* (ci), *Nematostella vectensis* (nv), *Caenorhabditis elegans* (ce), *Drosophila melanogaster* (dm), *Escherichia coli* (ec), *Thermus thermophilus* (tt) and *Bacillus subtilis* (bs).

Sequences of Tor1A and the ER luminal domains LAP1 and LULL1 were subjected to an HHpred search (http://toolkit.tuebingen.mpg.de/hhpred). The best models from the resulting alignments (3cfw_A, 1r6b_X and 3pxi_A) were used to create structural models using MODELLER (toolkit.tuebingen.mpg.de/modeller). Structural models were visualized using PyMOL.

### Plasmids, Protein Expression and Purification

Recombinant proteins were expressed in E. coli. $Tor1A_{51-332}$, $LAP1_{356-583}$ and $LULL1_{233-470}$ fragments were expressed from a modified ampicillin resistant pETDuet-1 (Novagen) vector as N-terminally 6×His-7xArg-tagged fusion proteins. A cleavage site for human rhinovirus 3C protease was inserted between the 7xArg tag and the Tor1A fragment. For coexpression of $Tor1A_{51-332}$ and $LAP1_{356-583}$ or $LULL1_{233-470}$, cells were co-transformed with the Tor1A vector described above and a second, modified kanamycin resistant pETDuet-1 vector containing untagged LAP1 or LULL1. Mutations were introduced by site-directed mutagenesis.

$Tor1A_{51-332}$, $LAP1_{356-583}$ and $LULL1_{233-470}$ fragments were expressed in E. coli BL21(DE3) RIL strains (Stratagene). The bacterial cultures were grown at 30°C to an optical density (OD) of 0.6. Then, the culture was shifted to 18°C for 30 min and induced overnight with 0.2 mM IPTG. $LAP1_{356-583}$ and $LULL1_{233-470}$ expressing cells were resuspended in lysis buffer A (50 mM potassium phosphate pH 8.0, 400 mM NaCl, 40 mM imidazole) and lysed. The lysate was supplemented with 1U/ml Benzonase (Sigma) 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 potassium phosphate pH 8.0, 150 mM NaCl, 250 mM imidazole). The eluted protein was purified by cation-exchange chromatography against a gradient of 0.150 - 2 M NaCl with10 mM potassium phosphate pH 8.0 followed by dialysis to low salt buffer (10 mM potassium phosphate pH 8.0, 150 mM NaCl) and tag removal with 3C protease followed by another round of cation-exchange chromatography to remove tag and protease. The flow-through from the cation-exchange chromatography was concentrated and purified via size exclusion chromatography on a Superdex S200 column (GE Healthcare) equilibrated in buffer (10 mM Tris/HCl pH 7.4, 150 mM NaCl). Individually expressed $Tor1A_{51-332}$ and complexes followed the same procedure with the following differences: (1) lysis buffer contained 5 mM $MgCl_2$, 1 mM ATP and Tris/HCl pH 7.4 instead of potassium phosphate, (2) elution buffer contained 5 mM $MgCl_2$, 1 mM ATP 10 mM Hepes/NaOH pH 8.0 instead of potassium phosphate, (3) ion-exchange buffer contained 5 mM $MgCl_2$, 0.5 mM ATP 10 mM Hepes/NaOH pH 8.0 and (4) size exclusion buffer contained 5 mM $MgCl_2$ and 0.5 mM ATP.

**Analytical Centrifugation**

$Tor1A_{51-332-E/Q}$-$LULL1_{233-470}$ complex was gel-filtered in 10 mM Tris/HCl pH 8.0, 150 mM NaCl immediately prior to the experiments. Analytical ultracentrifugation experiments were carried out using an An60Ti rotor in an Optima XL-A centrifuge. Samples were adjusted to an optical density of 0.6 and loaded into

Epon-charcoal two-channel centerpieces, fit with quartz windows, and spun at 42,000rpm. Absorbance data was collected at 280 nm at 5 replicates per 1 nm step. Sedimentation velocity data was fitted using the Ultrascan II software (http://ultrascan.uthscsa.edu).

**Electron Microscopy**

Tor1A$_{51\text{-}332\text{-}E/Q}$-LULL1$_{233\text{-}470}$ complex, purified via Ni-affinity chromatography, was negatively stained with uranyl acetate (1% v/v) on continuous carbon-film grids. Electron micrographs of single particles were recorded on a Tecnai Spirit electron microscope (Whitehead Institute) using an 80KeV beam. A total of 1,080 Tor1A$_{51\text{-}332\text{-}E/Q}$-LULL1$_{233\text{-}470}$ particles were then boxed into a single stack, from which the particle images were 2×-binned to 3.63 Å per pixel. The dataset was subjected to "direct classification" using PARTICLE (www.image-analysis.net/EM).

## REFERENCES


Basham, S.E., and Rose, L.S. (2001). The Caenorhabditis elegans polarity gene ooc-5 encodes a Torsin-related protein of the AAA ATPase superfamily. Development *128*, 4645–4656.

Basham, S.E., and Rose, L.S. (1999). Mutations in ooc-5 and ooc-3 Disrupt Oocyte Formation and the Reestablishment of Asymmetric PAR Protein Localization in Two-Cell Caenorhabditis elegans Embryos. Dev Biol *215*, 253–263.

Biter, A.B., Lee, S., Sung, N., and Tsai, F.T.F. (2012). Structural basis for intersubunit signaling in a protein disaggregating machine. Proc Natl Acad Sci USA *109*, 12515–12520.

Brunger, A.T., and DeLaBarre, B. (2003). NSF and p97/VCP: similar at first, different at last. 126th Nobel Symposium. Membrane Proteins: Structure, Function and Assembly *555*, 126–133.

Erzberger, J.P., and Berger, J.M. (2006). Evolutionary relationships and structural mechanisms of AAA+ proteins. Annu Rev Biophys Biomol Struct *35*, 93–114.

Glynn, S.E., Glynn, S.E., Martin, A., Martin, A., Nager, A.R., Nager, A.R., Baker, T.A., Baker, T.A., Sauer, R.T., and Sauer, R.T. (2009). Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. Cell *139*, 744–756.

Goodchild, R.E., and Dauer, W.T. (2005). The AAA+ protein torsinA interacts with a conserved domain present in LAP1 and a novel ER protein. J Cell Biol *168*, 855–862.

Granata, A., Granata, A., Warner, T.T., and Warner, T.T. (2010). The role of torsinA in dystonia. European Journal of Neurology *17*, 81–87.

Guenther, B., Guenther, B., Onrust, R., Onrust, R., Sali, A., Sali, A., O'Donnell, M., O'Donnell, M., Kuriyan, J., and Kuriyan, J. (1997a). Crystal structure of the delta' subunit of the clamp-loader complex of E. coli DNA polymerase III. Cell *91*, 335–345.

Guenther, B., Guenther, B., Onrust, R., Onrust, R., Sali, A., Sali, A., O'Donnell, M., O'Donnell, M., Kuriyan, J., and Kuriyan, J. (1997b). Crystal Structure of the δ′ Subunit of the Clamp-Loader Complex of E. coli DNA Polymerase III. Current Biology *91*, 335–345.

Hewett, J. (2000). Mutant torsinA, responsible for early-onset torsion dystonia, forms membrane inclusions in cultured neural cells. Hum Mol Genet *9*, 1403–1413.

Jeruzalmi, D., O'Donnell, M., and Kuriyan, J. (2001a). Crystal structure of the processivity clamp loader gamma (gamma) complex of E. coli DNA polymerase III. Cell *106*, 429–441.

Jeruzalmi, D., Yurieva, O., Zhao, Y., Young, M., Stewart, J., Hingorani, M., O'Donnell, M., and Kuriyan, J. (2001b). Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of E. coli DNA polymerase III. Cell *106*, 417–428.

Jokhi, V., Ashley, J., Nunnari, J., Noma, A., Ito, N., Wakabayashi-Ito, N., Moore, M.J., and Budnik, V. (2013). Torsin mediates primary envelopment of large ribonucleoprotein granules at the nuclear envelope. Cell Reports *3*, 988–995.

Jungwirth, M.T., Kumar, D., Jeong, D.Y., and Goodchild, R.E. (2011). The nuclear envelope localization of DYT1 dystonia torsinA-deltaE requires the SUN1 LINC complex component. BMC Cell Biol *12*, 24.

Jungwirth, M., Dear, M.L., Brown, P., Holbrook, K., and Goodchild, R. (2010). Relative tissue expression of homologous torsinB correlates with the neuronal specific importance of DYT1 dystonia-associated torsinA. Hum Mol Genet *19*,

888–900.

Lee, S., Lee, S., Sowa, M.E., Sowa, M.E., Watanabe, Y.-H., Watanabe, Y.-H., Sigler, P.B., Sigler, P.B., Chiu, W., Chiu, W., et al. (2003). The Structure of ClpB. Cell *115*, 229–240.

Naismith, T.V., Dalal, S., and Hanson, P.I. (2009). Interaction of TorsinA with Its Major Binding Partners Is Impaired by the Dystonia-associated GAG Deletion. Journal of Biological Chemistry *284*, 27866–27874.

Narlikar, G.J., Sundaramoorthy, R., and Owen-Hughes, T. (2013). Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. Cell *154*, 490–503.

Nery, F.C., Nery, F.C., Zeng, J., Zeng, J., Niland, B.P., Niland, B.P., Hewett, J., Hewett, J., Farley, J., Farley, J., et al. (2008). TorsinA binds the KASH domain of nesprins and participates in linkage between nuclear envelope and cytoskeleton. J Cell Sci *121*, 3476–3486.

Ozelius, L.J., Hewett, J., Kramer, P., Bressman, S.B., Shalish, C., de Leon, D., Rutter, M., Risch, N., Brin, M.F., Markova, E.D., et al. (1997a). Fine localization of the torsion dystonia gene (DYT1) on human chromosome 9q34: YAC map and linkage disequilibrium. Genome Res. *7*, 483–494.

Ozelius, L.J., Hewett, J.W., Page, C.E., Bressman, S.B., Kramer, P.L., Shalish, C., de Leon, D., Brin, M.F., Raymond, D., Corey, D.P., et al. (1997b). The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein. Nat Genet *17*, 40–48.

Schlee, S., Schlee, S., Groemping, Y., Groemping, Y., Herde, P., Herde, P., Seidel, R., Seidel, R., Reinstein, J., and Reinstein, J. (2001). The chaperone function of ClpB from Thermus thermophilus depends on allosteric interactions of its two ATP-binding sites. Journal of Molecular Biology *306*, 889–899.

Soding, J. (2005). Protein homology detection by HMM-HMM comparison.

Bioinformatics *21*, 951–960.

Soding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Research *33*, W244–W248.

Soding, J., Remmert, M., Biegert, A., and Lupas, A.N. (2006). HHsenser: exhaustive transitive profile search using HMM-HMM comparison. Nucleic Acids Research *34*, W374–W378.

Sosa, B.A., Rothballer, A., Kutay, U., and Schwartz, T.U. (2012). LINC Complexes Form by Binding of Three KASH Peptides to Domain Interfaces of Trimeric SUN Proteins. Cell *149*, 1035–1047.

Starr, D.A., Starr, D.A., Fridolfsson, H.N., and Fridolfsson, H.N. (2010). Interactions between nuclei and the cytoskeleton are mediated by SUN-KASH nuclear-envelope bridges. Annu. Rev. Cell Dev. Biol. *26*, 421–444.

Torres, G.E., Sweeney, A.L., Beaulieu, J.M., Shashidharan, P., and Caron, M.G. (2004). Effect of torsinA on membrane proteins reveals a loss of function and a dominant-negative phenotype of the dystonia-associated E-torsinA mutant. Proc Natl Acad Sci USA *101*, 15650–15655.

Vander Heyden, A.B., Naismith, T.V., Snapp, E.L., and Hanson, P.I. (2011). Static retention of the lumenal monotopic membrane protein torsinA in the endoplasmic reticulum. Embo J. *30*, 3217–3231.

Vander Heyden, A.B., Vander Heyden, A.B., Naismith, T.V., Naismith, T.V., Snapp, E.L., Snapp, E.L., Hodzic, D., Hodzic, D., Hanson, P.I., and Hanson, P.I. (2009). LULL1 retargets TorsinA to the nuclear envelope revealing an activity that is impaired by the DYT1 dystonia mutation. Mol Biol Cell *20*, 2661–2672.

Warner, T.T., Granata, A., and Schiavo, G. (2010). TorsinA and DYT1 dystonia: a synaptopathy? Biochem. Soc. Trans. *38*, 452.

Wendler, P., Wendler, P., Ciniawsky, S., Ciniawsky, S., Kock, M., Kock, M., Kube, S., and Kube, S. (2012). Structure and function of the AAA+ nucleotide binding pocket. Biochim Biophys Acta *1823*, 2–14.

Xia, D., Xia, D., Esser, L., Esser, L., Singh, S.K., Singh, S.K., Guo, F., Guo, F., Maurizi, M.R., and Maurizi, M.R. (2004). Crystallographic investigation of peptide binding sites in the N-domain of the ClpA chaperone. J. Struct. Biol. *146*, 166–179.

Zhao, C., Brown, R.S.H., Chase, A.R., Eisele, M.R., and Schlieker, C. (2013). Regulation of Torsin ATPases by LAP1 and LULL1. Proc Natl Acad Sci USA *110*, E1545–E1554.

Zhu, L., Millen, L., Mendoza, J.L., and Thomas, P.J. (2010). A unique redox-sensing sensor II motif in TorsinA plays a critical role in nucleotide and partner binding. Journal of Biological Chemistry *285*, 37271–37280.

Zhu, L., Zhu, L., Wrabl, J.O., Wrabl, J.O., Hayashi, A.P., Hayashi, A.P., Rose, L.S., Rose, L.S., Thomas, P.J., and Thomas, P.J. (2008). The torsin-family AAA+ protein OOC-5 contains a critical disulfide adjacent to Sensor-II that couples redox state to nucleotide binding. Mol Biol Cell *19*, 3599–3612.