Structural Insights into LINC Complexes

Brian A Sosa, Ulrike Kutay, and Thomas U Schwartz

1Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, USA 2Institute for Biochemistry, ETH Zurich, CH-8093 Zurich, Switzerland

Abstract

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 reported in the past year.

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 [1,2], 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 [3,4].

Mechanical communication has long been recognized for cells interacting with their surrounding [5]. That the nucleus is also mechanically tethered to its environment has only been uncovered more recently [6]. 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 [7]. 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 [8,9]. Third, these nucleocytoplasmic linkages can also be used to determine the position of specific nuclear structures, like the ends of paired chromosomes during meiosis [10,11]. 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 [12–14].

Physical interactions across the nuclear envelope are mediated via LInkers of Nucleoskeleton and Cytoskeleton (LINC) complexes [15,16]. The center of LINC complexes is the interaction of INM-resident SUN (Sad1 and UNC-84) proteins with ONM-resident KASH (Klarsicht, ANC-1 and SYNE/Nesprin-1 and -2 Homology) 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 [17].

**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 Sad1 from *Schizosaccharomyces pombe* and UNC-84 from *Caenorhabditis elegans* [18]. With increasing complexity of the organism, the number of SUN proteins also increases. While 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 [16]. 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 [19]. 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 [20,21]. Similarly, the two major mammalian SUN proteins, Sun1 and Sun2, are widely expressed in different cell types [15,22]. In contrast, the three additional genes coding for Sun3, Sun4, and Sun5, respectively, all show a much more restricted, testis-specific expression pattern [23]. In addition to the SUN domains at the C terminus, there are also SUN-like proteins [24,25] 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 [26,27], 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**

KASH proteins are so far 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) [28]. The most striking feature of the luminal KASH peptide is the terminal ‘PPPX’ motif, where X is always the very terminal residue.

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 [29] KASH5, which binds SUN1, was reported recently to be a germ cell-specific protein involved in meiotic homologue pairing [30]. The ONM-resident, lymphoid-restricted membrane protein (LRMP) is the latest characterized putative KASH proteins [31]. 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 quite satisfactorily [17]. 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.
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 [17]. 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 [17]. 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 [17,32,33], 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 [34]. 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 [35]. 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 [23,36].

**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 “TAN lines” parallel to the actin fibers to which they are presumably attached [37–39]. 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 [11,40–42], 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 [17]. 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 [43], while KASH proteins are tail anchored and might be integrated into the ER via the action of the GET complex [44]. 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.
It is equally important to understand how LINC complexes disassemble and whether they can be remodeled or must be degraded to disassemble? Since the nucleus breaks down during mitosis in multicellular organisms, LINC complexes need to be disassembled at least once each cell cycle. TorsinA, a poorly characterized ER-resident AAA+ ATPase, is an attractive candidate for catalyzing the LINC complex disassembly [45]. Indirect evidence suggests that TorsinA destabilizes selected LINC complexes at the NE, after being recruited to the NE by other proteins.

Conclusions

Recent structural characterization of the core of the LINC complex, the SUN-KASH interaction, has provided researchers interested in mechanotransduction across the NE with a rich platform for analyzing this process in molecular detail. Since LINC-complexes are involved in broadly divergent functions, this will be an exciting field of study for years to come. Given the speed of recent progress, we can expect answers to mechanistic questions (assembly/disassembly/targeting) in the foreseeable future. In addition to uncovering fundame biology, progress in the field will hopefully also lead to therapeutic strategies that for treating the devastating diseases associated with LINC complex malfunctions, and NE diseases in general. An encouraging development in this regard is the recent finding that Sun1 overaccumulates in murine models for Emery-Dreifuss Muscular Dystrophy and in Hutchinson-Gilford Progeria Syndrome and that reducing Sun1 levels markedly improves these diseases [46].

Acknowledgments

We would like to apologize to those colleagues whose work could not be cited due to space limitations. Research in the Schwartz lab was supported by grant R21 NS075883 from the National Institute of Neurological Disorders and Stroke and by a grant from the Dystonia Medical Research Foundation.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest


26**. Tapley EC, Ly N, Starr DA. Multiple mechanisms actively target the SUN protein UNC-84 to the inner nuclear membrane. Mol Biol Cell. 2011; 22:1739–1752. Together with [27] these papers are the first to untangle the intricate targeting mechanism of SUN proteins to the INM. [PubMed: 21411627]


36. Kracklauer, MP.; Link, J.; Alsheimer, M. LINCing the Nuclear Envelope to Gametogenesis. Elsevier Inc; 2013.


Highlights

- SUN- and KASH-proteins form mechanical connections across the nuclear envelope.
- Crystal structures reveal that these LINC complexes are built from SUN trimers interacting with three KASH peptides.
- LINC complexes might define the spacing of inner and outer nuclear membrane.
- LINC complexes are involved in a growing number of diverse biological functions.
Figure 1.
Multiple sequence alignment of KASH peptides from highly diverged eukaryotes, including vertebrates, various metazoa, yeast, and plants. The numbers above the sequence mark residues in human Nesprin2/Syne2 important for SUN binding according to [17].
Figure 2.
A) Overview of a human SUN2 protomer isolated from its Nesprin-2 binding partners in the trimeric SUN-KASH complex [17]. 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.
Figure 3.
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.