SURVEY AND SUMMARY

Nuclear speckles: molecular organization, biological function and role in disease

Lukasz Galganski, Martyna O. Urbanek and Wlodzimierz J. Krzyzosiak*

Department of Molecular Biomedicine, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland

Received May 18, 2017; Revised August 16, 2017; Editorial Decision August 17, 2017; Accepted August 18, 2017

ABSTRACT

The nucleoplasm is not homogenous; it consists of many types of nuclear bodies, also known as nuclear domains or nuclear subcompartments. These self-organizing structures gather machinery involved in various nuclear activities. Nuclear speckles (NSs) or splicing speckles, also called interchromatin granule clusters, were discovered as sites for splicing factor storage and modification. Further studies on transcription and mRNA maturation and export revealed a more general role for splicing speckles in RNA metabolism. Here, we discuss the functional implications of the localization of numerous proteins crucial for epigenetic regulation, chromatin organization, DNA repair and RNA modification to nuclear speckles. We highlight recent advances suggesting that NSs facilitate integrated regulation of gene expression. In addition, we consider the influence of abundant regulatory and signaling proteins, i.e. protein kinases and proteins involved in protein ubiquitination, phosphoinositide signaling and nucleoskeletal organization, on pre-mRNA synthesis and maturation. While many of these regulatory proteins act within NSs, direct evidence for mRNA metabolism events occurring in NSs is still lacking. NSs contribute to numerous human diseases, including cancers and viral infections. In addition, recent data have demonstrated close relationships between these structures and the development of neurological disorders.

INTRODUCTION

Less than 1.5% of the human genome consists of protein-coding sequences and the number of protein-coding genes is similar across most higher eukaryotes. Notably, differences in developmental programs arise from numerous gene expression regulatory mechanisms, which allow different cell types to respond adequately to specific environmental conditions. These mechanisms depend on the flexibility and dynamics of molecular interactions, which can be promoted or prevented by spatial organization in the nucleus (1–3). The enhancement of necessary interactions and reduction of undesired interactions are facilitated by reversible separation of specific molecules within a spatially restricted area. Indeed, numerous macromolecules from interchromatin regions of the nucleoplasm, predominantly proteins and RNAs, are gathered within nuclear subcompartments (or nuclear bodies), e.g. nucleoli, NSs, paraspeckles, Cajal (coiled) bodies, gemini of Cajal bodies (gems) and promye locytic leukemia (PML) bodies. An increasing number of additional nuclear domains have been described, including nuclear stress bodies, histone locus bodies, polycomb bodies, DNA damage foci, cleavage bodies, matrix-associated deacetylase bodies and clastosomes (1–6). As a consequence, in addition to the exchange of molecules between the nucleus and cytoplasm, the tightly controlled distribution and movement of factors within the nucleus is an important level of regulation in many nuclear pathways, including RNA maturation.

Because alternative pre-mRNA splicing greatly increases transcriptome diversity in higher eukaryotes, nuclear bodies involved in splicing regulation are key gene expression regulators. These bodies include NSs, which are also known as splicing speckles, B snurposomes, splicing factor compartments, SC-35 domains and interchromatin granule clusters. The first observations of stained NSs using light microscopy were made by Santiago Ramon y Cajal in 1910. Electron microscopy (EM) observations and RNA identification in NSs were made by Hewson Swift in 1959. Two years later, J. Swanson Beck used the term ‘speckles’ for the first time to describe these bodies (4). The earliest identification of splicing factors and small nuclear ribonucleoproteins (snRNPs)
in NSs uncovered connections between NSs and splicing (7–9). NSs were thought to play a role predominantly in regulating the availability of splicing factors at transcription sites because alteration of their function or composition led to changes in alternative pre-mRNA splicing. However, as research on NSs progressed, additional NS functions have been revealed and will be discussed in this review.

More recent studies have demonstrated that proteins involved in chromosome localization, chromatin modification, transcription, splicing, 3'end processing, mRNA modification, mRNA coating with proteins and messenger ribonucleoprotein (mRNP) export are assembled in NSs, supporting the hypothesis that NSs act as a hub to coordinate all of the nuclear gene expression regulation steps. Importantly, all of these steps are coupled with RNA polymerase II transcription, which occurs within perichromatin fibrils in close proximity to NSs (10). Despite many studies aimed at functionally characterizing NS proteins, the precise role of NSs requires further clarification. This need for additional studies applies to extensively explored processes, such as splicing, because in addition to the conventional view that NSs function in the assembly, modification, temporary storage and recycling of splicing factors, several reports have shown splicing activity within NSs (11,12). Moreover, the majority of NS proteins can also be found at other nuclear locations and their specific roles in NSs, interacting partners and post-translational modifications need to be elucidated.

In this review, we describe the involvement of NS proteins in various nuclear gene expression regulation pathways. We also review recent insights into the role of regulatory proteins, which are enriched in NSs; these proteins include protein kinases, cytoskeletal elements, factors involved in ubiquitination, SUMOylation and phosphoinositide (PI) signaling. Finally, we discuss the connection between NSs and human disease, with an emphasis on neurological disorders and the defects in RNA synthesis and metabolism that contribute to these disorders (13).

**MOLECULAR ORGANIZATION**

**General characteristics**

The human interphase nucleus contains 20–50 NSs measuring up to several micrometers in diameter. EM has revealed that a single NS is composed of spots (interchromatin granules) measuring 20–25 nm in diameter connected by fine fibrils to form a cluster (14). Their size or irregular shape can alter dynamically, may vary between different cell types and depend on numerous factors, including cellular ATP levels, the phosphorylation status of various proteins, transcription of stress-activated genes, SWI/SNF chromatin remodeling and RNA polymerase II transcription and splicing (15,16). Inhibition of RNA polymerase II transcription or splicing leads to accumulation of proteins in enlarged NSs, whose normal size can be restored upon elimination of the block (17).

**Forces that maintain NSs**

The biophysical properties of NSs and the nucleoplasm do not differ substantially. NSs are slightly denser than the surrounding nucleoplasm, and the protein concentrations (ranging from 115 to 162 mg/ml) are similar in both compartments (18) or are slightly higher in NSs (19). Although NSs are highly dynamic structures and their components are constantly in flux (20), NSs remain clearly separated from the nucleoplasm and when isolated from the nuclei of mouse liver cells by Spector and colleagues, they remained stable and resistant to subsequent steps in the purification procedure (21,22).

It is widely accepted that NS assembly and maintenance depend on interactions among NS components. In addition to structured protein domains, very flexible low-complexity regions (LCRs) play an important role in protein–protein and protein–RNA interactions and are over-represented in NS proteins (Supplementary Table S1). LCRs can change the properties of a protein permanently or transiently after post-translational modification or upon protein partner binding. Thus, proteins containing LCRs are regulated and ensure that cellular processes can be adjusted. Depletion or mutation of LCRs alters protein–protein interactions, protein–RNA interactions, protein functions and localization to NSs (23).

Multiple LCRs, e.g. serine/arginine-rich (SR) motifs and folded domains coexisting within individual NS proteins enable interactions with multiple proteins at the same time. Transient and frequent low-specificity interactions can explain why NS proteins move 100-fold more slowly than GFP in the nucleoplasm (20). Thus, separation of proteins in NSs allows escape from unfavorable interactions and ensures the integrity of NSs despite the lack of a lipid boundary.

Another source of the spatial distinction of NSs is cellular crowding; high macromolecule concentrations favor phase separation over the homogeneity of solutions with low macromolecule concentrations (24). Accordingly, liquid droplets with clear boundaries are spontaneously generated in aqueous solutions in vivo and in vitro when concentrated proteins with LCRs are incubated in them. Such phase separation can be affected by temperature, pH, ionic strength and LCR modification (24,25).

Post-translational protein modifications (phosphorylation, methylation and conjugation with PIs (26) or SUMO-1 (27)) within LCRs have been demonstrated to not only regulate protein–protein interactions and recruitment to NSs (28–30) but also to increase structural order (31). Disorder-promoting stretches of repetitive amino acids, such as five or more consecutive histidine residues (32,33), constitute NS-targeting signals. Thus, the level of disorder appears to be the main force regulating NS composition.

**Nuclear speckles and the cell cycle**

NSs are very stable during interphase. However, the self-organizing properties of their components must be specifically suppressed during cell division. Disruption of the nuclear envelope following mitosis initiation leads to NS disassembly (Figure 1) and cytoplasmic dispersion of NS proteins (34). In addition to diffuse patterns, NS proteins in the cytoplasm assemble into an increasing number of mitotic interchromatin granules (MIGs), which are observed from metaphase to telophase (34–36). Because pre-mRNA splicing factors were found to be in an active state immediately
Figure 1. Schematic representation of the dynamics of NS components during the cell cycle. The diverse forms of NS protein assemblies are indicated with arrows: black represents enlarged NSs (possibly caused by transcriptional inhibition), blue represents a diffuse pattern, red represents MIGs and green represents nucleolar organizing region-associated patches (NAP). Importantly, different sets of proteins exhibit diverse patterns of cell cycle-regulated localization and diverse timing of their entry into and their presence in assemblies of NS proteins. Different assemblies of NS proteins can coexist and those in the minority are indicated with dotted lines. Note that NSs and MIG, in contrast to NAP, contain poly(A)^+ RNA.

after entry into the nucleus (36), it has been postulated that MIGs, being structurally similar to NSs (35), are required for the modification, assembly and delivery of pre-mRNA processing complexes to transcription sites in daughter nuclei.

After the reconstitution of the nuclear envelope, most pre-mRNA splicing factors gradually relocalize from MIGs to the nucleus within 10 min, but some of them (e.g. SRSF2) can stay in MIGs until G1 (36). During telophase, before formation of NSs, serine/arginine-rich splicing factors (SRSFs) in daughter nuclei gather transiently (for 15–20 min) in the proximity of active nucleolar organizing regions (NORs), in so-called NOR-associated patches (NAPs). SRSFs colocalize with CLK1 kinase (37), which is responsible for SRSFs leaving NSs (29). The assembly of SRSFs into NAPs and MIGs confirms their strong self-organizing properties, but SRSFs alone are not sufficient to form NSs. Hence, additional factors are required to nucleate NSs.

Because the establishment of transcription in telophase precedes NS formation (38), it has been hypothesized that the recruitment of splicing and processing factors to nascent transcripts triggers the spatial enrichment of NS proteins, followed by NS nucleation in the vicinity of active transcription sites. Accordingly, the ability of nascent pre-mRNA to trigger NS nucleation de novo has been shown in vivo (39).

The gathering of NS proteins can be explained by the self-organization model, which is based on protein–protein and protein–RNA interactions. However, induction of NS disintegration during prophase requires additional unknown factors. Cyclins seem to be potential NS cell cycle regulators because, to date, cyclin L1 is the only immobile protein in interphase NSs (40). Overall, the mechanisms that orchestrate cell cycle-dependent construction and deconstruction of NSs are still enigmatic and require further investigation.

Protein composition

Similar to other membrane-less bodies with liquid-like properties, NSs are characterized by the dynamic exchange of components within the nucleoplasm (24,25) and they share some proteins with other nuclear bodies. For instance, spliceosomal snRNPs are assembled in Cajal bodies before their relocation to NSs (2) and regulate the maturation of 3′ ends of histone transcripts within histone locus bodies (U2 snRNP) (41). NSs contain the paraspeckle proteins PSF (42) and PSP2, and both bodies often localize in close proximity (43). Similarly, Pat1b is found in both NSs and PML nuclear bodies (44), which can merge (45). However, the
functional consequences of the close relationships between distinct nuclear bodies remain to be revealed.

The number of proteins found in NSs using proteomic approaches (21,22) has greatly added to the numerous proteins found in NSs based on microscopic studies. As listed in Table 1 and Supplementary Table S1, NS proteins are involved in multiple nuclear gene expression regulation steps, such as epigenetic regulation, transcription activator and repressor functions, transcription elongation and termination, splicing, 3’ end processing, mRNA modification and mRNA packing and export (Figure 2). Proteins involved in gene expression regulation will be described in the ‘Biological Function’ section. The localization, interactions and degradation of these proteins is regulated by NS regulatory proteins, which include protein kinases and proteins involved in PI signaling, cytoskeletal organization and ubiquitination and are described in more detail below.

Protein kinases. Post-translational protein modifications that can alter protein properties are among the most important NS functions. Based on their protein composition, NSs seem to be nuclear hubs for protein phosphorylation, methylation, acetylation, ubiquitination and SUMOylation. As more than 30% of human proteins can be phosphorylated and kinase-encoding genes constitute ~2.4% of protein-coding human genes, protein phosphorylation is a fundamental regulatory mechanism.

In total, 31 protein kinases have been found in NSs (Supplementary Table S1) and the substrates of many NS protein kinases have been identified (46). Because many NS proteins participate in various gene expression steps that occur in different cellular compartments, the control of shuttling of these proteins has emerged as an important instrument of gene expression regulation. Reversible protein phosphorylation plays an important role in proper cellular localization of NS proteins, including SRSFs (47). SRSF1 is targeted to the nucleus (and further to NSs) upon SRPK1-driven phosphorylation (48). The resultant hypophosphorylated RS domain of SRSF1 prevents SNRNP70 binding and the subsequent formation of the spliceosomal E complex (49). In contrast, SRSF1 hyperphosphorylation by CLK1 promotes spliceosome assembly and releases SRSF1 from NSs (50). SRSFs are dephosphorylated during splicing and then serve as adaptors for mRNA export (51) or translation initiation (52). In summary, SRSF phosphorylation/dephosphorylation serves as a critical mechanism affecting both splicing and protein assembly in NSs (30). Accordingly, enhanced activity or over-expression of many NS kinases induces NS disassembly, indicating a role for kinases in NS maintenance.

Numerous kinases can phosphorylate multiple sites within a single NS protein and conversely, individual NS protein kinases can phosphorylate multiple substrates. In addition to splicing- or transcription-specific kinases (e.g. NLK), the action of many NS kinases is not restricted to a single-specific process. Thus, nuclear processes are mutually regulated. NS kinases also ensure communication with the extranuclear and extracellular environment by acting as modules of cellular signaling cascades. Therefore, NS kinases are mediators that integrate and modulate various signals via nuclear gene expression regulation events to adjust cellular homeostasis.

PI signaling. Derivatives of phosphoinositol act as another major group of signaling molecules that are relevant to NS function. In the cell nucleus, PIs are involved in regulation of gene expression, including chromatin modification, pre-mRNA maturation and mRNP export (26). Numerous NS proteins are affected by PIs directly (upon PI binding) or indirectly by the activity of PI-dependent regulatory proteins (e.g. protein kinases or ubiquitin ligases).

Growing evidence has indicated that PIs are localized in NSs (53). Moreover, PI-modifying enzymes, which add or remove phosphate from PIs (kinases and phosphatases) or hydrolyze them (phospholipases C), interact with each other in NSs (54), suggesting that some PI derivatives are produced in NSs. To date, more than twenty proteins implicated in PI signaling have been found in NSs. Among them, speckle-targeted PIP5K1a-regulated poly(A) polymerase has important roles in PI-dependent pre-mRNA 3’ end maturation and in regulation of the expression of specific mRNAs. The complexity of PI signaling may result from regulation of individual proteins by various PIs. For instance, binding of PI(4,5)P2 or phosphatidylinositol (PI) nuclear receptor steroidogenic factor 1 (NR5A1) modifies NR5A1-dependent gene expression (27,55).

PIs may have a substantial impact on NS function due to the multifunctionality of their downstream signaling proteins, which include several prominent NS protein kinases. One of them is AKT1, which enhances SRPK2 kinase activity toward SRSF2 and ACIN1. Another substrate of AKT1 is ALY/REF. PIs target ALY/REF to NSs independently of AKT1-driven phosphorylation but can also affect ALY/REF-regulated mRNA nuclear export in cooperation with AKT1 (26).

Moreover, several NS proteins, e.g. PARD6A/TIP-40 and PDLIM7/ENIGMA, contain PIP2-binding PDZ domains, which mediate protein–protein interactions. Disturbances in the PIP2 level or in PIP2-PDZ domain interactions abolish syntenin-2 and ZO-2 guanylate kinase targeting to NSs (56,57). Importantly, impaired ZO-2 (56) or syntenin-2 expression (57) promotes the formation of more dispersed and smaller NSs.

Taken together, these facts reinforce the view that NSs are nuclear hubs for PI production and signaling, which are important for protein targeting to NSs and NS assembly, as well as for the regulation of transcription, transcript maturation, splicing and export (26,27,55,58). However, our understanding of the most basic mechanisms of PI generation and its wide impact on nuclear functions is at an early stage and requires further elucidation.

Nucleoskeletal organization. In proteomic studies (21,22), different types of structural proteins were identified in NSs, but our knowledge at that time was not sufficient to unambiguously assign lamins, myosins and tubulins to NSs. However, later studies revealed that some of these cytoskeletal proteins are indeed NS components. Cytoskeletal re-arrangements are regulated by NS proteins involved in PI signaling (PIPA1A, INPPPL1, PDLIM7/ENIGMA and profilin-1) and calcium signaling (L-plastin, PTK2B and
Figure 2. RNA processing, from the site of transcription to nuclear export, is regulated by multiple proteins localized within NSs. The RNA pathway in the nucleus starts with transcription initiation and multiple proteins, many of which are found in NSs, are responsible for the processing of primary transcripts, including splicing, m6A modification, 3' end processing and export. In the figure, examples of proteins with known functions in transcription and RNA maturation are presented (green dots). Additionally, a large group of other NS-associated proteins is indirectly involved in the precise regulation of RNA processing and their examples are shown (red dots).

EPB41). This suggests that the functional associations between PIs, calcium and cytoskeletal proteins in the cytoplasm are similar in NSs.

Proteins implicated in cytoskeletal organization contribute to nuclear assembly (59) but also regulate transcription (60). Disruption of actin delivery to RNA polymerase II (RNAPII) profoundly impairs transcription elongation (61). In addition to its role in transcriptional elongation, myosin Ic takes part in the formation of the first phosphodiester bond during transcription initiation (62). Furthermore, actin and myosins are cofactors for several transcription and chromatin remodeling complexes (e.g. PCAF, INO80 and BAF, numerous components of which have been found in NSs) and RNA-binding proteins.

Nuclear actin exists in monomers (G-actin) and short polymers (filaments, F-actin). Actin polymerization regulation is important for nuclear transport and transcription because NS reorganization upon RNAPII inhibition and RNAPII recruitment to activated promoters depend on actin polymerization (63). However, despite the fact that modifiers of actin polymerization localize in NSs, primarily monomeric actin was shown to be concentrated in NSs (64). Therefore, deeper insight into the role of actin in NSs is required and whether NSs play a role in actin polymerization needs to be determined.
attachment (66). Little is known regarding the fate of proteins that are ubiquitinated in NSs because, to date, deubiquitinating enzymes have not been identified in NSs. This lack of deubiquitinating enzymes could suggest that after ubiquitination, proteins escape from NSs to be deubiquitinated by deubiquitinating enzymes (E4s) and E4 ubiquitin ligases. NSs contain more ubiquitin-activating enzymes (E1s), ubiquitin-conjugating enzymes (E2s) and E3 ubiquitin ligases. NSs contain more than 20 E3 ligases or constituentsof E3 ligase multisubunit complexes. Although active proteasomes have a role in protein degradation in NSs and affect NS assembly (65), only a specific type of polyubiquitination (attachment of four or more ubiquitin molecules) targets a protein for proteasomal degradation. Other forms of polyubiquitination and monoubiquitination adjust the activities (function, cellular localization and protein–protein interactions) of the conjugated protein.

Ubiquitin-dependent regulation is an important mechanism for splicing control and ubiquitination facilitates protein–protein interactions required for spliceosome assembly (66). Little is known regarding the fate of proteins that are ubiquitinated in NSs because, to date, deubiquitinating enzymes have not been identified in NSs. This lack of deubiquitinating enzymes could suggest that after ubiquitination, proteins escape from NSs to be deubiquitinated or undergo proteasomal degradation in NSs. Interestingly, several NS protein complexes contain subunits responsible for interactions with ubiquitin-specific proteases.

In contrast, SUMOylation of NS proteins is better understood. Attachment of SUMO-1 is a typical signal to target proteins to NSs. The presence of deSUMOylating enzymes in NSs suggests that SUMO-1-dependent protein targeting to NSs may be reversible. SUMOylation is involved in the regulation of many other NS-associated processes. For example, CBX4 is the main regulator of the relocation of chromatin to the proximity of NSs upon transcriptional activation (67) and SUMOylation of TCEG1 inhibits its activity in transcription elongation (68).

Together, the presence of numerous components of the ubiquitination/SUMOylation machineries in NSs merits defining NSs as nuclear ubiquitination and SUMOylation centers. However, most of the ubiquitination/SUMOylation mechanisms in NSs have yet to be uncovered, e.g. how SUMOylation and ubiquitination can cooperatively modulate the functions of individual proteins.

**RNAs in nuclear speckles**

Numerous studies have revealed the presence of RNAs, including poly(A)^+ RNAs (69,70) and noncoding RNAs, in NSs. A significant portion of total nuclear poly(A)^+ RNA is localized within NSs (71). However, there is no substantial immobile fraction of poly(A)^+ transcripts (72,73), as these transcripts are freely exchanged between NSs and the nucleoplasm.

The enrichment of some transcripts in NSs mainly depends on the presence of introns (12) but intronless transcripts also localize to NSs (74); however, they need to contain either additional motifs (e.g. sequences promoting the alternative export or ALREX, pathway) or specific mRNA structures to direct trafficking of intronless mRNAs through NSs.
Nuclear export seems to be the most prominent pathway regulating mRNAs residing in NSs because depletion of transcription export (TREX) complex components leads to enhanced association of mRNAs with NSs (12). Moreover, redirecting mRNAs to the nuclear export pathway is energy-dependent; therefore, this process significantly differs from passive intranuclear diffusion (75).

Unlike the transcription-independent exchange of poly(A)+ RNA between NSs and the nucleoplasm (76), mRNA nuclear export depends on active transcription (70), complete polyadenylation and splicing (77). Because most exons are spliced cotranscriptionally at the transcription site (78,79), NSs seem to be the location of the transcript maturation that is needed for nuclear export (80). Nevertheless, slow splicing or splicing defects lead to enrichment of transcripts in NSs because they inhibit transcript export (80,81). Such ‘difficult’ posttranscriptional splicing might occur in NSs and several reports have suggested that splicing can occur within NSs (11,12). However, this possibility requires further investigation.

Noncoding RNAs present in NSs include spliceosomal small nuclear RNAs (snRNAs), 7SK RNA, whose downregulation leads to mislocalization of NS components (82) and the long noncoding RNA metastasis-associated lung carcinoma transcript 1 (MALAT1), which will be described in more detail in the following section.

**Role of MALAT1 RNA.** Long noncoding RNA MALAT1 is also known as nuclear-enriched abundant transcript 2 (NEAT2) (83,84). The 8.7 kb human MALAT1 gene locus is on chromosome 11, ∼60 kb downstream of the NEAT1 gene, which gives rise to another long noncoding RNA that is a well-established structural component of paraspeckles. The primary transcript of the MALAT1 gene is cleaved at its 3′ end by RNase P to generate the ∼7.5-kb long MALAT1 RNA, which is stabilized by a highly conserved triple helical structure instead of a poly(A) tail and is retained in the cell nucleus. The remaining short 61 nt RNA-like masclRNA is localized exclusively to the cytoplasm (85).

Multiple SRSFs associate with MALAT1, including SRSF1, SRSF2 and SRSF3, which directly bind to their recognition sites at the 5′ end of MALAT1. In addition, MALAT1 binds other proteins and its NS localization depends on RNPS1, SRm160 and IBP160 (86–88). Of the nuclear RNAs, MALAT1 interacts directly with U1 snRNA but is unlikely to interact similarly with other noncoding nuclear RNAs. Similarly, MALAT1 does not form direct RNA–RNA contacts with nascent pre-mRNAs; rather, it binds indirectly via protein mediators (89). MALAT1 also interacts with chromatin at actively spliced genes, close to polyadenylation sites, in a transcription-dependent manner, similar to patterns of RNAPII occupancy (89). It also recruits machinery for gene activation and the movement of active chromatin to NSs (see below).

MALAT1 influences alternative splicing by regulating the phosphorylation and resulting nuclear distribution of splicing factors. In addition, MALAT1 downregulation seems to increase the cytoplasmic pool of poly(A)+ RNAs (90). However, in vivo analyses of mice depleted of MALAT1 revealed that MALAT1 is not essential for mouse development, gene expression regulation, alternative splicing, splicing factor phosphorylation or the organization of NSs (91,92). This difference between human cellular and mouse models is surprising as MALAT1 shows strong sequence conservation across mammalian species and is a highly abundant transcript.

**BIOLoGICAL FUNCTION**

Kinases, cytoskeletal proteins and enzymes of ubiquitin and PI metabolism are prominent nuclear gene expression regulators with a known role in transcription and splicing regulation (Figure 3). However, they represent only a small fraction of the NS proteome: >50% of the NS proteins listed in Supplementary Table S1 are involved in transcription or splicing regulation. Over-representation of splicing proteins in NSs has been apparent from early studies, but the data from recent years have revealed that the largest group of NS proteins is involved in transcription and includes transcription factors and chromatin remodeling factors.

NS formation and function are tightly associated with active transcription and RNAPII integrates transcript synthesis with the DNA template and DNA-regulating proteins on one hand and RNA maturation and export on the other (Figure 3). Notably, mRNA 3′ end processing and polyadenylation, mRNA m6A methylation, mRNA coating with proteins, execution of mRNA nuclear export and chromatin regulation are also directly coupled with splicing. Accordingly, more than 30 NS proteins play important roles in both transcription and splicing. These processes are connected not only functionally but also by physical protein–protein interactions. In the following subsection, we discuss the crucial role of RNAPII transcription and splicing in integration of the biological function of NSs.

**RNAPII transcription and splicing**

NSs are formed in the direct neighborhood of RNAPII transcription sites and contain several RNAPII subunits. Despite the numerous transcription factors residing in NSs (Supplementary Table S1), other proteins that directly regulate RNAPII during transcription initiation are underrepresented. In contrast, components of protein complexes that regulate transcription elongation are present. These include subunits of the super elongation complex and the general transcription elongation factor SII (elongin), which increases the catalytic rate of RNAPII transcription by suppressing transient pausing. NSs also contain the multifunctional epigenetic regulator TRIM28, whose phosphorylation stimulates the transition to transcription elongation, while the non-phosphorylated form causes RNAPII pausing (93).

Direct coupling of multiple mRNA maturation steps with the transcription process occurs through recruitment of suitable factors by the C-terminal domain (CTD) of Rpb1, the largest RNAPII subunit (94). Protein recruitment to the CTD is governed by phosphorylation events on the conserved heptapeptide Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7, which is repeated 52–times in the human CTD (95). Ser5 phosphorylation is connected to synthesis of the 5′ end of nascent pre-mRNA, while specific phosphorylation of Ser2 is responsible for transcription elongation and engagement of the pre-mRNA splicing and processing machineries (96).
Figure 3. The role of proteins localized to NSs in pre-mRNA synthesis, maturation and DNA template regulation; impact of NS regulatory proteins on depicted protein complexes is indicated with specific symbols explained in the figure. DNA template is shown in red and primary transcript in blue, exon (dark blue) intron (light blue). TFs: transcription factors; RNAPII: RNA Polymerase II; PIs: phosphoinositides; CTD: C-terminal domain of RNAPII.

Among the CTD regulators, components of the Ser2-phosphorylating p-TEFb complex, i.e. CDK9 (as well as CDK12/13) and cyclins have been found in NSs. The CDK9-interacting proteins HEXIM1 and BRD4, which are negative and positive regulators of p-TEFb, respectively, are also localized to NSs (97,98). In contrast, kinases implicated in Ser5 phosphorylation are absent from NSs, supporting the role of NSs in processes coupled to transcription elongation.

Transcription requires the formation of a short hybrid of nascent pre-mRNA and an unwound DNA template, the so-called R loop, which is known to increase DNA instability and affect DNA integrity (99). R loop removal is promoted by mRNA packing with a protein coat (100), a high RNAPII elongation rate, efficient transcription termination or mRNA 3′ end processing. However, the NS-associated mechanisms preventing the harmful effects of R loops on genome stability involve not only tight temporal integration between co-transcriptional steps of pre-mRNA maturation but also numerous DNA repair proteins localized in NSs (Supplementary Table S1).

The spatial organization of chromatin within the cell nucleus serves as an important level of gene expression regulation that is linked to NSs and RNAPII transcription. Genes on different chromosomes can be spatially clustered due to random movement of chromosomes within the nucleoplasm and specific binding of decondensed stretches of chromatin by transcriptional machinery, which is differentially arranged in various cell types. Spatial association occurs within actively transcribed genes and co-regulated genes gather around shared NSs (38,101). This association of chromatin domains with NSs is correlated with gene expression upregulation (102,103). Notably, NS dispersion leads to a decrease in chromatin domains in shared neighborhoods (104), supporting the importance of NSs in the spatial organization of transcriptionally active chromatin.

Most human pre-mRNAs contain introns, which are excised during the splicing process. Splicing is a key step in pre-mRNA maturation and is typically coupled with transcription elongation (105). The excision of introns in eukaryotes is catalyzed and regulated by five spliceosomal snRNPs and more than 100 associated proteins (106), most of which are known NS proteins. The resultant mature transcript can be alternatively spliced in specific cells or environmental conditions due to exon skipping or extending or by intron retention.

Notably, splicing depends on transcription elongation and vice versa: active splicing is required for Ser2 phosphorylation (107) of the CTD. Ser2p- and p-TEFb-induced transcription elongation is positively correlated with splicing efficiency (108). The spatial proximity of NSs and active RNAPII transcription sites allows convenient access to splicing factors because splicing factors are continuously added to and released from RNAPII as RNAPII moves along a gene. In addition to SRSFs, which interact with the CTD in a phosphorylation-dependent manner (109), several other splicing factors have been shown to directly interact with the CTD. Among them, U2AF65 (complexed with PRP19C) is required for CTD-dependent splicing activation (110).

Splicing is thought to affect the rate of nascent RNA synthesis by promoting RNAPII pausing (111). Conversely, RNAPII slowing promotes RNAPII accumulation in the intronic regions flanking the alternative exons, which allows recruitment of activators or repressors of alternative splicing. These can promote exon inclusion or exon skipping, respectively. Pre-mRNA splicing and negative regulation of transcription elongation are coupled by NS-localized proteins: e.g. the splicing factor/transcriptional repressor SF1 interacts with the CTD-binding regulators of transcription elongation (112). Additionally, BRM (the ATPase subunit of the SWI/SNF chromatin remodeler) (113) and ZNF326...
Splicing factors to chromatin (117) and conversely, splicing (116). Histone modifications also guide the recruitment of interactions that promote relaxation of the chromatin structure (115) and exon skipping is triggered by histone modifications. DNA wrapping around nucleosomes induce exon inclusion (117) and conversely, splicing factors recruit histone-modifying enzymes in a pre-mRNA-dependent manner (118). Similarly, histone acetyltransferases and methyltransferases interact with U1 snRNPs (119). Consequently, splicing inhibition or splicing factor depletion leads to changes in histone modification patterns (111).

These close relationships between splicing and epigenetic mechanisms of gene expression regulation may underlie the robust enrichment of proteins involved in epigenetic regulation in NSs. Notably, histones, histone acetyltransferases, methyltransferases, deacetylases and HP1 protein, which are fundamental players in epigenetic regulation of gene expression, were identified in NSs (Supplementary Table S1). However, the potential function of HP1 in NSs and the role of NSs in histone modification has not been directly studied, despite the fact that particular core catalytic subunits of histone modifiers can localize to non-overlapping NSs, suggesting that specific multisubunit complexes regulate a distinct set of genes (120). Because multiple protein complexes involved in histone modification or histone variant deposition and protein partners of HP1 (including SRSF1/3 (117)) are found in NSs, future research on epigenetics in the context of NS localization may uncover novel mechanisms of gene expression regulation.

In short, transcription and splicing are largely implicated in the regulation of chromatin structure and function. The dependence of splicing on epigenetic regulation is not the only connection that occurs between splicing factors and DNA regulatory proteins because multiple examples of DNA repair and DNA stability regulation by splicing factors has been published recently (reviewed in (121)). Because many DNA-regulating proteins have been found in NSs (122), the role of splicing factors and NSs in DNA regulation is an attractive target for future research.

**mRNP maturation and export**

Proteins localized in NSs not only participate in numerous aspects of mRNA synthesis and mRNP maturation and are essential for harmonization of these nuclear processes but also influence cytoplasmic mRNA events.

The final step of mRNA synthesis in almost all eukaryotes is the addition of a poly(A) tail at the 3′ end. The most prominent factors regulating and catalyzing endonucleolytic cleavage and the subsequent non-template poly(A) tail addition are localized in NSs (Supplementary Table S1) and their activity depends on their interaction with the CTD (94). Up to 70% of human mRNA transcripts are thought to be subjected to alternative polyadenylation (123). NS-associated processes, including CTD phosphorylation, epigenetic regulation, N6-methylation of adenosine (m6A) and splicing, are implicated in site selection in alternative polyadenylation (124,125). Relationships between splicing and 3′ end processing have been widely studied and the studies have shown that 3′ end processing is promoted by the intron upstream of the polyadenylation signal, whereas the splicing efficiency of the terminal intron is reduced by mutations at the polyadenylation signal (126). This reciprocal coupling is thought to be realized via interactions of splicing factors (e.g. U2AF65) or spliceosome components (U1A of U1 snRNPs) with nuclear poly(A)-polymerase (127) or cleavage and polyadenylation specificity factor 160 (CPSF160) (128), respectively, among other mechanisms.

m6A mRNA modification has a strong impact on orchestration of mRNP maturation. NSs contain the critical m6A-specific methyltransferases METTL14 and METTL3 that associate with the regulatory Wilms’ tumor 1-associating protein (WTAP), whose protein partners are NS proteins. The m6A modification is enzymatically removed by the demethylase ‘eraser’ enzymes fat mass and obesity-associated protein (FTO) (129) and ALKBH5 (130), which are also localized in NSs. The presence of all crucial nuclear elements of the m6A modification system in NSs suggests a prominent role for NSs in m6A regulation. Moreover, the known nuclear readers of m6A YTHDC1 and hnRNPb1a are also localized in NSs. YTHDC1 promotes SRSF3 NS localization and RNA-binding affinity but represses those of SRSF10, thus promoting exon inclusion and inhibiting exon skipping (131). m6A modifies structure (base pairing) and interactions of RNAs with proteins and thus affects numerous aspects of RNA metabolism, as shown for regulation of MALAT1 binding by hnRNP C (132).

All the aforementioned steps of mRNA synthesis and maturation are closely associated and leave signatures on the mRNA (e.g. deposited proteins, poly(A) tail length or modified nucleotides) that regulate its subsequent nuclear export, subcellular localization, stability and translation. More than 1000 different proteins can be recruited to human mRNAs to form mRNP particles. Among them, the core components of the exon–junction complex (EJC) are recruited to the spliceosome during the early steps of splicing (133). The EJC is a key regulator of mRNP stability (134) and attracts multiple EJC peripheral factors, a substantial portion of which are NS proteins (135,136) that specifically enhance subsequent gene expression steps, such as adjacent splicing events, mRNA nuclear export or translation.

The TREX complex plays an essential role in coordinating transcription and splicing with mRNA export (137,138) and TREX can be recruited to mRNA through the EJC in a splicing-dependent manner (139). TREX serves as a platform for binding of the NXF1/NXT1 heterodimer (140), which is a general export receptor and is responsible for translocation of mRNPs through the nuclear pore. In addition to the association of exporting complexes with the EJC, NSs contain additional proteins that interact with NXF1/NXT1 in proximity to the 5′ and 3′ ends of transcripts, e.g. the cap-binding complex (141) and CPSF (142), respectively. Furthermore, several other adap-
tors localized in NSs, namely, CIP29, LUZP4, UIF, Dbp5 and the SRSFs, have been shown to mediate or modify NXF1/NXT1 recruitment. This observation indicates that multiple NXF1/NXT1 complexes are capable of binding various sites within individual mRNPs.

mRNP export is well known to be enhanced by splicing (12,143). UAP56, which is found in the EJC (133), is a splicing factor essential for spliceosome assembly and for the interaction of U2 snRNP with the branch point during splicing (144). Moreover, TREX components have been shown to interact directly with Ser2-phosphorylated CTD (145). TREX recruitment to Ser2-phosphorylated CTD is mediated by the PRP19 complex (146), which can be considered an important link integrating mRNP export with transcription elongation and splicing. Importantly, NS proteins involved in splicing regulation, i.e. SRSF3/7 (147), the PRP19 complex and U2AF2 (148), are also required for export of intronless mRNAs.

The above mentioned mutual relationship between mRNP maturation and export machineries can be considered a part of the mRNA quality control system. This system impacts nuclear retention of transcripts containing splicing-defective introns that form partial spliceosomes (81) and allows the export of only fully processed mRNPs. However, the protein composition of mRNPs controlled by NSs may also influence cytoplasmic RNA degradation, for example, through the nonsense-mediated mRNA decay pathway, which is regulated by proteins involved in the EJC (149). Thus, NSs regulate mRNP formation and subsequent export efficiency, influencing subcellular targeting of mRNPs and translation.

**ROLE OF NUCLEAR SPECKLES IN THE PATHOGENESIS OF HUMAN DISEASES**

Multiple diseases are associated with altered NSs or aberrant function of proteins found in these structures. Some rare disorders are caused directly by mutations in genes encoding proteins and noncoding RNAs found in NSs (150–152). Among these diseases are retinitis pigmentosa 13, 18, 33, 60, which is caused by mutation in *PRPF8*, *PRPF3*, *SNRPN200* and *PRPF6* genes; mandibulo-facial dysostosis, Guion-Almeida type, which is caused by mutation in the *EFTUD2* gene; and thrombocytopenia-absent radius syndrome, which is caused by mutation in the *RBM8A* gene. However, most diseases associated with NS malfunction, including cancers and viral diseases, are indirectly linked to these structures.

Carcinogenesis is strongly associated with alternative splicing aberrations that are caused by changed expression of proteins involved in alternative splicing that often result in NSs morphology disturbances. Overexpression of NS-forming proteins, including SR proteins, is observed in many types of cancer (153) and abnormally high levels of SRSF1 have been shown to be sufficient to induce tumorigenesis (154). Similar effects have been described for several other splicing and polyadenylation-associated proteins (155,156). In cancers, the altered polyadenylation pattern is explained, in part, by the tendency to shorten transcripts to reduce their ability to undergo regulation, e.g. by miRNAs (157). Moreover, single nucleotide mutations in pre-mRNA splicing factors may be associated with a gain-of-function mechanism (158).

Viruses use multiple host proteins during their life cycle, which includes viral RNA splicing and export (159). Both SRSFs and hnRNP proteins have been shown to control HIV RNA splicing and that of other viruses, thereby strongly influencing the translation of viral proteins (160–162). Consequently, viral infection triggers changes in the cellular levels and localization of splicing factors, e.g. their exclusion from NSs (163,164). Proteins participating in host mRNA 3’ end formation are also involved in the viral reproductive cycle (165). Interestingly, even viruses that replicate in the cytoplasm may affect NS function (166).

Emerging evidence supports the involvement of NSs in the pathogenesis of neurological diseases. NS proteins are involved in neuronal cell differentiation, as well as in the control of neuron-specific splicing and contribute to the abnormalities observed in neurological diseases, which is discussed below.

**Neurological diseases**

Neurodegeneration that occurs during the progression of neurological disorders is accompanied by disruption of nuclear bodies, including NSs. The involvement of NSs in neurodegeneration often expands into general splicing aberrations and dysfunction of NS proteins. In some neurological diseases, neurodegeneration is attributed to protein and/or RNA aggregates that are formed within the cell nucleus.

Alzheimer’s disease (AD) is associated with β-amyloid, which is formed from an amyloid precursor protein that is regulated by U1 snRNP (167). One of the proteolytic products of this protein is the Cγ fragment, whose phosphorylated form localizes to NSs and is thought to take part in pre-mRNA splicing (168). Neurodegeneration is also linked to NSs by the GSK-3 protein, which accumulates in the cytoplasm of neuronal cells in AD and other tauopathies (169). A lower GSK-3 level in the nucleus of neurons results in aberrant splicing of exon 10 in the τ transcript and enrichment of NSs with SRSF2 (170).

Aberrations in alternative splicing control that may be attributed to proteins found in NSs are also observed in other diseases. In Parkinson’s disease (PD), alternative splicing events, including exon skipping in the leucine-rich pentatricopeptide repeat-containing RNA transcript that binds hnRNP A1-associated poly(A) mRNAs and is regulated by hnRNP A1, may contribute to neurodegeneration (171). Aberrant alternative splicing of transcripts involved in PD pathogenesis has been extensively investigated and includes missplicing of genes directly involved in disease development, as well as downstream effects resulting in the inclusion of alternative exons that alter interactions with miRNAs and lncRNAs (172).

In frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS), which are closely related diseases, a general disruption of nuclear structure that may affect NSs has been observed. The TDP-43 protein has been shown to serve as a cellular scaffold for multiple nuclear bodies, including Cajal bodies, PML bodies, NSs and gems (173). In some FTD patients, TDP-43 formed intracellular inclusions that mostly localized within the cytoplasm, decreas-
Figure 4. (A) MBNL1 binds to expanded CUG RNA, resulting in aberrant alternative splicing in myotonic dystrophy type 1. (B) HTT mRNA splicing is disturbed by SRSF6 in a CAG length-dependent manner. (C) PABN1 aggregates bind pre-mRNA, which hinders splicing factor binding and results in aberrant splicing. (D) RNA foci harboring transcripts containing expanded simple repeats have different localization patterns in relation to NSs. Schematic representations of RNA foci localization are shown.

ing its level in the nucleus (174). Similar inclusions were also observed in ALS and other neurodegenerative diseases (175). Aberrations in TDP-43 may therefore lead to significant disruption of nuclear structure and affect the function of nuclear compartments. Decreases in TDP-43 were also shown to impair splicing (176, 177), triggering the inclusion of cryptic exons (178). Many splicing events of disease-associated transcripts involved in neurodegeneration were disturbed upon depletion of TDP-43. Interestingly, TDP-43 was also shown to bind and regulate MALAT1 levels and depletion of TDP-43 led to decreased expression of this NS-specific lncRNA (177). In addition to TDP-43, ALS is associated with fused in sarcoma (FUS) protein aggregation (179). C-terminal mutations in FUS that disrupt the nuclear localization signal result in cytoplasmic accumulation of the protein (180). Physiologically, FUS is found in NSs and decreased levels of FUS and mutations in its sequence may alter the function of these structures.

In spinal muscular atrophy, in which the SMN1 gene is inactive, a disturbance in nuclear structure has also been observed (181, 182). SMN1 is involved in protein transport to NSs and along with close interactions between various nuclear compartments, a lack of the SMN1 protein may lead to disturbance of NSs, especially because the SMN complex is also involved in spliceosome assembly (183).

For several neuropsychiatric diseases, known associations with NSs may explain the mechanisms of disease progression. Known point mutations in, and copy number variants of, genes encoding EJC proteins that are found in NSs are linked to intellectual disabilities, autism and schizophrenia (184). Moreover, mutations in the 5′ UTR of EIF4A3, which is associated with the EJC, are implicated in Richieri-Costa-Pereira syndrome, which causes learning deficits, among other symptoms (185). The EJC has been shown to play an important role in neurodevelopment, which highlights its potential role in neuropsychiatric disorders (186). In schizophrenia patients, aberrant alternative splicing, including splicing of transcripts such as the glutamate transporter and microcephalin, which are important for neuronal function, was observed (187, 188). Aberrant splicing was also found in autism and often involved microexons, which are atypical exons only 3–27 nt in length. Interestingly, microexon splicing events often occur in neuronal cells and the functions of proteins found in NSs may also be altered by such microexon-dependent mechanisms, as has been shown for SR100/SRRM4 (189, 190). Thus, the current knowledge of the scale of microexon involvement in diseases may be just the tip of the iceberg.

Role of nuclear speckles in simple repeat expansion diseases. Simple sequence repeats occur within numerous human genes and expansion of these sequences beyond a critical number of copies in specific genes causes incurable genetic diseases. In these disorders, pathogenic effects can be triggered by the toxic RNA transcribed from an expanded al-
Figure 5. NSs differ considerably among various human cell lines. Representative microscopy images of cells labeled for SRSF2 (constitutive protein of NSs) are presented. Due to the different expression levels of SRSF2 in different cell lines, the anti-SRSF2 antibody was used at different dilutions: 1:500 in fibroblasts, 1:500 in SK-MCs, 1:200 in lymphoblasts and 1:500 in SK-N-MCs. Bars = 10 μm (for the fibroblast, SK-MC and SK-N-MC) and 5 μm (for the lymphoblast).

lele, and nuclear foci formed by mutant RNAs have been found in multiple neurological diseases belonging to this class. The reasons for nuclear retention of mutant RNA are not clear; however, RNA foci localization within or in proximity to NSs suggests that proteins present in these structures may contribute to foci formation.

Expanded CUG repeats present in the 3′ UTR of the DMPK transcript in myotonic dystrophy type 1 (DM1) form RNA foci that localize on the periphery of NSs (191). CUG RNA foci are arrested at the border of NSs due to RNA structure and aberrant interactions with MBNL1 proteins. MBNL1 silencing released mutant RNA foci from the periphery of NSs, proving the involvement of MBNL1 in nuclear retention of mutant transcripts (192). On the other hand, tautomycin-mediated disintegration of NSs did not influence CUG RNA foci (193), showing their structural independence of these nuclear structures. Interestingly, DNA loci of normal and mutant DMPK genes also localized to the periphery of NSs, demonstrating that RNA foci are formed close to the site of mutant RNA transcription. RNA foci in DM1 contain spliced mRNA (194), which suggests that proper splicing of this transcript does not require NS entry and that another mechanism must be responsible for the arrest.

In cellular models of myotonic dystrophy type 2 (DM2), in which expanded CCUG repeats are present in the ORFs of various functionally unrelated genes, mutant RNA is either temporarily or is partially captured within NSs (193,195,196). The presence of mutations did not trigger significant changes in the number or size of NSs in fibroblasts, but disintegration of NSs induced by tautomycin resulted in disassembly of RNA foci (193). In Huntington's disease (HD), which is one of the polyglutamine disorders, aberrant alternative splicing events were observed, including aberrant splicing of the τ transcript (197). Moreover, aberrant splicing of transcripts formed from the mutant allele in HD was observed and attributed to misregulation of SRSF6-controlled splicing (Figure 4B) (198). Similar to CAG repeats, foci formed by expanded CGG repeats present in the 5′ UTR of FMR1 in FXTAS localize within NSs (199). The CGG foci also disassembled when cells were treated with tautomycin; however, the mechanism proposed implied an interaction between tautomycin and the mutant transcript, preventing foci formation.

In cellular models of myotonic dystrophy type 2 (DM2), in which expanded CUG repeats are present in the intronic sequence of the ZNF9 gene, there is no connection between NSs and RNA foci (192). Similar to repeats in DM2, RNAs with expanded GGGGCC repeats localized in the intronic sequences of C9orf72 RNA, which are found in a fraction of ALS patients, localize mostly outside NSs (200). These observations can be explained by both types of
repeat-containing RNAs found in RNA foci being excised introns, unlike mutant mRNAs in triplet repeat expansion diseases in which RNA foci are associated with NSs.

As mutant RNAs interact abnormally with various proteins (201), transcripts with expanded repeats may trigger protein dislocation. This dislocation can have a dual effect: the sequestered proteins cannot perform their functions in their normal location as well as they can perform their aberrant function in their new location, but loss of protein function is more common, as shown for the MBNL1 protein in DM1 (Figure 4A). Similar mechanism was recently shown for oculopharyngeal muscular dystrophy (OPMD), which is caused by a GCG expansion within the PABPN1 gene; the mutant PABPN1 protein delocalized transcripts from NSs, resulting in decreased activity of splicing factors residing in NSs toward these transcripts (Figure 4C) (202). Interestingly, in OPMD, a mutant protein is involved in NS aberrations, as the polyadenylation function of the mutant protein is altered (203).

An investigation of the subnuclear localization of foci formed by RNAs containing various expanded repeats showed that their localization differs; transcripts containing expanded CAG and CGG repeats localize within NSs, those containing CUG repeats localize on the periphery of the NSs and those containing GGGGCC and CCUG repeats localize outside of NSs (Figure 4D). A single mechanism that triggers the disturbances in nuclear export of all mutant RNAs is thus unlikely.

SUMMARY AND FUTURE PERSPECTIVES

NSs were initially thought to be sites for storage and modification of splicing factors but have been now recognized as nuclear bodies that facilitate integrated regulation of gene expression. In this review, we summarized the current knowledge regarding the molecular composition and structural organization of these nuclear bodies. We discussed their properties and maintaining forces, fate during the cell cycle, canonical and newly postulated physiological functions and role in human disease. In characterizing these nuclear structures, our insights frequently expanded beyond the borders of NSs themselves into their nuclear environment. In this section, we put studies on NSs in a still broader context of high resolution mapping of the entire nuclear structure and function in the dimensions of space and time, which is the scope of such megaprojects as the 4D Nucleome Project (https://www.4dnucleome.org/) and the Human Protein Atlas (204). One of the goals of these initiatives is to gain deeper insight into the structure and function of nuclear compartments and nuclear bodies, including NSs. The 4D Nucleome Project has great potential for developing new methods capable of determining the molecular composition of NSs and methods for probing protein–protein and RNA–protein interactions within these structures as well as interactions between components of NSs and chromatin.

Also of great importance are new labeling and imaging methods to visualize NS constituents at high resolution. Such methods include CRISPR/Cas9-based labeling systems (205), super-resolution microscopic approaches (206) and cryo-EM (cryo-EM) (207) with cryo-electron tomography (cryo-ET) (208) for three-dimensional (3D) reconstruction of biological structures. The latter methods are based on EM imaging of frozen samples and do not require fixation or crystallization and thus, cryo-EM/ET allows native structures to be preserved and avoids artifacts. These methods enabled determination of the structure of the 21 MDa supraspliceosome (four spliceosomes gathered on pre-mRNA, ~50 nm) (209) and the 100 MDa polyribosome complex consisting of 20 ribosomes bound to mRNA (about one order of magnitude smaller than NSs) (210). Recent progress in cryo-EM instrumentation (electron detectors) and image processing methods allowed the structure of native ribosomes (211) and spliceosomes (212) to be resolved at near-atomic-resolution.

Structural studies of highly dynamic assemblies such as NSs remain challenging because protein composition and the shape and size of distinct NSs differ within individual cells and between various analyzed human cell types (Figure 5). The presence of multiple LCRs and the flexibility of NS proteins adds additional difficulty. However, new developments in cryo-EM/ET allow analysis of difficult biological specimens. Single particle analysis on one hand and averaging, which is obtained by merging corresponding structures, on the other hand, increase structure resolution. Additionally, image classification methods aid in managing sample heterogeneity (213). Furthermore, and most importantly for structures as large as NSs (0.8–1.8 μm), cellular cryo-ET enables visualization of particles in the native state inside cells. Analysis of thick biological samples without biochemical purification is possible due to progress in non-disruptive preparation of thin sections of biological material (214). In the near future, progress in cryo-ET is expected to allow further development of visual proteomics, i.e. 3D reconstruction of the detailed localization, conformation and interactions of proteins in a complex environment based on fitting of known structures from cryo-electron microscopy and nuclear magnetic resonance or solving structures de novo. Recently, the 26S proteasomes in intact neurons were visualized using this technology (215). Therefore, it is reasonable to expect that application of these technologies will bring new insights into the molecular organization of NSs and clarify their role in cell physiology and the molecular basis of diseases.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

National Science Centre [2012/06/A/NZ1/00094 to W.J.K., 2015/19/B/NZ2/02453 to W.J.K., 2015/17/N/NZ3/03629 to M.O.U.]; L’Oreal Poland Women in Science Fellowship (to M.O.U.); Polish Ministry of Science and Higher Education (KNOW program for years 2014–2019). Funding for open access charge: National Science Centre—Poland [2015/19/B/NZ2/02453]. Conflict of interest statement. None declared.

REFERENCES


None declared.


F очень много текста, многие из них являются цитатами из других научных статей. Вот некоторые из них:

- **Quesnel-Valli`eres, M., Irimia, M., Cordes, S.P. and Blencowe, B.J. (2015)** Essential roles for the splicing regulator nSIR100 SRMR4 during nervous system development. *Genes Dev.*, 29, 746–759.

**Nucleic Acids Research, 2017, Vol. 45, No. 18**


