

A conserved charged single α -helix with a putative steric role in paraspeckle formation

LÁSZLÓ DOBSON,^{1,3} LÁSZLÓ NYITRAY,² and ZOLTÁN GÁSPÁRI¹

¹Pázmány Péter Catholic University, Faculty of Information Technology and Bionics, H-1083 Budapest, Hungary

²Eötvös Loránd University, Department of Biochemistry, H-1117 Budapest, Hungary

ABSTRACT

Paraspeckles are subnuclear particles involved in the regulation of mRNA expression. They are formed by the association of DBHS family proteins and the NEAT1 long noncoding RNA. Here, we show that a recently identified structural motif, the charged single α -helix, is largely conserved in the DBHS family. Based on the available structural data and a previously suggested multimerization scheme of DBHS proteins, we built a structural model of a (PSPC1/NONO)_n multimer that might have relevance in paraspeckle formation. Our model contains an extended coiled-coil region that is followed by and partially overlaps with the predicted charged single α -helix. We suggest that the charged single α -helix can act as an elastic ruler governing the exact positioning of the dimeric core structures relative to each other during paraspeckle assembly along the NEAT1 noncoding RNA.

Keywords: paraspeckle; charged single α -helix; coiled coil; multimerization; structural modeling

INTRODUCTION

Paraspeckles are subnuclear bodies in eukaryotic cells involved in gene regulation, although their exact physiological role is still elusive (Fox et al. 2002; Fox and Lamond 2010; Nakagawa and Hirose 2012). They were shown to sequester mRNA molecules containing A-to-I edited sites and inverted Alu-repeats (Chen et al. 2008; Chen and Carmichael 2009) and thus hinder their cytoplasmic transport and translation. The main protein components of paraspeckles assemble along the long isoform of the NEAT1 noncoding RNA (lncRNA) acting as an architectural scaffold (Clemson et al. 2009; Sasaki et al. 2009). The core proteins forming the paraspeckle, such as NONO, PSPC1, and SFPQ, are closely related and contain a characteristic arrangement of two RRM (RNA recognition motif) domains, a specific NOPS segment, and a coiled-coil region, together referred to as the DBHS (*Drosophila* behavior, human splicing) structural motif (Staub et al. 2004).

The charged single α -helix (CSAH) is a recently identified structural motif (Süveges et al. 2009; Gáspári et al. 2012). CSAHs exhibit high charge density and regularly alternating blocks of positively and negatively charged residues (Knight et al. 2005; Süveges et al. 2009; Swanson and Sivaramakrishnan 2014). When these blocks consist of lysine and glutamate, a directed interaction between them is responsible to

confer stability in the single helical form (Baker et al. 2015). Although many CSAHs contain more arginines than lysines (Süveges et al. 2009), we are not aware of any detailed study on arginine–glutamate interactions. CSAHs have been shown to act as unusual elastic elements by exerting a constant force during their progressive unfolding (Wolny et al. 2014). The exact functional role of CSAH motifs has only been characterized in myosin VI and X where their unique mechanical properties enable them to act as an extension to the lever arm (Spudich and Sivaramakrishnan 2010).

Our previous surveys of CSAH-containing proteins in SwissProt—using a strict consensus of the available methods to minimize the number of false-positive hits—indicated that NONO and some other proteins involved in paraspeckle formation, but not all of them, contain a CSAH motif (Süveges et al. 2009; Gáspári et al. 2012). However, both the increase in the number of available sequences and our observation that CSAHs are preferentially associated with RNA-binding proteins (Gáspári et al. 2012) justifies a more detailed investigation of proteins of the DBHS family to elucidate the occurrence and possible role of CSAHs in these proteins.

A complete atomic-level structural description of paraspeckles is not known in full detail, although the structure of a heterodimer formed by the NONO and PSPC1 proteins

³Present address: Institute of Enzymology, Research Center for Natural Sciences, Hungarian Academy of Sciences, H-1117 Budapest, Hungary

Corresponding author: gaspari.zoltan@itk.ppke.hu

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.053058.115>.

© 2015 Dobson et al. This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

has recently been solved (Passon et al. 2012). This structure revealed a previously unobserved relative orientation of the RRM domains and a noncanonical antiparallel right-handed coiled coil as part of the dimer interface. The amino- and carboxy-terminal disordered regions and the carboxy-terminal part of the predicted coiled coil including the previously identified CSAH segment in NONO are missing from this structure. Based on investigations of paraspeckle formation by truncated variants of PSPC1, the authors proposed a scheme for higher-order organization of the PSPC1/NONO dimers involving the extended coiled-coil region. In this region, they identified a hendecad repeat pattern distinct from the usual heptad repeat characteristic of left-handed coiled coils but similar to that in the right-handed coiled-coil region observed in the dimer structure. Recently, while our manuscript was in preparation, X-ray structures of human SFPQ (splicing factor proline-glutamine-rich) became available (Lee et al. 2015). This structure revealed a similar mode of dimerization as for the PSPC1/NONO heterodimer and the presence of an extended single α -helical region. In the crystals, additional canonical coiled-coil interactions were observed between neighboring dimers. These interactions are different from those proposed for the extended coiled-coil region of the PSPC1/NONO heterodimer. Other recent studies described the formation of SFPQ/NONO heterotetramers (Snijders et al. 2015).

In this study, we analyzed the DBHS protein family with bioinformatic methods and found that the presence of a CSAH segment is a general feature of the family. To assess the structural role of the CSAH region, we built a first-approximation structural model of a multimeric form of a PSPC1/NONO complex, based on the proposed mode of higher-order association of these structures (Passon et al. 2012). This model is consistent with the presence of a CSAH segment and might provide a plausible explanation of its structural role.

RESULTS

The charged single α -helix is a conserved feature of DBHS proteins

Our CSAH identification method relies on the consensus of two conceptually different methods, SCAN4CSAH and FT_CHARGE (Süveges et al. 2009; Gáspári et al. 2012). Detailed analysis led to an interesting observation—namely, that the more stringent of the two methods, FT_CHARGE (Gáspári et al. 2012), predicted CSAH segments in all three human paraspeckle proteins, PSCPC1, NONO, and SFPQ, whereas the other approach, SCAN4CSAH, did not. However, setting the minimum length of CSAHs to 30 residues for SCAN4CSAH yields consensus hits for all three proteins. The CSAH segments are located at the carboxy-terminal (distal) region of the predicted extended coiled-coil region, partially overlapping with it (Fig. 1). Notably, the predicted CSAH

segment in PSPC1 is shorter than in NONO and SFPQ. However, a single Gln \rightarrow Glu change at position 349 of PSPC1 results in a CSAH prediction of the same length as in the other two proteins even with default parameters (i.e., a 40-residue minimal CSAH length for SCAN4CSAH). Nevertheless, here we accepted the shorter predicted CSAH as a conservative estimate of the occurrence of this motif.

To assess the general occurrence of CSAH segments in DBHS proteins, we performed a standard BLASTP search with all three human DBHS proteins (NONO_HUMAN, PSPC1_HUMAN, and SFPQ_HUMAN) on NCBI's "nr" database (accessed on 19 September 2014). This search identified 753 different proteins altogether (Supplemental Table S1). These hits were filtered according to the following criteria: (i) Sequences should contain two RRM domains with hidden Markov model (HMM) alignment length between 55 and 80 residues and (ii) with a maximum interdomain distance of 50 residues, and (iii) sequences were also required to contain a coiled coil of either at least 30 or 60 residues long (see below) and (iv) not closer than 30 and not farther than 120 residues from these domains. Coiled coils predicted by either ncoils or paircoil2 were accepted. The filtered sequences, all of which also contained a predicted NOPS domain, were then analyzed for the presence of predicted CSAH segments and its dependence on the length of the coiled coil. We found that the vast majority of the identified proteins with a coiled coil longer than 60 residues contain a recognizable CSAH region (Table 1). Most of the identified CSAH segments are located after the identified coiled-coil regions, typically in partial overlap with its distal part (Fig. 1).

The extended coiled-coil region of NONO and PSPC1 can be modeled as a right-handed superhelix

To get closer to the understanding of the functional role of the conserved CSAH segment in paraspeckle-forming proteins, we built a structural model of the (PSPC1/NONO)_n multimer. As a first step, we modeled the extended coiled-coil region for which no experimental structure is available. The authors of the crystal structure of the NONO-PSPC1 dimer observed that (i) the predicted extended coiled-coil region of the proteins plays an important role the higher-order organization of the paraspeckle proteins and (ii) this segment exhibits a hendecad repeat similar to the observed coiled coil. Thus, according to the correspondence between repeat pattern and coiled-coil geometry (Offer et al. 2002; Gruber and Lupas 2003), it can be reasonably assumed that this part might form a structurally similar, noncanonical, right-handed coiled coil (Passon et al. 2012). Based on these considerations, we prepared antiparallel coiled-coil models corresponding to the appropriate segments of the human NONO and PSPC1 proteins. First, we generated a right-handed coiled-coil backbone that matches the geometry of the coiled coil in the 3SDE structure well. This resulted in a

consistent correspondence between the repeats and coiled-coil geometry (not shown). However, although side-chain-side-chain interactions in our model are not considered to be necessarily accurate, the general architecture of the multimeric assemblies could be reasonably well captured, as we argue below.

The paraspeckle multimer is an elongated assembly with no strict stoichiometry

To assess the role of the CSAH segments, we built multimeric complexes using the experimental structures of the PSPC1/NONO dimers and the modeled extended coiled coils. The approach is similar to the computer modeling of the full two-headed myosin II molecule (Offer and Knight 1996). Although the model proposed by Passon et al. (2012) allows for arbitrary stoichiometry, we have chosen a hexameric structure as the smallest one where the central PSPC1/NONO dimer (chains C and D in the hexamer, colored green and blue-green in Fig. 2) is paired with another dimer on both sides (consisting of chains A/B and E/F, colored red/yellow and cyan/blue in Fig. 2, respectively), thus it is potentially capable of revealing all possible additional interactions within the limitations of the model. Thus, the choice of a hexameric structure as a first model is purely technical and we are not aware of any observations of hexameric complexes *in vivo*. The hexamer forms an elongated structure with the PSPC1/NONO dimers pointing outward. Along the long axis of the hexamer, the dimers are rotated $\sim 120^\circ$ relative to each other and with an approximate translation of 140 Å along the axis of the extended coiled coil (Fig. 2). We note that these values are dependent on how we model the exact pairings (see above) in the extended coiled-coil region as it influences the exact orientation of the two constituent helices relative to each other. The hendecad repeat structure of the coiled coil in the core dimers (denoted ccAB, ccCD, and ccEF in the hexamer) is not continuous with the repeat in the extended coiled-coil region (forming coiled coils ccBC and ccDE; Fig. 1), which is reflected by the fact that ccAB, ccCD, and ccEF are not coaxial with ccBC or ccDE. Instead, the different coiled-coil interactions occur at the two opposite sides of the

helices leading to a kink in the extended helix relative to those in the core dimers. In addition, regions where three helices lie beside each other are formed, leading to a three-helix α -sheet structure according to the classification of the CC+ database (Testa et al. 2009). During an exploratory 20-nsec molecular dynamics simulation, the CSAH regions remained largely helical. The simulation also indicated the possible formation of more extended interactions between the PSPC1/NONO dimers (Supplemental Tables S3, S4; Supplemental Figs S2, S3).

Using the same principles as for building the hexameric structure, we generated large multimeric assemblies of PSPC1/NONO dimers. These are expected to correspond to a physiologically more relevant state than the hexamer model. The shape of these multimers is elongated and they display a regular arrangement of outward-pointing RRM domains (Fig. 3).

DISCUSSION

Our observations strongly suggest that a charged region with strong propensity to form CSAH segments is present in the DBHS family and is a characteristic part of these proteins. If this is true, then this is the first protein family in which a CSAH segment is recognized as a regularly occurring motif. In the PSPC1/NONO dimer, the predicted CSAH segment starts close to the carboxy-terminal residue observable in the X-ray structure and overlaps with the predicted extended coiled-coil region (Fig. 1). This is consistent with the recent experimental structure of SFPQ (Lee et al. 2015) where the presence of a long single α -helix was observed. Although CSAHs are expected to be stable even in solution and in the SFPQ structure this helix is likely stabilized by crystal contacts, the observed helical structure can still be regarded as a strong support for our claim.

No structure predictions can be 100% accurate, and it has been specifically noted that predicted CSAH segments often overlap with predicted coiled coils (Peckham and Knight 2009; Süveges et al. 2009). However, we proposed that this might indicate that the same segment can form either CSAHs or weak coiled coils depending on the environment (Süveges et al. 2009). In our structural model of the para-

speckle multimer, the overlapping part of the predicted CSAH segments with the ccBC and ccEF coiled-coil regions were modeled within the extended coiled coils. Because of the kink in the long helices, the CSAH segments point slightly away from the core dimers without causing steric clashes. When the geometry of the predicted CSAH segments is adjusted to match that of a standard α -helix (i.e., the coiled-coil specific supercoiling is “removed”), this feature is even more pronounced (not shown). The proximal part of the CSAH segments is recognized

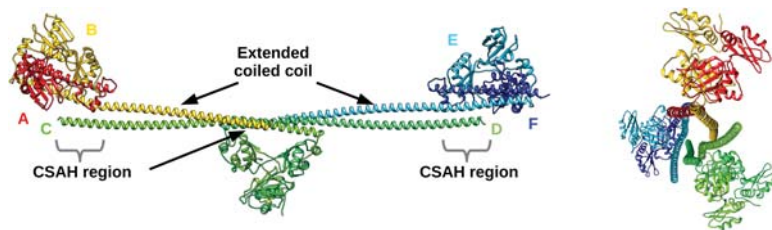


FIGURE 2. Side (*left*) and axis (*right*) views of the (PSPC1/NONO) hexamer model. The six chains are shown in different colors: chains A (red), C (green), and E (cyan) correspond to PSPC1, whereas chains B (yellow), D (blue-green), and F (blue) to NONO. Note that the first and last chains (red and blue) do not contain the extended coiled-coil and CSAH regions. Figure prepared with Chimera (Pettersen et al. 2004).



FIGURE 3. Different views of a large paraspeckle protein multimer model containing 22 core dimers and colored according to chain ID. The modeled arrangement results in a spiraling relative orientation of the subsequent tandem RRM regions. Figures prepared with Chimera (Pettersen et al. 2004).

by the SOCKET algorithm as part of the coiled coil, consistent both with the overlapping predictions and with our earlier suggestion that CSAH segments might form weak coiled coils under certain circumstances (Süveges et al. 2009). Per-residue energy for the modeled extended coiled-coil region estimated with Coilcheck+ is 0.643 kJ/mol, whereas this value is -2.677 kJ/mol for the coiled coil region in the PSPC1/NONO dimer, indicating an inherent instability for the interdimer superhelix. We propose that the transition from a weak coiled coil to a single helix along the structure is more or less continuous in the case of the paraspeckle multimer. This inherent instability is consistent with the notion that the presence of the NEAT1 lncRNA is required for paraspeckle assembly (i.e., multimerization does not occur in the absence of the scaffold RNA).

In our model of the (PSPC1/NONO)₃ hexamer, the carboxy-terminal ends of the CSAHs form an approximately planar four-stranded helical structure with the helices of the coiled coil of the central dimer, close to the cavity in the core PSPC1/NONO dimers, occurring between their coiled-coil and RRM regions. We note that this general arrangement is preserved even when the extended coiled-coil region is modeled with different registers and/or pairings. Thus, we believe that this type of robustness of our model might indicate that it reasonably reflects the general architecture of the paraspeckle assembly—provided our basic assumptions are valid—even if the details of side-chain interactions are not modeled accurately. We note that our model does not contain the amino- and carboxy-terminal disordered segments of PSPC1 and NONO, but the modeled CSAH region is ideally located to guide the interaction between one core dimer and the carboxy-terminal unstructured fragment of its neighbor. Our exploratory molecular dynamics calculations lend support both for the flexibility of the extended coiled coil and the ability for the CSAH to make and organize additional contacts within the multimer (Supplemental Fig. S2).

Our proposed mode of coiled-coil-mediated multimerization of PSPC1/NONO dimers is different from that observed for the SFPQ crystal structure (Lee et al. 2015). Our model is based on the considerations of Passon and coworkers and thus features a right-handed coiled coil with a hendecad

repeat, whereas Lee et al. (2015) observed a short canonical left-handed coiled coil with a heptad pattern. DBHS proteins are highly versatile and, although PSPC1 and NONO were proposed to form elongated multimers in paraspeckles (Passon et al. 2012) and assemble on the NEAT1 lncRNA, SFPQ has been shown to form stable heterotetramers with NONO (Snijders et al. 2015). Thus, it can be expected that the mode of interaction between dimers can be different in the two cases. In summary, the most likely explanation for the

apparent inconsistencies between observations on DBHS protein assembly is that the extended coiled-coil/CSAH segment of DBHS proteins is highly versatile and can mediate a number of different interactions and multimerization modes in the different proteins in a context-dependent manner, as exemplified by the behavior of NONO/SFPQ tetramers and multimers depending on the ionic strength applied (Snijders et al. 2015). Moreover, NONO can bind to specific DNA segments and is involved in transcriptional activation (Park et al. 2013; Yadav et al. 2014). NONO was shown to bind double-stranded DNA independently of RNA and single-stranded DNA (Yang et al. 1993). The exact structural forms involved in these interactions can be yet again different from those already investigated and might represent another extension to the structural versatility of DBHS proteins.

Paraspeckles are assembled along NEAT1 lncRNA molecules that might also provide additional stabilizing interactions for the complex ribonucleoprotein particle. We suggest that the CSAH segments might act as a “ruler” by guiding the positioning of the core dimers along the RNA molecule. In this respect, the mechanical properties of CSAH segments might provide the required elasticity for the assembly and could help to adapt to the shape and flexibility of the long RNA molecule scaffolds. The concept of a number of regularly spaced RNA interaction domains, as proposed by Passon et al. (2012) and modeled here, is compatible with the high fraction of repeat motifs in the NEAT1 lncRNA (Supplemental Table S2). This hypothesis assumes that the primary site of RNA–protein interactions within the paraspeckle are the RRM domains, although other interactions cannot be ruled out. It is unlikely, however, that the CSAH segments themselves would be involved in RNA binding; thus, as discussed earlier, their prevalence in RNA-binding proteins should be rationalized with a different functional role. It should nevertheless be noted that recent studies revealed that methylation of arginines in the coiled-coil/CSAH region of NONO influences its association with mRNA molecules (Hu et al. 2015). Two of the three positions where Arg \rightarrow Lys mutations caused similar effects (Arg357 and Arg365) are included in our model and point toward the surface of the modeled assembly. Interestingly, methylated

arginines identified in SFPQ are not localized to the predicted coiled-coil/CSAH region (Snijders et al. 2015).

The shape of the (PSPC1/NONO)₃ assembly is elongated and can still be lengthened by adding additional dimers. Thus, large assemblies with variable stoichiometry can be formed (Fig. 3). This is consistent with the elongated cylindrical shape with variable length but a fixed diameter observed for paraspeckle structures, although the dimensions of the full particles are an order of magnitude larger than those of our model of a single assembly. In the model proposed by Nakagawa and Hirose (2012), such assemblies could be placed either along the radius of the particle making contact with multiple sites of a single lncRNA or along the long axis, binding a number of RNA molecules at identical sites. It should also be noted that the orientation of the proteins within the particle might be much different than speculated here—for example, the flexibility and/or different pairings of the extended coiled coil might allow a curvature perpendicular to the long axis of our model.

DBHS proteins are highly versatile and can participate in a number of diverse interactions. For example, NONO can form heterotetramers with SFPQ (Snijders et al. 2015), the architecture of which can be markedly different from those proposed here allowing for undefined stoichiometry. Our study suggests an indispensable steric role for CSAH segments in the protein family, and their presence might be essential for the structural and functional versatility of DBHS proteins.

MATERIALS AND METHODS

Sequence analysis

Paraspeckle proteins were identified with a standard BLASTP search (Altschul et al. 1990) on NCBI's "nr" database complemented by domain composition analysis. Coiled-coil and CSAH regions were predicted with coils (Lupas et al. 1991) and paircoil2 (McDonnell et al. 2006), as well as with SCAN4CSAH and FT_CHARGE (Süveges et al. 2009; Gáspári et al. 2012). RRM and NOPS domains were identified with HMMER (Eddy 2011) using the appropriate profiles in Pfam (accessions PF00076.17 and PF08075.6). All methods were run with default parameters except that the minimum lengths of CSAH segments were set to 30 in SCAN4CSAH. Similarity filtering was performed using CD-hit (Li and Godzik 2006).

Structural modeling

Our basic scheme is similar to that implemented in the recently prepared CCBuilder web server (Wood et al. 2014), published after our modeling pipeline was finalized. Backbone coordinates of the extended coiled coil were generated using Crick's parametric equations as described by Offer and Sessions (1995). Coiled-coil parameters were estimated by running a grid search by iterating over values of R_0 (superhelix radius) from 4.5 Å to 6.0 Å by steps of 0.05 Å and P (pitch) from 250 Å to 310 Å by steps of 1 Å. One chain of the coiled coil was fit to chain B of 3SDE over residues 262–303 with lsqman (Kleywegt and Jones 1994) and the RMSD (root-mean-square deviation)

of the other chain of the superhelix with respect to chain A of 3SDE (residues 274–315) was monitored with an in-house script.

Side-chains were added based on a residue library generated from antiparallel coiled-coil structures identified with SOCKET (Walshaw and Woolfson 2001) from PDB SELECT (2012 November release, Griep and Hobohm 2010) after discarding transmembrane proteins listed in the PDBTM database release 2013 September 27 (Kozma et al. 2013). Residues with incomplete side-chains were excluded. Residues were extracted and stored in diamide form—that is, the α and carbonyl carbon atoms of the preceding residue and the amide nitrogen and α carbon atoms of the next residue were also stored. These atoms were used when fitting the side-chains to the coiled-coil backbone. Residues with matching type and SOCKET-assigned heptad position were selected randomly and fit to the backbone by lsqman. The final coiled-coil model was generated from the backbone coordinates of the model and the side-chain coordinates (starting with the β carbon) of the fitted diamides.

An antiparallel coiled coil of 2×122 residues was generated and the side-chains were added to its central region by ensuring their appropriate pairings. During the construction of the final model, residues in the polyglycine region, not part of the modeled sequence, were omitted.

To build the hexamer, three 3SDE dimers and two modeled extended coiled-coil regions were put together by fitting the first six residues of the coiled-coil chains to residues 313–318 of chain A and residues 297–302 of chain B in 3SDE. The structures were merged so that the last residues from the dimeric cores were 315 from chain A and 299 from chain B, coordinates for rest of the modeled chains were taken from the fitted coiled-coil structures. This resulted in an acceptable geometry of the peptide bonds at the site of merging, as checked by the PDB validation server.

Stability of the SOCKET-assigned coiled-coil regions was estimated using the Coilcheck+ web server (Sunitha et al. 2012). Secondary structure elements were identified with DSSPCont (Andersen et al. 2002) and intersubunit interfaces in the multimeric structure were analyzed using the PISA web service (Krissinel and Henrick 2007). Energy minimizations and molecular dynamics calculations were performed with GROMACS 4.5.5 (Hess et al. 2008).

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

This research was supported by the Hungarian Scientific Research Fund (OTKA grants NF104198 and K108437). The European Union and the European Social Fund provided financial support for the project under grant agreement no. TÁMOP 4.2.1/B-11/2/KMR-2011-0004. Support was also provided by grant 9876-1/2015/FEKUT.

Received June 26, 2015; accepted August 27, 2015.

REFERENCES

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* **215**: 403–410.
- Andersen CAF, Palmer AG, Brunak S, Rost B. 2002. Continuum secondary structure captures protein flexibility. *Structure* **10**: 175–184.

- Baker EG, Bartlett GJ, Crump MP, Sessions RB, Linden N, Faul CFJ, Woolfson DN. 2015. Local and macroscopic electrostatic interactions in single α -helices. *Nat Chem Biol* **11**: 221–228.
- Chen L-L, Carmichael GG. 2009. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol Cell* **35**: 467–478.
- Chen L-L, DeCerbo JN, Carmichael GG. 2008. Alu element-mediated gene silencing. *EMBO J* **27**: 1694–1705.
- Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, Lawrence JB. 2009. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell* **33**: 717–726.
- Eddy SR. 2011. Accelerated profile HMM searches. *PLoS Comput Biol* **7**: e1002195.
- Fox AH, Lamond AI. 2010. Paraspeckles. *Cold Spring Harb Perspect Biol* **2**: a000687.
- Fox AH, Lam YW, Leung AKL, Lyon CE, Andersen J, Mann M, Lamond AI. 2002. Paraspeckles: a novel nuclear domain. *Curr Biol* **12**: 13–25.
- Gáspári Z, Süveges D, Perczel A, Nyitray L, Tóth G. 2012. Charged single α -helices in proteomes revealed by a consensus prediction approach. *Biochim Biophys Acta* **1824**: 637–646.
- Griep S, Hobohm U. 2010. PDBselect 1992–2009 and PDBfilter-select. *Nucleic Acids Res* **38**: D318–D319.
- Gruber M, Lupas AM. 2003. Another 50th anniversary—new periodicities in coiled coils. *Trends Biochem Sci* **28**: 679–685.
- Hess B, Kutzner C, van der Spoel D, Lindahl E. 2008. GROMACS4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* **4**: 435–447.
- Hu S-B, Xiang J-F, Li X, Xu Y, Xue W, Huang M, Wong CC, Sagum CA, Bedford MT, Yang L, et al. 2015. Protein arginine methyltransferase CARM1 attenuates the paraspeckle-mediated nuclear retention of mRNAs containing IRAlus. *Genes Dev* **29**: 630–645.
- Kleywegt GJ, Jones TA. 1994. A super position. *CCP4/ESF-EACBM Newsletter Protein Crystallogr* **31**: 9–14.
- Knight PJ, Thirumurugan K, Xu Y, Wang F, Kalverda AP, Stafford WF III, Sellers JR, Peckham M. 2005. The predicted coiled-coil domain of myosin 10 forms a novel elongated domain that lengthens the head. *J Biol Chem* **280**: 34702–34708.
- Kozma D, Simon I, Tusnády GE. 2013. PDBTM: Protein Data Bank of transmembrane proteins after 8 years. *Nucleic Acids Res* **41**: D524–D529.
- Krissinel E, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**: 774–797.
- Lee M, Sadowska A, Bekere I, Ho D, Gully BS, Lu Y, Iyer KS, Trehwella J, Fox AR, Bond CS. 2015. The structure of human SFOQ reveals a coiled-coil mediated polymer essential for functional aggregation in gene regulation. *Nucleic Acids Res* **43**: 3826–3840.
- Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**: 1658–1659.
- Lupas A, Van Dyke M, Stock J. 1991. Predicting coiled coils from protein sequences. *Science* **252**: 1162–1164.
- McDonnell AV, Jiang T, Keating AE, Berger B. 2006. Paircoil2: improved prediction of coiled coils from sequence. *Bioinformatics* **22**: 356–358.
- Nakagawa S, Hirose T. 2012. Paraspeckle nuclear bodies—useful or useless? *Cell Mol Life Sci* **69**: 3027–3036.
- Offer G, Knight P. 1996. Structure of the head-tail junction of the myosin molecule. *J Mol Biol* **256**: 407–416.
- Offer G, Sessions R. 1995. Computer modelling of the α -helical coiled coil: packing of side-chains in the inner core. *J Mol Biol* **249**: 967–987.
- Offer G, Hicks MR, Woolfson DN. 2002. Generalized Crick equations for modeling noncanonical coiled coils. *J Struct Biol* **137**: 41–53.
- Park Y, Lee J-M, Hwang M-Y, Son G, Geum D. 2013. NonO binds to the CpG island of oct4 promoter and functions as a transcriptional activator of oct4 gene expression. *Mol Cells* **35**: 61–69.
- Passon DM, Lee M, Rackham O, Stanley WA, Sadowska A, Filipovska A, Fox AH, Bond CS. 2012. Structure of the heterodimer of human NONO and paraspeckle protein component 1 and analysis of its role in subnuclear body formation. *Proc Natl Acad Sci* **109**: 4846–4850.
- Peckham M, Knight PJ. 2009. When a predicted coiled coil is really a single helix, in myosins and other proteins. *Soft Matter* **5**: 2493–2503.
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* **25**: 1605–1612.
- Sasaki Y, Ideue T, Sabo M, Mituyama T, Hirose T. 2009. Men ϵ/β noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc Natl Acad Sci* **106**: 2525–2530.
- Snijders AP, Hautbergue GM, Bloom A, Williamson JC, Minshull TC, Phillips HL, Mihaylov SR, Gjerde DT, Hornby DP, Wilson SA, et al. 2015. Arginine methylation and citrullination of splicing factor proline- and glutamine-rich (SFPQ/PSF) regulates its association with mRNA. *RNA* **21**: 347–359.
- Spudich SA, Sivaramakrishnan S. 2010. Myosin VI: an innovative motor that challenged the swinging lever arm hypothesis. *Nat Rev Mol Cell Biol* **11**: 128–137.
- Staub E, Fizev P, Rosenthal A, Hinzmann B. 2004. Insights into the evolution of the nucleolus by an analysis of its protein domain repertoire. *Bioessays* **26**: 567–581.
- Sunitha M, Nair AG, Charya A, Jadhav K, Mukhopadhyay S, Sowdhamini R. 2012. Structural attributes for the recognition of weak and anomalous regions on coiled coils of myosins and other motor proteins. *BMC Res Notes* **5**: 530.
- Süveges D, Gáspári Z, Tóth G, Nyitray L. 2009. Charged single α -helix: a versatile protein structural motif. *Proteins* **74**: 905–916.
- Swanson CJ, Sivaramakrishnan S. 2014. Harnessing the unique structural properties of isolated α -helices. *J Biol Chem* **289**: 25460–25467.
- Testa D, Moutevelis E, Woolfson DN. 2009. CC+: a relational database of coiled-coil structures. *Nucleic Acids Res* **37**: D315–D322.
- Walshaw J, Woolfson DN. 2001. SOCKET: a program for identifying and analyzing coiled coil motifs within protein structures. *J Mol Biol* **307**: 1427–1450.
- Wolny M, Batchelor M, Knight PJ, Paci E, Dougan L, Peckham M. 2014. Stable single α -helices are constant force springs in proteins. *J Biol Chem* **289**: 27825–27835.
- Wood CW, Bruning M, Ibarra AA, Bartlett GJ, Thomson AR, Sessions RB, Brady RL, Woolfson DN. 2014. CCBUILDER: an interactive web-based tool for building, designing and assessing coiled-coil protein assemblies. *Bioinformatics* **30**: 3029–3035.
- Yadav SP, Hao H, Yang H-J, Kautzmann M-AI, Brooks M, Nellissery J, Klocke B, Seifert M, Swaroop A. 2014. The transcription-splicing protein NonO/p54nrb and three NonO-interacting proteins bind to distal enhancer region and augment rhodopsin expression. *Hum Mol Genet* **23**: 2132–2144.
- Yang Y-S, Hanke JH, Carayannopoulos L, Craft CM, Capra JD, Tucker PW. 1993. NonO, a non-POU-domain containing, octamer-binding protein, is the mammalian homolog of *Drosophila* nonAdiss. *Mol Cell Biol* **13**: 5593–5603.