Crystal structure of a SFPQ/PSPC1 heterodimer provides insights into preferential heterodimerization of human DBHS family proteins

Jie Huang, G. Patricia Casas Garcia, Matthew A. Perugini, Archa H. Fox, Charles S. Bond, and Mihwa Lee

From the Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria 3086 and the School of Human Sciences and the School of Molecular Sciences, The University of Western Australia, Crawley, Western Australia 6009, Australia

Members of the Drosophila behavior human splicing (DBHS) protein family are nuclear proteins implicated in many layers of nuclear functions, including RNA biogenesis as well as DNA repair. Definitive of the DBHS protein family, the conserved DBHS domain provides a dimerization platform that is critical for the structural integrity and function of these proteins. The three human DBHS proteins, splicing factor proline- and glutamine-rich (SFPQ), paraspeckle component 1 (PSPC1), and non-POU domain–containing octamer-binding protein (NONO), form either homo- or heterodimers; however, the relative affinity and mechanistic details of preferential heterodimerization are yet to be deciphered. Here we report the crystal structure of a SFPQ/PSPC1 heterodimer to 2.3-Å resolution and analyzed the subtle structural differences between the SFPQ/PSPC1 heterodimer and the previously characterized SFPQ homodimer. Analytical ultracentrifugation to estimate the dimerization equilibrium of the SFPQ-containing dimers revealed that the SFPQ-containing dimers dissociate at low micromolar concentrations and that the heterodimers have higher affinities than the homodimer. Moreover, we observed that the apparent dissociation constant for the SFPQ/PSPC1 heterodimer was over 6-fold lower than that of the SFPQ/NONO heterodimer. We propose that these differences in dimerization affinity may represent a potential mechanism by which PSPC1 at a lower relative cellular abundance can outcompete NONO to heterodimerize with SFPQ.

The Drosophila behavior human splicing (DBHS) family of proteins are ubiquitous nuclear proteins implicated in many aspects of nuclear functions. Defined by the highly conserved DBHS region that consists of two tandem RNA recognition motifs (RRMs), a NonA/paraspeckle domain (NOPS) and a C-terminal coiled-coil, the members of this protein family have been reported as multifunctional proteins, playing important roles in RNA biogenesis and transport, and subnuclear body (paraspeckle) formation via direction interaction with structural noncoding RNA as well as DNA damage repair.

The DBHS family of proteins are exclusively found in the animal kingdom; invertebrate or lower vertebrate typically encode a single DBHS protein, whereas higher vertebrates encode more than one. In humans, three DBHS proteins: SFPQ, PSPC1, and NONO, share greater than 70% identity within the DBHS region, flanked by low complexity N- and C-terminal domains that vary in sequence and length among the three paralogues (Fig. 1). DBHS proteins are obligatory dimers; the three human DBHS proteins form either homodimer or heterodimers among the three proteins in vitro (3–5) and all three possible heterodimers (SFPQ/NONO, SFPQ/PSPC1, and PSPC1/NONO) have been shown in mouse Sertoli cells.

The dimeric nature of the DBHS family of proteins has been confirmed by the crystal structures of human NONO/PSPC1 heterodimer (7), SFPQ homodimer (5), and Caenorhabditis elegans NONO-1 (2) in our previous studies. The crystal structures of the DBHS proteins consistently revealed an intimate dimer interface contributed by all four domains within the DBHS region. The dimer formation is largely driven by the hydrophobic interactions of the second RRM with the NOPS domain as well as the antiparallel coiled-coil interaction, whereas the first RRM is dispensable in dimer formation (2, 5, 7). Dimerization of the DBHS proteins is not only critical to their structural integrity but also to their functions such that dimerization incompetent DBHS proteins generated by mutation of key hydrophobic residues failed to form subnuclear bodies termed paraspeckles (7).

Combinatorial heterodimerization among the three DBHS proteins diversifies the functional repertoire and also contributes functional overlap and redundancy, evidenced by functional compensation of the loss of DBHS protein by the other. This work was supported by the Australian Research Council Discovery Early Career Research Award (ARC DECA) Fellowship DE150101243 (to M. L.) and National Health and Medical Research Council of Australia (NHMRC) Project Grants 1050585 and 513880 (to A. H. F. and C. S. B.). The authors declare that they have no conflicts of interest with the contents of this article.

This article contains Figs. S1–S3 and Tables S1 and S2.
The atomic coordinates and structure factors (code SWPA) have been deposited in the Protein Data Bank (http://wwpdb.org/).

1 To whom correspondence should be addressed. Tel.: 61-3-9479-3248; E-mail: mihwa.lee@latrobe.edu.au.

2 The abbreviations used are: DBHS, Drosophila behavior human splicing; NONO, non-POU domain–containing octamer-binding protein; NOPS domain, NonA/paraspeckle domain; PSPC1, paraspeckle protein component 1; RRM, RNA recognition motif; SFPQ, splicing factor proline- and glutamine-rich; r.m.s., root mean square; PDB, Protein Data Bank; TEV, tobacco etch virus.

This is an Open Access article under the CC BY license.
paralogues. Specifically, NONO deficiency in murine embryonic fibroblasts derived from NONO knockout mice led to spontaneous up-regulation of PSPC1, which formed a complex with SFPQ in place of NONO and compensated the loss of NONO in DNA repair (8). Similar up-regulation of SFPQ and PSPC1 was observed in patients with syndromic intellectual disability with little or no NONO protein expression due to the mutations in NONO (9). The data suggest that up-regulation of SFPQ and PSPC1 may compensate the down-regulated NONO at least partially; however, the distinct contributions of NONO as in SFPQ/NONO or PSPC1/NONO are irreplaceable by SFPQ/PSPC1 in regulation of synaptic transcription, leading to intellectual disability (9). Taken together, these observations indicate that dimerization among the three DBHS paralogues is likely interdependent as well as spatially and temporally dynamic. Thus, different DBHS heterodimers may have different tissue- or cell-type specific functions, contributing to the myriad of nuclear functions assigned to the DBHS proteins. However, the relative affinity and mechanistic details of the combinatorial dimer states are still unclear.

As a first step toward better understanding the relative affinity and mechanism of combinatorial dimerization of the three human DBHS proteins, we have carried out analytical ultracentrifugation on the SFPQ containing dimers (SFPQ homodimer, SFPQ/PSPC1, and SFPQ/NONO). In addition, we have determined the crystal structure of SFPQ/PSPC1 and analyzed the structural differences of the SFPQ/PSPC1 heterodimer with the previously characterized SFPQ homodimer.

Results

Preferential heterodimerization of DBHS family proteins

In our previous studies, we defined the dimerization domain of the DBHS proteins, encompassing the two RRMs, NOPS, and part of the coiled-coil domain (4, 5, 7) (Fig. 1). Using the truncated constructs containing the dimerization domain, we have shown that the three human DBHS proteins can form homodimers (SFPQ (5), NONO (10), and PSPC1 (PDB code 5IFN) as well as heterodimers (NONO/PSPC1 (7) and SFPQ/NONO (5) in vitro.

In this study, His-tagged PSPC1 (residues 61–320) was co-expressed with untagged SFPQ (residues 276–535) and the resulting SFPQ/PSPC1 heterodimer was successfully purified, confirming that the three human DBHS proteins can form all three possible heterodimers. One of the important observations from the course of purification of all six possible dimeric complexes (three homodimers and three heterodimers) is that human DBHS proteins display a preference for heterodimerization; when two DBHS proteins are co-expressed, the less expressed protein is exclusively found in the heterodimer (supporting Fig. S1), indicating that dimerization affinities of heterodimers are stronger than those of homodimers and/or heterodimers are more stable than homodimers.

Differential dissociation constants of SFPQ-containing dimers

As a first step toward characterizing the relative affinities for dimer formation among the SFPQ-containing dimers, we conducted sedimentation velocity experiments in the analytical ultracentrifuge comparing all three possible SFPQ-containing dimers: SFPQ homodimer, SFPQ/PSPC1, and SFPQ/NONO heterodimers at an initial total protein concentration of 5 μM. All three dimers have the same construct boundaries defined by the dimerization domain (RRM1, RRM2, NOPS, and partial coiled-coil) (Fig. 1). It should be noted that the exclusion of the extended coiled-coil domain beyond the dimerization domain was necessary to avoid complication of data analysis because of oligomerization mediated by the extended coiled-coil (5).

The data from sedimentation velocity experiments were fitted to a continuous size (c(s)) distribution model and the results demonstrated that all three samples at an initial concentration of 5 μM exist as two major species in solution with sedimentation coefficients of 2.3 and 3.8 S, consistent with monomer and dimer, respectively (Fig. 2, A–C, Table 1). There is also a small proportion of a higher order oligomer with a modal sedimentation coefficient of ~5.7 S, particularly in the SFPQ homodimer sample (Fig. 2A). More importantly, the relative proportion of dimers and monomers differed in each sample with significantly more dimeric species present for the heterodimeric complexes, particularly SFPQ/PSPC1 (Fig. 2C), as compared with the SFPQ homodimer, indicating that the dimerization affinities of the SFPQ-containing heterodimers are indeed stronger than that of the SFPQ homodimer.

This assertion was confirmed by performing sedimentation equilibrium experiments at multiple rotor speeds and initial protein concentrations. The data were analyzed by a global nonlinear regression method and best fitted to a monomer-dimer equilibrium for the SFPQ homodimer and to the 1:1 complex (i.e. A + B ⇌ AB) for the SFPQ-containing heterodimers. The resulting apparent dissociation constants showed that the SFPQ/PSPC1 heterodimer (KD,AB→A+B of 0.55 μM) is almost 20-fold tighter than the SFPQ homodimer (KD,2→1 of 10.54 μM) (Fig. 2, D–F, Table 1). By contrast, the SFPQ/NONO heterodimer (KD,AB→A+B of 3.65 μM) is ~3-fold tighter than the SFPQ homodimer (Table 1). A noticeable difference (>6-fold) in the apparent dissociation constant between the two SFPQ-containing heterodimers is also observed (Table 1).

Crystal structure of a SFPQ/PSPC1 heterodimer

To gain insight into the preferential heterodimerization of human DBHS proteins, we have crystallized the SFPQ/PSPC1 heterodimer encompassing the dimerization domain (Fig. 3A...
The structure of the SFPQ/PSPC1 heterodimer was solved by molecular replacement, and refined to 2.3-Å resolution (Fig. 3). Despite high sequence similarity between the two proteins within the dimerization domain (74 and 93% sequence identity and similarity, respectively) (supporting Fig. S2), several key residues that are different in the two proteins allowed unambiguous assignment of SFPQ and PSPC1 (supporting Fig. S3). The final model consists of one heterodimer in the asymmetric unit.

Consistent with the previously characterized structures of human DBHS proteins, the structure of the SFPQ/PSPC1 heterodimer displays a pseudosymmetrical dimer, reflecting high sequence identity within the dimerization domain (Fig. 3). The second RRM domain (RRM2), NOPs, and coiled-coil domain provide the major dimerization interface. Largely hydrophobic in nature, the interactions of the NOPs domain with the RRM2 and coiled-coil domain of the dimer partner are followed by the right-handed coiled-coil interactions between the two proteins.

The dimerization domain without the first RRM domain superposes well with each other (root mean square difference (r.m.s. deviation) of 1.05 Å for 146 common C\(^\pm\)H9251 positions), whereas the r.m.s. deviation for the overall superposition of the two protein chains is 2.53 Å (224 C\(^\pm\)H9251). The major deviation...
tion from the 2-fold symmetry is due to the arrangement of the two N-terminal RRM1 domains relative to each other. Although the RRM1 domains from each protein superpose well individually (r.m.s. deviation of 0.685 Å, 77 Cα), the observed rotation angle of the two RRM1 domains is 170.4°, breaking the 2-fold symmetry. In addition, the relative position of RRM1 in SFPQ to the dimer core is tilted by 27° toward the dimer core compared with that of RRM1 of PSPC1, contributing to further deviation from the 2-fold symmetry. This observation is consistent with the structure of a PSPC1/NONO heterodimer (PDB code 3SDE), the first DBHS protein structure we have characterized, where similar pseudo-symmetry was observed with the major deviation from the 2-fold symmetry contributed by the dispositions of the two RRM1 domains (7). In contrast, in the structure of the SFPQ homodimer (PDB code 4WII), the two RRM1 domains are positioned in a more symmetrical way (rotation angle of 178.8°) with no significant deviation in the dispositions of each RRM1 to the dimer core (Fig. 4A and B) (5). It is, therefore, not surprising that the use of the structure of PSPC/NONO (PDB code 3SDE) as a search model in phasing by molecular replacement found a straightforward structure solution, whereas a more complicated search strategy was required to produce readily interpretable electron density maps when the SFPQ homodimer (PDB code 4WII) was applied as a search model.

Structural comparison between the SFPQ/PSPC1 heterodimer and SFPQ homodimer

Due to high sequence similarity between SFPQ and PSPC1 within the dimerization domain (74 and 93% sequence identity and similarity, respectively), the overall structure of the SFPQ/PSPC1 heterodimer is indistinguishable from that of the SFPQ homodimer (PDB code 4WII) (Fig. 4A) (5). Analysis of interface using PISA (11) shows the common principal interface involving all four domains in both structures with over 40% of residues directly involved in the dimer interface (supporting Table S1). The principal dimer interface is largely driven by hydrophobic interactions evidenced by a highly negative gain of solvation energy on complex formation (Δ′G = −50.4 kcal/mol with p value 0.027 for the SFPQ/PSPC1 heterodimer; Δ′G = −46.1 kcal/mol with p value 0.061 for the SFPQ homodimer) as calculated by PISA (11).

Closer inspections on the amino acid sequences of the two proteins and superposition of the structures of the SFPQ-containing dimers, however, reveal subtle differences in local primary and tertiary structures. Although most of the primary sequence variations are scattered throughout the dimerization domain, four regions that have a stretch of amino acids with sequence variations are observed (supporting Fig. 2A). With exception of Region 2, these regions are involved in dimer interfaces with the resulting local structures showing subtle differences in the two structures; Region 1, located prior to the RRM1, contributes to the asymmetry of the two RRM1 arrangements described in the previous section, whereas Regions 3 and 4 are located in the NOPS and C-terminal coiled-coil, respectively, where further local structural differences are observed (Fig. 4A). Although most of the intermolecular interactions in these regions made by networks of hydrogen bonds and salt
bridges are conserved in both structures, subtle variations in hydrophobic interactions are observed (supporting Table S2) and detailed below.

Region 1 (residues 276–287 in SFPQ; 61–80 in PSPC1), showing the most sequence divergence between SFPQ and PSPC1, takes on a different secondary structure. In the struc-

Figure 4. Comparison of the SFPQ/PSPC1 heterodimer with the SFPQ homodimer. A, superposition of the SFPQ (blue)/PSPC1 (yellow) heterodimer with the SFPQ homodimer (brown, PDB code 4WII). Regions 1–4 harboring a stretch of amino acids with sequence variations and the associated local differences in structure are indicated. B, comparison of the two RRM arrangement in SFPQ/PSC1 (left, SFPQ in blue and PSPC1 in yellow) and SFPQ homodimer (right, brown). C, close-up view of the interface of the two RRMs in B; D and E, structural plasticity of the NOPS domain. NOPS domains are colored in yellow (PSPC1) and cyan (SFPQ), whereas interacting RRM2 from the partner subunit (RRM2) shown in blue (SFPQ) and pink (PSPC1) in D. The corresponding regions in the SFPQ homodimer (PDB 4WII) are shown in E, with the NOPS domains in salmon pink (Chain A) and gray (Chain B) and the interacting RRM2 in brown (Chain B) and green (Chain A). F, critical dimerization residues in the NOPS domain (boxed in A) show different side chain conformations in the SFPQ/PSPC1 heterodimer. The side chains of the conserved aromatic residues in NOPS are superposed; side chains for SFPQ shown in blue with its counterpart in dark gray with residue numbering for SFPQ. G–I, close-up views of the NOPS and RRM2 interactions in the SFPQ homodimer (G) and SFPQ/PSPC1 heterodimer (H and I) show variations in hydrophobic residues involved in the dimer interface. Italic residue numbers belong to PSPC1, whereas residue numbers with prime are for the partner subunit of the SFPQ homodimer.
Crystal structure of a SFPQ/PSPC1 heterodimer

ture of the SFPQ homodimer (PDB code 4WII), this region forms a short α-helical structure that is not involved in the dimer interface (5). The corresponding residues of PSPC1 in the structure of the SFPQ/PSPC1 heterodimer form a strand with no secondary structure. Interestingly, an additional hydrophobic residue (Phe-70) in Region 1 of PSPC1 participates in the dimer interface between the two RRM1 domains in the heterodimer, contributing to the asymmetry in the arrangement of the RRM1 domains (Fig. 4, B and C). The different hydrophobic environment associated with this asymmetry at the interface of the two RRM1 domains may contribute toward favoring heterodimerization.

Region 3 (residues 464–471 in SFPQ; 249–256 in PSPC1), located in the middle of the NOPS domain, appears to be flexible with different conformations observed in the two structures (Fig. 4, A, D, and E). The electron density of Region 3 of SFPQ (residues 462–474) is in fact disordered in the structure of the SFPQ/PSPC1 heterodimer and thus the residues in this region are not modeled in the final structure (Fig. 4D). Side chains of the majority of the residues in Region 3 of PSPC1 and the SFPQ homodimer are found in the solvent-accessible area, thus not contributing to the dimer interface. However, the first residue of Region 3 (Ala-464 in SFPQ and Met-249 in PSPC1) is directly involved in the dimer interface through hydrophobic interaction with a residue from RRM2 of the dimer partner (Fig. 4, D and E). Interestingly, the partner hydrophobic residue in RRM2 shows amino acid variation: Tyr-381 in SFPQ, whereas Val-166 in PSPC1. The resulting hydrophobic interaction in the heterodimer shows Met-249 of PSPC1 interacting with Tyr-381 from RRM2 of SFPQ, whereas Ala-464 in Region 3 of SFPQ (not observed in the SFPQ/PSPC1 structure) is likely to interact with Val-166 from RRM2 of PSPC1 when extrapolated from the structure of the SFPQ homodimer (Fig. 4, D and E).

Region 4 (residues 525–535 in SFPQ; 310–320 in PSPC1) is located at the end of anti-parallel coiled-coil interface. Although all of the residues of Region 4 in PSPC are observed, residues beyond 527 of SFPQ in the heterodimer structures are not observed with a similar observation made in the SFPQ homodimer. In addition, the majority of the residues in this region belong to the beginning of the “coiled-coil interaction motif” (residues 528–555 in SFPQ) characterized in our previous study, which mediates polymerization of the SFPQ dimers (5). Therefore, it is unlikely that the sequence variations in this region directly influence the dimer interface. Variations in these residues may have, if any, impacts on choice of polymerization partners among the DBHS dimers, which are yet to be characterized.

However, closer inspection of the dispositions of the distal coiled-coil domain reveals that highly conserved residues at the C-terminal end of the NOPS domain (Phe-486, Tyr-490, and Trp-494 for SFPQ; Phe-271, Tyr-275, and Trp-279 for PSPC1) take on drastically different side chain conformations in the two structures (Fig. 4, D–F). Although the structure of the SFPQ homodimer shows similar side chain conformations for these residues in each chain of the dimer (Fig. 4E), the side chains of the corresponding residues take on drastically different conformations in the heterodimer setting (Fig. 4, D and F). Importantly, two of these residues (Tyr-275 and Trp-279 for PSPC1 numbering) have been shown previously to be critical for DBHS dimerization; mutation of these residues to alanine in NONO and PSPC1 prevents homo- and heterodimer formation in yeast two-hybrid analyses (7). Similar observations on the different side chain conformations of these residues were also made in the structure of a NONO/PSPC1 structure. Thus, observation of different side chain conformations in the structure of the SFPQ/PSPC1 heterodimer further reinforces the idea that structural plasticity associated with these residues may influence choice of dimerization partner.

Due to the intimate interaction between NOPS and RRM2 of the partner DBHS protein (RRM2’), we further inspected the dimer interface between NOPS and RRM2’ and found a cluster of hydrophobic interactions between NOPS and RRM2’ that show differences in the amino acid identity (Fig. 4, G–I). The residues involved in this hydrophobic pocket belong to the RRM2’ (Val-435 for SFPQ numbering) and the distal end of NOPS (Ala-481, His-483, Tyr-488, and Ser-491 for SFPQ numbering). This region of NOPS also harbors the three conserved aromatic residues mentioned above that show different side chain conformations. The two clusters of hydrophobic residues are found in the different faces of the helix, however; whereas the three conserved aromatic residues face the distal coiled-coil domain of the partner protein (CC’), two of the residues (Tyr-488, and Ser-491 for SFPQ numbering) involved in the latter hydrophobic pocket are located in the outer face of the helix with the other two residues (Ala-481 and His-483 for SFPQ numbering) sitting on the hinge region before the beginning of the helix. With the exception of the alanine residue (Ala-481 for SFPQ; Ala-266 PSPC1), the amino acid identity of the four residues involved in this hydrophobic pocket varies among the three DBHS proteins. In combination with the three conserved aromatic residues, the variations of amino acid sequence in this hydrophobic pocket may contribute to the apparent differences in the dissociation constants of the homodimer and heterodimers.

In summary, it appears that preferential heterodimerization of the SFPQ-containing dimers is governed by synergistic effects of subtle variations in the local structures observed in the dimer interfaces encompassing all four domains (two RRM, NOPS, and coiled-coil) within the dimerization domain.

Discussion

Since the first identification and characterization of members of the DBHS protein family in the early 1990s, a myriad of nuclear functions have been designated to this family of proteins (1). Although exhibiting a large overlap of functions, the three human DBHS paralogues, SFPQ, PSPC1, and NONO, are incapable of complete functional compensation after loss of the other members in some cases (9). This suggests specific roles for each paralogue in different temporal and spatial contexts, presumably acting through the divergent N and C termini flanking the conserved DBHS domain. It has been established that the functional unit of the DBHS proteins is a dimer (3–5, 7); the three paralogues, a product of evolutionary duplication and divergence of an ancestral gene, homo- and heterodimerize with each other to add extra layers to the
functional complexities of the DBHS proteins in higher vertebrates.

In this study, built on our previous studies where preference to heterodimerization has been observed from the co-expression and purification of the three human DBHS proteins using a bacterial expression system, we confirmed through sedimentation analyses that the dimerization affinity for the SFPQ-containing heterodimers is indeed tighter than that of the SFPQ homodimer. Surprisingly, the apparent dissociation constant for the SFPQ/PSPC1 heterodimer is more than 19-fold lower than that of the SFPQ homodimer, whereas the SFPQ/NONO heterodimer is ∼3-fold tighter than the SFPQ-homodimer, suggesting that the majority of DBHS proteins are most likely to exist as heterodimers in vivo. This imposes significant implications in interpretation of the past literature on the human DBHS proteins where individual proteins have been addressed separately to probe their nuclear function. Thus, future work may need to revisit and validate the roles assigned to each individual paralogue taking into account their dimerization states.

The observed 19-fold difference in the apparent dissociation constant between the SFPQ/PSPC1 heterodimer and the SFPQ homodimer was unexpected given that the two paralogues share high sequence similarity within the dimerization domain. Reflecting this similarity, the structural determinants for preferential heterodimerization of SFPQ/PSPC1 compared with the homodimeric SFPQ cannot be attributed to a few key residues; contributed by all four domains (RRM1, RRM2, NOP5, and coiled-coil domain), synergistic effects from small variations in the identity or conformation of many residues largely in the hydrophobic pockets of the dimerization interface appear to favor heterodimerization.

Intrinsic differences in dimerization affinity among all six possible human DBHS dimers are most likely to be the major determinant of the dimerization partner choice in vivo. However, other factors including the differential expression level of the three paralogues and post-translational modifications may also contribute to the equilibrium and dynamics of homo- and heterodimerization in vivo.

Although all three paralogues are ubiquitously expressed in many cell types, relative expression levels are different; SFPQ is highly expressed across all tissue types, whereas varied expression levels of NONO and PSPC1 are observed in the Human Protein Atlas (www.proteinatlas.org)3 (12). In particular, the expression level of PSPC1 is highly tissue- and cell-type specific; very low expression of PSPC1 is observed in heart muscle and ovary tissue. In HeLa cells, PSPC1 is much less abundant than SFPQ and NONO (6), whereas a similar expression level of PSPC1 was observed in mouse Sertoli cell (13). In addition, the presence of multiple transcript variants producing isoforms of each paralogue adds extra layers of complexities in the combinatorial dimers of the three DBHS paralogues. Some of these variants lack a nuclear localization signal located at the C terminus (14, 15), suggesting that localization of DBHS proteins can be altered depending on the identity and isoform of the dimerization partner. This is particularly relevant with the emerging function and importance of the cytoplasmic pool of DBHS family proteins in neuronal development. Localization of the DBHS family proteins is predominantly nuclear; however, recent studies have shown that a non-nuclear pool of SFPQ is essential for axon motor development and viability (16, 17).

The effects of post-translational modifications on dimerization of DBHS proteins are yet to be characterized; however, various post-translational modifications on DBHS proteins including phosphorylation and methylation have been shown to alter their nuclear activities and localization of the proteins from the nucleus to the cytoplasm in some cases (18–20). It is plausible that the altered cellular activities and localization of DBHS proteins may directly reflect altered dimerization due to post-translational modifications.

In summary, we demonstrate that the apparent dissociation constants of the SFPQ-containing heterodimers are lower than that of the SFPQ homodimer, confirming that heterodimerization is indeed preferred to homodimerization among the SFPQ-containing dimers. A noticeable difference (>6-fold) in the apparent dissociation constant between the SFPQ/PSPC1 and SFPQ/NONO heterodimers is also observed. These differences in dimerization affinity may represent a potential mechanism by which less expressed PSPC1 can outcompete more abundant NONO to heterodimerize with the most abundant SFPQ in many cell and tissue types. Despite the noticeable difference in the apparent dissociation constants, very subtle differences in the dimerization interfaces are observed in the crystal structure of the SFPQ/PSPC1 heterodimer compared with that of the SFPQ homodimer, reinforcing the concept that the synergistic effects of subtle variations in the sequence and local structures result in stronger affinity in the DBHS heterodimers.

**Experimental procedures**

**Expression plasmid construction**

Human PSPC1 (residues 61–320) was cloned into the first multiple cloning site of pETDuet-1 (Novagen) using BamHI and EcoRI sites. A tobacco etch virus (TEV) protease cleavage site was engineered into the 5′ primer to enable removal of the His tag at a later stage of purification. Human SFPQ (residues 276–535) was cloned into pCDF-13 (EMBL) using NcoI and XhoI. The sequences were verified by DNA sequencing. Construction of pCDF-11-SFPQ(276–535) (5) and pETDuet-1-NONO(53–312) was previously described (7).

**Protein expression and purification**

For coexpression and purification of the SFPQ/PSPC1 heterodimer, pETDuet-1–PSPC1(61–320) and pCDF-13–SFPQ(276–535) were co-transformed into Rosetta2 (DE3) (Novagen). The cells were grown in LB growth medium supplemented with 100 μg ml−1 of ampicillin, 50 μg ml−1 of spectinomycin, and 50 μg ml−1 of chloramphenicol and protein expression was induced with isopropyl 1-thio-β-D-galactopyranoside (0.5 mM) at an OD of 0.8–1.0. The cells were harvested after further incubation at 25 °C for 16 h. The cell pellet was suspended in buffer A (20 mM Tris-HCl (pH 7.5), 250 mM NaCl, 1 mM D-

---

3 Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party hosted site.
Crystal structure of a SFPQ/PSPC1 heterodimer

10% glycerol) and disrupted using a TS series benchtop cell disruptor (Constant Systems Ltd.). The soluble fraction was applied to a nickel-chelating column (Hi-Trap HP, GE Healthcare) and the bound protein was eluted by applying a linear gradient of 25–500 mM imidazole in buffer A. Fractions containing His-PSPC1/SFPQ were pooled and subjected to TEV protease digestion with 1:100 ratio (w/w) at 4 °C overnight. Uncleaved heterodimer and His-TEV protease were removed by the second nickel-affinity chromatography. Flow-through fractions containing the cleaved PSPC1/SFPQ heterodimer were concentrated and applied to a HiLoad 16/600 Superdex 200 size exclusion column, pre-equilibrated with 20 mM Tris-HCl (pH 7.5), 0.25 M NaCl. Expression and purification of the SFPQ homodimer (residues 276–535) and SFPQ(276–535)/NONO(53–312) heterodimer was described previously (5). The purified proteins were typically concentrated to 7–10 mg/ml and stored at −80 °C until further use.

Sedimentation velocity

Sedimentation velocity experiments were conducted at 20 °C using a Beckman Coulter XL-A or XL-I analytical ultracentrifuge with an An-50 Ti 8-hole rotor using similar methods reported previously (21, 22). Briefly, double-sector quartz cells were loaded with 380 μl of protein samples at an initial concentration of 5 μM and 400 μl of reference buffer (20 mM Tris-HCl (pH 7.5), 500 mM NaCl). Data were collected at 230 nm and rotor speed of 40,000 rpm in continuous mode using a using a step size of 0.003 cm without averaging. The partial specific volume of SFPQ, SFPQ/NONO, and SFPQ/PSPC1, buffer density, and buffer viscosity were computed using SEDNTERP (23). Data were analyzed using the continuous size-distribution model employing the program SEDFIT (24–27) with a continuous size distribution model.

Sedimentation equilibrium

Sedimentation equilibrium experiments were performed in a Beckman Coulter XL-I analytical ultracentrifuge in an An-50 Ti 8-hole rotor using similar methods reported previously (21, 22). Briefly, 120 μl of protein samples (0.5–15 μM) and 140 μl of reference buffer (20 mM Tris-HCl (pH 7.5), 500 mM NaCl) were centrifuged at rotor speeds of 12,000 and 18,000 rpm until sedimentation equilibrium was attained (t = 36 and 48 h for each speed). Experiments were carried out using a step size of 0.001 cm with 10 averages at a wavelength optimized at 234 nm. Data were fitted to a monomer-dimer equilibrium model using SEDPHAT (21, 22, 28).

Crystallization and X-ray diffraction data collection

Crystals of the SFPQ/PSPC1 heterodimer were grown by hanging-drop vapor diffusion experiments at 20 °C. 1.5 μl of SFPQ/PSPC1 (10 mg/ml) was mixed with 1.5 μl of reservoir solution (50 mM MES (pH 5.5), 0.2 M NaCl, 3% (w/v) amino-hexanoic acid, 25% (w/v) PEG3350) and equilibrated against 0.5 ml of the reservoir solution. Crystals were transferred briefly to a drop containing artificial reservoir solution with 30% (w/v) glucose before cryocooling. Diffraction data were recorded on beamline MX2 at the Australian Synchrotron at a wavelength of 0.954 Å at 100 K. The data were processed with XDS (29), and merged and scaled with AIMLESS (30). Data collection and merging statistics are summarized in Table 2.

Table 2

| SFPQ/PSPC1 heterodimer |
|-------------------------|
| **Data collection** |
| Space group | C2 |
| Unit cell parameters (Å, °) | 89.2, 67.8, 92.2 β = 95.1 |
| Resolution (Å) | 26.2–2.29 (2.39–2.29) |
| No. of observations | 89,475 (9,327) |
| No. of unique reflections | 24,463 (2,534) |
| Completeness (%) | 98.6 (97.0) |
| Redundancy | 3.7 (3.7) |
| Rmerge (%) | 3.9 (29.0) |
| Rfree (%) | 3.6 (26.7) |
| CC1/2 | 0.998 (0.954) |
| Average I/σ(I) | 12.8 (2.5) |

| **Refinement** |
| R (%) | 21.9 (27.5) |
| Rfree (%) | 27.2 (29.9) |
| No. (%) of reflections in test set | 1223 (5.0) |
| No. of protein molecules per asymmetric unit | 2 (Chain A: SFPQ; Chain B: PSPC1) |
| R.m.s.d bond length (Å) | 0.01 |
| R.m.s.d bond angle (°) | 1.05 |
| Average B-factors (Å²) | 57.4 |
| Protein molecules | 57.6 |
| Water molecules | 51.4 |
| Ramachandran plot* |
| Residues other than Gly and Pro in:
Most favored regions (%) | 97.0 |
| Additional allowed regions (%) | 3.0 |
| Disallowed regions (%) | 0 |

PDB code | SWPA |
---|---|
*Calculated by BAVERAGE in CCP4 Suite (32). |
*Calculated using MolProbity (35). |

Structure solution and refinement

The crystal structure of the SFPQ/PSPC1 heterodimer was solved by molecular replacement using PHASER (31) within the CCP4 suite (32). The structure of the PSPC1/NONO heterodimer (PDB code 3SDE (7)) was used as search model after removing all nonprotein atoms. The structure of PSPC1/NONO found a solution with a log likelihood gain of 792 and a Z score of 25.3, locating one dimer in the asymmetric unit. Difference Fourier maps (Fo − Fc) from the first round of refinement with autoBUSTER (33) revealed several key residues that are different among SFPQ, PSPC1, and NONO, which allowed unambiguous assignment of SFPQ and PSPC1. At this stage, the correct amino acid sequence was incorporated, and iterative model building with COOT (34) and refinement with autoBUSTER was carried out. The final model consists of two chains with SFPQ in Chain A (residues 290–461 and 475–527) and PSPC1 in Chain B (residues 66–320), and 136 water molecules. The quality of the model was validated using MOLPROBITY (35). The refinement statistics are included in Table 2. The atomic coordinates have been deposited in the Protein Data Bank as entry SWPA.

Author contributions—A. F., C. B., and M. L. conceptualization; J. H., G. P. C. G., M. A. P., and M. L. data curation; J. H., G. P. C. G., M. A. P., and M. L. formal analysis; M. L. funding acquisition; J. H. and M. L. investigation; J. H., G. P. C. G., M. A. P., and M. L. methodology; M. A. P., A. F., C. B., and M. L. writing—original draft; M. L. project administration; M. A. P., C. B., and M. L. writing—review and editing; A. F., C. B., and M. L. resources.

---

6600 J. Biol. Chem. (2018) 293(17) 6593–6602
Crystal structure of a SFPQ/PSPC1 heterodimer

References

1. Knott, G. J., Bond, C. S., and Fox, A. H. (2016) The DBHS proteins SFPQ, NONO and PSPC1: a multipurpose molecular scaffold. Nucleic Acids Res. 44, 3898–4004 CrossRef Medline

2. Knott, G. J., Lee, M., Passon, D. M., Fox, A. H., and Bond, C. S. (2015) Caenorhabditis elegans NONO-1: insights into DBHS protein structure, architecture, and function. Protein Sci. 24, 2033–2043 CrossRef Medline

3. Fox, A. H., Bond, C. S., and Lamond, A. I. (2005) P54nrb forms a heterodimer with Psp1 that localizes to paraspeckles in an RNA-dependent manner. Mol. Biol. Cell 16, 5304–5315 CrossRef Medline

4. Lee, M., Passon, D. M., Hennig, S., Fox, A. H., and Bond, C. S. (2011) Construct optimization for studying protein complexes: obtaining diffusion-quality crystals of the pseudosymmetric PSPC1-NONO heterodimer. Acta Crystallogr. D. Biol. Crystallogr. 67, 981–987 CrossRef Medline

5. Lee, M., Sadowska, A., Bekere, I., Ho, D., Gully, B. S., Lu, Y., Yerer, K. S., Trewthella, J., Fox, A. H., and Bond, C. S. (2015) The structure of human SFPQ reveals a coiled-coil mediated polymer essential for functional aggregation in gene regulation. Nucleic Acids Res. 43, 3826–3840 CrossRef Medline

6. Kuwahara, S., Ikei, A., Taguchi, Y., Tabuchi, Y., Fujimoto, N., Ohinata, M., Uesugi, S., and Kurihara, Y. (2006) PSPC1, NONO, and SFPQ are expressed in mouse Sertoli cells and may function as coregulators of androgen receptor-mediated transcription. Biol. Reprod 75, 352–359 CrossRef Medline

7. Passon, D. M., Lee, M., Rackham, O., Stanley, W. A., Sadowska, A., Filipovska, A., Fox, A. H., and Bond, C. S. (2012) Structure of the heterodimer of human NONO and parasepeckle protein component 1 and analysis of its role in subnuclear body formation. Proc. Natl. Acad. Sci. U.S.A. 109, 4846–4850 CrossRef Medline

8. Li, S., Li, Z., Shu, F. J., Xiong, H., Phillips, A. C., and Dynan, W. S. (2014) Protein-protein interactions with global implicit mass conservation constraints: protein interactionswithglobalimplicitmassconservationconstraintsproviding infrastructure and expertise. Genes Dev. 7, 393–406 CrossRef Medline

Acknowledgments—We thank Dr. G. J. Knott for helpful discussion. Aspects of this research were undertaken on the Macromolecular Crystallography beamlines at the Australian Synchrotron (Victoria, Australia) and we thank the beamline staff for their professional support. We acknowledge the La Trobe University-Comprehensive Proteomics Platform (La Trobe University) and Macromolecular Interactions Facility at Bio21 Institute (University of Melbourne) for providing infrastructure and expertise.
Crystal structure of a SFPQ/PSPC1 heterodimer

32. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 235–242 CrossRef Medline

33. Bricogne, G., Brandl, B. E., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Smart, S. A. O. S., Vonrhein, C., and Womack, T. O. (2016) BUSTER version 2.10.2. Global Phasing Ltd., Cambridge, United Kingdom

34. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501 CrossRef Medline

35. Chen, V. B., Arendall W. B 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21 CrossRef Medline