The FEBS Journal

FEBS PRESS
science publishing by scientists

STATE-OF-THE-ART REVIEW

SERF, a family of tiny highly conserved, highly charged proteins with enigmatic functions

Bikash R. Sahoo and James C. A. Bardwell

Department of Molecular, Cellular and Developmental Biology, Howard Hughes Medical Institute, University of Michigan, Ann Arbor, MI, USA

Keywords
4F5; amyloid; amyloidosis; neurodegenerative diseases; protein aggregation; RNA binding protein; SERF

Correspondence
J. C. A. Bardwell, Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor, MI-48109, USA
Tel: +1 734 764 8028
E-mail: jbardwel@umich.edu

(Received 15 April 2022, revised 7 June 2022, accepted 10 June 2022)

doi:10.1111/febs.16555

Amyloid formation is a misfolding process that has been linked to age-related diseases, including Alzheimer's and Huntington's. Understanding how cellular factors affect this process *in vivo* is vital in realizing the dream of controlling this insidious process that robs so many people of their humanity. SERF (small EDRK-rich factor) was initially isolated as a factor that accelerated polyglutamine amyloid formation in a *C. elegans* model. SERF knockouts inhibit amyloid formation of a number of proteins that include huntingtin, α-synuclein and β-amyloid which are associated with Huntington's, Parkinson's and Alzheimer's disease, respectively, and purified SERF protein speeds their amyloid formation *in vitro*. SERF proteins are highly conserved, highly charged and conformationally dynamic proteins that form a fuzzy complex with amyloid precursors. They appear to act by specifically accelerating the primary step of amyloid nucleation. Brain-specific SERF knockout mice, though viable, appear to be more prone to deposition of amyloids, and show modified fibril morphology. Whole-body knockouts are perinatally lethal due to an apparently unrelated developmental issue. Recently, it was found that SERF binds RNA and is localized to nucleic acid-rich membraneless compartments. SERF-related sequences are commonly found fused to zinc finger sequences. These results point towards a nucleic acid-binding function. How this function relates to their ability to accelerate amyloid formation is currently obscure. In this review, we discuss the possible biological functions of SERF family proteins in the context of their structural fuzziness, modulation of amyloid pathway, nucleic acid binding and their fusion to folded proteins.

Introduction

The principle of homology transfer states that proteins showing recognizable sequence similarity very often have closely related functions. This is of tremendous practical value in biology because, once the function of one member of a family is well characterized, this function can be reasonably extrapolated to other homologues throughout evolution. However, there are a surprising number of protein families and protein domains where no function has yet been clearly determined for any one member [1,2]. For many of these proteins, text mining, machine learning, and bioinformatic analysis provides some clues about how these proteins are regulated, and how they interact with other biomacromolecules. However, establishing their

Abbreviations
Aβ, amyloid-beta; MOAG-4, modifier of aggregation-4; polyQ, polyglutamine; SERF, small EDRK-rich factor.

definitive role in the cell generally awaits experimental analysis [3–14]. The 4F5 family of proteins in the protein families database [15] is one particularly intriguing example. This family is broadly conserved, being found in essentially all eukaryotic organisms sequenced, its members are tiny proteins generally less than < 10 kDa in size, at least partially disordered and rich in charged residues with pI values > 10. One name for this family is SERF which stands for small ERDK-rich factor. Since the *in vivo* function of this family is still unclear, we will continue using this descriptive terminology until the families' function(s) are clearly established. An important hint of a role for the 4F5 family of proteins first identified to be linked to spinal muscular atrophy [16] came with the observations that *C. elegans* mutants in a member of this family form many fewer foci in a polyglutamine disease model and suppress amyloid proteotoxicity [17]. The protein was thus termed MOAG-4 (Modifier of aggregation-4) [17–23]. Subsequent work using the human and yeast homologues showed that the protein *in vitro* effectively accelerated β-amyloid and α-Synuclein aggregation [18,23,24]. These processes are respectively associated with Alzheimer's and Parkinson's disease, devastating age-related neurodegenerative diseases that affect millions worldwide [25]. Understanding how host factors affect these diseases might provide important clues toward their eventual treatment. Recently, it has been found that SERF proteins bind to RNA *in vitro* and localize to RNA-rich membrane-less compartments such as nucleoli *in vivo* [20]. The residues involved in RNA binding appear similar to those involved in accelerating amyloid formation, suggesting that the two functions may be linked, but the nature of this link remains mysterious [20]. Intriguingly, even though SERF family members are among some of the tiniest proteins known, often < 80 amino acids in total length, this protein appears to be is split up into at least two domains, a highly conserved at least partially disordered N-terminal domain that contain the sequences implicated in RNA binding and amyloid acceleration and a more helical, less conserved C-terminal domain. This N-terminal domain is independently found in a number of other proteins, generally fused onto the N-terminus of the protein although fusion to C-terminus or insertion into the middle of some proteins is also observed. The very short length of this domain (~ 30 residues) would seem to preclude a catalytic function, perhaps suggesting instead a function involved in binding or localization. RNA binding appears to be in at least some cases sufficient for localization of a protein to the nucleolus, plasma membrane, nucleus, and cytoplasm [26,27]. One simple hypothesis is that these sequences are involved in the localization to RNA-rich compartments such as the nucleolus.

Classification of SERF family proteins (4F5)

The SERF family is annotated in the Pfam database under the name 4F5, and the protein family database (PF04419) is large, consisting of ~ 3000 proteins which are present in ~ 1300 eukaryotic species. Extensive sequence data has currently been obtained for about the same number of eukaryotes, so it appears that SERF family members exist in the vast majority of eukaryotic species. SERF-related proteins are classified under at least over 40 different architectures. The most common architecture (Group-1) in this family of proteins (~ 60%, > 1700 proteins) has N-terminal domain of about ~ 35 residues that is shared by all 4F5 family members making it the signature N-SERF domain, fused to a second less conserved C terminal domain commonly ~ 25–45 amino acids in length (Fig. 1). Only a few SERF proteins including human SERF, yeast SERF and *C. elegans* MOAG-4 that belong to group-1 (Fig. 1) have been functionally characterized to date [17,18,20]. The remaining members of this group and all other groups of SERF-related proteins are almost completely uncharacterized. To date, only two SERF family protein structures are reported. The NMR model structure of the first SERF-like protein identified, namely the *C. elegans* MOAG-4 from group-1, shows a short helical central region (residues 45–70) flanked by disordered N- (residues 1–44) and C- (residues 70–82) terminal regions (Fig. 2) [21]. The N-terminal region is highly dynamic and characterized by a transiently populated α-helix, this region is known to mediate amyloid interaction and RNA binding. SERF binding partners reported to date such as α-synuclein and a 21-mer RNA oligo have been shown to form fuzzy complexes that apparently do not greatly alter the dynamics and conformational state that the N-SERF domain has on its own [20,21]. More work is needed to determine the range of physiological binding partners to SERF. The second major population of SERF-related proteins is comprised of the N-terminal SERF domain (N-SERF) connected to two different types of zinc-binding domains, to make up groups-2 and 3. Despite possessing two domains, the proteins found in these subgroups are generally no bigger than SERF proteins of group-1 with an average size of ~ 70 residues in length. Group-2 contains > 700 proteins with an N-SERF domain connected to a highly conserved family of metazoan zinc-binding [28]

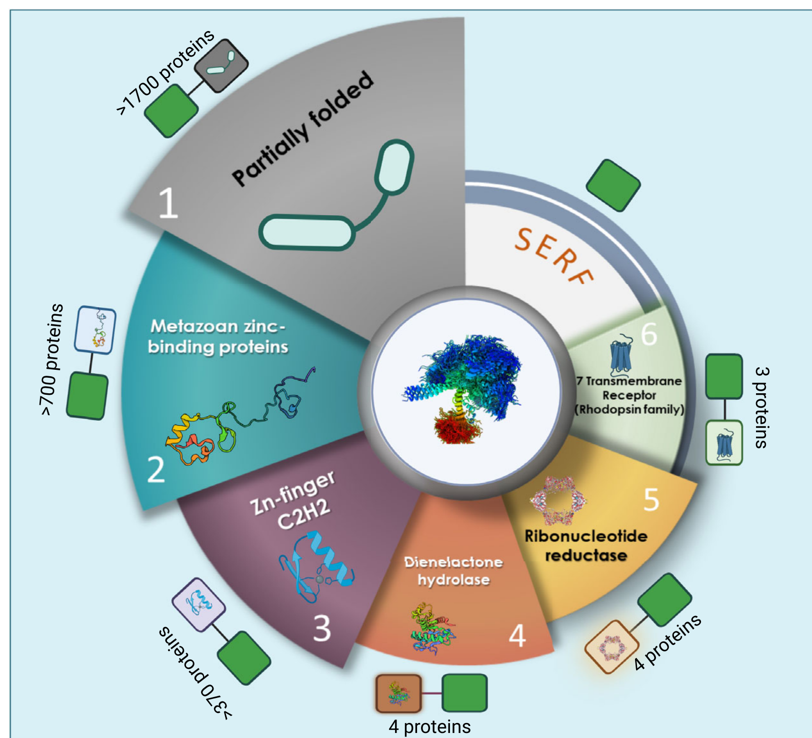

Fig. 1. An overview of SERF family proteins. The illustration shows six major classes of SERF and SERF-domain linked 4F5 family proteins (Pfam: PF04419). The 4F5 family includes 37 additional groups with distinct domain organization containing either 1 or 2 proteins and are not shown in the illustration. Readers are referred to check the Pfam database (PF04419) to retrieve information about these low abundance SERF domains containing proteins.

proteins (zf-met2) mostly from plants with a variable number of cysteine and histidine residues involved in zinc coordination. These proteins are mostly uncharacterized with the exception of At2g23090 from *Arabidopsis* whose NMR structure (PDB ID: 1WVK) was solved as part of a structural genomics effort, which shows it to be largely unstructured with a short helical region spanning residues 62–71 (Fig. 2, group-2). The third class of SERF-related proteins (Group-3) which contains > 370 proteins are characterized by a N-SERF domain fused to a Cys_2His_2 type zinc finger [29]. This class of zinc fingers are best known when present in other proteins for their roles in making DNA-binding proteins sequence-specific, but this type of zinc finger can also have a variety of other functions such as RNA binding or in mediating protein–protein interactions [30–35]. The SERF containing group-3 is very unusual within the Cys_2His_2 zinc finger containing proteins in that it just has a single zinc finger. The vast majority of proteins that contain Cys_2His_2 zinc fingers have multiple copies of these fingers to improve specificity or affinity. The proteins in group-3 are mostly derived from higher vertebrates including humans. The N-SERF domain is also found in rare instances fused to a variety of enzymes and other proteins in a variety of positions, N-terminus, C-terminus or within the middle of proteins which range in size from ~ 200–2000 amino acids. As illustrated in Fig. 1, a conserved N-SERF domain has been shown to be linked to dienelactone hydrolase, an enzyme known to degrade haloaromatic compounds [36]. Four ribonucleotide reductase enzymes are fused to a C-terminally located SERF domain. In addition to this, several well-studied enzymes, e.g., peptidylprolyl isomerase, dioxygenase, ATPase GET3, glycerol-3-phosphate dehydrogenase, ribonucleoside-diphosphate reductase, α-1,4-N-acetylglucosaminyltransferase, glycogen synthase kinase-3 and glucose-6-phosphate 1-epimerase contain SERF domains. SERF fusions are present in a limited number of instances within each of these protein families. Since most family members lack SERF-related domains, this suggests that these insertions have occurred recently in evolution and that the SERF domain fusion is unlikely to be absolutely essential for these enzyme's activities. The role of these SERF domains in modulating these protein's functions remains to be explored; one intriguing possibility given the known RNA binding activity of the SERF domain is that these fusions are being used by the cell to localize these various proteins adjacent to an RNA or within an RNA-rich compartment. Another possibility, given the known ability of SERF to modulate amyloid formation, is that these domains play a similar role in modulating the oligomerization status of the proteins to which they are linked to.

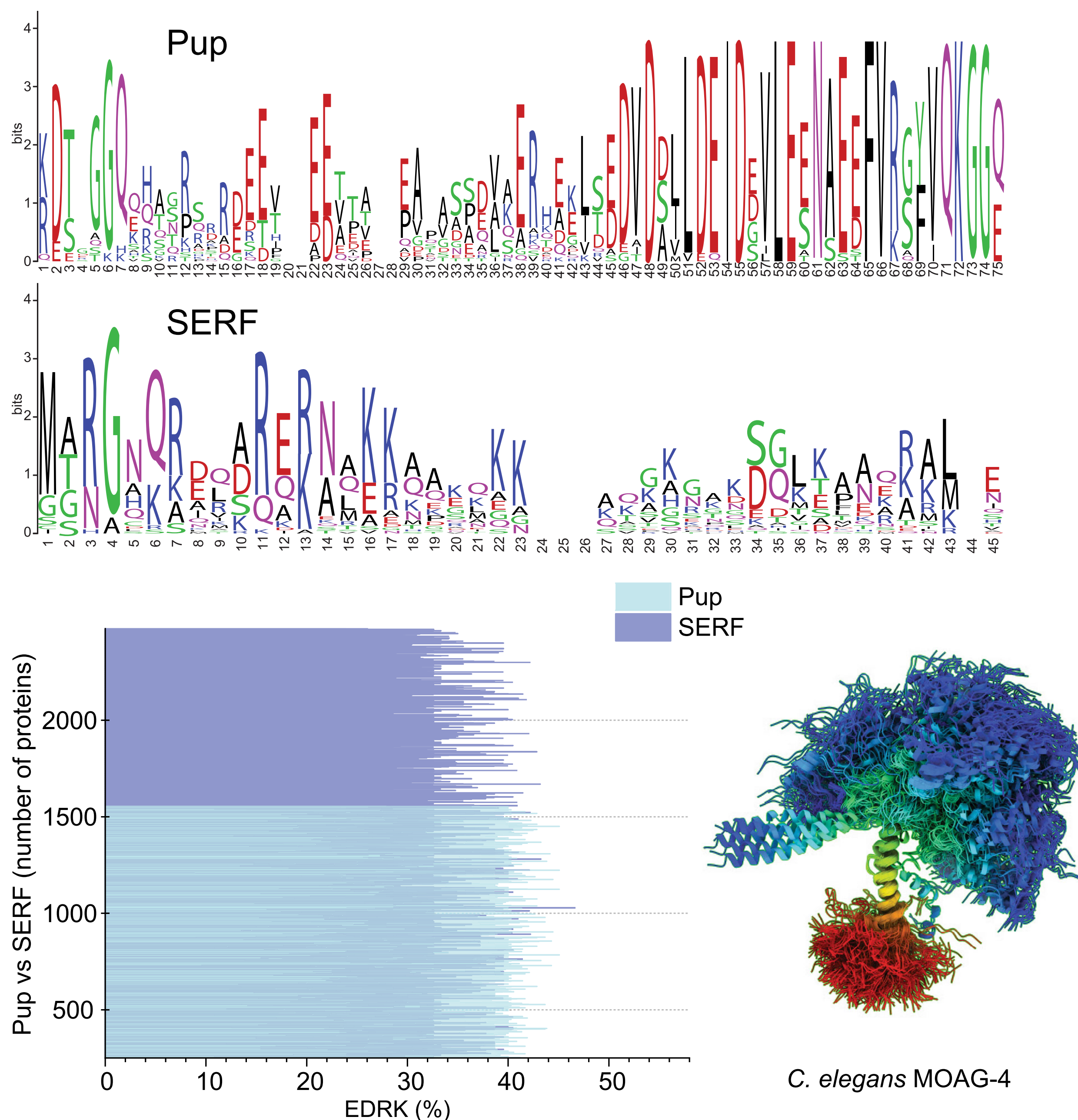

Fig. 2. Sequence logos of SERF and pup conserved motifs. The illustrated sequence logos are derived from 37 SERF and 23 pup motifs. The total charge distribution (EDRK%) for these motifs derived from a large number of SERF (2935) and pup (1301) proteins known to date are compared in the left bottom with the Y axis showing the number of proteins. The NMR model structure ensembles of *C. elegans* MOAG-4 (BMRB entry: 27058) is shown bottom right). The dynamic N-terminal SERF and more folded C-terminal α-helical domains are shown in blue-cyan and red-yellow, respectively.

SERF domain is highly charged

SERF family proteins, as reflected by their name (small EDRK-rich factor), are highly enriched in charged amino acids with an average pI > 10. As illustrated in Fig. 2A, our bioinformatic analysis that includes SERF motif sequences from all the subfamilies of SERF shown in Fig. 1 showed that the SERF domain has a high charge distribution. For the purposes of visualization of the variability of EDRK %

within a protein family we compared SERF family members to a prokaryotic ubiquitin-like protein (Pup) [37,38] used because it has a molecular size roughly similar to that of the SERF domain and like SERF is known to be at least partially disordered. SERF motif derived from > 2900 proteins shows an average 29.3% of EDRK amino acids which is similar to that of protein motif derived from ~ 1300 Pup proteins (Pfam: PF05639) that showed an average 33.2% of EDRK. Pup proteins were previously known to show a high percentage of charged residues and are intrinsically disordered [39] (Fig. 2). More specifically, amino acid decomposition in SERF motif yielded 4.2% glutamic acid, 3% aspartic acid, 7.5% arginine and 14.4% lysine making it particularly lysine rich, in comparison to the pup domain which shows an average of 11.2% glutamic acid, 13.7% aspartic acid, 4.2% arginine and 4.2% lysine. The high percentage of charged residues in SERF, in particular lysine, correlates to ribosomal proteins comprising of higher % of arginine and lysine residues (10–11%) as compared to non-ribosomal soluble proteins that on average have 4.7 and 5.9%, arginine and lysine residues, respectively, as calculated from over 500 species [40]. Importantly, SERF-related proteins have a higher percentage of positively charged amino acids as compared to acidic amino acids, this is also a characteristic feature for ribosomal proteins, DNA and RNA-binding proteins, and zinc-finger proteins, in particular those with a helix–loop–helix structural topology [41]. These positively charged residues are often involved in binding to negatively charged phosphates in RNA, as a result, RNA binding proteins have an average occurrence of 6.1 and 6.5 for lysine and arginine, respectively. Whereas all other amino acids in RNA binding protein interfaces are found at an average occurrence of < 4 [42]. Lysine and arginine are involved in the protein-RNA binding interfaces, contributing an average of ~ 26% to this interface. This high occurrence of lysine and arginine in SERF is consistent with the recent observation that human SERF1a interacts with RNA [20]. In general, non-ribosomal and soluble proteins have equivalent percentages (around 5%) of basic (Arg or Lys) and acidic (Asp or Glu) amino acids. SERF is particularly lysine rich. This is intriguing, as it has been shown that lysine-RNA interactions help drive liquid–liquid phase transitions [43] and as discussed below, SERF appears to localize to the nucleolus, an RNA rich liquid–liquid phase compartment. The distribution of these charged residues in SERF proteins is relatively uniform across the full length of the protein.

Physiological role of SERF: a tale of two stories

SERF accelerates amyloid aggregation and induce proteotoxicity

Misfolded proteins have several possible fates within a cell, being degraded by proteases or becoming part of an aggregate, an amyloid or a liquid–liquid phase transition compartment. It is still not very clear what determines these fates, and specifically what drives some misfolded proteins to self-assemble to form insoluble protein assemblies called amyloids [44,45]. Over a hundred proteins with intrinsic amyloid-forming propensities have been identified. These proteins *in vivo* tend to form highly ordered β-sheet protein fibres that comprise amyloid deposits as the end product [46,47]. As the amyloid formation process has been tightly linked a number of age-related diseases such as Alzheimer's and Huntington's diseases [48–52], it is of great interest to determine how this pathway is controlled in the cell, driven by the hope of eventually being able to modulate amyloid formation in order to alleviate the devastating effects of these diseases. Evidence suggests that the amyloid assembly pathway in the cell is modulated by several biomacromolecules including proteins and nucleic acids and smaller substances such as metals, metabolites, lipids, and synthetic molecules including nanoparticles [53–62]. Understanding the underlying mechanism of amyloid formation and how this molecular process happens in the crowded cellular environment is not covered here but is the subject of several recent reviews [52,53,63,64].

Proteins that interact with or affect the polymerization of the Alzheimer's disease-associated amyloid-precursor proteins Aβ-40 or Aβ-42 include apolipoprotein E, scrapie prion protein (PrPSc), p75 neurotrophin receptor (p75NTR), metabotropic glutamate receptors (mGluR5), immunoglobin receptors (FccRIIb and PirB), scavenger receptors (SCARA-1, SCARB-2, MARCO, RAGE, and CD36), toll-like receptors (TLR2 and TLR4), G-protein coupled receptors (FPR2 and CMKLR1), proteolytic enzymes (IDE, NEP, ECE1, ECE2, MMP2 and MMP9) and heat shock proteins such as Hsp60, Hsp70, Hsp90, in addition to the SERF-related proteins (SERF1, SERF2, and MOAG-4) which are the focus of this review [65–72]. These macromolecules appear to regulate amyloid deposition using different pathways that include proteolysis, transport, induction of apoptosis, alteration in cellular localization, endocytosis, and in the case of SERF, acceleration of amyloid aggregation. CRAM-1 a *C. elegans* protein that may be distantly related to

SERF2 ortholog [19], small molecules [73,74], and polyamines or amine derivatives [75–77] which like SERF are highly positively charged have been also shown to accelerate amyloid formation. Evidence is accumulating that amyloids themselves are not likely the toxic species but instead small oligomers that occur along the pathway to amyloid [78,79]. One attractive hypothesis, that has yet to be tested, is that by rushing amyloid-prone proteins through the initial stages of aggregation, SERF acts to reduce the concentration of these toxic species.

The SERF ortholog MOAG-4 was initially identified in *C. elegans* by its ability to regulate the amyloid aggregation of a polyglutamine (polyQ) protein construct that contains a 40-glutamine (Q40)-rich polypeptide fused to the yellow fluorescent protein (YFP) [17]. Point mutations, genetic deletion and RNAi silencing of MOAG-4 are reported to suppress polyQ aggregation both in terms of the number of bright yellow foci formed and in the reduction of the amount of SDS-resistant amyloids present in lysates (Fig. 3) [17]. In an Alzheimer's model that expresses β-amyloid, the MOAG-4 deletion was shown to substantially reduce the population (> 40%) of paralysed worms (Fig. 3B) [17]. Decline in motility was reported in a Parkinson's worm model expressing α-Synuclein and an Amyotrophic Lateral Sclerosis (ALS) model expressing a mutant superoxide dismutase [80]. The deletion of MOAG-4 significantly reduced the toxicity of α-Synuclein, but not in worms expressing mutant SOD. These pioneering *C. elegans* studies have been extended to humans and yeast. Human neuroblastoma

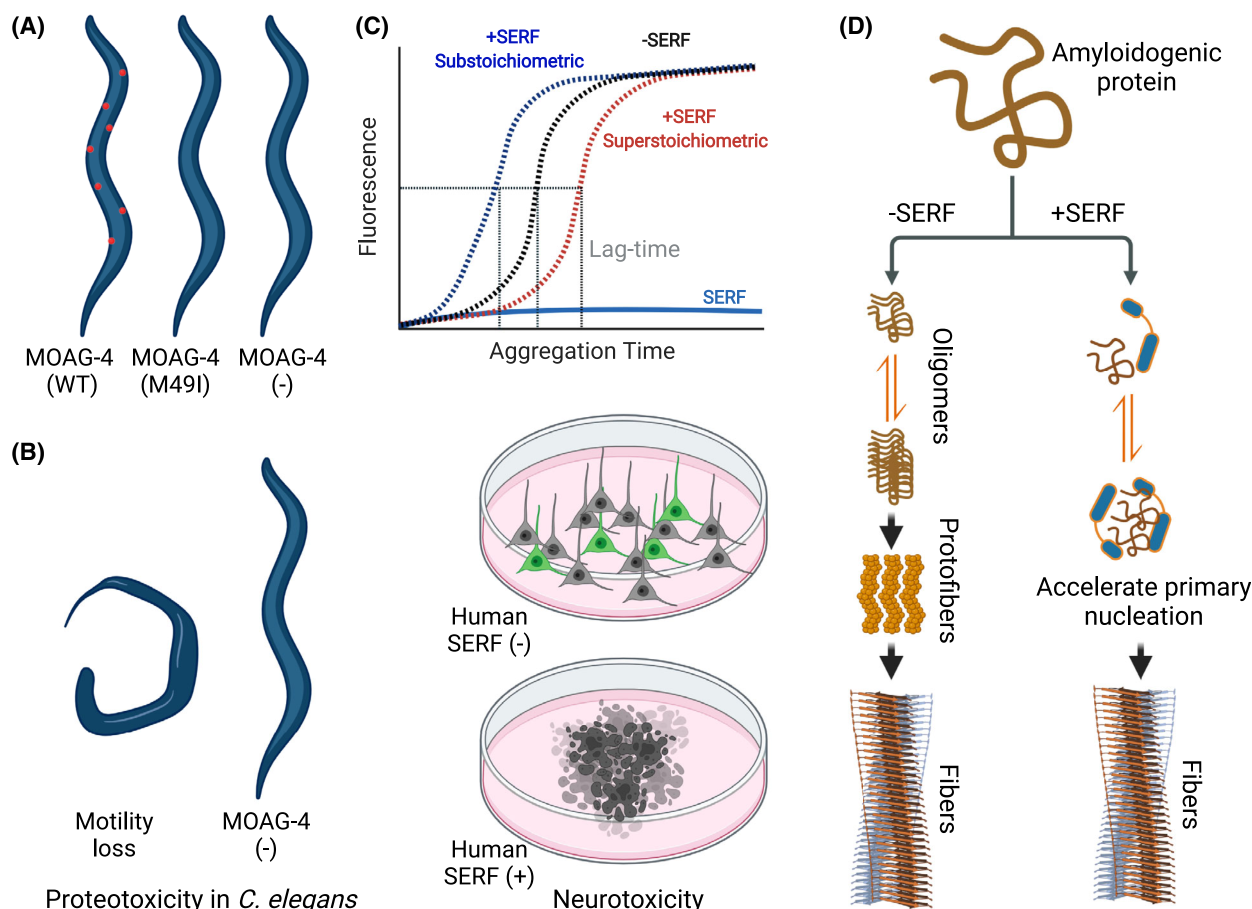

Fig. 3. SERF modulates amyloid aggregation and toxicity. (A) MOAG-4 in *C. elegans* is shown to induce the formation of protein amyloids (red dots) in an amyloid disease model. Point mutations or a deletion of the gene for MOAG-4 (SERF) suppresses amyloid formation [17]. (B) MOAG-4 mutant strains reduce the proteotoxicity induced by overexpression of polyQ in *C. elegans* and SERF1a and SERF2 in human neuroblastoma cells [17]. (C) SERF modulates β-amyloid aggregation kinetics in a fluorescence assay. When present in substoichiometric quantities, SERF accelerates amyloid aggregation; however, when mixed with a high, likely non-physiological concentration of yeast SERF, β-amyloid aggregation is inhibited [24]. (D) Mechanistic studies disclosed SERF forms a fuzzy complex with amyloid proteins and accelerate their aggregation by acting on the primary nucleation step in the aggregation pathway.

cells overexpressing human SERF1a or SERF2 were identified to significantly increase the polyQ (Q74) aggregation and Q74-induced cell death; on the contrary, RNAi silencing rescues the toxic phenotypes through suppressing the Q74 aggregation (Fig. 3B) [17].

Possible anti-aggregation and protective roles for SERF

In contrast to SERF's reported role in accelerating amyloid aggregation, a recent study has reported a protective role vis-à-vis protein aggregation. Human SERF2 protein, referred by these workers as Hero7, along with other five proteins are shown to remain soluble after high-temperature treatment and were thus termed heat-resistant obscure (Hero) proteins [81]. Very different from the polyQ amyloid acceleration effect shown for the *C. elegans*, SERF homologue MOAG-4, GFP fused to SERF2/Hero7 was shown to suppress TAR DNA-binding protein 43 (TDP-43) aggregation *in vitro* and also suppress the toxicity in tissue culture cells that express TDP43, a protein involved in ALS and frontotemporal dementia. Monitoring *Drosophila* eye pigment is considered to be a model system for neurodegenerative diseases [81]. TDP43 aggregation is linked to eye degeneration in *Drosophila*, but co-expression with SERF2/Hero7 suppresses this degeneration arguing for a possible anti-aggregation behaviour for SERF, when it is present in high amounts [24]. Though this latter result that is consistent with the anti-aggregation behaviour of yeast SERF on Aβ *in vitro* when it is present in high amounts (Fig. 3C), it is still unclear if SERF reduces or enhances proteotoxicity *in vivo* or if it does so in a client selective fashion. In addition, the abundance of SERF proteins varies significantly from organism to organism and even cell line to cell line (https://www.pax-db.org/) [24]. There can be several copies of SERF-related proteins within one organism, humans have SERF1a, SERF2 and ZNF706 for instance, but in addition each of these genes can encode multiple isoforms human SERF1a has two major isoforms and SERF2 has four. These complications limit our current understanding of SERF-related protein's normal and pathological functions.

Underlying mechanism of SERF-induced amyloid aggregation

The mechanism of action of SERF proteins in driving amyloid fibrillation has been studied independently by two groups using different amyloid and SERF systems. Falsone et al. [18] reported Human SERF1a protein acts *in vitro* to accelerate the kinetics of aggregation of an array of amyloid-forming proteins including α-Synuclein, Htt, β-amyloid, and PrP. Structurally, SERF1a was reported to be at least partially disordered and dynamic as evidenced from circular dichroism and NMR measurements. Fluorescence experiments (Fig. 3C) using the amyloid-specific dye (Thioflavin-T) revealed that α-Synuclein fibres are generated faster and in greater amounts in the presence of an equimolar amount of SERF1a than in its absence, aggregation of non-amyloidogenic proteins such as actin, insulin and citric synthase was not affected by SERF1a [18].

Charge generally plays an important role in amyloid formation [79,82,83]. For example, neutralization of amyloidogenic proteins by various means including mutation, chemical agents and interfering biomolecules accelerates fibrillation [82–85]. SERF is a highly charged protein, so it is not surprising to observe its electrostatically driven interaction with amyloidogenic proteins [21,22]. Pras et al. [22], using a peptide array-based screening approach, found that peptides binding to SERF were enriched in the acidic residues (aspartic acid and glutamic acid) and non-SERF binding peptide fragments were enriched in the basic amino acids arginine and lysine. A triple mutation in SERF2 that neutralized K16, K17 and K23 substantially reduced its binding to the peptide arrays. The highly charged SERF1a is hypothesized to directly interact with the C-terminal acidic region of α-Synuclein, partially unfolding it and allowing it to self-associate [18]. Structural studies identified the existence of a similar binding interface between the disordered C-terminus of α-Synuclein and a transiently populated α-helix present in the N-terminal 22 residues of *C. elegans* MOAG-4 [21]. Abolishment of SERF1a and α-Synuclein interaction is seen when the basic amino acids in SERF1a, i.e., K13/K16/K17 were simultaneously mutated to the acidic residue glutamic acid, implicating a role of charge in driving SERF1a and α-Synuclein interaction. A role of charge–charge interaction in driving *in vivo* amyloid aggregation by SERF has been demonstrated by showing that SERF-driven protein aggregation was reduced in lysine mutants in SERF2 and MOAG-4 [22].

The aggregation pathway, kinetics and mechanism of SERF action were further studied by Meinen et al. [24]. Yeast SERF was tested against two well-studied human amyloid proteins α-Synuclein and β-amyloid. This study shows that SERF primarily influenced the aggregation pathway of amyloid formation by accelerating primary nucleation (Fig. 3D) [24]. Native mass-

spectrometry revealed that SERF acts on aggregation primarily by interacting with monomeric α-Synuclein populations either maintaining or increasing the degree of disorder in α-Synuclein. SERF is shown to bind the N- and C-terminus of β-amyloid or α-Synuclein, respectively, that are characterized by negative charge residues. A fuzzy complex state of SERF1a with α-Synuclein was also proposed by Merle et al [23]. with no significant alteration in the degree of α-Synuclein disorder occurring upon complexion. Human SERF1 and α-Synuclein were verified to interact in human neuroblastoma cells by tracking their colocalization post transfection.

SERF: a nucleic-acid binding protein

Although the pathology-associated function of SERF-linked species may well be related to its ability to facilitate amyloid formation, if SERF is important for neural-specific amyloid formation, one might expect brain-specific SERF knockouts to exhibit differences in amyloid formation within neural tissues when amyloid prone proteins are expressed there. A mouse SERF knockout is reported to alter intracellular Aβ accumulation and is shown to have a higher plaque deposition when relative to a wild-type mouse [86]. Although minor differences in amyloid's dye-binding characteristics have been observed, change in Aβ production, or the Aβ levels or processing of the amyloid precursor protein have not been observed [86]. Whole-body SERF knockouts in C57BL/6 N background mice are perinatally lethal and show developmental defects including behavioural and neurological deficits in mice implying that SERF does play an important role in development; however, the connection of these observations to SERF's role in amyloid formation and pathology of neurodegenerative diseases is still unclear [87]. What then is the normal physiological function of SERF or the N terminal SERF domains that are found fused to a variety of other proteins, predominately zinc-binding domains? One clue comes out of a recent report that SERF binds to a 21-mer RNA with relatively low affinity (K_d ~ 5 μM measured at 20 mM NaCl) but fails to interact with a DNA molecule of the same sequence or polyanions like heparin [20]. Other RNA molecules such as tRNA and yeast total RNA were found to bind to SERF1a slightly more tightly with affinities of around K_d ~ 1–2 μM. In comparison, site-specific RNA binding proteins such as helicases, heterogeneous nuclear ribonucleoproteins (hnRNP), eIF4, TDP-43, FUS and ribonucleoprotein (RNP1 and RNP2) tend to interact much more tightly, with affinities in the low nM range [88–90]. The physiological significance or specificity of this SERF RNA binding is unclear at this time. Interestingly, eGFP-directed pulldowns using nuclei from SERF1a-eGFP expressing cells are highly enriched in RNA-binding proteins including proteins involved in ribosomal synthesis and RNA splicing, DNA/RNA helicases, mRNA splicing factors, mRNA processing factors, translation initial factors, and ribosomal proteins [91–102]. It seems unlikely that the tiny SERF protein is able to specifically interact directly with all these different proteins; it is more likely that SERF simply

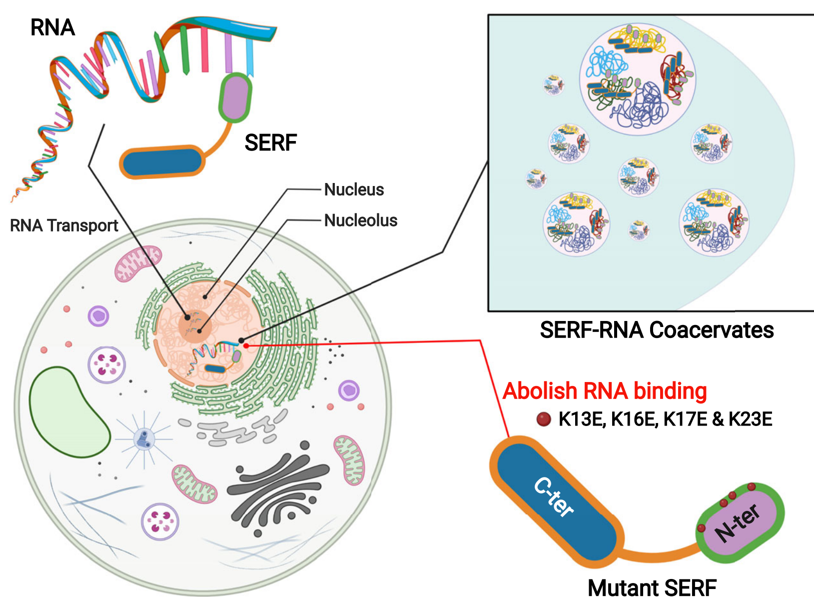

Fig. 4. Human SERF1a binding to RNA may promote nucleolar localization, RNA transportation and RNA-granule formation in non-pathological conditions [20]. SERF1a is shown to localize in nucleus and more diffusely in cytoplasm and is postulated to assist in RNA transport to nucleolus [20]. Under stress conditions SERF1a is postulated to exit nucleus and exert its pathological, amyloid-related function. Introduction of N-terminal lysine mutations is shown to abolish SERF's RNA binding and nucleolar localization function similar to the effects these mutations have on SERF's ability to bind amyloid proteins [22].

binds to the RNAs that are bound to these proteins. SERF1a binding to RNA drives the formation of phase-separated condensate or liquid droplets *in vitro*, a property shared by a variety of other at least partially disordered proteins, and colocalizes with RNA-rich compartments *in vivo* such as the nucleolus, and nucleus [26] which provides some evidence that the RNA interactions observed *in vitro* may have physiological significance (Fig. 4). SERF1a has been proposed to transport RNA to the liquid-like nucleolus (Fig. 4). Intriguingly, Lys17, a highly conserved lysine that is required for SERF1a to facilitate amyloid formation, when mutated to glutamic acid also strongly affects SERF's ability to interact with RNA [20], providing preliminary evidence that the two functions of SERF might be linked. Both α-Synuclein and RNA are shown to bind the same interface of SERF1 at similar binding affinities. Is thus unclear under what circumstances that SERF would interact with amyloid precursors which are presumably less abundant than RNA species. It has been proposed that though the pathological function of SERF may be to facilitate amyloid formation, the normal function of SERF may have more to do with its ability to interact with and possibly transport RNA [20]. This hypothesis is based mainly on affinity measurements *in vitro* and *in vivo* colocalization studies but both the pathological and normal roles of this class of tiny well-conserved proteins remain to be established.

Concluding remarks

SERF or 4F5 are a family of well-conserved tiny, highly charged and at least partially disordered proteins that are so far relatively poorly uncharacterized. Though initially characterized as being involved in accelerating amyloid formation, their normal *in vivo* role which may involve RNA binding, remains unclear. Independent of their actual *in vivo* function the biophysical amenability of this tiny class of proteins makes them an attractive model system to answer several fundamental biological questions concerning the role of protein charge and disorder in nucleic acid and amyloidogenic protein binding, protein solubilization, subcellular and nucleolar localization, and phase separation.

Acknowledgements

JCAB is a Howard Hughes Medical Institute Investigator and this study was supported by funds from HHMI.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

Conceptualization: JCAB, BRS; manuscript writing: BRS, JCAB; supervision: JCAB; funding acquisition: JCAB.

Data availability statement

Data used in the review article are derived from public domain resources.

References

1 Ijaq J, Chandrasekharan M, Poddar R, Bethi N, Sundararajan VS. Annotation and curation of uncharacterized proteins- challenges. *Front Genet*. 2015;**6**:119.

2 Veuthey A-L, Pillet V, Yip YL & Ruch P (2009) Text mining for swiss-Prot curation: a story of success and failure. *Nat Prec*. https://doi.org/10.1038/npre.2009.3166.1

3 Papanikolaou N, Pavlopoulos GA, Theodosiou T, Iliopoulos I. Protein-protein interaction predictions using text mining methods. *Methods*. 2015;**74**:47–53.

4 Muzio G, O'Bray L, Borgwardt K. Biological network analysis with deep learning. *Brief Bioinform*. 2021;**22**:1515–30.

5 Lei Y, Li S, Liu Z, Wan F, Tian T, Li S, et al. A deep-learning framework for multi-level peptide–protein interaction prediction. *Nat Commun*. 2021;**12**:5465.

6 Lv Z, Ao C, Zou Q. Protein function prediction: from traditional classifier to deep learning. *Proteomics*. 2019;**19**:e1900119.

7 Bonetta R, Valentino G. Machine learning techniques for protein function prediction. *Proteins*. 2020;**88**:397–413.

8 Saha S, Prasad A, Chatterjee P, Basu S, Nasipuri M. Protein function prediction from dynamic protein interaction network using gene expression data. *J Bioinform Comput Biol*. 2019;**17**:1950025.

9 Hong J, Luo Y, Zhang Y, Ying J, Xue W, Xie T, et al. Protein functional annotation of simultaneously improved stability, accuracy and false discovery rate achieved by a sequence-based deep learning. *Brief Bioinform*. 2019;**21**:1437–47.

10 Frasca M, Bianchi NC. Multitask protein function prediction through task dissimilarity. *IEEE/ACM Trans Comput Biol Bioinform*. 2019;**16**:1550–60.

11 Kulmanov M, Khan MA, Hoehndorf R. DeepGO: predicting protein functions from sequence and interactions using a deep ontology-aware classifier. *Bioinformatics*. 2018;**34**:660–8.

12 Saha S, Chatterjee P, Basu S, Nasipuri M, Plewczynski D. Funpred 3.0: improved protein function prediction using protein interaction network. *PeerJ*. 2019;**2019**: e6830.

13 Kulmanov M, Hoehndorf R. DeepGOPlus: improved protein function prediction from sequence. *Bioinformatics*. 2020;**36**:1187.

14 Hakala K, Kaewphan S, Bjorne J, Mehryary F, Moen H, Tolvanen M, et al. Neural network and random Forest models in protein function prediction. *IEEE/ACM Trans Comput Biol Bioinform*. 2020;**19**:1772–81.

15 Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: the protein families database in 2021. *Nucleic Acids Res*. 2021;**49**:D412–9.

16 Scharf JM, Endrizzi MG, Wetter A, Huang S, Thompson TG, Zerres K, et al. Identification of a candidate modifying gene for spinal muscular atrophy by comparative genomics. *Nat Genet*. 1998;**20**:83–6.

17 van Ham TJ, Holmberg MA, van der Goot AT, Teuling E, Garcia-Arencibia M, Kim HE, et al. Identification of MOAG-4/SERF as a regulator of age-related proteotoxicity. *Cell*. 2010;**142**:601–12.

18 Falsone SF, Meyer NH, Schrank E, Leitinger G, Pham CLL, Fodero-Tavoletti MT, et al. SERF protein is a direct modifier of amyloid fiber assembly. *Cell Rep*. 2012;**2**:358–71.

19 Balasubramaniam M, Ayyadevara S, Shmookler Reis RJ. Structural insights into pro-aggregation effects of *C. elegans* CRAM-1 and its human ortholog SERF2. *Sci Rep*. 2018;**8**:14891.

20 Meyer NH, Dellago H, Tam-Amersdorfer C, Merle DA, Parlato R, Gesslbauer B, et al. Structural fuzziness of the RNA-organizing protein SERF determines a toxic gain-of-interaction. *J Mol Biol*. 2020;**432**:930–51.

21 Yoshimura Y, Holmberg MA, Kukic P, Andersen CB, Mata-Cabana A, Fabio Falsone S, et al. MOAG-4 promotes the aggregation of α-synuclein by competing with self-protective electrostatic interactions. *J Biol Chem*. 2017;**292**:8269–78.

22 Pras A, Houben B, Aprile FA, Seinstra R, Gallardo R, Janssen L, et al. The cellular modifier MOAG-4/SERF drives amyloid formation through charge complementation. *EMBO J*. 2021;**40**:e107568.

23 Merle DA, Witternigg A, Tam-Amersdorfer C, Hartlmüller C, Spreitzer E, Schrank E, et al. Increased aggregation tendency of alpha-synuclein in a fully disordered protein complex. *J Mol Biol*. 2019;**431**:2581–98.

24 Meinen BA, Gadkari VV, Stull F, Ruotolo BT, JCA B. SERF engages in a fuzzy complex that accelerates primary nucleation of amyloid proteins. *Proc Natl Acad Sci USA*. 2019;**116**:23040–9.

25 Gammon K. Neurodegenerative disease: brain windfall. *Nature*. 2014;**515**:299–300.

26 Wang B, Zhang L, Dai T, Qin Z, Lu H, Zhang L, et al. Liquid–liquid phase separation in human health and diseases. *Signal Transduct Target Ther*. 2021;**6**:290.

27 Guillen-Chable F, Bayona A, Rodríguez-Zapata LC, Castano E. Phase separation of intrinsically disordered nucleolar proteins relate to localization and function. *Int J Mol Sci*. 2021;**22**:13095.

28 Albà MM. Zinc-finger domains in metazoans: evolution gone wild. *Genome Biol*. 2017;**18**:168.

29 Fedotova AA, Bonchuk AN, Mogila VA, Georgiev PG. C2H2 zinc finger proteins: the largest but poorly explored family of higher eukaryotic transcription factors. *Acta Naturae*. 2017;**9**:47–58.

30 Brown RS. Zinc finger proteins: getting a grip on RNA. *Curr Opin Struct Biol*. 2005;**15**:94–8.

31 McBryant SJ, Veldhoen N, Gedulin B, Leresche A, Foster MP, Wright PE, et al. Interaction of the RNA binding fingers of Xenopus transcription factor IIIA with specific regions of 5 S ribosomal RNA. *J Mol Biol*. 1995;**248**:44–57.

32 Gamsjaeger R, Liew CK, Loughlin FE, Crossley M, Mackay JP. Sticky fingers: zinc-fingers as protein-recognition motifs. *Trends Biochem Sci*. 2007;**32**:63–70.

33 Mackay JP, Crossley M. Zinc fingers are sticking together. *Trends Biochem Sci*. 1998;**23**:1–4.

34 Shi Y, Berg JM. Specific DNA-RNA hybrid binding by zinc finger proteins. *Science*. 1995;**268**:282–4.

35 Brayer KJ, Segal DJ. Keep your fingers off my DNA: protein-protein interactions mediated by C2H2 zinc finger domains. *Cell Biochem Biophys*. 2008;**50**:111–31.

36 Ngai KL, Schlomann M, Knackmuss HJ, Ornston LN. Dienelactone hydrolase from *Pseudomonas* sp. strain B13. *J Bacteriol*. 1987;**169**:699–703.

37 Cappadocia L, Lima CD. Ubiquitin-like protein conjugation: structures, chemistry, and mechanism. *Chem Rev*. 2018;**118**:889–918.

38 Herrmann J, Lerman LO, Lerman A. Ubiquitin and ubiquitin-like proteins in protein regulation. *Circ Res*. 2007;**100**:1276–91.

39 Liao S, Shang Q, Zhang X, Zhang J, Xu C, Tu X. Pup, a prokaryotic ubiquitin-like protein, is an intrinsically disordered protein. *Biochem J*. 2009;**422**:207–15.

40 Lott BB, Wang Y, Nakazato T. A comparative study of ribosomal proteins: linkage between amino acid distribution and ribosomal assembly. *BMC Biophys*. 2013;**6**:13.

41 Bartas M, Červeň J, Guziurová S, Slychko K, Pečinka P. Amino acid composition in various types of nucleic acid-binding proteins. *Int J Mol Sci*. 2021;**22**:922.

42 Krüger DM, Neubacher S, Grossmann TN. Protein–RNA interactions: structural characteristics and hotspot amino acids. *RNA*. 2018;**24**:1457–65.
43 Ukmar-Godec T, Hutten S, Grieshop MP, Rezaei-Ghaleh N, Cima-Omori MS, Biernat J, et al. Lysine/RNA-interactions drive and regulate biomolecular condensation. *Nat Commun*. 2019;**10**:2909.
44 Soto C, Pritzkow S. Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases. *Nat Neurosci*. 2018;**21**:1332–40.
45 Aguzzi A, O'Connor T. Protein aggregation diseases: pathogenicity and therapeutic perspectives. *Nat Rev Drug Discov*. 2010;**9**:237–48.
46 Varadi M, de Baets G, Vranken WF, Tompa P, Pancsa R. AmyPro: a database of proteins with validated amyloidogenic regions. *Nucleic Acids Res*. 2018;**46**:D387–92.
47 Eisenberg D, Jucker M. The amyloid state of proteins in human diseases. *Cell*. 2012;**148**:1188–203.
48 Ghiso J, Frangione B. Amyloidosis and Alzheimer's disease. *Adv Drug Deliv Rev*. 2002;**54**:1539–51.
49 Gosselet F, Saint-Pol J, Candela P, Fenart L. Amyloid-β peptides, Alzheimer's disease and the blood-brain barrier. *Curr Alzheimer Res*. 2013;**10**:1015–33.
50 Scherzinger E, Sittler A, Schweiger K, Heiser V, Lurz R, Hasenbank R, et al. Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. *Proc Natl Acad Sci USA*. 1999;**96**:4604–9.
51 McGowan DP, van Roon-Mom W, Holloway H, Bates GP, Mangiarini L, Cooper GJS, et al. Amyloid-like inclusions in Huntington's disease. *Neuroscience*. 2000;**100**:677–80.
52 Ke PC, Zhou R, Serpell LC, Riek R, Knowles TPJ, Lashuel HA, et al. Half a century of amyloids: past, present and future. *Chem Soc Rev*. 2020;**49**:5473–509.
53 Nguyen PH, Ramamoorthy A, Sahoo BR, Zheng J, Faller P, Straub JE, Dominguez L, Shea JE, Dokholyan N v., de Simone A, Ma B, Nussinov R, Najafi S, Ngo ST, Loquet A, Chiricotto M, Ganguly P, Mccarty J, Li MS, Hall C, Wang Y, Miller Y, Melchionna S, Habenstein B, Timr S, Chen J, Hnath B, Strodel B, Kayed R, Lesné S, Wei G, Sterpone F, Doig AJ & Derreumaux P (2021) Amyloid oligomers: a joint experimental/computational perspective on Alzheimer's disease, Parkinson's disease, type II diabetes, and amyotrophic lateral sclerosis. *Chem Rev* **121,** 2545–2647.
54 Bradley-Whitman MA, Timmons MD, Beckett TL, Murphy MP, Lynn BC, Lovell MA. Nucleic acid oxidation: an early feature of Alzheimer's disease. *J Neurochem*. 2014;**128**:294–304.
55 Adlard PA, Bush AI. Metals and Alzheimer's disease: how far have we come in the clinic? *J Alzheimers Dis*. 2018;**62**:1369–79.
56 Das N, Raymick J, Sarkar S. Role of metals in Alzheimer's disease. *Metab Brain Dis*. 2021;**36**:1627–39.
57 Chew H, Solomon VA, Fonteh AN. Involvement of lipids in Alzheimer's disease pathology and potential therapies. *Front Physiol*. 2020;**11**:598.
58 Kao YC, Ho PC, Tu YK, Jou IM, Tsai KJ. Lipids and alzheimer's disease. *Int J Mol Sci*. 2020;**21**:1505.
59 Pettegrew JW, Panchalingam K, Hamilton RL, Mcclure RJ. Brain membrane phospholipid alterations in Alzheimer's disease. *Neurochem Res*. 2001;**26**:771–82.
60 Zaman M, Khan AN, Wahiduzzaman ZSM, Khan RH. Protein misfolding, aggregation and mechanism of amyloid cytotoxicity: an overview and therapeutic strategies to inhibit aggregation. *Int J Biol Macromol*. 2019;**134**:1022–37.
61 Giorgetti S, Greco C, Tortora P, Aprile FA. Targeting amyloid aggregation: an overview of strategies and mechanisms. *Int J Mol Sci*. 2018;**19**:2677.
62 Sahoo BR, Genjo T, Bekier M, Cox SJ, Stoddard AK, Ivanova M, et al. Alzheimer's amyloid-beta intermediates generated using polymer-nanodiscs. *Chem Commun*. 2018;**54**:12883–6.
63 Knowles TPJ, Vendruscolo M, Dobson CM. Erratum: the amyloid state and its association with protein misfolding diseases (nature reviews molecular cell biology (2014) 15 (384-396)). *Nat Rev Mol Cell Biol*. 2014;**15**:496.
64 Iadanza MG, Jackson MP, Hewitt EW, Ranson NA, Radford SE. A new era for understanding amyloid structures and disease. *Nat Rev Mol Cell Biol*. 2018;**19**:755–73.
65 Tahara K, Kim HD, Jin JJ, Maxwell JA, Li L, Fukuchi KI. Role of toll-like receptor signalling in Aβ uptake and clearance. *Brain*. 2006;**129**:3006–19.
66 Kakimura J-I, Kitamura Y, Takata K, Umeki M, Suzuki S, Shibagaki K, et al. Microglial activation and amyloid-β clearance induced by exogenous heat-shock proteins. *FASEB J*. 2002;**16**:601–3.
67 Fernandez-Vizarra P, Lopez-Franco O, Mallavia B, Higuera-Matas A, Lopez-Parra V, Ortiz-Muñoz G, et al. Immunoglobulin G fc receptor deficiency prevents Alzheimer-like pathology and cognitive impairment in mice. *Brain*. 2012;**135**:2826–37.
68 Costantini C, Della-Bianca V, Formaggio E, Chiamulera C, Montresor A, Rossi F. The expression of p75 neurotrophin receptor protects against the neurotoxicity of soluble oligomers of β-amyloid. *Exp Cell Res*. 2005;**311**:126–34.
69 Wisniewski T, Drummond E. APOE-amyloid interaction: therapeutic targets. *Neurobiol Dis*. 2020;**138**:104784.

70 Linden R, Martins VR, Prado MAM, Cammarota M, Izquierdo I, Brentani RR. Physiology of the prion protein. *Physiol Rev*. 2008;**88**:673–728.
71 Evans CG, Wisén S, Gestwicki JE. Heat shock proteins 70 and 90 inhibit early stages of amyloid β-(1-42) aggregation in vitro. *J Biol Chem*. 2006;**281**:33182–91.
72 Campanella C, Pace A, Bavisotto CC, Marzullo P, Gammazza AM, Buscemi S, et al. Heat shock proteins in Alzheimer's disease: role and targeting. *Int J Mol Sci*. 2018;**19**:2603.
73 Xu Y, Maya-Martinez R, Guthertz N, Heath GR, Manfield IW, Breeze AL, et al. Tuning the rate of aggregation of hIAPP into amyloid using small-molecule modulators of assembly. *Nat Commun*. 2022;**13**:1040.
74 Xie W, Kim KH, Vince R, More SS. The amyloid aggregation accelerator diacetyl prevents cognitive decline in Alzheimer's mouse models. *Chem Res Toxicol*. 2021;**34**:1355–66.
75 Luo J, Yu CH, Yu H, Borstnar R, Kamerlin SCL, Gräslund A, et al. Cellular polyamines promote amyloid-Beta (Aβ) peptide fibrillation and modulate the aggregation pathways. *ACS Chem Nerosci*. 2013;**4**:454–62.
76 Limbocker R, Chia S, Ruggeri FS, Perni M, Cascella R, Heller GT, et al. Trodusquemine enhances Aβ 42 aggregation but suppresses its toxicity by displacing oligomers from cell membranes. *Nat Commun*. 2019;**10**:225.
77 Sahoo BR, Genjo T, Nakayama TW, Stoddard AK, Ando T, Yasuhara K, et al. A cationic polymethacrylate-copolymer acts as an agonist for β-amyloid and an antagonist for amylin fibrillation. *Chem Sci*. 2019;**10**:3976–86.
78 Petersen RC, Aisen P, Boeve BF, Geda YE, Ivnik RJ, Knopman DS, et al. Mild cognitive impairment due to Alzheimer disease in the community. *Ann Neurol*. 2013;**74**:199–208.
79 Schmittschmitt JP, Scholtz JM. The role of protein stability, solubility, and net charge in amyloid fibril formation. *Protein Sci*. 2009;**12**:2374–8.
80 Gidalevitz T, Krupinski T, Garcia S, Morimoto RI. Destabilizing protein polymorphisms in the genetic background direct phenotypic expression of mutant SOD1 toxicity. *PLoS Genet*. 2009;**5**:e1000399.
81 Tsuboyama K, Osaki T, Matsuura-Suzuki E, Kozuka-Hata H, Okada Y, Oyama M, et al. A widespread family of heat-resistant obscure (hero) proteins protect against protein instability and aggregation. *PLoS Biol*. 2020;**18**:e3000632.
82 Chiti F, Calamai M, Taddei N, Stefani M, Ramponi G, Dobson CM. Studies of the aggregation of mutant proteins in vitro provide insights into the genetics of amyloid diseases. *Proc Natl Acad Sci USA*. 2002;**99**:16419–26.
83 Chiti F. Relative importance of hydrophobicity, net charge, and secondary structure propensities in protein aggregation. In: Uversky VN, Fink AL, editors. Protein misfolding, aggregation, and conformational diseases. Volume **4**. Boston: Springer; 2007. p. 43–59.
84 Kim Y, Park JH, Lee H, Nam JM. How do the size, charge and shape of nanoparticles affect amyloid β aggregation on brain lipid bilayer? *Sci Rep*. 2016;**6**:19548.
85 Pálmadóttir T, Malmendal A, Leiding T, Lund M, Linse S. Charge regulation during amyloid formation of α-synuclein. *J Am Chem Soc*. 2021;**143**:7777–91.
86 Stroo E, Janssen L, Sin O, Hogewerf W, Koster M, Harkema L, Youssef SA, Beschorner N, Wolters AHG, Bakker B, Thathiah A, Foijer F, van de Sluis B, van Deursen J, Jucker M, de Bruin A & Nollen EAA (2021) Deletion of Serf2 shifts amyloid conformation in an Aβ amyloid mouse model. *bioRxiv*. https://doi.org/10.1101/2021.01.05.423442. [PREPRINT]
87 Cleverley K, Lee WC, Mumford P, Collins T, Rickman M, Cunningham TJ, et al. A novel knockout mouse for the small EDRK-rich factor 2 (Serf2) showing developmental and other deficits. *Mamm Genome*. 2021;**32**:94–103.
88 Kuo PH, Doudeva LG, Wang YT, Shen CKJ, Yuan HS. Structural insights into TDP-43 in nucleic-acid binding and domain interactions. *Nucleic Acids Res*. 2009;**37**:1799–808.
89 Wang X, Schwartz JC, Cech TR. Nucleic acid-binding specificity of human FUS protein. *Nucleic Acids Res*. 2015;**43**:7535–43.
90 Görlach M, Burd CG, Dreyfuss G. The determinants of RNA-binding specificity of the heterogeneous nuclear ribonucleoprotein C proteins. *J Biol Chem*. 1994;**269**:23074–8.
91 Yu S, Li X, Liu G, Han J, Zhang C, Li Y, et al. Extensive nuclear localization of α-synuclein in normal rat brain neurons revealed by a novel monoclonal antibody. *Neuroscience*. 2007;**145**:539–55.
92 Maroteaux L, Campanelli JT, Scheller RH. Synuclein: a neuron-specific protein localized to the nucleus and presynaptic nerve terminal. *J Neurosci*. 1988;**8**:2804–15.
93 Miraglia F, Ricci A, Rota L, Colla E. Subcellular localization of alpha-synuclein aggregates and their interaction with membranes. *Neural Regen Res*. 2018;**13**:1136–44.
94 Gal J, Zhang J, Kwinter DM, Zhai J, Jia H, Jia J, et al. Nuclear localization sequence of FUS and induction of stress granules by ALS mutants. *Neurobiol Aging*. 2011;**32**:2323.e27–40.
95 Shiihashi G, Ito D, Yagi T, Nihei Y, Ebine T, Suzuki N. Mislocated FUS is sufficient for gain-of-toxic-function amyotrophic lateral sclerosis phenotypes in mice. *Brain*. 2016;**139**:2380–94.

96 Vance C, Scotter EL, Nishimura AL, Troakes C, Mitchell JC, Kathe C, et al. ALS mutant FUS disrupts nuclear localization and sequesters wild-type FUS within cytoplasmic stress granules. *Hum Mol Genet*. 2013;**22**:2676–88.

97 Strong MJ, Volkening K, Hammond R, Yang W, Strong W, Leystra-Lantz C, et al. TDP43 is a human low molecular weight neurofilament (hNFL) mRNA-binding protein. *Mol Cell Neurosci*. 2007;**35**:320–7.

98 Archbold HC, Jackson KL, Arora A, Weskamp K, Tank EMH, Li X, et al. TDP43 nuclear export and neurodegeneration in models of amyotrophic lateral sclerosis and frontotemporal dementia. *Sci Rep*. 2018;**8**:4606.

99 Tziortzouda P, van den Bosch L, Hirth F. Triad of TDP43 control in neurodegeneration: autoregulation, localization and aggregation. *Nat Rev Neurosci*. 2021;**22**:197–208.

100 Sun Y, Zhao K, Xia W, Feng G, Gu J, Ma Y, et al. The nuclear localization sequence mediates hnRNPA1 amyloid fibril formation revealed by cryoEM structure. *Nat Commun*. 2020;**11**:6349.

101 Uversky VN. Protein intrinsic disorder-based liquid–liquid phase transitions in biological systems: complex coacervates and membrane-less organelles. *Adv Colloid Interface Sci*. 2017;**239**:97–114.

102 Posey AE, Holehouse AS, Pappu RV Phase separation of intrinsically disordered proteins. *Methods Enzymol*. 2018;**611**:1–30.