Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Cell Biology

GFAP-isoforms in the nervous system: Understanding the need for diversity

Alexandra J. E. M. de Reus[1], Onur Basak[1], Werner Dykstra[1], Jessy V. van Asperen[2], Emma J. van Bodegraven[1,a] and Elly M. Hol[1,a]

Abstract

Glial fibrillary acidic protein (GFAP) is an intermediate filament (IF) protein expressed in specific types of glial cells in the nervous system. The expression of GFAP is highly regulated during brain development and in neurological diseases. The presence of distinct GFAP-isoforms in various cell types, developmental stages, and diseases indicates that GFAP (post-)transcriptional regulation has a role in glial cell physiology and pathology. GFAP-isoforms differ in sub-cellular localisation, IF-network assembly properties, and IF-dynamics which results in distinct molecular interactions and mechanical properties of the IF-network. Therefore, GFAP (post-)transcriptional regulation is likely a mechanism by which radial glia, astrocytes, and glioma cells can modulate cellular function.

Addresses

[1] Department of Translational Neuroscience, Brain Center, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
[2] Institut NeuroMyoGène (INMG), Unité Physiopathologie et Génétique du Neurone et du Muscle, Unversité Claude Bernard Lyon 1 CNRS UMR 5261, INSERM U1315, Lyon, France

Corresponding author: Hol, Elly M. (e.m.hol-2@umcutrecht.nl)
[a] Shared last authorship.

Current Opinion in Cell Biology 2024, **87**:102340

This review comes from a themed issue on **Cell Architecture 2023**

Edited by **Patrick Lusk** and **John Eriksson**

For complete overview of the section, please refer the article collection - Cell Architecture 2023

Available online 23 February 2024

https://doi.org/10.1016/j.ceb.2024.102340

Introduction

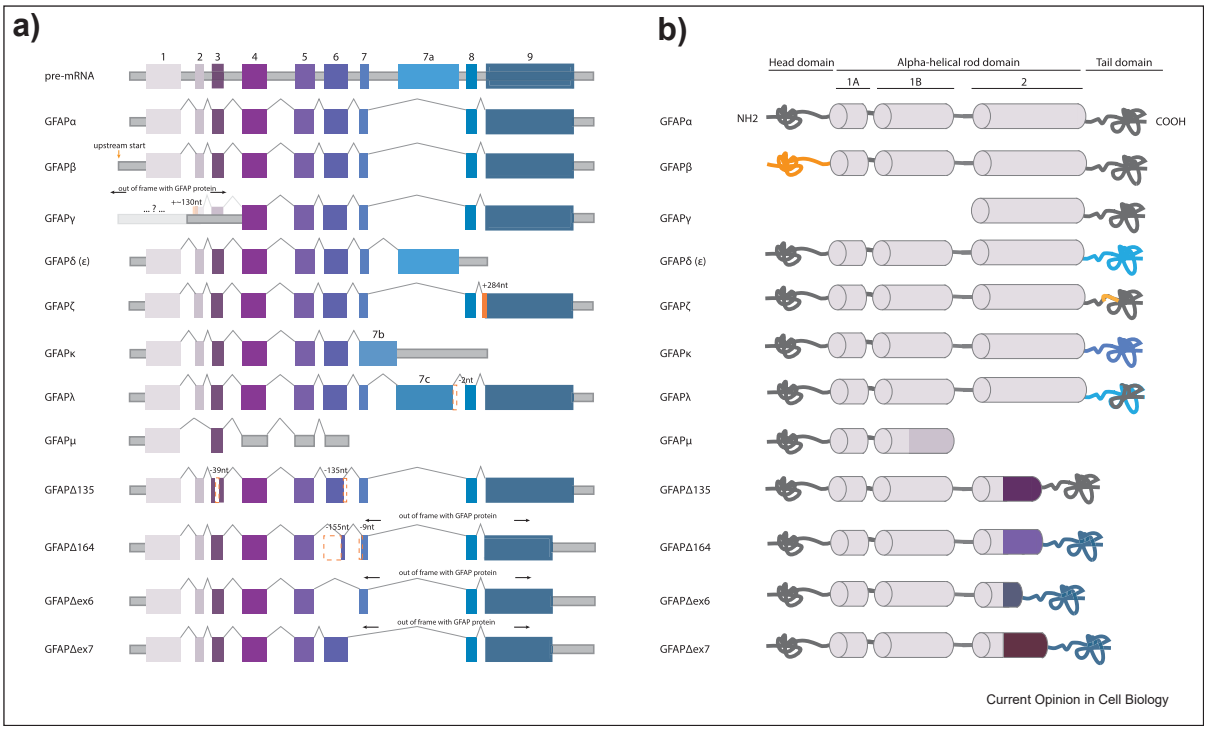
Glial fibrillary acidic protein (GFAP) is an intermediate filament (IF) protein and forms, together with other IF proteins, microtubules, and actin, the cytoskeleton of radial glia, astrocytes, and glioma cells [1]. The protein forms complex filamentous structures with vimentin, synemin, and nestin, which are cell-type and cell-state dependent. GFAP-isoform expression adds another level of complexity to these filamentous structures. Different mRNA products of the same gene can arise due to transcriptional events, such as alternative transcription start- or end-sites, or post-transcriptional events, such as alternative splicing and alternative 3′-end polyadenylation [2]. These mRNA variants may in turn, but not necessarily, give rise to different proteins. Alternative splicing is a common phenomenon. Around 95% of human genes with multiple exons undergo alternative splicing [3]. Isoforms contribute to tissue- and cell-type specificity and play important roles during development and in disease [3–5]. Alternative splicing is prevalent in the brain across species [4,5], and especially genes with cytoskeleton-related functions (based on gene ontology) are alternatively spliced [6]. In this review, we will discuss the occurrence and impact of (post-)transcriptional regulation of GFAP, an IF gene that is known to produce a high number of isoforms. Why cells need so many GFAP-isoforms remains an open question. Here, we will discuss new findings on alternative splicing and GFAP-isoforms in the context of IF biology and provide an outlook on what new technologies can bring us to resolve this question.

GFAP-isoforms, alternative transcription- and translation-start sites and splicing

To this day, twelve human and seven murine GFAP-isoforms have been described [7], with GFAPα as the canonical isoform and GFAPλ and -μ as the newest isoforms [8,9] (Figure 1a + Table 1). Both GFAPβ and GFAPγ originate from alternative transcription start sites, the other 10 isoforms result from alternative splicing (Table 1). Most GFAP-isoforms are predicted to have the typical IF structure of an amino(N)-terminal head domain, a central α-helical rod, and a carboxy(C)-terminal tail [10], although not every isoform seems to be translated into protein (Figure 1b + Table 1). The complete crystal structure of IF proteins is difficult to acquire due to their intrinsic physical flexibility and disordered head and tail domains [11]. However, the crystal structure of the 1B-rod domain of GFAP, which is shared by most of the GFAP isoforms, has been obtained [12].

Differential translation initiation- and stop-sites add another level to protein diversification. The 80S

Figure 1



GFAP-isoforms and hypothesised protein structure.
a) Pre-mRNA of GFAP and GFAP-isoforms are presented. Exons are numbered and shown with coloured boxes (exons and introns are not drawn to scale). Dark grey boxes represent untranslated regions. Insertions into exons are presented with orange-filled boxes, and deletions in exons with unfilled, orange-dotted boxes. The amount of inserted or deleted nucleotides (nt) is indicated above the orange boxes. The orange arrow in GFAPβ indicates an alternative transcription start site. The transcription start site for GFAPγ is unknown, as indicated by the question mark. GFAPγ includes part of intron 1, however, this inclusion leads to out-of-frame protein with canonical GFAP, which is represented with out-of-frame arrows and transparent intron 1 to exon 3 boxes. The translation of this mRNA would potentially result in an entirely different protein than GFAP. The longest open reading frame in-frame with canonical GFAP starts with an initiation codon present in exon 4. Deletions of GFAPΔ164, GFAPΔex6 and GFAPΔex7 are also out-of-frame, indicated with the out-of-frame arrows. It is likely that an alternative polyadenylation signal within the 3′UTR of exon 3 of GFAPμ is used, causing the removal of the exon 3-exon 4 junction. **b)** The hypothesised protein structures of the GFAP-isoforms are shown here. Proteins have a typical intermediate filament structure with an N-terminal-head, alpha-helical rod and a C-terminal-tail domain. Different colours indicate alterations in domains. The C-termini of GFAPα, GFAPβ, GFAPγ, and GFAPΔ135 are the same and are represented by the same colour. The N terminal peptide sequence of GFAP from Schwann cells, which is suggested to be GFAPβ, is different from the canonical head domain of GFAPα, as indicated with a different colour. GFAPγ in-frame with canonical GFAP has an initiation codon present in exon 4. This would result in a protein lacking the head domain and a part of the rod domain. Since the head domain is essential for the assembly of filaments, it is likely that this protein would not form filaments or integrate into the IF-network. GFAPδ and GFAPκ have unique C-termini. GFAPζ likely has an alteration in its C-terminus, due to the inclusion of part of intron 8. GFAPλ will likely contain a combination of the C-terminal regions of GFAPα and GFAPδ. Skipping of exon 2 in GFAPμ induces a frameshift and a premature termination codon in exon 3, resulting in an alteration in coil 1B and a shortened protein. The GFAP+1 isoforms have alterations in their coil 2, indicated with different colours. GFAPΔ164, GFAPΔex6, and GFAPΔex7 share the same C-termini.

initiation complex may assemble at an up- or downstream translation initiation site (TIS) compared to the annotated TIS or terminate at an alternative stop codon. Thus, one transcript can result in multiple proteins or polypeptides. Ribosome footprinting revealed a reduction in the utilization of the annotated *Gfap* TIS following KCl depolarisation in a neuron-glia culture. Additionally, a subtle increase was observed in the usage of an out-of-frame downstream TIS [13]. This demonstrates that the TISs of *Gfap* are regulated and suggests that this alternate use has a functional significance. Overall, the transcriptomic and proteomic variants that arise from the GFAP gene are substantial. We are just beginning to understand the complexity of the regulation of GFAP expression and the function of the GFAP-isoforms.

Regulation of GFAP mRNA splicing

Many transcription factors control GFAP expression [14]. However, the factors that regulate GFAP alternative splicing are still largely unknown. Splicing of the GFAPδ variant is dependent on splicing factor SR6 and histone acetylation [15]. Low-level inhibition of histone deacetylases (HDAC) increases GFAPδ and GFAPκ expression in an SR6-dependent manner. HDAC inhibition leads to a collapse of the GFAP and vimentin

Table 1

Overview of GFAP-isoforms.

Isoform	Species (mRNA)	Alternative splicing or alternative transcription start site	Differences compared to canonical form	Observed in (tissue/cell type):	Validated protein expression?
GFAPα	Human, mouse, rat, pig.	The canonical form consists of nine exons.	–	The highest expressed isoform predominantly expressed in CNS astrocytes. Also expressed in radial glia and glioma cells.	Yes, in human, mouse, and rat. Human sequence results in a 432 aa protein.
GFAPβ	Human, mouse, rat.	Alternative transcription start site.	169 or 79 nt upstream start site, uncertainty about where translation starts.	Present in rodent Schwann cells, hepatic stellate cells and rodent brain. Also in human lymphocytes and gliomas.	No, tryptic mapping showed the absence of an N-terminal peptide sequence in GFAP from Schwann cells, which is suggested to be GFAPβ.
GFAPγ	Human, mouse.	Alternative transcription start site.	TSS is unknown, either upstream to canonical or within intron 1. mRNA excludes exon 1 and includes the last ~130 nt of intron 1, leading to an out-of-frame GFAP protein. The longest open reading frame in-frame with canonical GFAP, starts with an ATG in exon 4.	Present in mouse brain, bone marrow, and spleen, and in adult human brain.	No specific antibody available.
GFAPδ (=GFAPε)	Human, mouse, other mammals.	Alternative splicing.	Has an extra exon, i.e. 7a (part of canonical intron 7) with an alternative polyadenylation signal. Lacks exons 8 and 9.	In adult human brain: expressed in a subset of SVZ neural stem cells, in subpial astrocytes, in astrocytes adjoining the RMS, and in the olfactory bulb. Also expressed in pathological conditions; ischemic stroke, vanishing white matter disease, AxD in Rosenthal fibers, spinal cord astrocytoma and glioma.	Yes, in human and mouse. Human sequence results in a 431 aa protein, where the last 41aa are different from canonical GFAP protein.
GFAPζ	Human, mouse.	Alternative splicing.	Includes the last 284 nt of intron 8.	Human and mouse control/wild-type and AD brains.	No specific antibody available.
GFAPκ	Human, mouse, pig.	Alternative splicing.	Lacks exon 8 and 9, but contains exon 7b, which consists of exon 7 and intron 7a. Has alternative polyadenylation signal.	Expressed in enteric glial cells and upregulated in patients with Parkinson's disease. Also expressed in human, mouse, and pig brains; in AD mouse and brain tissue; Rosenthal fibers of AxD patients' brains; human gliomas.	No proof, there are specific antibodies available.
GFAPλ	Human	Alternative splicing.	Contains exon 7c, which includes exon 7 and shares the same splice acceptor site as exon 7a, but employs a splice donor site 2 nt 5′ of the stop codon of GFAPδ.	AxD's mutations result in overexpression of this isoform.	Yes, predicted protein consists of the first 430 aa of GFAPδ followed by the terminal 42 aa of GFAPα.
GFAPμ	Human	Alternative splicing.	Skips exon 2, which results in an out-of-frame transcript with a premature termination codon in exon 3, causing a short coding sequence.	Human healthy brain, glioma cell lines and primary glioma cells.	Yes, results in a shorter protein of 179 aa.
GFAPΔ135	Human, not found in mouse.	Alternative splicing.	A 39 nt in-frame deletion in exon 3 and a 135 nt in-frame deletion in exon 6.	Expressed in specific mature human astrocyte subtype and in AD and focal lesions in chronic epilepsy.	No proof as GFAPpan antibodies cannot discriminate between canonical GFAP and this isoform.

(*continued on next page*)

Table 1. (*continued*)

Isoform	Species (mRNA)	Alternative splicing or alternative transcription start site	Differences compared to canonical form	Observed in (tissue/cell type):	Validated protein expression?
GFAPΔ164	Human, not found in mouse.	Alternative splicing.	A 155 nt and 9 nt out-of-frame deletion in exon 6 and exon 7 respectively.	Expressed in a specific mature human astrocyte subtype, and in AD and focal lesions in chronic epilepsy. Determined by GFAP+1 antibody, directed against the C-terminus of GFAPΔex6, GFAPΔex7 and GFAPΔex164, which is similar.	Maybe, but GFAP+1 antibody cannot discriminate between GFAPΔ164, - ex6 and ex7. Leads to an out-of-frame protein which has canonical N-terminus but altered C-terminus
GFAPΔex6	Human, not found in mouse.	Alternative splicing.	Exon 6 deletion.	Expressed in a specific mature human astrocyte subtype, in AD and in focal lesions in chronic epilepsy.	See GFAPΔ164.
GFAPΔex7	Human, mouse.	Alternative splicing.	Exon 7 deletion.	Low expression levels in mouse and human hippocampus and in AD mouse models and human AD samples.	See GFAPΔ164.

Information in this table is based on reviews [1,7] that include the original references of the studies. For more information see ENSEMBL for human (GFAP ENSG00000131095), mouse (Gfap ENSMUSG00000020932) and rat (Gfap ENSRNOG00000002919) specific isoforms. aa = amino acids, AD = Alzheimer's disease, AxD = Alexander Disease, CNS = central nervous system, nt = nucleotides, RMS = rostral migratory stream, SVZ = subventricular zone, TSS = transcription start site.

network, which is rescued by SR6 depletion, demonstrating the impact of GFAP-isoforms on the organisation of the IF network [15]. GFAPδ depends on regulated usage of the alternative exon 7a and the alternative polyadenylation site in this exon, whereas GFAPκ is the result of an inefficient usage of the splice acceptor site upstream exon 7a [16]. GFAP splicing is also affected by mutations in the GFAP gene, that cause Alexander Disease (AxD), a rare neurological disorder [17]. The 1289G > A and 1290C > A GFAP heterozygous mutations both result in an upregulation of the GFAPλ isoform in 3 AxD cases [8]. In addition, others discovered that a heterozygous c.619-1G > A GFAP mutation in the canonical splice acceptor site of intron 3 resulted in very late-onset AxD via activation of a cryptic splice site, leading to loss of Glu207 at the start of exon 4 [18]. These studies provide examples of how GFAP alternative splicing can be regulated or affected in disease, but at the same time illustrate the complexity of the regulatory mechanisms.

GFAP-isoforms in cells and tissues: Where and when are they expressed?

GFAP-isoforms in development and aging

GFAP is predominately expressed in the central and peripheral nervous system [1]. Nowadays, with the use of single-cell RNA sequencing (scRNAseq), unprecedented resolution for identifying alternative isoform usage can be achieved [19]. Using this technology, Moratibo and colleagues determined gene co-expression modules of radial glial cells (RGL) in the early postnatal mouse hippocampus and found that *Gfap*-isoforms, which they annotated as *Gfap-0* to *Gfap-5*, were differentially expressed across the modules [20]. *Gfap-0* and *-1* were high in astrocyte-specific RGL. In contrast, *Gfap-3/-4/-5* were high in all RGL and were downregulated during differentiation. Looking at their structure and the isoform-level information [21], *Gfap-1* and *Gfap-5* likely correspond to *Gfapα* and *Gfap-3* to *Gfapδ*. Both *Gfap-0* and *Gfap-2* are, in our opinion, undocumented isoforms. This shift in alternative splicing of GFAP-isoforms may be involved in restricting the fate of RGLs or result from their timing of differentiation and migration during the complex hippocampal development. When looking at GFAP during human development, it is worth noting that GFAP expression is not restricted solely to the astroglia lineage and GFAP amount is very variable in dividing ventricular zone progenitors [22,23]. New evidence is beginning to demonstrate a role for other IFs during cell differentiation and lineage determination [24]. Thus, it is likely that the variable amount and isoforms of GFAP also play a role.

Oshea et al. 2022 investigated how different CNS environments influence the gene expression and differentiation fate of mouse neural progenitor cells (NPCs)

[25]. NPCs exposed to different cell-fate-determining factors, either promoting astroglial or oligodendroglial lineages, showed ∼44 kDA *Gfap* variants in cells with long filamentous appearances. ∼50 kDA *Gfap* variants were enriched in NPCs cultured in the presence of fetal bovine serum or under spontaneous differentiation conditions and displayed a flatter or more polygonal-shaped morphologies [25]. The identity of these *Gfap* variants was not further investigated but the differences in molecular weight could arise from differential post-translational modifications or alternative splicing. It does show that GFAP variants are associated with morphological differences in differentiating NPCs suggesting that GFAP variants are involved in the establishment of certain developmental cell fates.

Gómez-Gálvez et al. 2020 found that *Gfapδ* expression is increased in the substania nigra of healthy middle-aged rats compared to young and old rats [26]. Furthermore, proteomic and bioinformatic analyses revealed that there are proteomic changes in the substania nigra of healthy rats during aging, that are related to protein aggregation, oxidative stress, and astrocyte function and that *Gfap* is an important hub network gene, that is connected with many proteins that are altered during aging [26]. The role of GFAP during aging still needs further investigation, but as this protein is also upregulated in neurodegenerative diseases, it is likely to play a part in (ab)normal aging.

GFAP-isoforms in disease

Besides the presence of the canonical GFAPα in all GFAP-expressing cells, the GFAP-isoforms show region-, cell type- or disease-specific expression (summarised in Table 1). For instance, there is an increase in GFAPλ expression in AxD and GFAPδ and GFAPκ are enriched in Rosenthal fibers of post-mortem AxD patients' brain [27]. In the human adult brain, GFAPδ is preferentially expressed in a subset of astrocytes identified as neural stem cells (NSCs) in the subventricular zone (SVZ). GFAPδ is also expressed in pathological conditions such as vanishing white matter (VWM) disease, spinal cord astrocytoma and, together with GFAPμ, in glioma [7,9,28–31]. Gliomas are diffusely infiltrating brain tumors derived from glial cells or their precursors, which are categorised in different grades of malignancy. Interestingly, proteomic experiments showed decreased N-terminal and C-terminal peptide coverage in normal brain tissue compared to glioblastoma tissue, indicative of differential expression of GFAP isoforms [32]. Furthermore, in grade IV glioma we observed higher levels of GFAPδ/α ratio compared to lower grade glioma [28,33,34]. *In vitro* studies in glioma cell models show that a higher GFAPδ/α, due to a decrease in GFAPα, induces molecular alterations which impact the interaction of the cell with the extracellular matrix and equip the cell for invasion, a key characteristic of high-grade glioma [33–35]. Indeed, glioma cells with a different GFAPδ/α ratio show different 3D growth and invasion dynamics in *ex vivo* organotypic brain slices and in the *in vivo* mouse brain [28]. The GFAPδ/α ratio is also affected in VWM disease, due to an increase in GFAPδ, whereas GFAPα levels remained unchanged [36]. This increase in GFAPδ/α corresponds to less mature astrocytes, which likely contributes to the VWM pathology [36].

GFAP alternative splicing as a tool to modulate cellular function

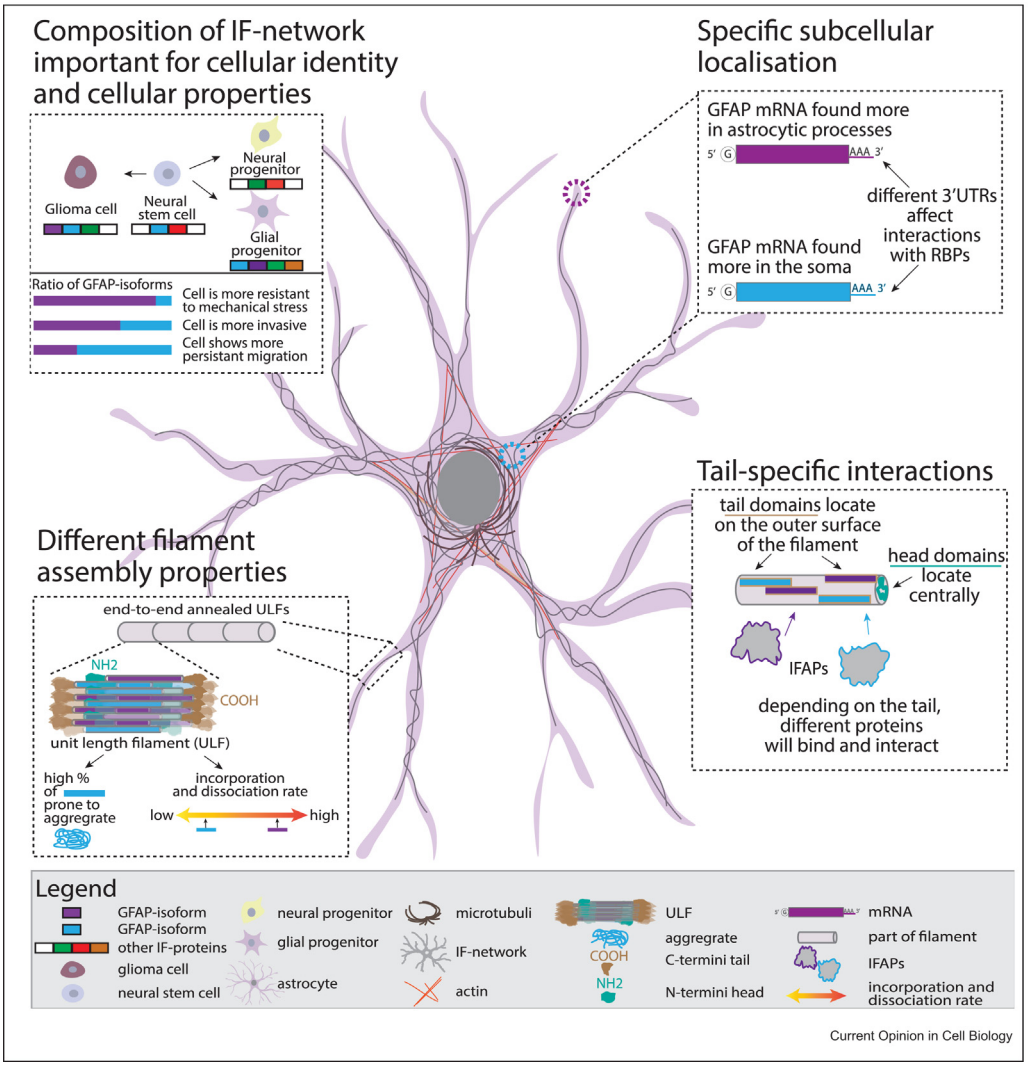
Alternative splicing contributes to the diversity in cellular functions [3,5,6]. In case of IFs, the composition of the IF-network is associated with specific cellular processes, and evidence is increasing that GFAP alternative splicing provides an additional layer of modulation to adapt cellular function in the CNS (Figure 2).

Modulation of IF-network characteristics

GFAP-isoforms mainly differ in their C-terminal tail and coil 2 rod domain (Figure 1b). IF proteins form coiled-coil dimers, and two dimers form anti-parallel tetramers: the building blocks of a unit-length filament (ULF). ULFs anneal longitudinally to form mature filaments [10]. The differences in GFAP-isoform protein structures can affect filament assembly and filament-filament interactions and thus the IF-network characteristics. GFAPα can self-assemble into filaments, but GFAPδ, GFAPκ, GFAPλ, and GFAPμ need GFAPα or vimentin to form co-assembled filaments [9,27,37]. IFs only tolerate low levels of these assembly compromised isoforms before it leads to a collapse of the network [9,27,37–40]. In mutated GFAP dimers, such as in AxD, stabilising interactions are lost and tetramer formation is affected, causing a similar disruption of GFAP filament assembly [12].

Low-level incorporation of assembly incompetent isoforms is allowed but will change filament properties and thereby characteristics of the IF-network. For example, GFAPδ shows a slower incorporation and dissociation rate compared to GFAPα [38], which is likely to affect the exchange dynamics of IF subunits [38]. Different cellular processes might depend on higher IF-network turnover rates and GFAP alternative splicing could be a mechanism by which the cell modulates its filament network dynamics. Indeed, the altered IF-network dynamics changes cell shape and focal adhesion size [38]. In addition, GFAP alternative splicing might lead to a change in the mechanical properties of the IF-network. IFs and their composite networks are well-known for their unique mechanical properties. The intrinsic physical properties of each filament and the interactions between filaments when they organise into a network make the IF-network flexible, stretchable, and highly resistant to breakage

Figure 2



GFAP alternative splicing as a cellular tool to modulate function: a hypothesis. Astrocyte with cytoskeleton consisting of actin, microtubules and IF-network is shown. GFAP-isoforms are presented in blue and purple colours, for simplicity only two isoforms are depicted. Left upper box: composition of IF-network likely contributes to cell fate during development and disease. Different compositions of the IF-network are represented with the coloured boxes underneath the cells. Depending on the composition of the IF-network, cells have different properties. For example, they are more invasive or more resistant to mechanical stress. This is also influenced by differences in GFAP-isoform compositions. Lower left box: unit-length filaments (ULFs) are shown to anneal longitudinally to form mature filaments. The differences in GFAP-isoform protein structures can affect filament assembly and filament-filament interactions and thus the IF-network characteristics. For example, certain GFAP-isoforms are more prone to aggregate and a high percentage of this isoform in the network will lead to a collapse of the IF network. Isoforms with slower incorporation and dissociation rates will likely affect the exchange dynamics of IF subunits. Upper right box, the 3′untranslated region (UTR) is important for subcellular localisation, which is likely affected by different interactions with RNA binding proteins (RBPs). Lower right box: the tail domains are located on the outer surface of the filament, exposed to interact with intermediate filament-associated proteins (IFAPs). Therefore, isoforms with different tails have different protein interactions and could impact cellular functions.

[41,42]. Cell mechanical properties result from the fundamental structure of IFs and the interactions between IF proteins within and between filaments and are known to change during development and in pathology. The C-terminal tail is important to withstand high stresses in the IF-network [41,42] and the extent to which the IF monomers and polymers engage in hydrophobic and electrostatic interactions, and thereby their mechanical properties, is determined by the IF's amino acid sequences [43]. As GFAP-isoforms have different amino-acid sequences and C-terminal tails, it is likely that GFAP-isoforms can change the mechanical properties of the IF-network. Furthermore, GFAPα and GFAPδ have different predicted kinase binding sites and phosphorylation residues [44] that can contribute to the mechanical properties of the IF-network [45] and thereby modulate cellular function.

Modulation of filament-protein interactions

Based on the 3D organisation of vimentin [46], the C-terminal tails of GFAP are likely on the outside of the filaments, where they are available for interaction with other proteins. The incorporation of GFAP-isoforms in the IF-network therefore will affect the variety of protein interactions. Indeed, GFAPδ has a decreased affinity for peripherin, internexin, and desmosomal proteins compared to GFAPα, and specifically binds to presenilin-1, a protein that is part of the γ-secretase complex [40]. Thus, the specific C-terminus of GFAP-isoforms leads to different protein-protein interactions and could impact cellular functions.

Specialised subcellular domains in astrocytes

Recent studies have shown sub-cellular localisation of mRNAs and that local translation occurs in astrocytes [47–50]. Astrocytes have specialized domains, with processes that affect synaptic function and endfeet that are involved in blood-brain barrier integrity. Distinct *Gfap* mRNA localisations in astrocytes are indicative of subcellar translation of specific *Gfap*-isoforms. *Gfap*α is more abundant in the processes, while *Gfap*δ is found more in the soma [47,49]. These observations suggest a role for *Gfap*-isoforms in supporting a unique subcellular environment in astrocytes. In Alzheimer's disease mice, *Gfap* mRNAs are redistributed, which is even more pronounced in plaque-associated astrocytes [49]. This indicates that sub-cellular localisation of GFAP-isoform mRNAs changes in disease and implies that astrocyte mRNA transport is dysregulated in Alzheimer's. MicroRNAs (miRNAs) and RNA binding proteins (RBPs), such as quaking (QKI), typically interact with 3′UTRS and are involved in mRNA transport, translation, and stability [14]. There are indications that miRNAs and QKI are involved in post-transcriptional regulation of GFAP expression [14]. The inclusion of exon 7a combined with a different polyadenylation signal in GFAPδ, changes the 3′UTR sequences and can affect interactions with RBPs, which can explain the different mRNA subcellular localisation. Overall, subcellular localisation of GFAP-isoform mRNAs likely permits spatially controlled protein expression, which could function for efficient local translation and fast cellular responses, essential for a rapid cell-state-specific reaction to pathological conditions. However, whether subcellular GFAP-networks can have different compositions and whether local IF subunit exchange or polymerisation occurs remains to be investigated.

Resolving GFAP-isoform diversity and function: The way forward

We hypothesise that GFAP alternative splicing provides a mechanism to adapt to cellular functions in the CNS, as alternative splicing results in differential subcellular localisation of GFAP-isoforms, changes IF- and protein-interactions, which in turn may result in different cellular properties and functions. We just started to uncover the functional complexity of GFAP-isoforms, see box 1 for the main remaining questions. Before we can answer these questions, some critical issues need to be resolved first. One main obstacle is the annotation of GFAP-isoforms in ENSEMBL (human GFAP ENSG00000131095). As the list of isoforms is not validated and incomplete; a total of 38 human GFAP transcripts are listed which is not supported yet by experimental evidence. Most information is based on exon abundance and exon-exon splicing events detected by short-read sequencing. However, it is impossible to reliably detect multiple splicing events on the same mRNA molecules. Recent advances in long-read sequencing will help to conclusively show which GFAP-isoforms are expressed, including the detection of the transcription start site, poly adenylation sites, intron inclusion and alternative splicing events. Combined with scRNAseq, this method allows for cell-type specific isoform detection [19,51]. Furthermore, the recently published human pan-genome [52] better represents the diverse human population and may lead to the identification of new splicing variants present in only a fraction of the population.

Box 1. Remaining questions about GFAP-isoform diversity

How many GFAP mRNA isoforms are expressed and which of these encode for functional proteins?

What regulates alternative splicing and transcription start sites in the GFAP gene?

What are the transcription start sites of GFAPβ and GFAPγ?

What are the differential translation initiation- and stop-sites of GFAP isoforms?

What is the mechanism of GFAP-isoform subcellular localisation?

How do the different GFAP-protein structures affect filament assembly?

Does the incorporation of GFAP-isoforms change the properties of the filaments?

Does GFAP-isoform expression impact the mechanical properties of cells?

How are the GFAP N-terminal heads and C-terminal tails arranged within the filament?

How do different GFAP heads and tails affect protein-protein interactions?

What are the post-translational modifications of the GFAP variants?

What is the interactome of the different GFAP variants?

What is the function of GFAP-isoforms during development, ageing, and disease?

There is an inconsistency on the identity of GFAP-isoforms between different sources, which is likely due to differences in isoform annotation. For instance, Joglekar et al. 2021 [21] used long-read sequencing, that can detect isoforms at single cell level, which may reveal *Gfap* isoforms not detected by other studies. Moreover, tissue-specific isoform expression may bias results. Systematic screening of published and new transcriptomic and proteomic datasets with isoform-specific bioinformatics pipelines, as well as experiments designed to reveal GFAP isoforms are essential to discover new GFAP isoforms and elucidate their properties. For example, probe-based imaging and sequencing (e.g., Nanostring CosMx, Nanostring GeoMx system) methods [53] will help to detect different GFAP exon-exon boundaries and can be employed to better quantify GFAP-isoforms via sequencing. Furthermore, single molecular fluorescent *in situ* hybridisation (FISH) probes designed for specific exons of *Gfap* can reveal cell-cell differences in isoform usage in a spatially informed manner [49]. Finally, antisense oligonucleotides and CRISPR technology now provide tools to manipulate *Gfap* isoforms *in vivo* [28,54]. To conclude, these techniques provide interesting opportunities to unravel more about GFAP-isoforms and hopefully result in fruitful investigations.

Declaration of competing interest

Nothing to declare.

Data availability

No data was used for the research described in the article.

Acknowledgements

This work was supported by NWO "VI.Veni.212.117" (E.J.v.B.), UMCU Brain Center "Rudolf Magnus Young Talent Fellowship (E.J.v.B.), and the Dutch Cancer Society "KWF 101123" (E.M.H.). NWO Gravitation program BRAINSCAPES: Roadmap from Neurogenetics to Neurobiology, "NWO: 024.004.012" (E.M.H. and O.B.). The European Union's Horizon 2020 research and innovation programme under the "Marie Skłodowska-Curie grant agreement No 101109363" (J.V.v.A.).

References

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

* of special interest
** of outstanding interest

1. Messing A, Brenner M: **GFAP at 50**. *ASN Neuro* 2020, **12**, 175909142094968.

2. Mitschka S, Mayr C: **Context-specific regulation and function of mRNA alternative polyadenylation**. *Nat Rev Mol Cell Biol* 2022, **23**:779–796.

3. Wright CJ, Smith CWJ, Jiggins CD: **Alternative splicing as a source of phenotypic diversity**. *Nat Rev Genet* 2022, **23**: 697–710.

4. Nikom D, Zheng S: **Alternative splicing in neurodegenerative disease and the promise of RNA therapies**. *Nat Rev Neurosci* 2023, https://doi.org/10.1038/s41583-023-00717-6.

5. Mazin PV, Khaitovich P, Cardoso-Moreira M, Kaessmann H: **Alternative splicing during mammalian organ development**. *Nat Genet* 2021, **53**:925–934.

6. Rodriguez JM, Pozo F, di Domenico T, Vazquez J, Tress ML: **An analysis of tissue-specific alternative splicing at the protein level**. *PLoS Comput Biol* 2020, **16**, e1008287.

7. Van Asperen JV, Robe PAJT, Hol EM: **GFAP alternative splicing and the relevance for disease – a focus on diffuse gliomas**. *ASN Neuro* 2022, **14**, 175909142211020.

8. Helman G, Takanohashi A, Hagemann TL, Perng MD, Walkiewicz M, Woidill S, Sase S, Cross Z, Du Y, Zhao L, *et al.*: **Type II Alexander disease caused by splicing errors and aberrant overexpression of an uncharacterized GFAP isoform**. *Hum Mutat* 2020, **41**:1131–1137.

9. * Bodegraven EJ, Sluijs JA, Tan AK, Robe PAJT, Hol EM: **New GFAP splice isoform (GFAPμ) differentially expressed in glioma translates into 21 kDa N-terminal GFAP protein**. *Faseb J* 2021:35.
This is the first study showing a very short GFAP transcript, which codes for a novel 21 kDa GFAP protein. A specific peptide sequence, identifying this isoform has been detected in a proteomics study and is annotated as GFAP-B1DIR4 in proteomicsdb.org. The GFAP isoform misses a large part of the rod domain and has no tail domain. The function of this isoform, that cannot self-assemble, is still elusive. This study shows that validation of annotated mRNA and protein isoforms is needed.

10. Eldirany SA, Lomakin IB, Ho M, Bunick CG: **Recent insight into intermediate filament structure**. *Curr Opin Cell Biol* 2021, **68**: 132–143.

11. Vermeire P-J, Stalmans G, Lilina AV, Fiala J, Novak P, Herrmann H, Strelkov SV: **Molecular interactions driving intermediate filament assembly**. *Cells* 2021, **10**:2457.

12. ** Kim B, Kim S, Jin MS: **Crystal structure of the human glial fibrillary acidic protein 1B domain**. *Biochem Biophys Res Commun* 2018, **503**:2899–2905.
Intermediate filaments are notoriously difficult to crystalize due to the biochemical and biophysical properties of the proteins. This is the first crystal structure of the coiled-coil 1B domain of GFAP. This study shows that GFAP filaments assemble via lateral association of two parallel coiled-coil dimers with their 1B domains arranged in an anti-parallel manner, similar to vimentin. This information is important to understand how GFAP assembles and interacts with other proteins.

13. Sapkota D, Lake AM, Yang W, Yang C, Wesseling H, Guise A, Uncu C, Dalal JS, Kraft AW, Lee J-M, *et al.*: **Cell-type-specific profiling of alternative translation identifies regulated protein isoform variation in the mouse brain**. *Cell Rep* 2019, **26**: 594–607.e7.

14. Brenner M, Messing A: **Regulation of GFAP expression**. *ASN Neuro* 2021, **13**, 175909142098120.

15. Kanski R, Sneeboer MAM, Van Bodegraven EJ, Sluijs JA, Kropff W, Vermunt MW, Creyghton MP, De Filippis L, Vescovi A, Aronica E, *et al.*: **Histone acetylation in astrocytes suppresses GFAP and stimulates a re-organization of the intermediate filament network**. *J Cell Sci* 2014, https://doi.org/10.1242/jcs.145912.

16. Blechingberg J, Lykke-Andersen S, Jensen TH, Jorgensen AL, Nielsen AL: **Regulatory mechanisms for 3'-end alternative splicing and polyadenylation of the Glial Fibrillary Acidic Protein, GFAP, transcript**. *Nucleic Acids Res* 2007, **35**: 7636–7650.

17. Brenner M, Johnson AB, Boespflug-Tanguy O, Rodriguez D, Goldman JE, Messing A: **Mutations in GFAP, encoding glial fibrillary acidic protein, are associated with Alexander disease**. *Nat Genet* 2001, **27**:117–120.

18. * Amano E, Yoshida T, Mizuta I, Oyama J, Sakashita S, Ueyama S, Machida A, Yokota T: **Activation of a cryptic splice site of *GFAP* in a patient with adult-onset alexander disease**. *Neurol Genet* 2021, **7**:e626.
This case report shows that an intronic mutation in the GFAP gene in an Alexander disease patient results in aberrant splicing of GFAP. This study contributes to earlier findings of mutations causing changes in GFAP splicing in Alexander patients (see reference 8).

19. Hardwick SA, Hu W, Joglekar A, Fan L, Collier PG, Foord C, Balacco J, Lanjewar S, Sampson MM, Koopmans F, *et al.*: **Single-nuclei isoform RNA sequencing unlocks barcoded exon connectivity in frozen brain tissue**. *Nat Biotechnol* 2022, **40**: 1082–1092.

20. ** Morabito S, Reese F, Rahimzadeh N, Miyoshi E, Swarup V: **hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data**. *Cell Reports Methods* 2023, **3**:100498.
This study describes a novel bioinformatics pipeline, to perform gene network analyses on scRNAseq and spatial RNAseq data. It is an R package to perform high-dimensional weighted gene co-expression network analysis (hdWGCNA). This pipeline goes beyond analysis of conventional single-cell RNAseq, and is able to include information on isoforms, as it can work with long-read single-cell RNAseq data. Therefore, hdWGCNA will be highly instrumental in determining single-cell expression of GFAP isoforms.

21. ** Joglekar A, Prjibelski A, Mahfouz A, Collier P, Lin S, Schlusche AK, Marrocco J, Williams SR, Haase B, Hayes A, *et al.*: **A spatially resolved brain region- and cell type-specific isoform atlas of the postnatal mouse brain**. *Nat Commun* 2021, **12**:463.
This study describes a first mouse brain atlas of region and cell type-specific expression of gene isoforms. These isoforms are quantified, and the dataset includes Gfap isoforms.

22. Holst CB, Brøchner CB, Vitting-Seerup K, Møllgård K: **Astrogliogenesis in human fetal brain: complex spatiotemporal immunoreactivity patterns of GFAP , S100, AQP 4 and YKL -40**. *J Anat* 2019, **235**:590–615.

23. Arellano JI, Morozov YM, Micali N, Rakic P: **Radial glial cells: new views on old questions**. *Neurochem Res* 2021, **46**: 2512–2524.

24. Redmond CJ, Coulombe PA: **Intermediate filaments as effectors of differentiation**. *Curr Opin Cell Biol* 2021, **68**:155–162.

25. * O'Shea TM, Ao Y, Wang S, Wollenberg AL, Kim JH, Ramos Espinoza RA, Czechanski A, Reinholdt LG, Deming TJ, Sofroniew MV: **Lesion environments direct transplanted neural progenitors towards a wound repair astroglial phenotype in mice**. *Nat Commun* 2022, **13**:5702.
In this study different *Gfap* variants induced in neural progenitors are described based on bands of different molecular weights on a Western blot. The presence of cell-fate determining factors or fetal bovine serum resulted in the expression of different *Gfap* variants. Whether these are caused by (post-)transcriptional mechanisms or posttranslational modifications was not determined.

26. Gómez-Gálvez Y, Fuller HR, Synowsky S, Shirran SL, Gates MA: **Quantitative proteomic profiling of the rat substantia nigra places glial fibrillary acidic protein at the hub of proteins dysregulated during aging: implications for idiopathic Parkinson's disease**. *J Neurosci Res* 2020:98.

27. ** Lin N, Yang A, Chang C, Perng M: **Elevated GFAP isoform expression promotes protein aggregation and compromises astrocyte function**. *Faseb J* 2021:35.
This study shows the expression of GFAPδ and GFAPκ in mouse astrocytes from mice with several copies of the GFAP transgene. The higher expression of the isoforms is linked to GFAP aggregation and impaired glutamate transport, thereby providing evidence for a functional change induced by GFAP isoforms.

28. ** Uceda-Castro R, van Asperen JV, Vennin C, Sluijs JA, van Bodegraven EJ, Margarido AS, Robe PAJ, van Rheenen J, Hol EM: **GFAP splice variants fine-tune glioma cell invasion and tumour dynamics by modulating migration persistence**. *Sci Rep* 2022, **12**:424.
Here it is shown that human glioma cells expressing either mainly GFAPα or mainly GFAPδ (engineered by CRISPR editing), perform differently when studying cell invasion and migration. This study shows the potential of differential splicing of GFAP to affect cell behaviour.

29. Van Den Berge SA, Middeldorp J, Zhang CE, Curtis MA, Leonard BW, Mastroeni D, Voorn P, Van De Berg WDJ, Huitinga I, Hol EM: **Longterm quiescent cells in the aged human subventricular neurogenic system specifically express GFAP-δ: GFAP-δ in aged human SVZ stem cells**. *Aging Cell* 2010, **9**: 313–326.

30. Zhou L, Li P, Chen N, Dai L, Gao K, Liu Y, Shen L, Wang J, Jiang Y, Wu Y: **Modeling vanishing white matter disease with patient-derived induced pluripotent stem cells reveals astrocytic dysfunction**. *CNS Neurosci Ther* 2019, **25**: 759–771.

31. Bodegraven EJ, Asperen JV, Robe PAJ, Hol EM: **Importance of GFAP isoform-specific analyses in astrocytoma**. *Glia* 2019, **67**:1417–1433.

32. ** Zeneyedpour L, Stingl C, Kros JM, Sillevis Smitt PAE, Luider TM: **Novel antibody–peptide binding assay indicates presence of immunoglobulins against EGFR phospho-site S1166 in high-grade glioma**. *IJMS* 2022, **23**:5061.
In this study, the sequence coverage of GFAP protein is identified in a proteome analysis. Surprisingly, the 3 normal brain tissue samples lacked sequence coverage in the first 35 and last 30 amino acids of the head and tail domain, respectively. This in sharp contrast with 3 glioblastoma tissue samples, indicating a difference in GFAP isoforms expression between normal brain tissue and glioblastoma.

33. Stassen OMJA, van Bodegraven EJ, Giuliani F, Moeton M, Kanski R, Sluijs JA, van Strien ME, Kamphuis W, Robe PAJ, Hol EM: **GFAPδ/GFAPα ratio directs astrocytoma gene expression towards a more malignant profile**. *Oncotarget* 2017, **8**:88104–88121.

34. Bodegraven EJ, Asperen JV, Sluijs JA, Deursen CBJ, Strien ME, Stassen OMJA, Robe PAJ, Hol EM: **GFAP alternative splicing regulates glioma cell–ECM interaction in a DUSP4-dependent manner**. *Faseb J* 2019, **33**:12941–12959.

35. Moeton M, Kanski R, Stassen OMJA, Sluijs JA, Geerts D, Tijn P, Wiche G, Strien ME, Hol EM: **Silencing GFAP isoforms in astrocytoma cells disturbs laminin-dependent motility and cell adhesion**. *Faseb J* 2014, **28**:2942–2954.

36. Dooves S, Bugiani M, Postma NL, Polder E, Land N, Horan ST, van Deijk A-LF, van de Kreeke A, Jacobs G, Vuong C, *et al.*: **Astrocytes are central in the pathomechanisms of vanishing white matter**. *J Clin Invest* 2016, **126**:1512–1524.

37. Perng M-D, Wen S-F, Gibbon T, Middeldorp J, Sluijs J, Hol EM, Quinlan RA: **Glial fibrillary acidic protein filaments can tolerate the incorporation of assembly-compromised GFAP-delta, but with consequences for filament organization and alphaB-crystallin association**. *Mol Biol Cell* 2008, **19**: 4521–4533.

38. Moeton M, Stassen OMJA, Sluijs JA, van der Meer VWN, Kluivers LJ, van Hoorn H, Schmidt T, Reits EAJ, van Strien ME, Hol EM: **GFAP isoforms control intermediate filament network dynamics, cell morphology, and focal adhesions**. *Cell Mol Life Sci* 2016, **73**:4101–4120.

39. Roelofs RF, Fischer DF, Houtman SH, Sluijs JA, Van Haren W, Van Leeuwen FW, Hol EM: **Adult human subventricular, subgranular, and subpial zones contain astrocytes with a specialized intermediate filament cytoskeleton**. *Glia* 2005, **52**: 289–300.

40. Nielsen AL, Jørgensen AL: **Self-assembly of the cytoskeletal glial fibrillary acidic protein is inhibited by an isoform-specific C terminus**. *J Biol Chem* 2004, **279**:41537–41545.

41. Block J, Schroeder V, Pawelzyk P, Willenbacher N, Köster S: **Physical properties of cytoplasmic intermediate filaments**. *Biochim Biophys Acta Mol Cell Res* 2015, **1853**:3053–3064.

42. Van Bodegraven EJ, Etienne-Manneville S: **Intermediate filaments from tissue integrity to single molecule mechanics**. *Cells* 2021, **10**:1905.

43. Lorenz C, Forsting J, Style RW, Klumpp S, Köster S: **Keratin filament mechanics and energy dissipation are determined by metal-like plasticity**. *Matter* 2023, **6**:2019–2033.

44. Boyd SE, Nair B, Ng SW, Keith JM, Orian JM: **Computational characterization of 3′ splice variants in the GFAP isoform family**. *PLoS One* 2012, **7**, e33565.

45. Kraxner J, Lorenz C, Menzel J, Parfentev I, Silbern I, Denz M, Urlaub H, Schwappach B, Köster S: **Post-translational modifications soften vimentin intermediate filaments**. *Nanoscale* 2021, **13**:380–387.

46. ** Eibauer M, Weber MS, Kronenberg-Tenga R, Beales CT, Boujemaa-Paterski R, Turgay Y, Sivagurunathan S, Kraxner J, Köster S, Goldman RD, *et al.*: **Vimentin filaments integrate low complexity domains in a highly complex helical structure**. *bioRxiv* 2023, https://doi.org/10.1101/2023.05.22.541714.

In this landmark study, Eibauer et al. showed, using cryo-focused ion beam milling combined with cryo-electron tomography, an unprecedented view of a 3D vimentin filament structure. This is highly informative for GFAP, as both proteins can co-assemble and have a high similarity. In this study, it is shown that the head domains of vimentin form a luminal fiber and that the tail domains form contact sites between protofibrils.

47. Mazaré N, Oudart M, Moulard J, Cheung G, Tortuyaux R, Mailly P, Mazaud D, Bemelmans A-P, Boulay A-C, Blugeon C, *et al.*: **Local translation in perisynaptic astrocytic processes is specific and changes after fear conditioning**. *Cell Rep* 2020, **32**:108076.

48. Boulay A-C, Saubaméa B, Adam N, Chasseigneaux S, Mazaré N, Gilbert A, Bahin M, Bastianelli L, Blugeon C, Perrin S, *et al.*: **Translation in astrocyte distal processes sets molecular heterogeneity at the gliovascular interface**. *Cell Discov* 2017, **3**:17005.

49. * Oudart M, Tortuyaux R, Mailly P, Mazaré N, Boulay A-C, Cohen-Salmon M: **AstroDot: a new method for studying the spatial distribution of mRNA in astrocytes**. *J Cell Sci* 2020, https://doi.org/10.1242/jcs.239756.

This paper describes an Image J plugin AstroDot and an R package AstroStat to determine the spatial distribution of mRNAs in astrocytes. As a proof of concept, the spatial distribution of GFAPα and GFAPδ is determined, corroborating earlier data of Thomsen et al., 2013 (PLoS ONE 8, e72110), that GFAPα is more expressed in the processes and GFAPδ more in the soma.

50. Sakers K, Lake AM, Khazanchi R, Ouwenga R, Vasek MJ, Dani A, Dougherty JD: **Astrocytes locally translate transcripts in their peripheral processes**. *Proc Natl Acad Sci USA* 2017, **114**.

51. ** Glinos DA, Garborcauskas G, Hoffman P, Ehsan N, Jiang L, Gokden A, Dai X, Aguet F, Brown KL, Garimella K, *et al.*: **Transcriptome variation in human tissues revealed by long-read sequencing**. *Nature* 2022, **608**:353–359.

This study presents a large human long-read RNAseq dataset with over 70,000 novel transcripts for annotated genes, including splice variants, alternative 3′ polyadenylation, and alternative start- or end-sites. The technology of long-read RNAseq is essential for mapping the full extent of GFAP isoforms.

52. Liao W-W, Asri M, Ebler J, Doerr D, Haukness M, Hickey G, Lu S, Lucas JK, Monlong J, Abel HJ, *et al.*: **A draft human pangenome reference**. *Nature* 2023, **617**:312–324.

53. Kleino I, Frolovaitė P, Suomi T, Elo LL: **Computational solutions for spatial transcriptomics**. *Comput Struct Biotechnol J* 2022, **20**:4870–4884.

54. Hagemann TL, Powers B, Mazur C, Kim A, Wheeler S, Hung G, Swayze E, Messing A: **Antisense suppression of glial fibrillary acidic protein as a treatment for Alexander disease: GFAP ASO Therapy in AxD**. *Ann Neurol* 2018, **83**:27–39.