UCLouvain

AGRO LOUVAIN

Faculté des bioingénieurs

Structural study of the endophilin A isoforms and their role in clathrin-independent endocytosis

Auteur : Ian Coupez

Promoteur : Pr. Pierre Morsomme (UCL/IBST)

Lecteurs : Pr. Bernard Knoops (UCL/IBST)

Pr. Henri-François Renard (UNamur/NARILIS)

Année académique 2020-2021

Mémoire de fin d'études présenté en vue de l'obtention du diplôme de Bioingénieur en chimie et bio-industries

Remerciements

Avant toute chose, je tiens à remercier toutes les personnes qui, de prêt ou de loin, ont contribué à la réalisation de ce mémoire.

A commencer par mon promoteur, Pierre Morsomme, qui m'a donné l'opportunité d'intégrer une équipe de recherche motivée et accueillante. Je lui suis extrêmement reconnaissant pour son encadrement et son aide durant toute cette année.

Je voudrais remercier tout particulièrement François Tyckaert, mon encadrant, pour son aide sans faille aussi bien durant la réalisation des manips que lors de la rédaction du mémoire. C'est certainement la personne dont j'ai le plus appris cette année et je lui suis infiniment reconnaissant de m'avoir consacré autant de temps. Son sens du détail et son expertise m'ont poussé à acquérir une rigueur dans mon travail qui me manquait parfois. D'une simple relation de travail est né une relation d'amitié qui m'a beaucoup apporté durant cette année.

Je remercie également toute l'équipe qui m'a accueilli cette année et qui m'a permis d'apprendre énormément dans une ambiance chaleureuse et conviviale. Je pense avoir eu l'occasion de discuter au moins une fois avec chaque personne et j'ai découvert des gens géniaux et passionnés par ce qu'ils font, toujours prêt à aider en cas de besoin.

Sans oublier évidement les autres mémorants, Andréa, Antoine, Apolline, Bastien, Charlotte, Félicie. Merci pour tous ces moments, que ça soit pendant les galères en labo ou les soirées entre nous, c'était toujours un plaisir d'être entouré de gens comme vous.

Enfin, je souhaiterais remercier tous mes proches, famille ou amis, qui m'ont soutenu et encouragé durant cette année.

Ce mémoire fut pour moi une expérience extrêmement enrichissante dont je ne tirerai que du positif, aussi bien d'un point de vue professionnel que personnel.

Abstract

Among the wide variety of membrane bending proteins, the superfamily of BAR domain proteins has been one of the most extensively studied in the recent years. BAR stands for Bin/Amphiphysin/Rvs167 which are the first three described proteins of this superfamily. These proteins have been shown to function in many cellular processes requiring local membrane deformation. One of particular interest is endocytosis, the process by which cells internalize their nutrients but also recycle their plasma membrane components. Many BAR domain proteins were shown to act in different stages of endocytic processes. In particular, members of the endophilin A BAR subfamily held our attention for their strong involvement in these mechanisms.

The endophilin A subfamily is composed of three isoforms, endophilin A1, A2 and A3. They all share a common structure composed of four distinct domains: a N-terminal amphipathic helix, a BAR domain, a SH3 domain and a linker region connecting the last two. This global structure makes endophilins A remarkable proteins to build endocytic vesicles. Indeed, endophilin A were shown to be involved in well-described endocytic mechanisms such as clathrin-mediated endocytosis or fast endophilin-mediated endocytosis. In addition, recent research in my lab highlighted a new endocytic modality in which the A3 isoform plays a central role, but neither the A1 or A2 isoforms are involved. The first identified cargo of this new mechanism is the activated leukocyte cell adhesion molecule (ALCAM), also called CD166. The involvement of endophilin A3 in this mechanism but not of A1 and A2 raised the question of the origin of this isoform specificity.

To answer this question, we developed a strategy composed of four distinct steps. We started with the obtention of a library of plasmids encoding for mutant proteins fused with GFP and constructed using a domain swapping approach with endophilin A2 and A3. We then used these plasmids to assess the colocalization between CD166 and the mutants overexpressed in HeLa cells. We next asked whether the overexpression of these mutants could rescue CD166 endocytosis in HeLa cells previously knocked-down for endophilin A3. Finally, we performed pull-down experiment with GST fused with CD166 cytosolic tail as bait protein to highlight potential interaction between CD166 and the mutants.

The results from the three experiments all confirmed the specificity of endophilin A3 for CD166 endocytosis unlike endophilin A2. The linker-SH3 region appeared crucial for cargo recognition and the linker region's variability between endophilin isoforms appeared to be not sufficient to explain the specificity of endophilin A3 in CD166 endocytic mechanism. Unfortunately, further investigation are still required to conclude exactly which domain(s) of endophilin A3 provide(s) its specificity for this mechanism.

Table of content

Abbreviations

ALCAM	Activated Leukocyte Cell Adhesion Molecule
AP2	Adaptor Protein 2
Arp 2/3	Actin related protein 2/3
BAR	Bin1/Amphiphysin/Rvs167
BSA	Bovine Serum Albumin
CALM	Clathrin Assembly Lymphoid Myeloid leukemia
CD166-CT	CD166 Cytosolic Tail
CHC	Clathrin Heavy Chain
CIE	Clathrin-Independent Endocytosis
CLC	Clathrin Light Chain
CLIC	Clathrin-Independent Carriers
CME	Clathrin-Mediated Endocytosis
CMV	Cytomegalovirus
CRD	Carbohydrate Recognition Domain
CRM1	Chromosomal Maintenance 1
DAPI	4’,6-diamidino-2-phenylindole
F-BAR	Fes/CIP4 Homology-BAR
FBS	Fetal Bovine Serum
FDS	Friction Driven Scission
FEME	Fast Endophilin-Mediated Endocytosis
Gal3	Galectin-3
Gal8	Galectin-8
GAP	GTPase Activating Protein
GEEC	Glycosylphosphatidylinositol-anchored protein Enriched Compartment
GEP	Guanine nucleotide Exchange Factor
GFP	Green Fluorescent Protein

GPCR	G-Protein Coupled Receptor
GSL	Glycosphingolipid
GST	Glutathione S-Transferase
GTP	Guanosine Triphosphate
HeLa	Henrietta Lacks
I-BAR	IMD/Inverted-BAR
NPF	Nucleation Promoting Factor
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PFA	Paraformaldehyde
PH	Pleckstrin Homology
PIP_1	Phosphatidylinositol monophosphate
PIP_2	Phosphatidylinositol biphosphate
PIP_3	Phosphatidylinositol triphosphate
PRD	Proline Rich Domain
PX	Phox Homology
ROI	Region Of Interest
SH3	Src Homology 3
SJ1-145	145 kDa isoform of Synaptojanin
SNX	Sorting Nexin
SVE	Synaptic Vesicle Endocytosis
WASP	Wiskott-Aldrich Syndrome Protein

Introduction

In the introduction of this Master thesis, three topics will be developed. The first section will start with a general overview on proteins from the BAR domain superfamily, which act as sensors and inducers of membrane curvature in cells. In the second part, we will focus on the endophilin A BAR subfamily, focusing on the structure, expression and functions of its members. Finally, the involvement as well as the specificity of these proteins in distinct endocytic mechanisms will be discussed.

I. BAR domain proteins

I.1 General characteristics

At the interface between actin cytoskeleton and cell membrane, a wide variety of proteins sits, playing diverse roles such as initiation of actin polymerization, modulation of membrane curvature, or cell signaling. Lately, proteins from the BAR domain superfamily have been shown to act as key players in these functions. BAR is the acronym for Bin1/Amphiphysin/Rvs167, the first proteins of this superfamily to be identified by sequence homology back in the 90's (Sakamuro, 1996). Since then, 64 proteins have been assigned to this superfamily in *homo sapiens*.

The membership in this superfamily is strictly based on the presence of a BAR domain, which consists into a sequence of ~200-280 amino acids that forms three kinked antiparallel α-helices. Those three helices will homodimerize in an antiparallel fashion, adopting a characteristic crescent shape dimer. The dimer exposes positively charged residues on the concave face that will interact with negatively charged phospholipids such as phosphatidylserine and phosphatidylinositol enriched in the inner leaflet of the cell membrane (**Figure 1A**). BAR domain proteins will therefore act as membrane deformation sensors or inducers through the intrinsic curvature of their dimer (Carman et Dominguez. 2018).

Actually, proteins from the BAR domain superfamily are able to modulate membrane shape and curvature through three distinct mechanisms. The first one simply involves the rigidity and intrinsic curvature of the protein dimer that is imposed to the membrane through electrostatic interactions (**Figure 1A**). Some BAR domain proteins harbor amphipathic helices on their concave face, capable of inserting into the membrane, further amplifying the curvature (**Figure 1B**). Finally, oligomerization of BAR domain proteins through various type of interactions (see section I.3.2 of introduction) leads to the formation of a tubular protein scaffold deforming the membrane at a larger scale (**Figure 1C**) (Mim et Unger, 2015) (Suetsugu, 2016).

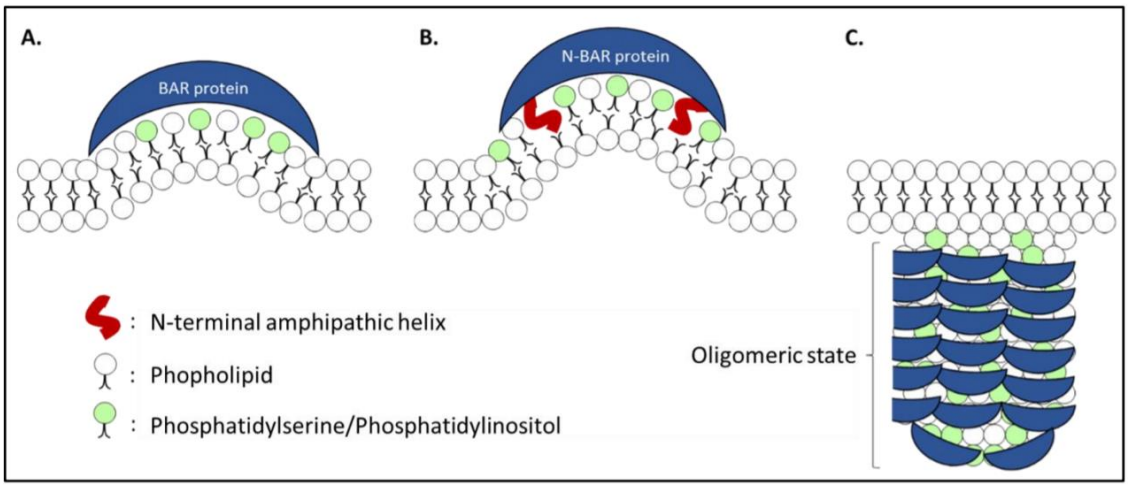


Figure 1. Mechanisms of membrane curvature by BAR domain proteins. (A) The BAR domain dimer imposes its intrinsic curvature to the plasma membrane through electrostatic interactions. **(B)** The insertion of amphipathic helices stabilizes the interaction with the membrane and amplify its curvature. **(C)** Through oligomerization, BAR domain dimers assemble helical coat or lattice at the cell membrane, leading to the formation of tubules.

I.2 Cellular functions

Unsurprisingly, BAR domain proteins were shown to play a crucial role in a variety of cellular processes requiring local membrane deformation. These include endocytosis, autophagy, organelle trafficking, cell motility or T-tubule biogenesis in muscles (reviewed in Frost et al., 2009). While every member of this family share the characteristic BAR domain, this great diversity in functions implies the existence of differentiating factors within the family. Indeed, variability in the BAR domain folding itself or the presence of additional domains will provide each BAR domain protein its singularity and specificity for the mechanism it is involved in.

I.3 Diversity in the BAR domain superfamily

I.3.1 BAR domain

Based on their sequence homology and structural differences, BAR domain proteins have been classified into three distinct subfamilies : (1) the classical, shorter (~167 Å) and more curved crescent-shape BAR (Peter et al., 2004), (2) the longer (~225 Å) and less curved Fes/CIP4 homology-BAR (F-BAR) (Frost et al., 2007), (3) and the Inverted-BAR (I-BAR) in which positive amino acids are exposed on the convex face of the dimer, therefore sensing/generating negative curvature (Lee et al., 2007) (**Figure 2**). Even within these subfamilies, substantial variability in the monomer intrinsic curvature as well as in the dimerization angle leads to very specific membrane curvature recognition and induction.

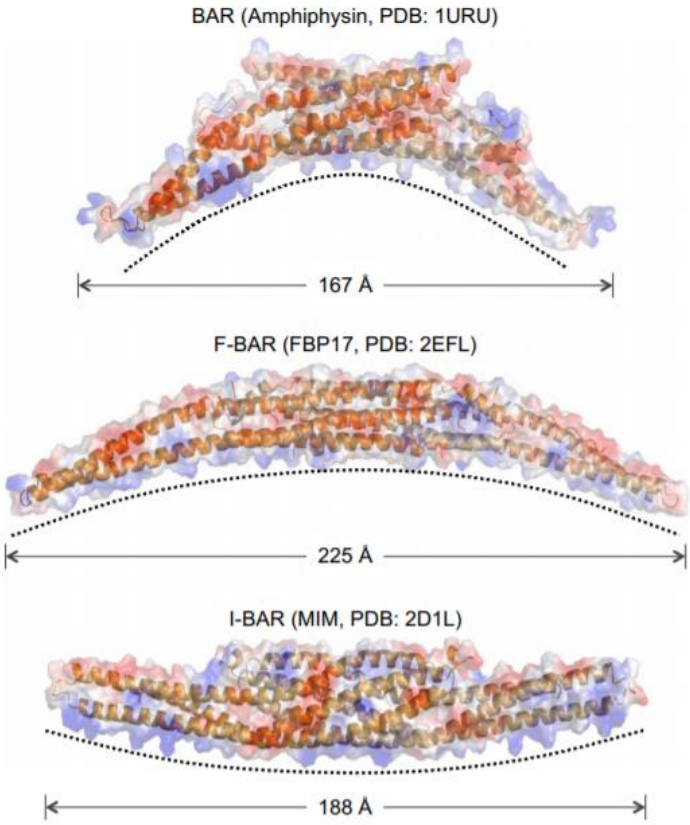


Figure 2. Structure of representative members of the three BAR domain protein subfamilies (Carman et Dominguez, 2018). The classical and shorter BAR is represented by amphiphysin. The longer and less curved F-BAR is represented by Formin Binding Protein 17 (FBP17). The I-BAR exposing positive amino acids at the convex face of the dimer is represented by the protein Missing in Metastasis (MIM).

Notably, some classical BAR domain proteins such as endophilins or amphiphysin possess additional amphipathic helical extensions capable of inserting into the membrane, further modifying their affinity for curved membrane and their ability to deform it. These additional helices can be amino-terminal next to the BAR domain (so-called N-BAR domains), or even embedded within the BAR domain itself, as it is the case for endophilin (see section II.1.1 of introduction) (Mim et al., 2012).

I.3.2 Oligomerization and type of coat/lattice formed

As mention above, BAR domain proteins can oligomerize and assemble tubular coats on the membrane leading to its tubulation. Surprisingly, 3D reconstruction of these tubular scaffolds formed by F-BAR and N-BAR domain proteins appeared to assemble through different type of interactions. CIP4 and FBP17, both members of the F-BAR subfamily, assemble helical coats on membrane through lateral and tip-tip interactions. On the other hand, N-BAR proteins such as endophilins or amphiphysin form lattices held together through antiparallel interactions between their amphipathic helices inserted into the membrane. As a direct consequence, N-BAR scaffolds expose larger areas of the membrane surface, which might be critical for the access of proteins such as the GTPase dynamin or the phosphatase Synaptojanin (**Figure 3**) (more details in section III.2 of introduction) (Mim et al., 2012) (Suetsugu. 2016).

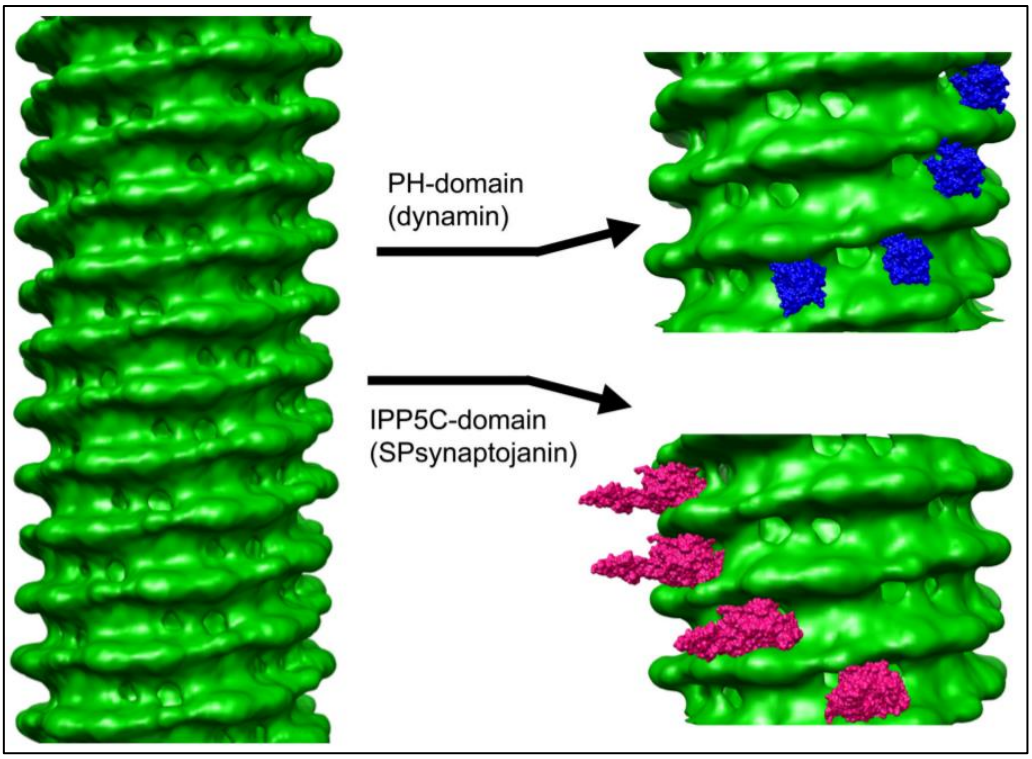


Figure 3. N-BAR scaffolds compartmentalize the membrane surface for interactions with downstream effectors (Mim et al., 2012). N-BAR scaffolds expose larger membrane surface areas compare to F-BAR scaffolds. These areas match the size of membrane binding domain from downstream effector such as the PH-domain of dynamin **(Blue)** or the inositol polyphosphate phosphatase catalytic domain (IPP5C) of Synaptojanin **(red)**.

I.3.3 Auxiliary domains

A last factor providing variability within the BAR domain superfamily is the presence of functional auxiliary domains located at the N or C-terminal part of the protein.

SH3 domain

The Src Homology 3 domain (SH3) is the most frequent domain in BAR proteins present in almost 50% of them (Carman et Dominguez. 2018). This domain is capable of binding to Proline Rich Domains (PRD) that are present in a large subset of proteins. Among the classical SH3-binding proteins, the GTPase dynamin is worth to mention for its essential role in the scission of endocytic vesicle at cell surface (see section III.2 of introduction) as well as nucleation promoting factors (NPFs) of the Arp2/3 complex that control the branching of actin filaments (Suetsugu et Gautreau, 2012).

GAP and GEP

BAR domain proteins can also regulate actin cytoskeleton assembly through the presence of GTPase activating proteins (GAPs) and guanine nucleotide exchange factors (GEFs) domains. These domains act as regulator of Rho GTPase activity. Rho GTPases are master regulators of actin cytoskeleton dynamics, acting as molecular on/off switches of GTP hydrolysis. Approximately 35% of BAR domain proteins contain at least one GAP or GEF domain, making them an essential link between membrane trafficking and actin cytoskeleton dynamics (Aspenström. 2019).

Other domains

Some auxiliary domains can also confer lipid specificity such as phox homology (PX) (Yarar et al., 2008) or pleckstrin homology (PH) domains (Wang et Shaw, 1995). For example, members of the sorting nexin (SNX) subfamily possess a PX domain that binds specifically to phosphatidylinositol 4,5-bisphosphate ($PI(4,5)P_2$), allowing their targeting to sorting endosomes or forming endocytic pits (van Weering et al., 2010).

II. Endophilin-A subfamily

From all the BAR domain proteins, the N-BAR endophilin-A subfamily has been one of the most extensively studied in the recent years, especially for its role in endocytic processes, which is of particular interest in our case. This family is composed of three isoforms, endophilin A1, A2 and A3 of about 350 amino acids in length (~40 kDa) and differentially expressed in tissues (Kjaerulff et al., 2011). Following their discovery in 1997, these proteins have since been considered as redundant, owing to their high sequence similarity (~80%). However, recent studies started to notice isoform-specific functions in some mechanisms, calling for a better individual characterization of the three endophilin-A isoforms.

Of note, a second endophilin subfamily also exists, the endophilin B, composed of two isoforms, B1 and B2. Their structure is relatively similar to endophilins A but they act in different mechanisms such as apoptosis or autophagy (Cuddeback et al., 2001) (Pierrat et al., 2001). This subfamily will not be treated in this document, meaning that every time endophilins will be mentioned, it will automatically refers to the endophilin A subfamily.

II.1 Structural analysis : domains, sequences and functions

All three endophilin isoforms share a common structure composed of four different domains: the N-terminal amphipathic helix (H0), the BAR domain that contains an additional amphipathic insert (H1I), a SH3 domain and a linker region that connects the last two domains (**Figure 4A**).

II.1.1 N-terminal amphipathic helix H0 and the additional insert H1I

- *Structure*

Endophilins are member of the N-BAR subfamily, meaning that they harbor an amino-terminal amphipathic helix (residues 2-23 : H0 on **Figure 4**) capable of inserting into the plasma membrane. Notably, on the first α-helix of their BAR domain, they possess another insert (residues 60-88 : H1I on **Figure 4**) that protrudes from the concave face of the dimer, forming an additional amphipathic helix. An endophilin A dimer therefore possesses four of those amphipathic helices penetrating in the inner leaflet of the membrane (Kjaerulff et al., 2011).

The N-terminal amphipathic helix is an asymmetrically solvated α-helix in which one face is exposed to the membrane (green region on **Figure 5**) and the other to the solvent (blue region on **Figure 5**). This conformation makes it disordered in solution and ordered upon membrane binding (Gallop et al., 2006). On the other hand, the additional H1I helix possesses a hydrophobic side that is flanked on both sides by positively charged residues. This seems to indicate a penetration of the membrane by the hydrophobic side coupled with electrostatic interactions with the charged lipid headgroups (Gallop et al., 2006).

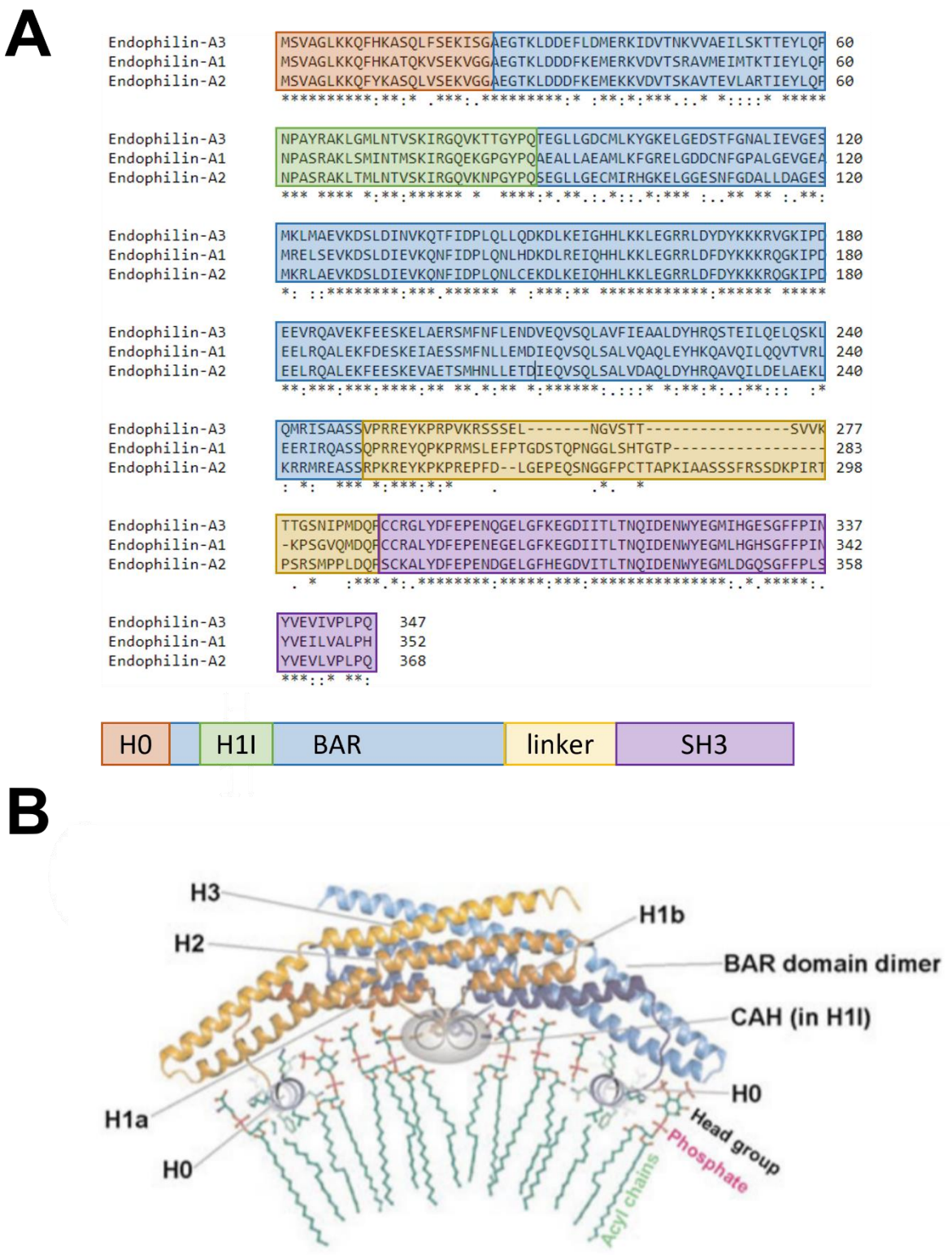


Figure 4. Sequences and structure of the three endophilin A isoforms. (A) Multiple sequences alignment obtained with Clustal Omega. **Orange:** N-terminal amphipathic helix. **Blue:** BAR domain. **Green:** additional amphipathic helix inserted into the first α-helix of the BAR domain. **Yellow:** linker region between BAR and SH3 domains **Purple:** SH3 domain. **(B)** Binding of the endophilin N-BAR domain to the cytoplasmic leaflet of the plasma membrane. **H0:** N-terminal amphipathic helix. **H1a, H1b, H2 and H3:** The three α-helices of the BAR domain. **CAH(inH1I):** additional amphipathic insert located on the first α-helix of the BAR domain.

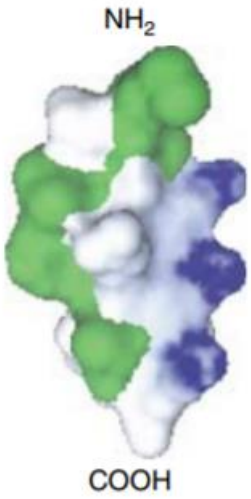


Figure 5. 3D structure of residues 1-16 of the amphipathic helix H0 (Gallop et al., 2006). Hydrophobic residues are colored in green and cytosol-exposed residues in blue.

- *Sequence's comparison*

When looking at the sequence's comparison (**Table 1**, **Figure 4A**), it seems that both amphipathic regions are very well conserved. Therefore, it seems unlikely that endophilin isoforms specificity comes from their amphipathic helices.

Table 1. Sequence identity and similarity between paired endophilin N-terminal domains H0 and additional amphipathic helices H1I amino acid sequences. Obtained using the EMBOSS Needle tool with the EBLOSUM62 Matrix (Madeira et al., 2019).

N-terminal amphipathic helix H0			
	A1/A2	A1/A3	A2/A3
Identity [%]	85.7	76.2	81.0
Similarity [%]	95.2	85.7	90.5
Additional amphipathic helix H1I			
	A1/A2	A1/A3	A2/A3
Identity [%]	82.8	75.9	86.2
Similarity [%]	93.1	82.8	86.2

- *Functions*

According to mutation and deletion studies, the H0-helix and additional insert H1I were shown to participate, along with the BAR domain, to tubule formation at the cell surface. These studies also showed that the additional insert is crucial for stabilizing the dimer. Additional interactions of those two domains with membranous proteins is not excluded but need further investigation (Gallop et al., 2006).

More advanced structural studies further demonstrated that the N-BAR domain was mostly responsible for the formation of the scaffold around the tubule, a function that is conserved among the three isoforms. The antiparallel interaction between two H0-helices from N-BAR domains in adjacent row of the lattice allows the stabilization of

membrane curvature. A very surprising observation was that interactions between H0-helices seemed to be non-specific, meaning that they could occur between two different N-BAR proteins such as endophilin and amphiphysin. This observation changed the idea that N-BAR lattices around membrane tubules were very ordered arrays, but are in fact more variable and flexible scaffolds (Mim et al., 2012).

Finally, the insertion of amphipathic helices into the membrane bilayer was shown to promote membrane vesiculation *in vitro* (Boucrot et al., 2012). An explanation for this phenomenon is that shallow hydrophobic insertions such as with epsin or endophilin amphipathic helices span mainly the polar head region of membrane monolayers. These insertions would expand the bilayer surface compared to bilayer midplane, resulting in a destabilization of the membrane stability that can possibly lead to vesicle scission (Boucrot et al., 2012). Another suggested mechanism relies on pulling forces provided by dynamic actin polymerization and microtubule-based molecular motors which, together with the assembly of a rigid protein coat around membrane tubules, would enable scission through a generic friction based mechanism, so called friction-driven scission (FDS) (more details in section III.2.2 of the Introduction) (Simunovic et al., 2017).

- *Summary*

In short, the N-terminal amphipathic helix of endophilins plays a crucial role in the formation of the protein lattice by interacting with other amphipathic helices in a non-specific fashion, generating a flexible scaffold. The additional insert H1I might also contribute to scaffold assembly, as well as it contributes to the stabilization of the dimer. Both amphipathic helices are capable of destabilizing the membrane, therefore promoting vesicle scission. In conclusion, the amphipathic helices seems to be crucial components of membrane tubulation and vesiculation, which are both key events during endocytosis.

II.1.2 BAR domain

- *Structure*

The BAR domain of endophilins is classically composed of three kinked antiparallel helices forming a crescent shape dimer that generates membrane positive curvature. The crystal structure of mouse endophilin A1 BAR domain (residues 1-256) was solved from a selenomethionine-containing crystal using single wavelength anomalous diffraction method with a 2.3 Å resolution. The distance between the two distal ends of the dimer is ~114 Å and is predicted to fit a curved membrane with a ~30 nm diameter (Wesseinhorn, 2005). *In vitro* studies performed on liposome showed that rat endophilin-A1 deforms lipid bilayer into tubules with outer diameter ranging between 20-100 nm (Farsad et al., 2001).

Helix 1 of the BAR domain is divided into two fragments (residues 30-59 : H1a and 89-104 residues : H1b on **Figure 4B**), interrupted by the amphipathic insert. The helix 1b extends along the same axis as helix 1a. Helix 1b is connected by a short loop

(residues 105-109) to the slightly bent helix 2 (residues 110-173: H2 on **Figure 4B**). The helix 2 is followed by another loop (residues 174-181) that reverses the chain, leading to helix 3 (residues 182-247 : H3 on **Figure 4B**) which also bends. The carboxy-terminal end of H1 and H3 as well as the amino-terminal of H2 contribute to dimerization through interactions between hydrophobic amino acid (Weissenhorn, 2005).

- *Sequence comparison*

The BAR domain is the region presenting the greatest variation between the three endophilin isoforms (aside from the linker region), even if the similarity remains relatively high (ranging from 84 to 90%) (**Tableau 2**, **Figure 4A**). While the crystal structure of the three isoforms is believed to be very similar, this slight variability could be at the origin of minor structural differences between the isoforms therefore explaining partially the isoform-specific functions observed within this protein subfamily (see section III of introduction).

Table 2. Sequence identity and similarity between paired endophilin BAR domains amino acid sequences. Obtained using the EMBOSS Needle tool with the EBLOSUM62 Matrix (Madeira et al., 2019).

BAR domain			
	A1/A2	A1/A3	A2/A3
Identity [%]	76.3	67.0	69.3
Similarity [%]	89.7	84.0	85.6

Punctuated mutations within the BAR domains could for example lead to a slightly different curvature degree and dimerization angle. The increased variability observed at the end of the third helix H3 of the BAR domain (**Figure 4A**) could possibly explain the specificity for dimerization but also modify the orientation of the linker, therefore modifying the position of the SH3 domain. Finally, single mutations can also add or remove additional interaction regions such as Ca^{2+} binding sites, modifying the way the different isoform functions are regulated.

All these possibilities remain quite hypothetic and the precise understanding of the endophilin A isoforms specificity obviously calls for a better individual characterization of their structure.

- *Functions*

As explained in the first section of this document, the main function of the BAR domain is to induce or sense membrane curvature through electrostatic interactions (Carman et Dominguez. 2018). Furthermore, endophilins have the ability to form a lattice around tubular membrane invaginations, therefore explaining their involvement in endocytic mechanisms. While the N-BAR domain stabilizes the lattice through antiparallel interactions between successive rows, the BAR domain is the functional unit stabilizing membrane curvature and therefore dictating the tubule diameter (Mim et al., 2012).

Interestingly, the crystal structure of the human endophilin A1 BAR domain revealed several Cd^{2+} binding sites, some of them localized in the BAR domain α-helices, and some in the variable linker region. Those sites could correspond to Ca^{2+} binding sites according to their similar ionic radii (0.97 fo Cd and 0.99 for Ca) (Weissenhorn, 2005). This observation was consistent with a previously observed Ca^{2+}-dependent interaction between rat endophilin A2 and voltage-gated Ca^{2+} channels. The hypothesis states that a high concentration of Ca^{2+} in brain cells (>1 µM) would induce the binding of Ca^{2+} to endophilin Glu267 that is believed to promote autoinhibition by the SH3 domain (see section II.1.3 of introduction) and therefore inhibiting endophilin-Ca^{2+} channel complex formation (Chen et al., 2003). This mechanism is believed to be a major regulator of synaptic vesicle endocytosis (SVE), a mechanism in which endophilin (mostly A1) was shown to be determining (see section II.2 of introduction). However, these studies worked with endophilin isoforms coming from different species. Since this kind of mechanism can be amino acid specific, a single mutation could therefore completely modify the binding of Ca^{2+} to the protein. This once again call for a better individual characterization of each human endophilin isoform to understand more precisely which isoform is subjected to this regulatory mechanism.

- *Summary*

The BAR domain is the region that provides endophilin their main function, i.e. membrane curvature recognition and induction. Its combination with the N-terminal amphipathic helix leads to the formation of a protein lattice on the cell membrane, stimulating its tubulation. The presence of the SH3 domain also provides an interaction dimension, allowing endophilins to act into complex and coordinated processes (more details in section II.1.3 of introduction). Finally, the presence of Ca^{2+} binding sites along the BAR domain might suggest a regulatory mechanism of autoinhibition by the SH3 domain. However, a more precise characterization of each isoform individually may bring more light on the specificity of these mechanisms.

II.1.3 SH3 domain

- *Structure*

All three endophilin isoforms possess a carboxy-terminal SH3 domain adopting a β-barrel core with a hydrophobic groove that typically recognizes proline rich domain (PRD) (SH3 domain of rat endophilin A2 on **Figure 6**) (Loll et al., 2008).

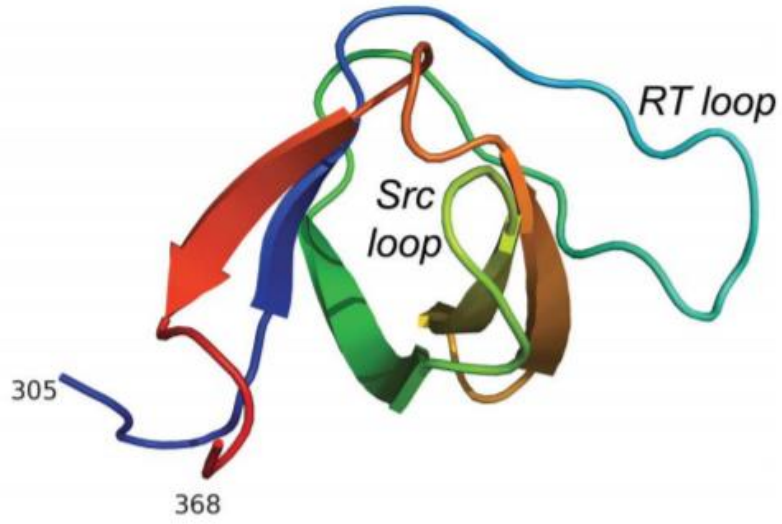


Figure 6. Crystal structure of rat endophilin A2 SH3 domain (Loll et al., 2008). The color scheme runs from blue at the N-terminus (residue 305 in the endophilin sequence) to red at the C-terminus (residue 368) of the domain. In this view, the RT loop can be seen at the far right side of the molecule (cyan) and the Src loop extends from the plane of the page directly toward the viewer (yellow–green). The ligand-binding groove lies between these two loops.

- *Sequence comparison*

Although the SH3 domain seems to be the most conserved region among the three endophilin isoforms (**Table 3**, **Figure 4A**), it does not always show the same affinity for different PRD-containing proteins (Tian et al., 2012) (Boucrot et al., 2015), making it an interesting candidate for isoform specificity. This observation suggests that the amino acids located around or in the PRD influence the affinity for SH3 domain and vice versa.

Table 3. Sequence identity and similarity between paired endophilin SH3 domains amino acid sequences. Obtained using the EMBOSS Needle tool with the EBLOSUM62 Matrix (Madeira et al., 2019).

SH3 domain			
	A1/A2	A1/A3	A2/A3
Identity [%]	79.7	86.4	79.7
Similarity [%]	89.9	93.2	91.5

- *Functions*

As mentioned above, this domain provides endophilin the ability to bind several PRD-containing proteins. Nevertheless, the SH3 domain of the three endophilins isoforms does not show the same affinity for all the PRD-containing proteins. Indeed, endophilin A1 and A2 isoforms showed a greater interaction with Ca^{2+} channel in a Ca^{2+}-dependent fashion, while endophilin A3 does not present this characteristic. Conversely, endophilin A3 shows greater affinity for dynamin than A1 or A2 (Tian et al., 2012).

Further to simple affinity differences, some PRD-containing proteins are capable of binding to one specific endophilin isoform, while they present no detectable interactions with the others. It is the case for various G-protein-coupled receptors

(GPCRs) that can specifically bind to endophilin A1, A2, or both, but not A3 (see section II.2.2 of introduction) (Boucrot et al., 2015). Another example is the activated leukocyte cell adhesion molecule (ALCAM/CD166) which binds specifically with endophilin A3 but not A1 or A2 (see III.2.3 of introduction) (Renard et al., 2020).

Interestingly, the SH3 domain was suggested to have an autoinhibitory effect by interacting with proline rich regions in the endophilin itself, preventing its interactions with other proteins and its recruitment to the plasma membrane (Meinecke et al., 2013). As mentioned above, this mechanism could be Ca^{2+} dependent (Chen et al., 2003). However, whether this mechanism is redundant in all three isoforms and efficiently regulates all endophilin isoforms functions needs further investigation.

- *Summary*

To summarize, the SH3 domain provides endophilins the ability to interact with different PRD-containing proteins, allowing their recruitment in complex machineries requiring coordination. It might also regulate endophilin activity through auto-inhibition in a Ca^{2+} and isoform-dependent fashion. Despite the high sequence homology in the three isoform SH3 domains, differences in affinity or even in specificity for different binding partners were observed and could be a potential explanation for mechanism specificity.

II.1.4 Linker

- *Structure*

The BAR and SH3 domains are connected by a linker region variating greatly in size between the different isoforms. Little is known about the conformation adopted by this linker and therefore on the potential direct function it could have (Kjaerulff et al., 2011).

- *Sequence comparison*

The linker is clearly the most variable region, as much as in size than in amino acids composition, between the three isoforms (**Tableau 4**, **Figure 4A**). A1 and A3 possess a shorter linker of ~45 amino acids while endophilin A2 has a longer one of ~60 residues. The linker of endophilin A1 displays many potential phosphoresidues, making it a potential target for post-translational modifications, while A2 and A3 linkers do not (Kjaerulff et al., 2011). Those differences make the linker region another interesting candidate for explaining isoform specificity.

Table 4. Sequence identity and similarity between paired endophilin linker regions amino acid sequences. Obtained using the EMBOSS Needle tool with the EBLOSUM62 Matrix (Madeira et al., 2019)

Linker region			
	A1/A2	A1/A3	A2/A3
Identity [%]	37.9	42.3	30.3
Similarity [%]	47.0	51.9	39.4
Gap [%]	30.3	25.0	37.9

- *Functions*

While no direct function related to this domain has been identified so far, it has been demonstrated that the linker is an important region in regulating receptor-mediated endocytosis. Indeed, endophilin A3 has been shown to be highly expressed in olfactory nerves and to have an inhibitory effect on dopamine D2 receptors, facilitating dopamine induced differentiation of those cells. This inhibitory effect, quite opposed to the role of endophilin A1 in clathrin-mediated endocytosis was proven linker-dependent through a domain swapping approach (Sugiura et al., 2004). This observation seems quite surprising for a simple linker region. A possible explanation could be that linker specificity comes from the way this region regulates the position of the SH3 domain and therefore the way it can interacts with other proteins.

- *Summary*

Given the shortness and variability of this region in the three isoforms, it is quite unlikely that it has a function on its own. However, the linker might regulate the position of the SH3 domain relatively to the BAR domain and therefore completely change the way endophilins interact with proteins, or even with itself.

II.2 Expression and cellular localization

II.2.1 Expression

The genes coding for endophilin A1, A2 and A3 are located on the chromosomes 9, 19 and 15 respectively. Endophilin A2 expression is ubiquitous, with a slight increase in pancreas, placenta, prostate, testis and uterus. Endophilin A1 is almost exclusively expressed in the brain and endophilin A3 is mostly expressed in brain and testis. More precisely, during fetal development, endophilin A1 is produced at higher level in fetal cerebellum, whereas endophilin A3 is expressed more importantly in fetal brain. This difference seems to disappear in adult brain (Giachino et al., 1997).

II.2.2 Subcellular localization

The subcellular localization of endophilins has been mostly studied in neurons. All three isoforms present a diffuse cytosolic background with a punctate pattern superimposed upon it. Endophilin highly concentrate at presynaptic nerve terminals, especially in response to synaptic stimulation, in agreement with the Ca^{2+}-dependent interaction with Ca^{2+} channel that mostly localizes in this region. In neurons, endophilins colocalize with dynamin I, synaptojanin and amphiphysin I (Ringstad et al.,

1997). This observation is not surprising since endophilins have been shown to interact directly with these proteins during clathrin-mediated endocytosis (see section III.2.1 of introduction).

In opposition with the mostly cytosolic localization of all the three isoforms in neurons, endophilin A2 present a predominant nuclear localization in haemopoietic, fibroblasts and epithelial cells. The subcellular localization of endophilin A2 in these cells was shown to be regulated by CRM1 dependent nuclear export, which could explain the cell-context-dependent subcellular localization of this protein (Cheung et al., 2004). Furthermore, endophilin A2 presented a very dynamic localization and high expression level in different phases of the mitotic cell. Of particular interest is the localization of endophilin A2 at late mitosis in the midzone and midbody of the cell, a pattern that is very similar to the spatial distribution of proteins involved in cytokinesis. This observation could give a glimpse to a potential role of endophilin A2 in late mitosis during cell replication (Cheung et al., 2004). It shows that in addition to have isoform-dependent functions, endophilins (at least the A2 isoform) could as well have cell-context-dependent functions.

In overexpression studies performed on HeLa cells, localization of exogenous endophilin A3 was shown to be almost exclusively cytosolic, sometimes presenting very dynamic filamentous strands in addition to the previously observed punctuated pattern. Those strands did not colocalize so much with the actin cytoskeleton. However, they showed a significant colocalization with a particular subset of microtubules (Hughes et al., 2004). This observation is in adequation with more recent studies demonstrating that the combination of a pulling force, provided by motors such as dynein on microtubules, and a BAR-domain scaffold enable the scission of the tubule from the membrane through the mechanism of friction driven scission (FDS) (more details in section III.2.2 of introduction) (Simunovic et al., 2017).

II.3 Endophilin : a perfect protein for endocytic vesicle building

Looking at their global structure, endophilins appear to be the perfect proteins to build endocytic vesicles. Their SH3 domain allows the interaction with a wide range of proteins. This may be useful for the recruitment of specific cargoes as well as other proteins involved in endocytic machineries such as dynamin or the actin cytoskeleton. Their BAR domain drives membrane curvature which is key for the generation of endocytic pits. Finally, their multiple amphipathic helices contribute to the formation of tubular scaffold around the endocytic pit and stabilize membrane curvature in addition to supporting the membrane scission (Boucrot et al., 2015). Taken together, it is not surprising that endophilins were quickly identified as important actors in various endocytic mechanisms, which are of particular interest here and will be the topic of the following section.

III. Endocytosis and the role of endophilins

III.1 Introduction to endocytosis

Endocytosis is an essential cellular process required for the uptake of nutrients from cell environment and turnover of plasma membrane components. It can be considered as the opposite mechanism of exocytosis, which describes the fusion of internal vesicles with the plasma membrane. Together, these two mechanisms modulate the composition of the cell membrane and therefore are key regulators of many aspects of the cell such as signaling, adhesion, mobility or polarity (Doherty and McMahon, 2009).

The first endocytic mechanism identified was the clathrin-mediated endocytosis (CME) back in the 60's. This constitutively active and highly conserved mechanism is by far the best characterized to date. Later, in the 90's, several evidences highlighted the existence of other endocytic processes in animal cells, therefore categorized as clathrin-independent endocytosis (CIE). These mechanisms have their own machinery, involving a subset of diverse proteins, and allow the uptake of very specific cargoes in response to different signals (**Figure 7**) (Casamento et Boucrot, 2020). The function redundancy of these distinct pathways allows to question their biological relevance. As usually in the living realm, the most likely explanation is that the existence of all these mechanisms provides flexibility and adaptability to the cell in the way it can regulate the composition of its membrane and its effectors.

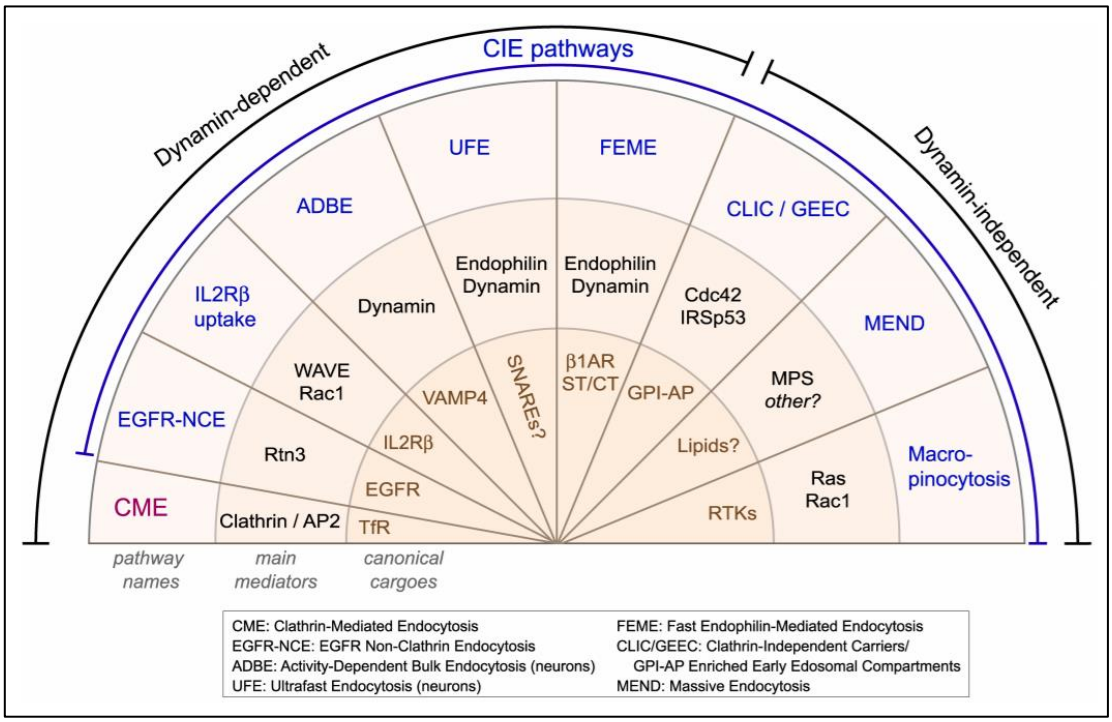


Figure 7. Overview of the currently characterized endocytic pathways, the main mediator of their machinery and canonical cargoes (Casamento et Boucrot, 2020). Endocytic pathways have been categorized according to the main proteins involved in their machinery. The first major distinction is made between mechanisms involving clathrin or not during the construction of the endocytic vesicle (CME/CIE). The second major distinction comes from the involvement of the GTPase dynamin during vesicle scission (Dynamin-dependent/independent).

Despite their physiological differences, many CIE mechanisms present similarities regarding the proteins involved in their machinery. This is for instance the case of the GTPase dynamin which is the major actor of vesicle scission in a majority of endocytic processes (**Figure 7**). It is also worth to mention members of the BAR domain superfamily that are found in many of endocytic processes acting as driver of membrane curvature. Therefore, the different endocytic modalities must not be seen as independent pathways, but more as a particular organization of redundant actors that coordinate themselves differently according to the situation (cell type, signal, protein levels, cargo, …).

III.2 Endophilin in endocytic mechanisms

As mentioned in the previous section, endophilins present all the characteristics required to build endocytic vesicles. It is therefore not surprising that endophilins were initially identified as a component of CME, more specifically during vesicle scission (Gad et al., 2000). Given its high expression level in the brain, endophilin A1 appeared as a crucial component of synaptic vesicle scission in neurons (Verstreken et al., 2002). More recently, new CIE mechanisms were discovered in which some endophilin isoforms play a central role. These gather the fast endophilin-mediated endocytosis (FEME) (Boucrot et al., 2015) and galectin-8/endophilin A3-mediated endocytosis (Renard et al., 2020).

III.2.1 Clathrin mediated endocytosis

Clathrin-mediated endocytosis (CME) is the first endocytic mechanism to be discovered. For a long time, it was considered as the sole endocytic pathway and its extensive study led to a very precise understanding of the molecular mechanisms that occur during the formation of clathrin-coated vesicles. Still, further investigation is required to fully understand the coordination between all the components of this complex machinery (Kaksonen et Roux. 2018).

CME holds its name from the central role played by clathrin. Clathrin is a dimer formed by two proteins: Clathrin Heavy Chain (CHC, ~190 kDa) and Clathrin Light Chain (CLC, ~ 25 kD). The association of three dimers forms a triskelion that is the functional unit of clathrin scaffold. These triskelia oligomerize at plasma membrane to form a polygonal coat around nascent endocytic pits, together with the help of adaptor proteins (AP) complexes such as AP2 (Royle, 2012). Without going into details, CME can be simplified as a succession of 5 major steps that are highly conserved among eukaryotic cells (**Figure 8**).

- *Initiation*

The assembly of the clathrin coat at plasma membrane initiates the formation of the pit. The coat contains as well adaptor proteins such as the AP2 complex (Kelly et al., 2014), epsins (Ford et al., 2002) and clathrin assembly lymphoid myeloid leukaemia protein (CALM) family (Miller et al., 2015). These adaptor proteins recruit clathrin at

the nascent endocytic pit and form the ‘pioneer module’ responsible for initiating the endocytic process (Cocucci et al., 2012) .

- *Cargo recruitment*

The specific recruitment of diverse cargo proteins resides in different kind of interactions between components of the clathrin coat and binding sites located in the cytosolic part of cargo proteins (Traub, 2009).

- *Membrane bending*

The generation of a fully-formed clathrin-coated vesicle is driven by two coordinated processes. First, the polymerization of clathrin into a spherical coat imposing its curvature to the membrane (Kirchausen et Harrison, 1981). And secondly, the polymerization of actin filaments at the endocytic pit, regulated by members of the Wiskott-Aldrich syndrome protein (WASP) family and the complex of actin-related protein 2/3 (Arp 2/3) (Merrifield et al., 2004) (Sirotkin et al., 2005).

- *Scission*

The scission consists in the separation of the forming vesicle from the plasma membrane. This mechanism is characterized by the successive recruitment of diverse BAR domain proteins. Starting with the F-BAR protein FBP17, followed by intermediate curvature BAR proteins like sorting nexin 9 (SNX9), finally leading to the recruitment of N-BAR proteins such as endophilin and amphiphysin. This coordinated recruitment of BAR domain proteins is believed to be regulated by two parameters: (1) the increasing curvature of the membrane (Itoh et al., 2005) and (2) different phosphorylation states of phosphatidylinositol (PIPs) that are sequentially generated during vesicle formation (Posor et al., 2013). The SH3 domains of endophilin and amphiphysin then recruit dynamin by interacting with its PRD. The scaffold built by N-BAR domain proteins provides a template for dynamin oligomerization, which finally uses GTP hydrolysis energy to constrict the neck and release the vesicle from the plasma membrane (Antonny et al., 2016).

- *Uncoating*

Finally, proteins from the uncoating module disassemble the coat around the vesicle. This module contains chaperones, protein kinases and lipid phosphatases (Kaksonen et Roux, 2018). Notably, endophilin was shown to recruit the 145 kDa isoform of the phosphatase synaptojanin (SJ1-145) that dephosphorylates $PI(4,5)P_2$, stimulating uncoating (Perera et al., 2006).

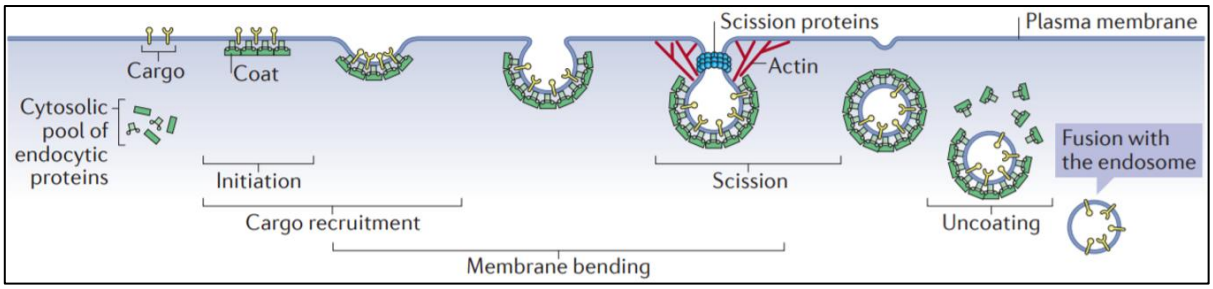


Figure 8. Schematic representation of the different molecular modules involved in the different steps of CME (Kaksonen et Roux, 2018). Initiation: The assembly of the clathrin coat along with adaptor proteins initiate the endocytic event. **Cargo recruitment:** Specific cargoes are recruited to the forming pit by direct or indirect interaction with proteins from the pioneer module. **Membrane bending:** The polymerization of clathrin along with coordinated actin polymerization lead to the formation of the endocytic vesicle. **Scission:** Successive recruitment of BAR domain proteins and dynamin along with actin polymerization lead to the scission of the endocytic vesicle from the plasma membrane. **Uncoating:** Proteins from the uncoating module disassemble the clathrin-coat.

III.2.2 Fast Endophilin Mediated Endocytosis (FEME)

Although endophilin was initially identified as a CME component, several observations, such as the fact that endophilin was not detected in every forming clathrin-vesicles (Taylor et al., 2011) suggested a peripheral role of endophilin in this mechanism. Interestingly, in 2015, Boucrot and his team highlighted a new CIE modality responsible for the uptake of a subset of ligand-activated receptors in which endophilin plays a central role (Boucrot et al., 2015). Since then, this new pathway referred as fast endophilin-mediated endocytosis (FEME) has been strongly investigated providing a detailed understanding of the molecular mechanisms underlying it. Just like CME, the mechanism of FEME carrier formation can be divided in a succession of steps (**Figure 9**) precisely reviewed in Casamento et Boucrot, 2020, in which endophilin systematically appears, highlighting its central role in the pathway.

- *Priming*

FEME carrier formation initiates with the pre-enrichment of endophilin molecules into small patches, which last 5 to 15 seconds, mostly located at the leading edge of cells (Boucrot et al., 2015). These patches are formed through the successive recruitment of different proteins. It starts with the active GTP loaded Cdc42 Rho GTPase, bound to $PI(3,4,5)P_3$ at the cell surface, which recruits FBP17 and CIP4 (Chan Wah Hak et al., 2017). The SH3 domains of these F-BAR proteins then recruit the SHIP1/2 phosphatase and lamellipodin. SHIP1/2 hydrolyzes $PI(3,4,5)P_3$ into $PI(3,4)P_2$, stabilizing the lamellipodin through its PH domain. Finally, the multiple PRD in lamellipodin lead to the accumulation of endophilin at the endocytic pit (Vehlow et al., 2013). Upon activation of the receptors recognized by FEME, these patches initiate carrier formation. In the absence of signal, the priming complex simply disassembles and a new patch appears at another location (I. priming on **Figure 9**).

It is interesting to note the likeness in the successive recruitment of BAR domain proteins (F-BAR followed by N-BAR) with the vesicle scission step during CME. Although both mechanisms are characterized by specific actors, the redundancy

observed underlies the fact that endocytic mechanisms are not individual pathways but more a combination of different sub mechanisms that gives rise to a specific process according to the situation.

- *Cargo selection and carrier formation*

So far, 16 FEME cargoes have been identified among GPCRs, receptor tyrosine kinases, cytokine receptors and axon guidance receptors. All of these cargoes present the characteristic of binding directly or indirectly to at least one endophilin isoform (mostly A1 and A2) through their SH3 domain (II. Receptor activation and III. Cargo sorting on **Figure 9**) (Boucrot et al., 2015).

Of note, this pathway can also be highjacked by the cholera and shiga toxins for infection (Renard et al., 2015). In this context, recruitment of endophilin to infection sites does not involve the SH3 domain but rather its ability to sense the curvature generated by toxin binding.

However, the mechanism by which the uptake of these cargoes only occurs upon ligand binding as well as the initiation of membrane tubulation remain elusive. A model involving microtubules has been suggested to explain the elongation of the tubules. Dynein and dynactin would interact with the plasma membrane and adjacent microtubules to promote extension of nascent tubules (Day et al., 2015). This elongation process would be coordinated with the polymerization of endophilin around the tubule to stabilize membrane curvature (IV.FEME carrier formation on **Figure 9**). However, further investigations are required to understand more precisely the molecular coordination that occurs during the formation of these invaginations.

- *Membrane scission*

In addition to the well-known constriction role of dynamin recruited by endophilin, other mechanisms have been shown to promote membrane scission such as local actin polymerization as well as friction driven scission (FDS) (Renard et al., 2015). FDS is a generic mechanism in which the friction imposed by the membrane bound protein (here endophilin) to the underlying membrane leads to scission. It requires two mechanical conditions: (1) an external force elongating the tubule, which is provided by dynein and microtubules, and (2) a membrane-bound endophilin scaffold that generates friction through the interaction of its four amphipathic helices and BAR domain with the underlying phospholipids (Simunovic et al., 2017).

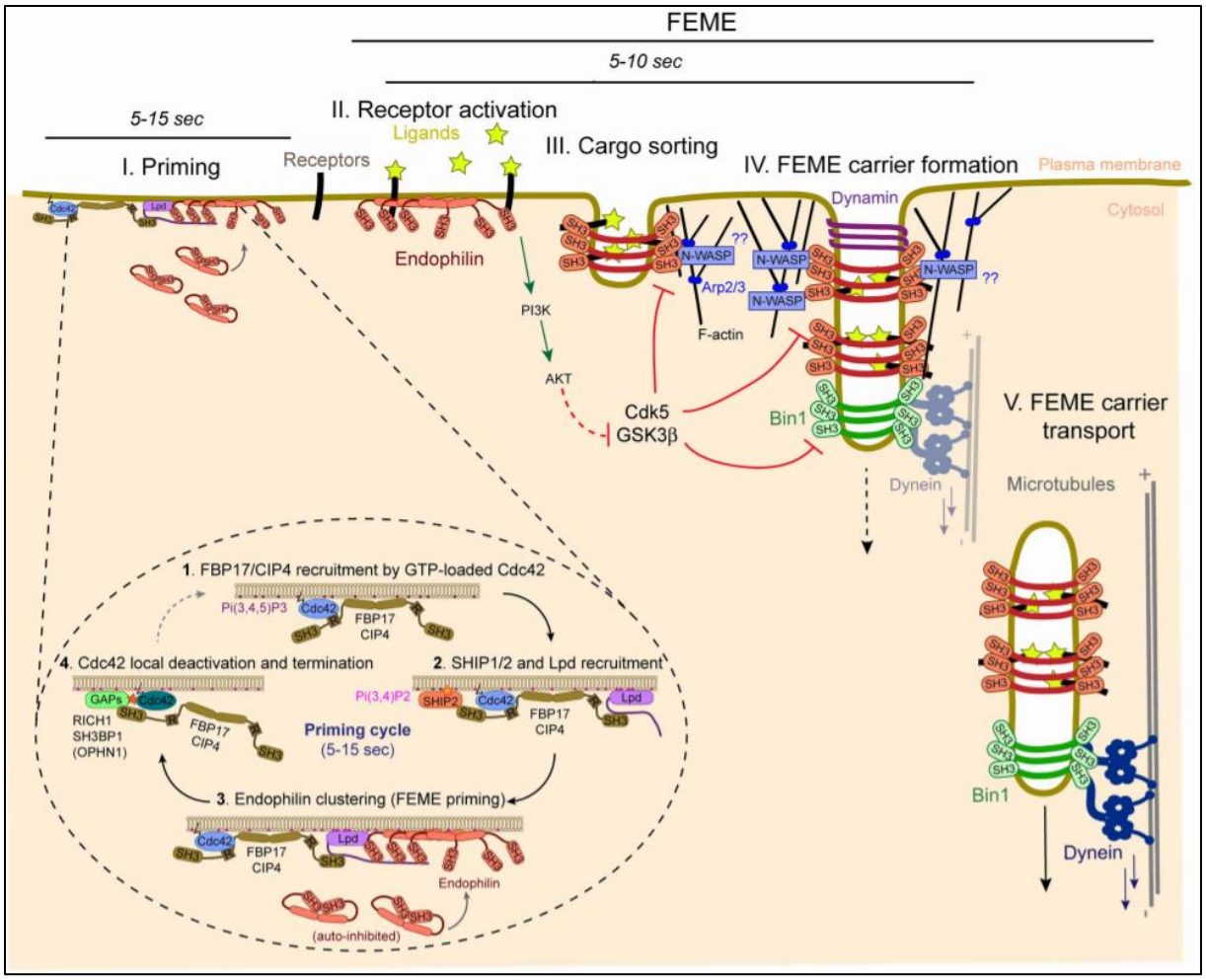


Figure 9. Schematic representation of the different molecular modules involved in the different steps of FEME (Casamento et Boucrot, 2020). I. Priming: successive recruitment of diverse proteins leads to the formation of endophilin patches at the cell surface that initiate endocytosis upon specific receptor activation by their ligand. In absence of signal, the priming complex disassembles. **II. Receptor activation:** the activation of receptors stabilizes pre-enriched endophilin patches and initiates their internalization. **III. Cargo sorting:** cargoes that present direct or indirect interaction with endophilin are further recruited to nascent FEME pits. **IV FEME carrier formation:** actin and endophilin polymerization lead to tubulation of the membrane. The coordinated action of dynamin, microtubule-based molecular motors and friction-driven scission leads to the release of the vesicle from the plasma membrane. **V. FEME carrier transport:** after budding, FEME carrier are transported on microtubules by dynein.

III.2.3 Galectin-8/Endophilin A3-Mediated Endocytosis

Although CME and FEME are both well-characterized endocytic mechanisms involving endophilins, the specificity of the three isoforms in these processes is not always clear or detailed, most probably due to the high sequence homology within the endophilin A subfamily. However, very recently, a new endocytic pathway entirely specific to the A3 isoform called Galectin-8/Endophilin A3-mediated endocytosis has been described in my lab (Renard et al., 2020), questioning isoform redundancy and specificity in endocytic mechanisms.

- *Activated leukocyte cell adhesion molecule (ALCAM) is endocytosed in an endophilin A3-dependent manner*

In a quantitative proteomic screening to search for membrane proteins endocytosed by CIE, activated leukocyte cell adhesion molecule (ALCAM/CD166) appeared as an interesting hit since its surface level was decreased upon inhibition of CME. CD166 is a type I transmembrane protein member of the immunoglobulin (Ig) superfamily. It possesses a short cytoplasmic tail and five extracellular domains : two amino-terminal variable domains (V), followed by three constant one (C) (van Kempen et al., 2001). Heterophilic interactions between CD166 and CD6, a protein found at the surface of T-cells, has been shown to be crucial for T-cell proliferation and activation (Zimmerman et al., 2006). In addition to its role in immunity, CD166 is also a regulator of cell adhesion through transient homophilic interactions between V domains. CD166 is also capable of cis-oligomerization through lateral interactions of the three C domains, forming CD166 clusters that reinforce the adhesiveness of the cell at these spots (von Lersner et al., 2019). Of note, dysregulation in CD166 surface level emerged from multiple independent studies as a contributor to tumor progression through coordinated changes in cell adhesion (von Lersner et al., 2019).

Since CD166 external level was decreased upon CME inhibition, it was hypothesized that CD166 was endocytosed by a clathrin-independent mechanism which switches its post-endocytic trafficking from recycling to lysosomal degradation when CME is blocked, as it has been proposed for other clathrin-independent cargoes (Dutta et Donaldson., 2015). Further colocalization and functional assays later confirmed the clathrin-independent nature of CD166 endocytosis (Renard et al., 2020).

Interestingly, depletion of endophilin A3 strongly inhibited CD166 uptake in HeLa cells, while no effect was observed when depleting A1 or A2 isoforms. In addition, pull-down experiments using CD166 cytosolic tail as a bait revealed transient interaction with endophilin A3 but not with A2 (Renard et al., 2020). These observations suggest that CD166 is not endocytosed by the ‘classical’ FEME, since this mechanism is mostly endophilin A2 specific, and reinforce the hypothesis that endophilin isoforms are not functionally redundant.

- *Clustering of CD166 by galectin-8 promotes its internalization*

Although endophilin A3 presented a very short lifetime at the cell surface (~90% of spot lasted less than 30s at the cell surface), the clustering of CD166 at specific spots clearly decreased endophilin A3 dynamic (~80% of spot lasted more than 50s). This is an interesting similarity with FEME in which endophilin forms discrete clusters at the plasma membrane that are stabilized only upon receptor activation by their ligand, leading to the initiation of the carrier formation (Casamento et Boucrot., 2020). Although CD166 is not a receptor that is activated by a soluble ligand, glycosylated CD166 was shown to bind extracellular galectin-8 (Gal8), leading to the formation of CD166 clusters (Fernandez et al., 2016). Gal8 is a protein from the mammalian lectin family made of a tandem repeat of two carbohydrate recognition domains (CRD). The two CRDs are joined together by a linker peptide and confer binding specificity for glycosylated proteins.

Interestingly, the inhibition of Gal8-CD166 interaction by two strategies (interaction competitor and glycosphingolipids depletion) resulted in a strong decrease in CD166 uptake, suggesting that clustering of CD166 by Gal8 promotes its internalization, similarly to the GL-Lect hypothesis. According to this hypothesis, the binding of lectins to glycosphingolipids (GSLs) and/or glycosylated proteins at the cell surface induces membrane invagination by generating an asymmetric stress (**Figure 10**) (Johannes et al., 2015). This is notably the case for CD44 and β1-integrin uptake by the GLIC/GEEC pathway (another CIE mechanism) that is initiated by their binding to extracellular galectin-3 (Gal3) (Lakshminarayan et al., 2014).

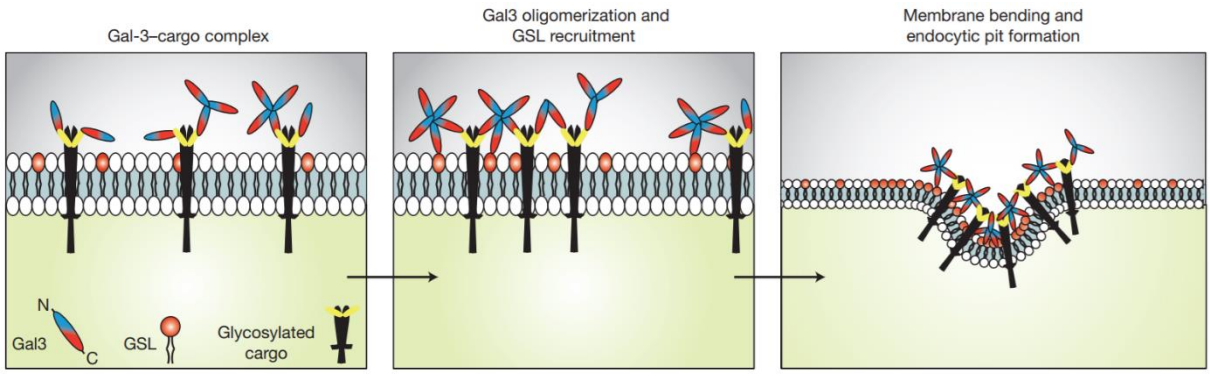


Figure 10. Binding of lectins to GSLs and/or glycosylated proteins induces membrane remodeling and endocytosis initiation according to the GL-Lect hypothesis (Lakshminarayan et al., 2014). A lectin, in this case Gal3, binds glycosylated lipids and proteins at the outer leaflet of the plasma membrane. The oligomerization of Gal3 leads to the formation of clusters, enriched in GSLs and glycosylated proteins. This clustering leads to an asymmetric stress that initiate membrane curvature and the formation of an endocytic pit.

Colocalization assays and combined depletion of endophilin A3 and Gal8 further confirmed that CD166 clustering by Gal8 and its endocytosis mediated by endophilin A3 are two related mechanisms (Renard et al., 2020).

- *Summary*

So far, experimental results indicate that the hetero-binding of Gal8 to membranous CD166 induces the formation of clusters containing a high concentration of GSLs and

CD166. These clusters would induce the recruitment and the stabilization of endophilin A3 at the cell surface through direct or indirect interaction between the two proteins, leading to the formation of a tubular endocytic pits and the internalization of CD166. This new mechanism is currently named galectin-8/endophilin A3-mediated endocytosis due to the major involvement of these two proteins in the process. A schematic representation of this mechanism compared to FEME is shown on **Figure 11**. Further investigation is still required to determine whether other molecular actors involved in later stages of FEME such as dynein and microtubules are also involved in this new mechanism.

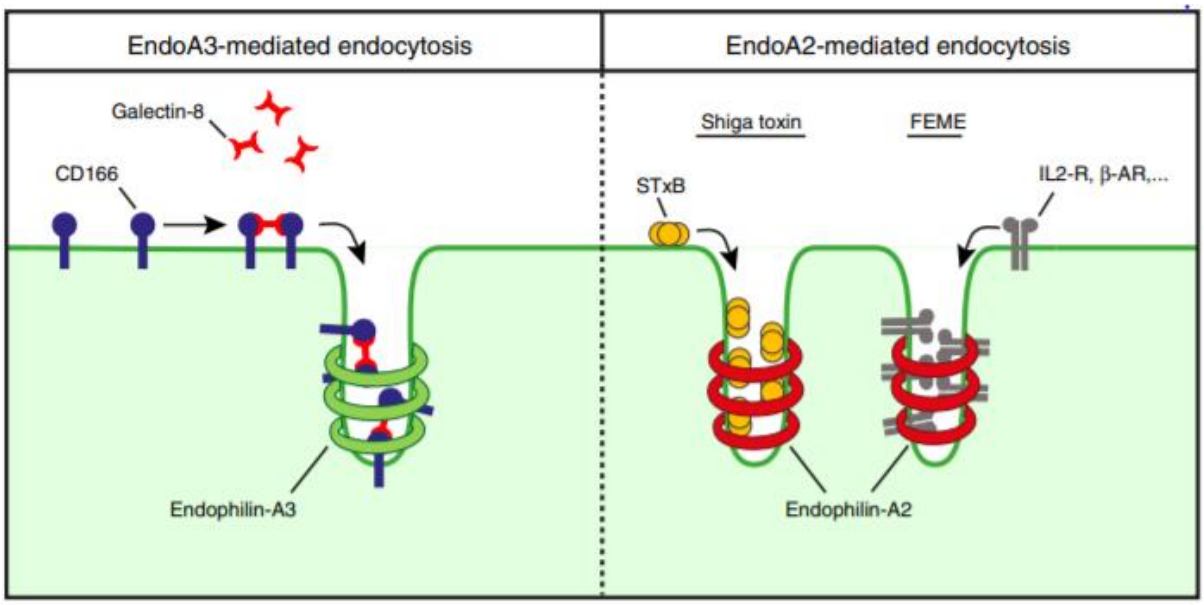


Figure 11. Schematic representation of galectin-8 and endophilin A3-mediated endocytosis and FEME (Renard et al., 2020). EndoA3-mediated endocytosis: Gal8 binds to glycosylated CD166 molecules, inducing the formation of clusters that stabilize pre-enriched endophilin A3 at the cell surface and that initiate the formation of an endocytic pit. **EndoA2-mediated endocytosis**: Membrane curvature induced by the binding of Shiga toxin or the activation of FEME cargoes such as IL2-R or β-AR by their ligand induce the stabilization of pre-enriched endophilin A2 at the cell surface and the initiation of endocytic pit formation.

Material and Methods

I. Molecular biology techniques

I.1 General techniques

I.1.1 *E. coli* transformation by thermic choc

Every plasmid used during this master thesis were amplified by thermo-competent bacteria *E. coli* (TOP10 strain).

Bacteria were thawed on ice. Then 2-5 µL of plasmid containing solution were added to 50 µL of bacteria. After 15 minutes incubation on ice, the thermic choc at 42°C was applied to the cells for 45 seconds before getting back on ice for 5 minutes. The mix was diluted in 450 µL of LB medium and incubated for 1h at 37°C. 100 µL are finally spread out on a petri dish containing ampicillin and incubated at 37°C overnight.

I.1.2 Plasmid extraction

Isolation of the plasmid transformed in *E. coli* was done using the "Pure Yield™ Plasmid Miniprep System" (Promega) and "Plasmid Plus Midi Kit" (Qiagen) following the suppliers' recommendations.

I.1.3 DNA electrophoresis

DNA electrophoresis was performed in gel containing 1% agarose diluted in TBE buffer (890 mM Tris, 890 mM boric acid, 2 mM EDTA). 2 µL of Midori Green (Fastgene®) for DNA detection were added to 30 µL of gel before agarose polymerization. DNA migration was performed under a voltage of 130 V for 30 minutes and revealed under UV (λ=265 nm).

I.1.4 DNA purification on gel

The different DNA fragments were purified from gel using the kit "QIAquick® Gel Extraction Kit" (Qiagen) following the supplier recommendations.

I.2 Mutant sequences obtention

I.2.1 Extraction of total RNA from HeLa and U2OS cells

To obtain human endoA2 cDNA sequence, total RNA was isolated from HeLa and U2OS cells grown in culture flasks. RNA extraction was performed using TRIzol™ reagent (ThermoFisher), following the supplier recommendations. Total RNA was then stored at -70°C before being used for cDNA obtention by reverse transcriptase.

I.2.2 cDNA synthesis by reverse transcriptase

Total cDNA was obtained from the total RNA of U2OS cells previously isolated using the "QuantiTect® Reverse Transcription Kit" (Qiagen), following the supplier recommendations. cDNA was stored at -20°C before being used for endophilin A2 sequence PCR amplification.

I.2.3 PCR reactions

Every fragment required for the construction of the different mutants were obtained by PCR. The PCR mixes were performed in a final volume of 50 µL containing:

- 1 µL of template (diluted to a concentration of ~100 ng.µL^{-1} for plasmid and ~5 ng.µL^{-1} for insert) ;
- 0.5 µM of each primer forward (F) and reverse (R) ;
- 0.2 mM of each dNTP ;
- 1 µL of Phusion High-Fidelity DNA polymerase (Thermo Scientific) ;
- 10 µL of buffer Phusion GC 5x concentrated (Thermo Scientific) ;
- X µL of water to reach a final volume of 50 µL.

- *Endophilin A2 full-length*

The first step consisted in amplifying endoA2 sequence from total cDNA of U2OS cells. Primers and cyclisation conditions are shown in **Table 5 and 6** respectively.

Table 5. Primers for endophilin A2 PCR amplification

Primer	Sequence
F-BAR EndoA2	5'-ATGTCGGTGGCGGGGCT-3'
R-SH3 EndoA2	5'-CTGCGGCAGGGGCACAA-3'

Table 6. Cyclization condition for endophilin A2 PCR amplification

Step (35 cycles)	Temperature [°C]	Time [min', sec'']
Initial denaturation	98	30''
Denaturation	98	10''
Annealing	67	30'
Extension	72	40''
Final extension	72	10'
Conservation	4	∞

Once the endoA2 sequence obtained, the fragment served as template to be re-amplified by PCR using other primers to be cloned by the Gibson method. The primers and cyclization conditions are shown in **Table 7 and 8** respectively.

Table 7. Primers for endophilin A2 (Gibson) PCR amplification

Primer	Sequence
F-BAR EndoA2 (Gibson)	5'-CGGCCTAGCTAGCGCTTAAGATGTCGGTGGCGGGGCTGAA-3'
R-SH3 EndoA2 (Gibson)	5'-GACTGCAGAATTCGAAGCTTCTGCGGCAGGGGCACAAGCA-3'

Table 8. Cyclization condition for endophilin A2 (Gibson) PCR amplification

Step (35 cycles)	Temperature [°C]	Time [min', sec'']
Initial denaturation	98	30''
Denaturation	98	10''
Annealing + Extension	72	35''
Final extension	72	10'
Conservation	4	∞

- *EndoA3-ΔlinkerSH3*

The DNA fragment coding for EndoA3-ΔlinkerSH3 was amplified by PCR using as template the plasmid coding for the full-length endoA3. Primers are shown in **Table 9**.

Table 9. Primers for EndoA3-ΔlinkSH3 PCR amplification

Primer	Sequence
F-BAR EndoA3 (Afl-II)	5'-GCGCTTAAGATGTCGGTGGCCGG-3'
R-BAR EndoA3 (Eco-R1)	5'-CCGGAATTCGACTGGATGCAGCTGATATTC-3'

Cyclization conditions were similar to those used for endoA2 (Gibson) PCR amplification (**Table 8**).

- *BAR A2-SH3 A3*

The DNA fragment corresponding to the BAR domain of endoA2 was amplified using the fragment of full-length endoA2 as template. On the other hand, the fragment corresponding to the linker and SH3 domain of endoA3 was amplified using the plasmid coding for full-length endoA3 as template (already available at the lab). Primers and cyclization conditions are shown in **Table 10 and 11** respectively.

Table 10. Primers for BAR-EndoA2 and linkerSH3-EndoA3 fragments PCR amplification

Primer	Sequence
BAR-EndoA2	
F-BAR EndoA2 (Afl-II)	5'-GCGCTTAAGATGTCGGTGGCGGGGCT-3'
R-BAR EndoA2 (homology linker EndoA3)	5'-CGTCTGGGGACTGAGGAAGCTTCCCGCAT-3'
LinkerSH3-EndoA3	
F-linker EndoA3 (homology BAR EndoA2)	5'-GGAAGCTTCCTCAGTCCCCAGACGAGAATACA-3'
R-SH3 EndoA3 (Eco-R1)	5'-CCGGAATTCGAAGCTTCTGAGGTAAAGGCACGA-3'

Table 11. Cyclization condition for BAR-EndoA2 and linkSH3-EndoA3 fragments PCR amplification

Step (35 cycles)	Temperature [°C]	Time [min', sec'']
Initial denaturation	98	30''
Denaturation	98	10''
Annealing + Extension	72	25''
Final extension	72	10'
Conservation	4	∞

The fragments were then assembled in a triple PCR reaction with equimolar concentrations of each fragment, using the primers corresponding to the full-length chimeric protein extremities. Cyclization conditions were similar to those used for the fragment obtention, except for the annealing and extension time which was 35 s instead of 25.

- *BAR A3-SH3 A2*

The procedure to obtain the BAR A3-SH3 A2 mutant was similar to the one for BAR A2-SH3 A3, except for the templates which were obviously exchanged and the primers which are shown in **Table 12**.

Table 12. Primers for BAR-EndoA3 and linkerSH3-EndoA2 fragments PCR amplification

Primer	Sequence
BAR-EndoA3	
F-BAR EndoA3 (Afl-II)	5'- GCGCTTAAGATGTCGGTGGCCGG -3'
R-BAR EndoA3 (homology linker EndoA2)	5'-CCGCTTAGGGCGACTGGATGCAGCTGATAT-3'
LinkerSH3-EndoA2	
F-linker EndoA2 (homology BAR EndoA3)	5'-TGCATCCAGTCGCCCTAAGCGGGAGTATA-3'
R-SH3 EndoA2 (Eco-R1)	5'-CCGGAATTCGAAGCTTCTGCGGCAGGGGCACAA-3'

- *BAR A2-linker A3-SH3 A2*

Once again, the procedure to obtain BAR A2-linker A3-SH3 A2 (B2L3S2) was similar to the one for the two previous chimeric proteins. Except that this time, three fragments were required instead of two and a plasmid coding for full-length endophilin A2 were used instead of the fragment alone as template.

Like full-length endophilin A2, the fragment was designed to be cloned by the Gibson method instead of restriction ligation. The primers designed for this purpose are shown in **Table 13**.

Table 13. Primers for BAR-EndoA2, linker-EndoA3 and SH3-EndoA2 fragments PCR amplification

Primer	Sequence
BAR-EndoA2	
F-BAR EndoA2 (homology plasmid)	5'-CGGCCTAGCTAGCGCTTAAGATGTCGGTGGCGGGGC TGAA-3'
R-BAR EndoA2 (homology linker EndoA3)	5'-TTCTCGTCTGGGGACTGAGGAAGCTTCCCG-3'
Linker-EndoA3	
F-linker EndoA3 (homology BAR EndoA2)	5'-CGGGAAGCTTCCTCAGTCCCCAGACGAGAA-3'
R-linker EndoA3 (homology SH3 EndoA2)	5'-GCTCGGCTGGTCCAGGGGAATGTTAGAACC-3'
SH3-EndoA2	
F-SH3 EndoA2 (homology linker EndoA3)	5'-GGTTCTAACATTCCCCTGGACCAGCCGAGC-3'
R-SH3 EndoA2 (homology plasmid)	5'-GACTGCAGAATTCGAAGCTTCTGCGGCAGGGGCACAA GCA-3'

The three fragments were assembled together in a single triple PCR reaction with equimolar concentrations of each fragment, using the primers corresponding to the full-length chimeric protein extremities. Once again, cyclization conditions were similar to those used for the fragment obtention, except for the annealing and extension time which was 35 s instead of 25.

I.3 Cloning

I.3.1 Restriction-ligation method

The constructions Endo A3-ΔlinkerSH3, BAR A2-SH3 A3 and BAR A3-SH3 A2 were cloned into the plasmid pIRES and fused at their N-terminal with eGFP using the restriction-ligation method. The extremity primers for these constructions presented above have been designed to contain the required restriction sites in a way that puts the mutant protein and the GFP sequences in frame.

Each DNA fragments (the plasmid pIRES containing full-length endophilin A3 and the sequence of each mutants) were digested by the restriction enzymes Afl-II and Eco-R1 (New England Biolabs). The plasmid was digested in a final volume of 20 µL containing ~6 µg of plasmid, 1 µL of each restriction enzyme, 1 µL of Calf Intestine Phosphatase, 2 µL of buffer Cutsmart 10x concentrated (New England Biolabs) and complemented with water. The inserts were digested in a final volume of 40 µL containing ~300 ng of insert, 1 µL of each restriction enzyme, 4 µL of buffer Cutsmart 10x concentrated and complemented with water. Digestion were performed at 37°C for 2h.

Ligation were performed with a 5x molar excess of insert compared to plasmid in a solution of 20 µL containing 1 µL of T4 DNA ligase, 2 µL of buffer DNA ligase and

complemented with water. After 1h incubation at room temperature, 2 µL of ligation solution were transformed into *E. coli* (see section I.1.1 of Material and methods).

I.3.2 Gibson method

The full-length fragment of endoA2 and BAR 2-linker A3-SH3 A2 were cloned using the Gibson method. The extremity primers displayed above have been designed to contain homology sequences of ~20 nucleotides with the terminal sequences of the plasmid in a way that put the insert and GFP sequence in frame.

Equimolar quantities of insert and plasmid were mixed with 10 µL of Gibson mix (NEBuilder® HiFi DNA Assembly Master Mix), complemented with water to reach a final volume of 20 µL. The mix was incubated at 50°C for 1h. 5 µL were than transformed into *E. coli* (see section I.1.1 of Material and methods).

I.3.3 Sequencing

After *E. coli* transformation, positive clones were verified by restriction using the restriction enzymes Pst I or Sma I (New England Biolabs). Clones that presented the expected pattern were finally sequenced by the society Microsynth. The mixes contained 3 µL of primer corresponding to one extremity of the fragment, ~18 ng of DNA/100 bp complemented with water to reach a final volume of 15 µL.

II. Cellular biology techniques

II.1 General techniques

II.1.1 antibodies

Table 14 gathers each antibody used for immunofluorescence and western blot analysis.

Table 14. List of antibodies used during the Master thesis

Antibodies	Origin	Supplier
Immunofluorescence		
anti-human CD166	Mouse	Bio-Rad
anti-mouse Alexa546	Goat	ThermoFisher
Western Blot		
anti-EndoA1	Rabbit	Cell Signaling Technology®
anti-EndoA2	Mouse	Santa Cruz Biotechnology Inc.
anti-EndoA3	Rabbit	Sigma
anti-CHC	Mouse	Bioscience
anti-GFP	Mouse	Roche
HRP anti-rabbit	Goat	Agilent
HRP anti-mouse	Goat	Sigma

II.1.2 Cell culture

HeLa cells and U2OS cells were grown at 37°C under 5 % CO_2 in DMEM high glucose glutamax (Gibco) supplemented with 10% fetal bovine serum (FBS), 100 $U.mL^{-1}$ penicillin, 100 $\mu g.mL^{-1}$ streptomycin and 1mM sodium pyruvate.

II.1.3 DNA transfection

Chemical transfection

For immunofluorescence experiments, plasmids encoding for our different mutants were transfected in HeLa cells using FuGENE® HD transfection reagent (Promega). More specifically, a transfection mix containing 50 µL of Opti-Mem® (Gibco), 3 µL of FuGENE® and ~1 µg of plasmid DNA was added to 500 µL of medium containing 100 000-120 000 cells. Cells were used for experiment 16-24h after transfection.

Stable expression

For the GST pull-down with the full-length endophilin A3, a cell line stably expressing endophilin A3 fused to GFP was already available.

Nucleofection

For the GST pull-down with the construct Endo A3-ΔlinkerSH3, the transfection of two 10 cm plates would have required too much product for chemical transfection. Nucleofection was therefore preferred. HeLa cells were transfected using the “Amaxa®

Cell line Nucleofector® Kit R" (Lonza) according to the manufacturer's recommendations. Cells were used for experiment 16-24h after nucleofection.

II.1.4 RNA interference

For the rescue and triple knockdown experiments, siRNAs were transfected using HiPerfect® transfection reagent (Qiagen) or Lipofectamine® RNAiMAX reagent (ThermoFisher). Cells were initially plated in 6 well plates with 300 000 cells per well and a final volume of 1.6 mL. The transfection mix contained 400 µL of Opti-Mem®, 4 µL of HiPerfect or 1 µL of Lipofectamine and a total of 8 µL of siRNA containing solution (10µM). After 48h at 37°C, cells were transferred into 4 well plates in a final volume of 500 µL of medium. For flow cytometry experiments, 120 000 cells were plated per well. For immunofluorescence, 80 000-100 000 cells were plated per well with a previously introduced coverslip to allow cell fixation. Additional wells containing 120 000 cells were also prepared for western blot analysis.

The different siRNA sequences used during the manipulations are shown in **Table 15**.

Table 15. List of siRNA used for the rescue and triple knockdown experiments. (*) indicate siRNA targeting non-coding sequences of the gene.

Name	Protein	Target sequence	Supplier
Hs_SH3GL2_5	Endophilin A1	5'-AAAGACTCTTTGGACATAGAA-3'	Qiagen
Hs_SH3GL1_5	Endophilin A2	5'-CACCAGCAAGGCGGTGACAGA-3'	Qiagen
Hs_SH3GL3_5	Endophilin A3	5'-CCAGACGAGAATACAAGCCAA-3'	Qiagen
Hs_SH3GL3_9*	Endophilin A3	Check the supplier	Qiagen
Hs_SH3GL3_10*	Endophilin A3	Check the supplier	Qiagen

II.2 Colocalization assays

80 000-100 000 cells were incubated in 500 µL DMEM + HEPES + 10% FBS at 37°C for 20 minutes. Cells were then incubated with 250 µL of anti-CD166 primary antibody solution (diluted 250x in the previous solution) for 15 minutes at 37°C. After three washes with PBS^{2+}, cells were fixed with paraformaldehyde 4 % (PFA) for 20 minutes at 37°C. Cells were then quenched for 5 minutes with 500 µL NH_4Cl 50 mM and permeabilized for 30 minutes with 500 µL of saponine solution (0.02% saponine and 0.2% BSA in PBS) at room temperature (RT). Following permeabilization, cells were incubated with A546-labeled secondary antibody (diluted 250x in the saponine solution) for 45 minutes at RT and in the dark. Finally, samples were washed 3 times with 500 µL of saponine solution, and mounted on microscopy slides with 8 µL of Fluoromount™ Aqueous Mounting (Sigma Aldrich) containing 2 $\mu g.mL^{-1}$ of DAPI (Sigma).

II.3 Rescue and triple knockdown assays

II.3.1 CD166 surface level - flow cytometry

To measure CD166 surface level for the rescue and triple knockdown assays, 120 000 cells were incubated in 500 µL DMEM + HEPES + 10% FBS at 37°C for 20 minutes. After 3 washes with ice cold PBS^{2+}, cells were incubated on ice for 10 minutes in PBS^{2+} containing 0.2% BSA to fix the cells. Cells were then incubated with the primary antibody solution (diluted 500x in the previous solution) for 40 minutes on ice. After 3 washes with the same buffer, the A546-labeled secondary antibody solution was added (diluted 250x still in the same buffer) for 40 minutes on ice. Finally, after 3 washes with PBS, cells were incubated in 500 µL of PBS containing EDTA (4 mM) on ice for 30 minutes and then detached manually to be analyzed by flow cytometry.

All measurements were obtained with a Guava easyCyte (Merck-Millipore) flow cytometer. Fluorescence intensity of 10.000 cells were measured for each condition with an excitation wavelength of 488 nm. A background measure performed on cells non-incubated with primary antibody (unlabeled) was subtracted from all other conditions. Finally, each median intensity was normalized relatively to control condition median fixed as 100%.

II.3.2 CD166 internalized level - confocal microscopy

80 000-100 000 cells were incubated in 500 µL DMEM + HEPES + 10% FBS at 37°C for 20 minutes. Cells were then incubated with 250 µL of anti-CD166 primary antibody solution (diluted 250x in the previous solution) for 15 minutes at 37°C. After three successive washes with ice cold PBS^{2+}, cells were washed 3 times with an acidic solution (300 mM NaCl, 5 mM KCl, 1 mM CaCl2, 1 mM MgCl2, 0.2 mM acetic acid, pH 2.5) for 20 seconds and 3 times again with cold PBS^{2+} before fixation with PFA on ice. These acidic washes aim at removing surface-bound antibodies corresponding to non-internalized CD166. The quenching, permeabilization and secondary antibody labeling steps were performed similarly to the colocalization assays (see section II.2 of Material and methods)

II.4 GST pull-down assays

II.4.1 Column preparation

2L cultures of BL21 *E. coli* transformed with GST-CD166 tail and GST constructs previously available at the lab were incubated for 3h at 37 °C with 100µM IPTG for induction. After centrifugation, cell pellets were resuspended in a solution of 150 mM NaCl, 25 mM HEPES, pH 7.4, and 2 mM EDTA complemented with 1 mM PMSF, 2 mM DTT, 100 $\mu g.mL^{-1}$ lysozyme, and protease inhibitor cocktail. After incubation on ice for 30 min, cells were lyzed by snap freezing in liquid nitrogen, quick thawing, and brief sonication. Lysates were clarified by ultracentrifugation in a Beckman Ti45 rotor for 1 h at 41,000 rpm (~131,000 × g). Following centrifugation, the supernatant was bound to 1 mL column bed of glutathione beads (GE Healthcare). The beads were

washed 5 times with 10 column volumes of wash buffer (500 mM NaCl, 25 mM HEPES, pH 7.4, and 2 mM EDTA). Finally, protein-bound beads were resuspended in 1 ml of lysis solution (150 mM NaCl, 25 mM HEPES, pH 7.4, and 2 mM EDTA).

II.4.2 Pull-down assays

HeLa cells stably expressing GFP-tagged endophilin A3 and HeLa cells transfected by nucleofection with EndoA3-ΔlinkerSH3 construct were quickly washed with ice-cold PBS, lysed in ice-cold lysis buffer (150 mM NaCl, 25 mM HEPES, pH 7.4, and 2 mM EDTA complemented with 1 mM PMSF, 1% NP40, protease, and phosphatase inhibitor cocktail (PhosSTOP roche)), sonicated, and spun at 14,000 × g for 15 min. 100 µl of GST/GST-CD166 tail-bound beads were then exposed to cell lysates for 2h at 4 °C, pelleted in a cooled benchtop centrifuge, and washed five times in lysis buffer complemented with 1% NP40 only. Beads were then boiled for 10 min in 100 µl of sample buffer for elution, and 8 µl of sample were run on sodium dodecyl sulfate polyacrylamide gel electrophoresis, ("input" lanes correspond to 1.5% of cell extract (final concentration of ~0.04 $mg.ml^{-1}$)). The proteins were transferred and immunoblotted using anti-endoA3 for endophilin A3 and anti-GFP for endoA3-ΔlinkerSH3 antibodies. Finally, membranes were stained with Coomassie blue to control the amount of GST fragments in each sample.

III. Imaging and quantification

III.1 Fluorescence microscopy

The previously immobilized samples were observed and imaged with the confocal microscope Zeiss LSM 710 equipped with 63X oil objective. This microscope possesses 7 laser lines (405, 440, 458, 488, 514, 543 and 633 nm). Only the 488 (green channel: GFP) and 543nm (Red channel: Alexa fluor 546) were used for these experiments.

III.2 Image processing and quantification

III.2.1 Colocalization

Colocalization was quantified with an object based quantification method using the plugin JACoP from the program ImageJ/Fiji. The two channels (488 and 546nm) from one image were splited. The most intense spots were separated from the background noise by setting manually for each channel the threshold with the “Find Maxima” tool. Each representative spot were then enlarged to the size of 2 pixels. Finally, the results were expressed as the percentage of CD166 spots colocalizing (with a 2 pixels tolerance level) with GFP spots per cells. 5 cells were quantified per replicate and 3 replicates were made for each mutant.

III.2.2 Internalized signal

The program Icy was used to quantify the amount of internalized anti-CD166 antibody in the rescue and triple knockdown experiments. Regions of interest (ROIs) were drawn manually around each cells (cells that presented a sufficient GFP signal for the rescue conditions). In all of these regions, bright spots corresponding to endocytic structures were detected by the plugin “spot detector” based on a manually fixed threshold. The same threshold was used for every conditions within the same replicate. Icy then calculates the number of endocytic structures per cell as well as their respective fluorescent intensity. These results are summed up for each cells to obtain the total amount of anti-CD166 antibody internalized, and divided by the surface level obtained by flow cytometry for each conditions. Finally, the results were normalized relatively to the median of the control condition, fixed as 100%. 30 to 60 cells were quantified for each replicate.

Objective and Strategy

I. Context and objectives

As introduced above, proteins from the BAR domain superfamily, and more precisely members of the endophilin A subfamily, act in key steps of endocytic processes. Surprisingly, despite sharing high sequence similarity, endophilin A isoforms were shown to control the CIE of distinct cargoes, calling for a better individual characterization of these proteins.

One of the main goal of my host laboratory is to understand the role of BAR domain proteins in unconventional endocytic mechanisms. Of note, colleagues recently identified a new endocytic modality mediated specifically by the endophilin A3 isoform, which controls the uptake of the tumor marker CD166. Consequently, in this master thesis, my objective consists in deciphering which domain of endophilin isoforms is responsible for cargo specificity.

II. Strategy

To answer this question, we developed a strategy based on the use of chimeric endophilin constructs generated by a domain swapping approach. The goal is to assess the functionality of these chimeric proteins compared to full-length native endophilin proteins in the context of CD166 endocytosis. If some mutants behave more like endophilin A2 while others behave similarly to A3, we should be able to decipher which domain of endophilin A3 is responsible for CD166 specificity.

Our experimental strategy consists in four steps:

1. Generation of GFP-tagged endophilin A2 & A3 mutants by a domain swapping approach.
2. Colocalization assay - check the association between the mutants and CD166 in HeLa cells.
3. Rescue experiment - assess the ability of mutants to rescue the endocytosis of CD166 in HeLa cells knocked-down for endoA3.
4. Pull-down assay - highlight direct or indirect interactions between the mutants and CD166 cytosolic tail.

Finally, we performed a side experiment questioning more deeply isoform redundancy in CD166 uptake. The internalized level of CD166 was compared between cells knocked-down for endoA3 alone and cells knocked-down for the three endophilin isoforms (triple knockdown, TKD). The results of this experiment are shown in **Annex 4**.

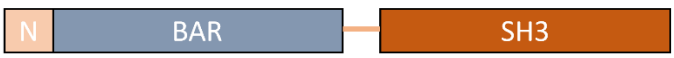
Results

I. Generation of endoA2 and A3 mutants by a domain swapping approach

The first part of my master thesis consisted in creating a library of plasmids encoding for various GFP-tagged endophilin mutants built using a domain swapping approach. To reduce the number of mutants in the library, we decided to focus on the A2 and A3 isoforms, which were shown to control the endocytosis of distinct CIE cargoes. All mutant proteins were cloned into the mammalian expression vector pIRES under the control of the strong constitutive promoter from Cytomegalovirus (CMV), and fused at their C-terminal with enhanced GFP. Several endophilin constructs were already available in the lab, including the full-length endoA3 and the two SH3 domains (EndoA2/A3-ΔBAR). The human cDNA coding for endoA2 was extracted from U2OS cells and inserted into the GFP plasmid as detailed in section I.2.1 and I.2.2 of material and methods. A mutated version of endoA3 lacking its linker and SH3 domains (EndoA3-ΔlinkerSH3) was also produced. In this project, three chimeric proteins composed of swapped domains of endoA2 and A3 were generated by triple PCR. The first one is composed of the BAR domain of endoA2 and the linker-SH3 domains of A3 (BAR A2-SH3 A3), while the second one is the reverse situation (BAR A3-SH3 A2). Finally, the last chimeric mutant consists in the full-length endoA2 sequence where the linker region is replaced by the one of endoA3 (BAR A2-linker A3-SH3 A2 or B2L3S2) (see section I of material and methods). More details concerning the construction and cloning of these mutants can be found in **Annex 1**.

The 8 endophilin constructs of the library are listed in the table below.

Table 16. Endophilin plasmid library. Domains coming from endoA3 or endoA2 sequences are labeled in **blue** and **orange**, respectively. **(*)** Indicates the plasmids that were already available at the start of the master thesis.

Name	Construction
EndoA3 *	
EndoA3-ΔlinkerSH3	
EndoA3-ΔBAR *	
EndoA2	
EndoA2-ΔBAR *	
BAR A2-SH3 A3	
BAR A3-SH3 A2	
BAR A2-linker A3-SH3 A2	

II. Colocalization between CD166 and endophilin mutants

Once the plasmid library completed, we measured the colocalization levels between our GFP-tagged mutants and CD166 in HeLa cells, using this data as a first readout of cargo specificity. Five constructs of the library were used in this experiment: endoA3, endoA2, endoA3-ΔlinkerSH3, BAR A2-SH3 A3 and BAR A3-SH3 A2. Full-length endoA3 acts here as a control/reference, given its well established association with CD166, and endoA2 as a negative control. All plasmids were transfected in HeLa cells and incubated for 15 minutes with anti-CD166 antibody before fixation. Anti-CD166 was then detected using Alexa546-labelled secondary antibody. Fifteen confocal images were acquired for each construct (**Figure 12A**). Finally, the percentage of CD166-positive structures colocalizing with the different endophilin constructs (**Figure 12B**) as well as the number of colocalization events per cell (**Figure 12C**) were calculated from confocal images using an object-based quantification approach (see section III.2.1 of material and methods). This quantification method allows us to focus specifically on vesicular structures that are of particular interest when studying endocytosis.

As illustrated in **Figure 12A**, GFP-tagged endophilin constructs (green signal) display a strong punctuated pattern, with a more diffuse signal for the full-length endoA2. On the other hand, CD166 (red signal) is particularly enriched at cell-cell contacts, which is in accordance with its cellular function. Nevertheless, CD166 punctuated structures can be observed towards the inside of the cell, possibly representing endocytic events.

When quantifying colocalization levels, some interesting differences between endophilin constructs can be noticed. In fact, endoA2 as well as endoA3-ΔlinkerSH3 constructs show significantly smaller colocalization levels with CD166-positive structures (14% & 19%) than the control (endoA3-28%) (**Figure 12B**). Surprisingly, both chimeric protein (BAR A2-SH3 A3 and BAR A3-SH3 A2) display similar colocalization levels to endoA3. These results obtained for endoA2, endoA3-ΔlinkerSH3 and the chimera BAR A2-SH3 A3 may suggest a role of the linker or SH3 domains in the specificity for CD166. However, swapping the linker-SH3 domains of endoA3 with the one of endoA2 did not reduce colocalization level to the one of full-length endoA2.

In addition, we took a look at the total number of colocalization events per cell (**Figure 12C**). Interestingly, this data clearly show that despite similar colocalization levels between the control and chimeric proteins, cells transfected with endoA3 show much more colocalization events per cell (~100 events/cell), suggesting a higher level of internalized anti-CD166. Of note, endoA2 and endoA3-ΔlinkerSH3 constructs show once again much less colocalization events than the control (~30><100 events/cell), supporting a role of the linker or SH3 domains in cargo recognition.

Altogether, these results are not robust enough to highlight a specific domain of endophilin responsible for CD166 specificity. In addition, in view of the data displayed on the two graphs (**Figure 12B&C**), it is important to remind that colocalization does not equal function. Further experiments, as shown in the following sections, will target more specifically the functional effect of these endophilin constructs on CD166 uptake.

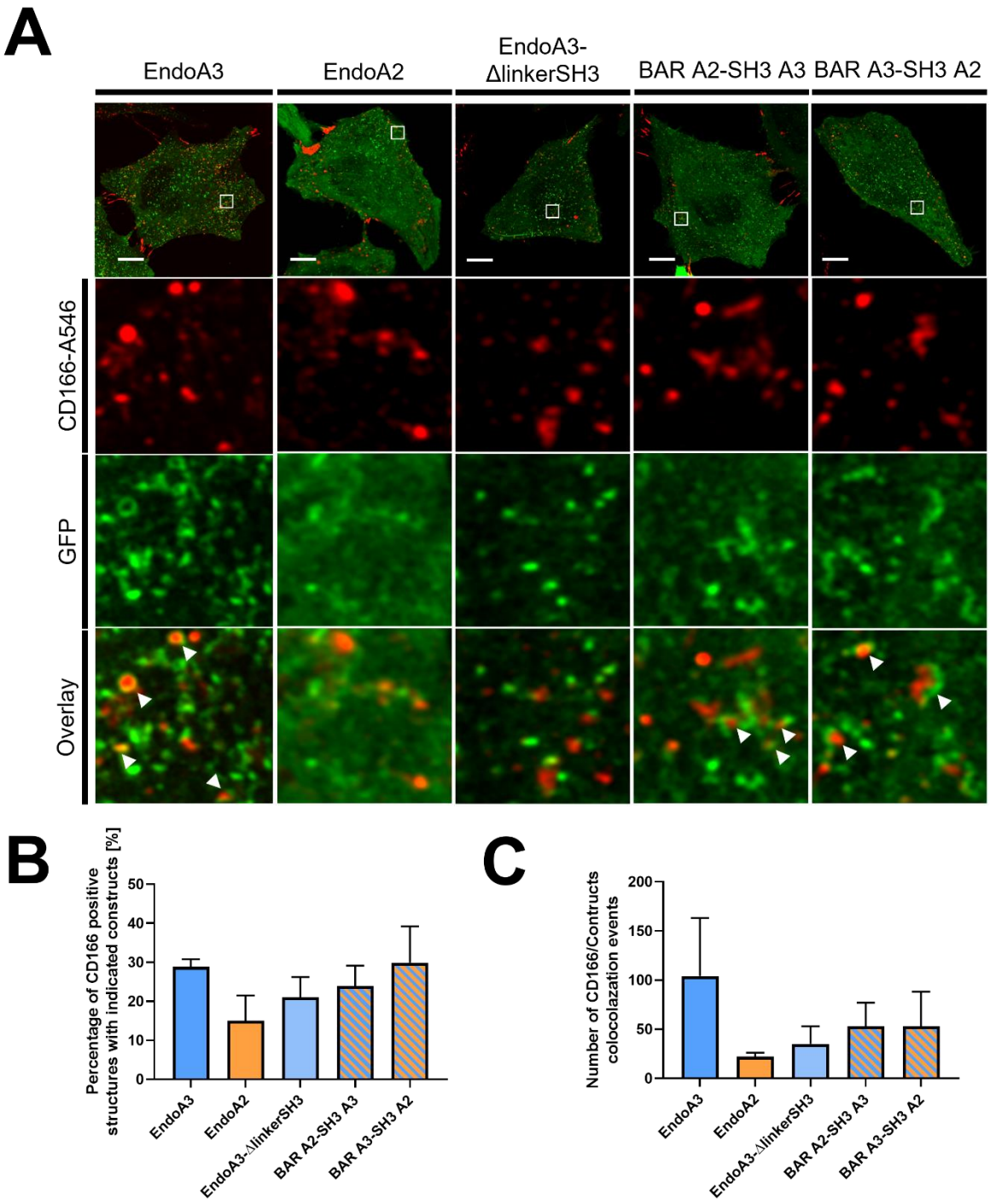


Figure 12. Colocalization of CD166-positive structures with full-length and mutated endophilins. (A) Continuous uptake of anti-CD166 antibody (red) in HeLa cells transfected with GFP-tagged endophilin mutants (green). Colocalization events are highlighted with white arrows. Scale bar = 10 µm **(B)** Percentage of CD166-positive structures colocalizing with the different endophilin constructs. CD166 shows ~30% of colocalization with endoA3, as well as with the two chimeric proteins BAR A2-SH3 A3 and BAR A3-SH3 A2. Full-length endoA2 and endoA3-ΔlinkerSH3 show ~2 and ~1.5 times less colocalization than the control, respectively. **(C)** Average number of colocalization events per cell for each endophilin construct. Note the higher number of colocalization events for the full-length endoA3, which suggests a higher level of internalized anti-CD166 when compared to the other conditions. **(B&C)** 5 cells/replicate, three independent replicates. NS, not significant, * $P < 0.05$, *** $P < 0.001$, **** $P < 0.0001$ (Kruskal–Wallis test with Dunn's multiple comparison post hoc test, two-sided). Error bars in graphs are median with 95% CI.

III. Rescue of CD166 uptake by the different mutants

For the next experiment, we asked whether the overexpression of our mutants could rescue CD166 endocytosis in cells knocked-down for endoA3. To silence endoA3 endogenous expression, HeLa cells were transfected with two siRNA sequences targeting different non-coding loci of endoA3 mRNA sequence (see section II.1.4 of material and methods). The individual silencing efficiency of each siRNA as well as their combination was verified by western blot (**Figure 13A**). Cells transfected with non-targeting siRNA were used here as control (siCtrl). Two days after silencing, cells were transfected with the different endophilin constructs. 16-24h after, cells were treated to quantify internalized anti-CD166 by confocal microscopy, while additional samples were used to quantify CD166 surface level by flow cytometry (see section II.3.1 and II.3.2 of material and methods). In order to label specifically CD166 endocytosed signal, cells were washed 3 times with an acidic solution to remove membrane bound anti-CD166. Internalized anti-CD166 signal in efficiently transfected cells (cells presenting a GFP signal) was captured by confocal microscopy and quantified using the spot detector plugin in Icy software (see section III.2.2 of material and methods). Finally, the internalized signal of each condition was divided by its relative surface level (measured by flow cytometry) compared to the control condition (siCtrl).

To validate our approach, we started with a proof of concept experiment (1 replicate) by measuring the rescue of CD166 uptake when overexpressing the full-length endoA3 or a cytosolic GFP (empty plasmid-negative control) (**Figure 13B**). As expected, the knockdown of endogenous endoA3 strongly decreased CD166 endocytosis by 50%. This effect can be visualized looking at the endocytosed CD166 signal (red) which appears much more punctuated in control cells compared to a weaker and more diffuse signal in endoA3 depleted cells (**Figure 13C**). Of note, expression of cytosolic GFP did not affect the internalized level of CD166 (**Figure 13B**). Conversely, expression of exogenous endoA3 clearly rescued CD166 uptake, with an almost two times fold increase in internalized signal compare to the control condition (siCtrl). This rescue of CD166 endocytosis by endoA3 overexpression is illustrated in the **Figure 13D**.

Once the approach validated, the experiment was repeated with the mutants (3 replicates). Results are shown on **Figure 13E** and additional images can be found in (**Annex 3**). Unsurprisingly, endoA2 did not succeed to rescue CD166 endocytosis as much as endoA3, underlying the specificity of the A3 isoform in this mechanism. EndoA3-ΔlinkerSH3 did not succeed either to rescue CD166 endocytosis, suggesting an important role of the linker-SH3 domains for the specificity of endoA3. The chimeric protein corresponding to endoA2 with endoA3's linker (B2L3S2) also failed at rescuing CD166 endocytosis. This suggests that, despite its great variability between the three endophilin isoforms, the linker is not responsible for endoA3 specificity in this mechanisms. Surprisingly, no significant difference can be observed between the two

chimeric proteins BAR A2-SH3 A3 and BAR A3-SH3 A2 that both succeeded to rescue CD166 endocytosis almost as much as full-length endoA3. Unfortunately, these results does not allow us to conclude which part of the protein regulates endoA3 specificity in the context of CD166 endocytosis.

It is worth to mention that results from the rescue assays suffer from a great variability between the replicates, which makes it difficult to draw solid conclusions. The different sources of variability as well as suggestions to avoid it will be more deeply discussed in the discussion section.

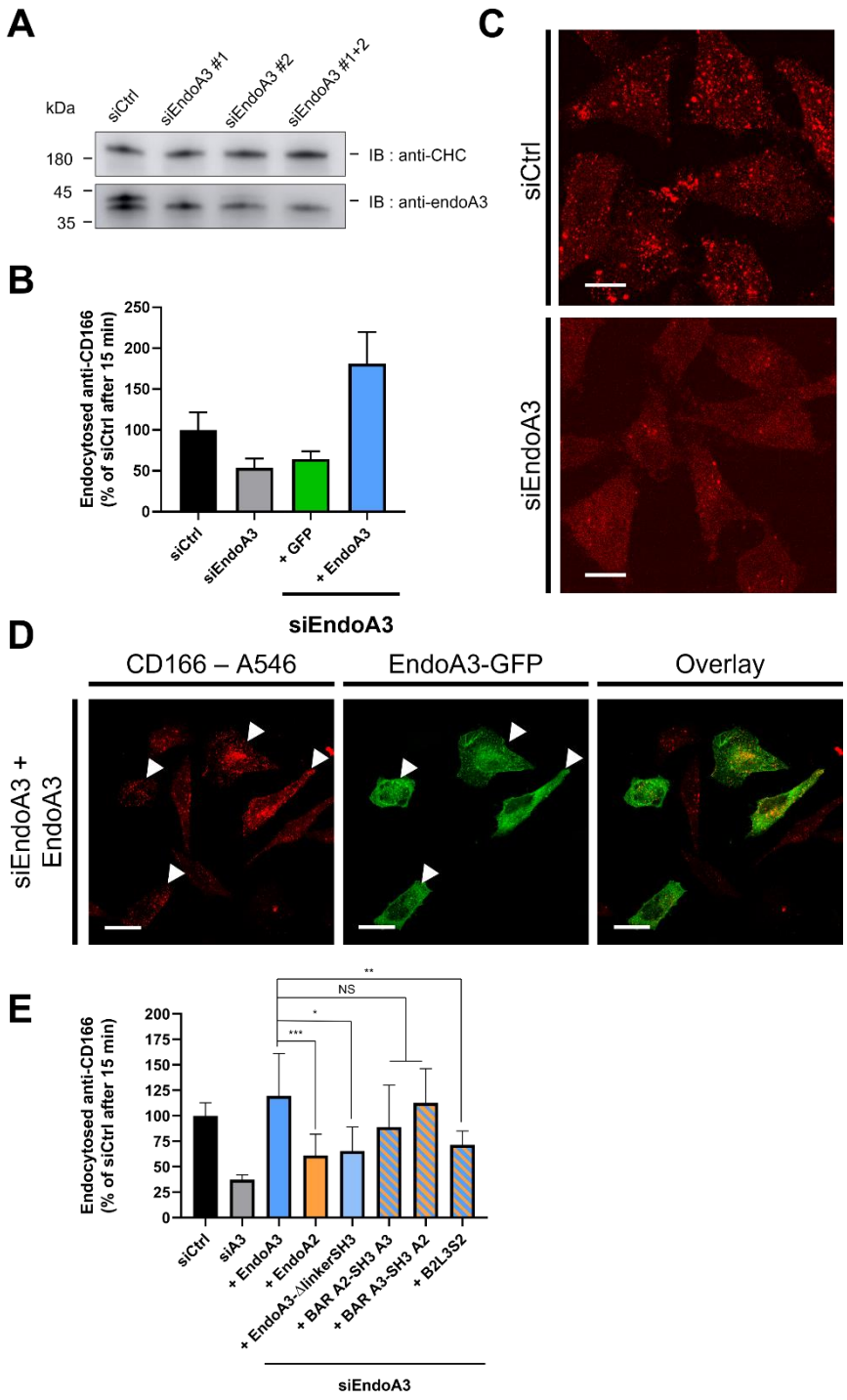


Figure 13. Rescue of CD166 uptake by endophilin mutants. (A) Western blot analysis of HeLa cells transfected with two different siRNA targeting endogenous endoA3 individually (#1 and #2) or combined (#1+2). Note the efficient knockdown in each condition **(B)** Proof of concept rescue experiment. EndoA3 knockdown effectively reduces anti-CD166 internalization by 50%. Expression of the empty GFP plasmid does not alter CD166 endocytosis in contrast to the overexpression of full-length endoA3 that fully restores CD166 uptake. 30-60 cells/replicate, one replicate. **(C)** Representative images of internalized CD166 signal in HeLa cells transfected with non-targeting siRNA (siCtrl) and endoA3-targeting siRNA (siEndoA3). The punctuated red signal observed in the control condition corresponding to CD166 endocytic vesicles is weaker and more diffuse in the endoA3 knockdown condition. Scale bar = 15 µm **(D)** Cells effectively transfected with GFP-tagged endoA3 (green signal and white arrows) show stronger endocytosed CD166 signal compared to non-transfected cells. Scale bar = 25 µm **(E)** Rescue experiment with the various endophilin constructs. Internalized signal is normalized relatively to the median of the siCtrl condition, considered here as the 100% of endocytosis. 30-60 cells/replicate, 3 independent replicates. NS, not significant. * $P < 0.05$. ** $P < 0.01$. *** $P < 0.001$ (Kruskal-Wallis test with Dunn's multiple comparisons post hoc test, two-sided).

IV. Pull-down assay - assessing CD166 interaction with endophilin mutants

The last objective of my master thesis consisted in assessing *in vitro* endoA3 specificity for CD166. This interaction was previously demonstrated in the lab (Renard et al, 2020), although the domain responsible for specificity is still unknown. To answer this question, we performed pull-down assays using Glutathione S-transferase (GST) fused to CD166 cytosolic tail (CD166-CT) that we coupled to glutathione beads to act as bait protein. GST alone was used here as a control condition. Cell lysates of HeLa cells expressing different endophilin constructs were then incubated with each resin. After five successive washes, western blot were carried out to compare the fraction of endophilin interacting with both resins.

In view of the colocalization and rescue data, we first compared the affinity for the resins of full-length endoA3 and the mutant that lacks the linker-SH3 domains (endoA3-ΔlinkerSH3). For lysate preparation, a cell line stably expressing GFP-tagged endoA3 was already available at the lab, giving a homogenous expression of the construct within the cell population. Regarding the endoA3-ΔlinkerSH3 mutant, nucleofection was preferred to chemical transfection for more efficiency.

As expected, full-length endophilin A3 was strongly retained by the GST-CD166-CT resin, in contrast to GST only (5X enrichment), confirming the interaction between the two proteins (**Figure 14A**). Conversely, no enrichment was observed for the endoA3-ΔlinkerSH3 construct, suggesting that the linker-SH3 region is indeed responsible for specificity (**Figure 14B**). Nevertheless, the western blot signal obtained for endoA3-ΔlinkerSH3 was overall very weak, most probably due to partial transfection within the cell lysate. Therefore, if this experiment confirmed what has previously been demonstrated, it still needs further investigation to confirm the results obtained with the endoA3-ΔlinkerSH3 construct.

Of note, due to lack of time, we did not perform pull-down assays for the remaining mutants of the library, and only one replicate was obtained for endoA3 and endoA3-ΔlinkerSH3. No doubt that the protocol still needs to be optimized, especially for the mutant constructs that require a better transfection yield. Suggestions for improvement will be discussed in the following section.

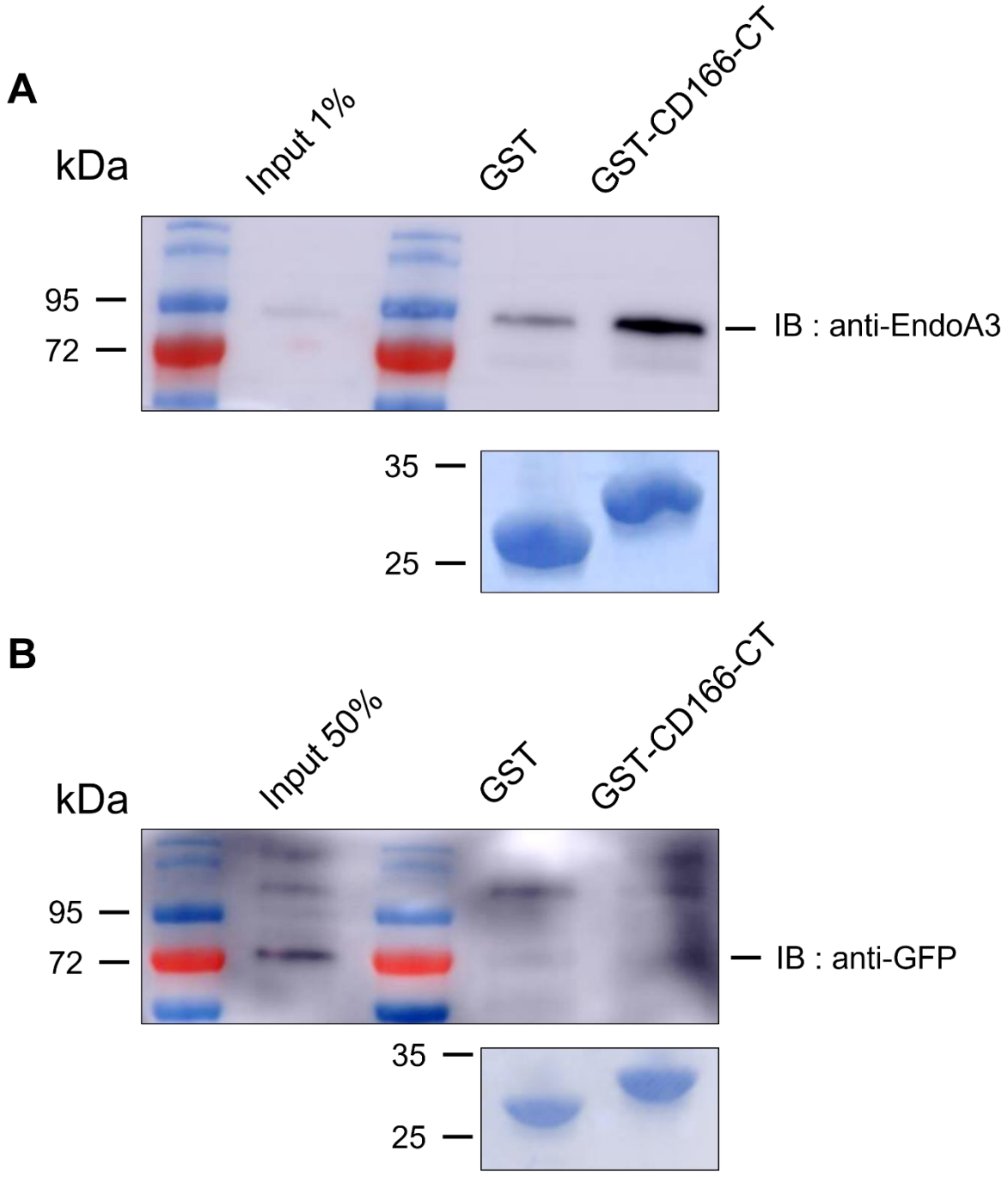


Figure 14. Pull-down experiment using GST-coupled CD166 cytosolic tail exposed to cell lysates expressing various GFP-tagged endophilin constructs. (A) GFP-tagged endoA3 was detected with a monoclonal anti-endoA3 antibody. GST and GST-CD166-CT were detected using Coomasie blue staining and used as loading controls. Note the clear enrichment of GFP-tagged endoA3 in the GST-CD166-CT condition compared to GST alone, which confirms the previously described interaction between endoA3 and CD166. **(B)** GFP-tagged EndoA3-ΔlinkerSH3 was detected using anti-GFP antibody. Note the absence of enrichment in the GST-CD166-CT condition.

Discussion

I. Colocalization between CD166 and endophilin mutants

The objective of this first experiment consisted in comparing colocalization levels of CD166 positive structures with the different endophilin constructs. Colocalization events were easily identifiable for cells transfected with full-length endoA3 (more images in **Annex 2.1**). As expected, full-length endoA2 colocalized twice less with CD166 than endoA3, confirming the specificity of the A3 isoform for this endocytic mechanism. However, it is important to notice that cells transfected with endoA2 displayed a much more diffuse cytosolic pattern (more images in **Annex 2.2**), which makes the detection of clear colocalization events more difficult. Since the threshold used to distinguish endophilin spots from the background signal is set manually, this might have account for some variability in the results between the different constructs.

EndoA3-ΔlinkerSH3 colocalization with CD166 was also lower than the full-length protein, highlighting the importance of the linker-SH3 region for cargo specificity. Concerning the two chimeric proteins BAR A2-SH3 A3 and BAR A3-SH3 A2, some colocalization events that seemed to correspond to CD166 endocytic vesicles coated with the mutants could be observed in both conditions (more images in **Annex 2.4** & **2.5**). The percentage of colocalization confirms that there is no significant differences between full-length endoA3 and both of these constructs. However, a few things need to be taken into account before affirming that these constructs are both capable of inducing CD166 endocytosis. Firstly, it is likely that each chimeric proteins are capable of forming heterodimers with endogenous full-length proteins, BAR A2-SH3 A3 with endoA2 and BAR A3-SH3 A2 with endoA3. Therefore, some colocalization events observed with the BAR A3-SH3 A2 mutant might come from heterodimers containing one full-length endophilin A3, therefore distorting the conclusion we can draw (this statement is also valid for endoA3-ΔlinkerSH3). In addition, the fact that the number of colocalization events per cell is approximately twice higher when cells are transfected with full-length endoA3 compared to both chimeric proteins might indicate that none of the two mutants are as effective as endoA3 to induce CD166 endocytosis. Regarding the last chimeric protein B2L3S2, this construct was not yet available at the start of this experiment.

Obviously, working in overexpression conditions is always a source of bias, sometimes making the comparison between the different conditions difficult. Unfortunately, since we are working with chimeric proteins that do not exist endogenously, labelling them by immunofluorescence is impossible and overexpression is the only way to perform this kind of experiment. One possible but time-consuming improvement would be to obtain cell lines stably expressing each of our mutant, which might remove some variability coming from the transfection step. Another improvement would be to remove the capacity of heterodimerization between endogenous endophilin and our mutants by knocking-down endogenous endoA3 expression (as we did for the rescue assays).

Anyway, these assays were mostly performed to obtain preliminary information regarding the way our mutant behave in overexpression conditions. In order to draw more solid conclusions, these results have to be combined with data coming from the rescue and GST pull-down assays.

II. Rescue of CD166 uptake by the different mutants

This second experiment aimed at determining if our mutants could rescue CD166 endocytosis in HeLa cells knocked-down for endoA3. Before discussing the data, it is worth mentioning that this experiment suffer from great variability due to several reasons. Firstly, the efficiency of siRNA-based silencing can vary between the cells, even within a single replicate. Consequently, some cells transfected with siRNAs targeting endoA3 presented a pattern pretty similar to cells from the control condition, while other presented a clear decrease in the amount of CD166 internalized (more images in **Annex 3.1**). This variability obviously affected our results because it was not possible to confirm whether a rescue phenotype observed in a cell transfected with one mutant was actually caused by this mutant or simply due to less effective endoA3 knockdown. Secondly, some transfected cells expressed endophilin constructs very strongly while others displayed a very weak GFP signal. It seems logical to think that the expression level of the constructs modifies the impact it can have on CD166 endocytosis. Therefore, cells transfected with the same endophilin mutant might show very different rescue phenotypes, increasing the variability of our results.

Despite these sources of variability, there are still clear conclusions that can be drawn from our data. As expected, the overexpression of full-length endoA3 clearly succeeded to rescue CD166 endocytosis in the majority of transfected cells (more images in **Annex 3.2**). These results, combined with the colocalization assays support the strong implication of endoA3 in CD166 endocytic mechanism that was previously highlighted in my lab. Regarding endoA2, the results are consistent with the colocalization assays and clearly show that this protein does not rescue CD166 uptake as much as endoA3 does. It is however important to notice that endoA2 overexpression succeeded to rescue partially CD166 endocytosis compared to the siEndoA3 condition, suggesting a potential redundancy between endophilin isoforms in CD166 endocytosis. To verify if this redundancy does happen endogenously, we compared CD166 internalization in cells knocked-down for endoA3 alone (siEndoA3) and for the three endophilin isoforms simultaneously (TKD) using the same immunofluorescence approach as in rescue assays with 15 minutes internalization. The results are shown in **Annex 4** and clearly demonstrate that the combined depletion of the three endophilin isoforms does not further decrease CD166 endocytosis compared to endoA3 knockdown alone. These results are not in agreement with what we observe in endoA2 rescue assays. One possible explanation would be that the overexpression of endoA2 might upregulate global endocytosis levels in the cell and indirectly increase CD166 internalization.

In accordance with colocalization data, the rescue experiment with endoA3-ΔlinkerSH3 highlighted the crucial role of the SH3 domain for endoA3 recruitment in CD166 endocytosis. The partial rescue when compared to the siEndoA3 condition is hardly explainable. One hypothesis could be that in cells where endoA3 depletion was not complete, endoA3-ΔlinkerSH3 was able to dimerize with the remaining endogenous protein, forming dimers containing one SH3 domain and therefore partially rescuing CD166 uptake. The fact that the BAR A2-linker A3-SH3 A2 mutant did not rescue either CD166 endocytosis demonstrates that the specificity of endoA3 in this mechanism cannot be exclusively attributed to the linker region, suggesting that the SH3 domain is the one responsible for endoA3 specificity.

Nevertheless, the two chimeric proteins, BAR A2-SH3 A3 and BAR A3-SH3 A2 managed to rescue CD166 endocytosis almost as much as full-length endoA3, with BAR A3-SH3 A2 slightly more than BAR A2-SH3 A3. If the specificity of endoA3 for CD166 only came from the SH3 domain, we would have expected the mutant BAR A2-SH3 A3 to rescue CD166 uptake as much as full-length endoA3, and conversely the BAR A3-SH3 A2 to behave more like endoA2. This is however not what we observe.

Overall, these results point to a role of the SH3 domain in CD166 recognition by endoA3, as it is the case for endoA2 and several FEME cargoes (Boucrot et al., 2015). However, the linker and BAR domain might regulate the position of the SH3 domain relatively to the rest of the protein, therefore modifying the way it can interacts with its cargo. As mentioned above, the results suffer from different sources of variability. Some aspects can be improved. Similarly to colocalization assays, it could be very interesting to obtain cell lines stably expressing each of our mutants to remove the variability coming from the transfection step. More biological replicates could as well be added to refine the results. Finally, it could be interesting to enlarge the scope of investigation by performing similar experiments but looking at endoA2 specific cargoes such as Interleukin-2 receptor (IL2-R) or β-adrenergic receptor (β-AR).

III. GST Pull-down assay with CD166 cytosolic tail

The last objective of my master thesis consisted in assessing *in vitro* endoA3 specificity for CD166 using our different mutants. Unfortunately, we only had the time to perform this experiment for the full-length endoA3 and endoA3-ΔlinkerSH3 constructs. The clear enrichment of endoA3 in the GST-CD166-CT condition is consistent with previous results (Renard et al., 2020). However, if we could not see an enrichment of endoA3-ΔlinkerSH3 signal in the GST CD166-CT condition, the western blot signal was overall very weak, probably due to the low percentage of transfected cell in the cell lysate. Indeed, the absence of stable cell lines expressing our various endophilin mutants, like we had for full-length endoA3, forced us to use nucleofection to express our protein of interest.

Sadly, the results from this last experiment only confirm what was already knew, and did not bring any further information concerning CD166 interaction with the different endophilin isoforms. With more time, it would have been interesting to perform this experiment with a more optimized protocol and for the other mutants. Once again, cell lines stably expressing each of our mutant would be of great help to perform these experiments, allowing us to bypass the transfection step and have 100% of positive cells.

Conclusion and Perspectives

The main objective of my master thesis was to assess which domain of endophilin A3 was responsible for cargo specificity in the context of CD166 endocytosis.

The first step of the project consisted in the generation of a plasmid library encoding for different endophilin mutants using a domain swapping approach. We then used these constructs to assess their association with CD166 endocytic structures and to see whether they were able to rescue CD166 uptake in HeLa cells depleted with endogenous endophilin A3. The results from these two experiments confirmed the specific implication of endophilin A3 in CD166 endocytosis, in contrast to the A2 isoform. Additional results from the triple knockdown experiment further demonstrated that no functional redundancy between the three endogenous endophilin isoforms could be observed in the context of CD166 endocytosis. Overall, our data highlight the importance of the linker-SH3 region for cargo recognition. However, it is still difficult to point out one specific domain of the protein.

We also performed GST pull-down assays to assess *in vitro* the interaction of our mutants with CD166 cytosolic tail. Results from this experiment only allowed us to confirm the previously identified interaction between CD166 and endoA3. As a perspective, it could be of great of interest to complete these experiments with all the mutants in the library. The data from these interaction assays could help interpreting the results from the colocalization and rescue assays and therefore allows us to draw more solid conclusions.

However, it might be possible that the specificity cannot be assigned to one single domain. Therefore, our domain swapping approach might not be robust enough, and a more complex structural analysis of the three endophilin isoforms would be necessary to answer the question of endophilin A3 specificity.

Bibliography

Antonny, B., Burd, C., De Camilli, P., Chen, E., Daumke, O., Faelber, K., Ford, M., Frolov, V. A., Frost, A., Hinshaw, J. E., Kirchhausen, T., Kozlov, M. M., Lenz, M., Low, H. H., McMahon, H., Merrifield, C., Pollard, T. D., Robinson, P. J., Roux, A., & Schmid, S. (2016). Membrane fission by dynamin : What we know and what we need to know. *The EMBO Journal*, *35*(21), 2270-2284. https://doi.org/10.15252/embj.201694613

Aspenström, P. (2019). BAR Domain Proteins Regulate Rho GTPase Signaling. *Advances in Experimental Medicine and Biology*, *1111*, 33-53. https://doi.org/10.1007/5584_2018_259

Boucrot, E., Ferreira, A. P. A., Almeida-Souza, L., Debard, S., Vallis, Y., Howard, G., Bertot, L., Sauvonnet, N., & McMahon, H. T. (2015). Endophilin marks and controls a clathrin-independent endocytic pathway. *Nature*, *517*(7535), 460-465. https://doi.org/10.1038/nature14067

Boucrot, E., Pick, A., Çamdere, G., Liska, N., Evergren, E., McMahon, H. T., & Kozlov, M. M. (2012). Membrane fission is promoted by insertion of amphipathic helices and is restricted by crescent BAR domains. *Cell*, *149*(1), 124-136. https://doi.org/10.1016/j.cell.2012.01.047

Carman, P. J., & Dominguez, R. (2018). BAR domain proteins-a linkage between cellular membranes, signaling pathways, and the actin cytoskeleton. *Biophysical Reviews*, *10*(6), 1587-1604. https://doi.org/10.1007/s12551-018-0467-7

Casamento, A., & Boucrot, E. (2020). Molecular mechanism of Fast Endophilin-Mediated Endocytosis. *The Biochemical Journal*, *477*(12), 2327-2345. https://doi.org/10.1042/BCJ20190342

Chan Wah Hak, L., Khan, S., Di Meglio, I., Law, A.-L., Lucken-Ardjomande Häsler, S., Quintaneiro, L. M., Ferreira, A. P. A., Krause, M., McMahon, H. T., & Boucrot, E. (2018). FBP17 and CIP4 recruit SHIP2 and lamellipodin to prime the plasma membrane for fast endophilin-mediated endocytosis. *Nature Cell Biology*, *20*(9), 1023-1031. https://doi.org/10.1038/s41556-018-0146-8

Chen, Y., Deng, L., Maeno-Hikichi, Y., Lai, M., Chang, S., Chen, G., & Zhang, J. (2003). Formation of an endophilin-Ca2+ channel complex is critical for clathrin-mediated synaptic vesicle endocytosis. *Cell*, *115*(1), 37-48. https://doi.org/10.1016/s0092-8674(03)00726-8

Cheung, N., So, C. W., Yam, J. W. P., So, C. K. C., Poon, R. Y. C., Jin, D.-Y., & Chan, L. C. (2004). Subcellular localization of EEN/endophilin A2, a fusion partner gene in leukaemia. *The Biochemical Journal*, *383*(Pt 1), 27-35. https://doi.org/10.1042/BJ20040041

Cocucci, E., Aguet, F., Boulant, S., & Kirchhausen, T. (2012). The first five seconds in the life of a clathrin-coated pit. *Cell*, *150*(3), 495-507. https://doi.org/10.1016/j.cell.2012.05.047

Cuddeback, S. M., Yamaguchi, H., Komatsu, K., Miyashita, T., Yamada, M., Wu, C., Singh, S., & Wang, H. G. (2001). Molecular cloning and characterization of Bif-1. A novel Src homology 3 domain-containing protein that associates with Bax. *The Journal of Biological Chemistry*, *276*(23), 20559-20565. https://doi.org/10.1074/jbc.M101527200

Day, C. A., Baetz, N. W., Copeland, C. A., Kraft, L. J., Han, B., Tiwari, A., Drake, K. R., Luca, H. D., Chinnapen, D. J.-F., Davidson, M. W., Holmes, R. K., Jobling, M. G., Schroer, T. A., Lencer, W. I., & Kenworthy, A. K. (2015). Microtubule Motors Power Plasma Membrane Tubulation in Clathrin-Independent Endocytosis. *Traffic*, *16*(6), 572-590. https://doi.org/10.1111/tra.12269

Doherty, G. J., & McMahon, H. T. (2009). Mechanisms of endocytosis. *Annual Review of Biochemistry*, *78*, 857-902. https://doi.org/10.1146/annurev.biochem.78.081307.110540

Dutta, D., & Donaldson, J. G. (2015). Sorting of Clathrin-Independent Cargo Proteins Depends on Rab35 Delivered by Clathrin-Mediated Endocytosis. *Traffic (Copenhagen, Denmark)*, *16*(9), 994-1009. https://doi.org/10.1111/tra.12302

F, M., Ym, P., J, L., N, B., T, G., N, M., P, B., Arn, T., Sc, P., Rd, F., & R, L. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, *47*(W1), W636-W641. https://doi.org/10.1093/nar/gkz268

Farsad, K., Ringstad, N., Takei, K., Floyd, S. R., Rose, K., & De Camilli, P. (2001). Generation of high curvature membranes mediated by direct endophilin bilayer interactions. *The Journal of Cell Biology*, *155*(2), 193-200. https://doi.org/10.1083/jcb.200107075

Fernández, M. M., Ferragut, F., Cárdenas Delgado, V. M., Bracalente, C., Bravo, A. I., Cagnoni, A. J., Nuñez, M., Morosi, L. G., Quinta, H. R., Espelt, M. V., Troncoso, M. F., Wolfenstein-Todel, C., Mariño, K. V., Malchiodi, E. L., Rabinovich, G. A., & Elola, M. T. (2016). Glycosylation-dependent binding of galectin-8 to activated leukocyte cell adhesion molecule (ALCAM/CD166) promotes its surface segregation on breast cancer cells. *Biochimica Et Biophysica Acta*, *1860*(10), 2255-2268. https://doi.org/10.1016/j.bbagen.2016.04.019

Ford, M. G. J., Mills, I. G., Peter, B. J., Vallis, Y., Praefcke, G. J. K., Evans, P. R., & McMahon, H. T. (2002). Curvature of clathrin-coated pits driven by epsin. *Nature*, *419*(6905), 361-366. https://doi.org/10.1038/nature01020

Frost, A., De Camilli, P., & Unger, V. M. (2007). F-BAR proteins join the BAR family fold. *Structure (London, England: 1993)*, *15*(7), 751-753. https://doi.org/10.1016/j.str.2007.06.006

Frost, A., Unger, V. M., & De Camilli, P. (2009). The BAR domain superfamily : Membrane-molding macromolecules. *Cell*, *137*(2), 191-196. https://doi.org/10.1016/j.cell.2009.04.010

Gad, H., Ringstad, N., Löw, P., Kjaerulff, O., Gustafsson, J., Wenk, M., Di Paolo, G., Nemoto, Y., Crum, J., Ellisman, M. H., De Camilli, P., Shupliakov, O., & Brodin, L. (2000). Fission and Uncoating of Synaptic Clathrin-Coated Vesicles Are Perturbed by Disruption of Interactions with the SH3 Domain of Endophilin. *Neuron*, *27*(2), 301-312. https://doi.org/10.1016/S0896-6273(00)00038-6

Gallop, J. L., Jao, C. C., Kent, H. M., Butler, P. J. G., Evans, P. R., Langen, R., & McMahon, H. T. (2006). Mechanism of endophilin N-BAR domain-mediated membrane curvature. *The EMBO Journal*, *25*(12), 2898-2910. https://doi.org/10.1038/sj.emboj.7601174

Giachino, C., Lantelme, E., Lanzetti, L., Saccone, S., Bella Valle, G., & Migone, N. (1997). A novel SH3-containing human gene family preferentially expressed in the central nervous system. *Genomics*, *41*(3), 427-434. https://doi.org/10.1006/geno.1997.4645

Hughes, A. C., Errington, R., Fricker-Gates, R., & Jones, L. (2004). Endophilin A3 forms filamentous structures that colocalise with microtubules but not with actin filaments. *Molecular Brain Research*, *128*(2), 182-192. https://doi.org/10.1016/j.molbrainres.2004.06.016

Itoh, T., Erdmann, K. S., Roux, A., Habermann, B., Werner, H., & De Camilli, P. (2005). Dynamin and the actin cytoskeleton cooperatively regulate plasma membrane invagination by BAR and F-BAR proteins. *Developmental Cell*, *9*(6), 791-804. https://doi.org/10.1016/j.devcel.2005.11.005

Johannes, L., Parton, R. G., Bassereau, P., & Mayor, S. (2015). Building endocytic pits without clathrin. *Nature Reviews. Molecular Cell Biology*, *16*(5), 311-321. https://doi.org/10.1038/nrm3968

Kaksonen, M., & Roux, A. (2018). Mechanisms of clathrin-mediated endocytosis. *Nature Reviews. Molecular Cell Biology*, *19*(5), 313-326. https://doi.org/10.1038/nrm.2017.132

Kelly, B. T., Graham, S. C., Liska, N., Dannhauser, P. N., Höning, S., Ungewickell, E. J., & Owen, D. J. (2014). Clathrin adaptors. AP2 controls clathrin polymerization with a membrane-activated switch. *Science (New York, N.Y.)*, *345*(6195), 459-463. https://doi.org/10.1126/science.1254836

Kirchhausen, T., & Harrison, S. C. (1981). Protein organization in clathrin trimers. *Cell*, *23*(3), 755-761. https://doi.org/10.1016/0092-8674(81)90439-6

Kjaerulff, O., Brodin, L., & Jung, A. (2011). The structure and function of endophilin proteins. *Cell Biochemistry and Biophysics*, *60*(3), 137-154. https://doi.org/10.1007/s12013-010-9137-5

Lakshminarayan, R., Wunder, C., Becken, U., Howes, M. T., Benzing, C., Arumugam, S., Sales, S., Ariotti, N., Chambon, V., Lamaze, C., Loew, D., Shevchenko, A., Gaus, K., Parton, R. G., & Johannes, L. (2014). Galectin-3 drives glycosphingolipid-dependent

biogenesis of clathrin-independent carriers. *Nature Cell Biology*, *16*(6), 592-603. https://doi.org/10.1038/ncb2970

Lee, S. H., Kerff, F., Chereau, D., Ferron, F., Klug, A., & Dominguez, R. (2007). Structural Basis for the Actin-Binding Function of Missing-in-Metastasis. *Structure*, *15*(2), 145-155. https://doi.org/10.1016/j.str.2006.12.005

Loll, P. J., Swain, E., Chen, Y., Turner, B. T., & Zhang, J. (2008). Structure of the SH3 domain of rat endophilin A2. *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications*, *64*(Pt 4), 243-246. https://doi.org/10.1107/S1744309108007574

Meinecke, M., Boucrot, E., Camdere, G., Hon, W.-C., Mittal, R., & McMahon, H. T. (2013). Cooperative Recruitment of Dynamin and BIN/Amphiphysin/Rvs (BAR) Domain-containing Proteins Leads to GTP-dependent Membrane Scission* [S]♦. *Journal of Biological Chemistry*, *288*(9), 6651-6661. https://doi.org/10.1074/jbc.M112.444869

Merrifield, C. J., Qualmann, B., Kessels, M. M., & Almers, W. (2004). Neural Wiskott Aldrich Syndrome Protein (N-WASP) and the Arp2/3 complex are recruited to sites of clathrin-mediated endocytosis in cultured fibroblasts. *European Journal of Cell Biology*, *83*(1), 13-18. https://doi.org/10.1078/0171-9335-00356

Miller, S. E., Mathiasen, S., Bright, N. A., Pierre, F., Kelly, B. T., Kladt, N., Schauss, A., Merrifield, C. J., Stamou, D., Höning, S., & Owen, D. J. (2015). CALM Regulates Clathrin-Coated Vesicle Size and Maturation by Directly Sensing and Driving Membrane Curvature. *Developmental Cell*, *33*(2), 163-175. https://doi.org/10.1016/j.devcel.2015.03.002

Mim, C., Cui, H., Gawronski-Salerno, J. A., Frost, A., Lyman, E., Voth, G. A., & Unger, V. M. (2012). Structural basis of membrane bending by the N-BAR protein endophilin. *Cell*, *149*(1), 137-145. https://doi.org/10.1016/j.cell.2012.01.048

Mim, C., & Unger, V. M. (2012). Membrane curvature and its generation by BAR proteins. *Trends in Biochemical Sciences*, *37*(12), 526-533. https://doi.org/10.1016/j.tibs.2012.09.001

Perera, R. M., Zoncu, R., Lucast, L., De Camilli, P., & Toomre, D. (2006). Two synaptojanin 1 isoforms are recruited to clathrin-coated pits at different stages. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(51), 19332-19337. https://doi.org/10.1073/pnas.0609795104

Peter, B. J., Kent, H. M., Mills, I. G., Vallis, Y., Butler, P. J. G., Evans, P. R., & McMahon, H. T. (2004). BAR domains as sensors of membrane curvature : The amphiphysin BAR structure. *Science (New York, N.Y.)*, *303*(5657), 495-499. https://doi.org/10.1126/science.1092586

Pierrat, B., Simonen, M., Cueto, M., Mestan, J., Ferrigno, P., & Heim, J. (2001). SH3GLB, a new endophilin-related protein family featuring an SH3 domain. *Genomics*, *71*(2), 222-234. https://doi.org/10.1006/geno.2000.6378

Posor, Y., Eichhorn-Gruenig, M., Puchkov, D., Schöneberg, J., Ullrich, A., Lampe, A., Müller, R., Zarbakhsh, S., Gulluni, F., Hirsch, E., Krauss, M., Schultz, C., Schmoranzer, J., Noé, F., & Haucke, V. (2013). Spatiotemporal control of endocytosis by phosphatidylinositol-3,4-bisphosphate. *Nature*, *499*(7457), 233-237. https://doi.org/10.1038/nature12360

Renard, H.-F., Simunovic, M., Lemière, J., Boucrot, E., Garcia-Castillo, M. D., Arumugam, S., Chambon, V., Lamaze, C., Wunder, C., Kenworthy, A. K., Schmidt, A. A., McMahon, H. T., Sykes, C., Bassereau, P., & Johannes, L. (2015). Endophilin-A2 functions in membrane scission in clathrin-independent endocytosis. *Nature*, *517*(7535), 493-496. https://doi.org/10.1038/nature14064

Renard, H.-F., Tyckaert, F., Lo Giudice, C., Hirsch, T., Valades-Cruz, C. A., Lemaigre, C., Shafaq-Zadah, M., Wunder, C., Wattiez, R., Johannes, L., van der Bruggen, P., Alsteens, D., & Morsomme, P. (2020). Endophilin-A3 and Galectin-8 control the clathrin-independent endocytosis of CD166. *Nature Communications*, *11*(1), 1457. https://doi.org/10.1038/s41467-020-15303-y

Ringstad, N., Nemoto, Y., & De Camilli, P. (1997). The SH3p4/Sh3p8/SH3p13 protein family : Binding partners for synaptojanin and dynamin via a Grb2-like Src homology 3 domain. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(16), 8569-8574.

Royle, S. J. (2006). The cellular functions of clathrin. *Cellular and molecular life sciences : CMLS*, *63*(16), 1823-1832. https://doi.org/10.1007/s00018-005-5587-0

Sakamuro, D., Elliott, K. J., Wechsler-Reya, R., & Prendergast, G. C. (1996). BIN1 is a novel MYC–interacting protein with features of a tumour suppressor. *Nature Genetics*, *14*(1), 69-77. https://doi.org/10.1038/ng0996-69

Simunovic, M., Manneville, J.-B., Renard, H.-F., Evergren, E., Raghunathan, K., Bhatia, D., Kenworthy, A. K., Voth, G. A., Prost, J., McMahon, H. T., Johannes, L., Bassereau, P., & Callan-Jones, A. (2017). Friction Mediates Scission of Tubular Membranes Scaffolded by BAR Proteins. *Cell*, *170*(1), 172-184.e11. https://doi.org/10.1016/j.cell.2017.05.047

Sirotkin, V., Beltzner, C. C., Marchand, J.-B., & Pollard, T. D. (2005). Interactions of WASp, myosin-I, and verprolin with Arp2/3 complex during actin patch assembly in fission yeast. *The Journal of Cell Biology*, *170*(4), 637-648. https://doi.org/10.1083/jcb.200502053

Suetsugu, S. (2016). Higher-order assemblies of BAR domain proteins for shaping membranes. *Microscopy (Oxford, England)*, *65*(3), 201-210. https://doi.org/10.1093/jmicro/dfw002

Suetsugu, S., & Gautreau, A. (2012). Synergistic BAR–NPF interactions in actin-driven membrane remodeling. *Trends in Cell Biology*, *22*(3), 141-150. https://doi.org/10.1016/j.tcb.2012.01.001

Sugiura, H., Iwata, K., Matsuoka, M., Hayashi, H., Takemiya, T., Yasuda, S., Ichikawa, M., Yamauchi, T., Mehlen, P., Haga, T., & Yamagata, K. (2004). Inhibitory Role of Endophilin 3 in Receptor-mediated Endocytosis *. *Journal of Biological Chemistry*, *279*(22), 23343-23348. https://doi.org/10.1074/jbc.M312607200

Taylor, M. J., Perrais, D., & Merrifield, C. J. (2011). A High Precision Survey of the Molecular Dynamics of Mammalian Clathrin-Mediated Endocytosis. *PLOS Biology*, *9*(3), e1000604. https://doi.org/10.1371/journal.pbio.1000604

Tian, Q., Zhang, J.-F., Fan, J., Song, Z., & Chen, Y. (2012). Endophilin isoforms have distinct characteristics in interactions with N-type Ca2+ channels and dynamin I. *Neuroscience Bulletin*, *28*(5), 483-492. https://doi.org/10.1007/s12264-012-1257-z

Traub, L. M. (2009). Tickets to ride : Selecting cargo for clathrin-regulated internalization. *Nature Reviews. Molecular Cell Biology*, *10*(9), 583-596. https://doi.org/10.1038/nrm2751

van Kempen, L. C., Nelissen, J. M., Degen, W. G., Torensma, R., Weidle, U. H., Bloemers, H. P., Figdor, C. G., & Swart, G. W. (2001). Molecular basis for the homophilic activated leukocyte cell adhesion molecule (ALCAM)-ALCAM interaction. *The Journal of Biological Chemistry*, *276*(28), 25783-25790. https://doi.org/10.1074/jbc.M011272200

van Weering, J. R. T., Verkade, P., & Cullen, P. J. (2010). SNX-BAR proteins in phosphoinositide-mediated, tubular-based endosomal sorting. *Seminars in Cell & Developmental Biology*, *21*(4), 371-380. https://doi.org/10.1016/j.semcdb.2009.11.009

Vehlow, A., Soong, D., Vizcay-Barrena, G., Bodo, C., Law, A.-L., Perera, U., & Krause, M. (2013). Endophilin, Lamellipodin, and Mena cooperate to regulate F-actin-dependent EGF-receptor endocytosis. *The EMBO Journal*, *32*(20), 2722-2734. https://doi.org/10.1038/emboj.2013.212

Verstreken, P., Kjaerulff, O., Lloyd, T. E., Atkinson, R., Zhou, Y., Meinertzhagen, I. A., & Bellen, H. J. (2002). Endophilin mutations block clathrin-mediated endocytosis but not neurotransmitter release. *Cell*, *109*(1), 101-112. https://doi.org/10.1016/s0092-8674(02)00688-8

von Lersner, A., Droesen, L., & Zijlstra, A. (2019). Modulation of cell adhesion and migration through regulation of the immunoglobulin superfamily member ALCAM/CD166. *Clinical & experimental metastasis*, *36*(2), 87-95. https://doi.org/10.1007/s10585-019-09957-2

Wang, D. S., & Shaw, G. (1995). The association of the C-terminal region of beta I sigma II spectrin to brain membranes is mediated by a PH domain, does not require membrane proteins, and coincides with a inositol-1,4,5 triphosphate binding site. *Biochemical and Biophysical Research Communications*, *217*(2), 608-615. https://doi.org/10.1006/bbrc.1995.2818

Weissenhorn, W. (2005). Crystal structure of the endophilin-A1 BAR domain. *Journal of Molecular Biology*, *351*(3), 653-661. https://doi.org/10.1016/j.jmb.2005.06.013

Yarar, D., Surka, M. C., Leonard, M. C., & Schmid, S. L. (2008). SNX9 Activities are Regulated by Multiple Phosphoinositides Through both PX and BAR Domains. *Traffic*, *9*(1), 133-146. https://doi.org/10.1111/j.1600-0854.2007.00675.x

Zimmerman, A. W., Joosten, B., Torensma, R., Parnes, J. R., van Leeuwen, F. N., & Figdor, C. G. (2006). Long-term engagement of CD6 and ALCAM is essential for T-cell proliferation induced by dendritic cells. *Blood*, *107*(8), 3212-3220. https://doi.org/10.1182/blood-2005-09-3881

Annexes

Annex 1. Genetic constructs

1.1 Obtention and cloning of endoA3-ΔlinkerSH3, BAR A2-SH3 A3 and BAR A3-SH3 A2 by restriction ligation

1.1.1 Fragment obtention

- *EndoA3-ΔlinkerSH3*

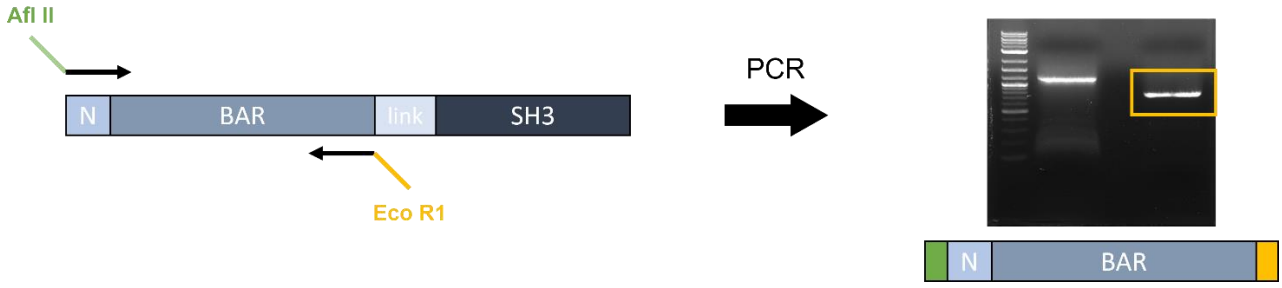


Figure S1. Obtention of endoA3-ΔlinkerSH3 sequence. EndoA3-ΔlinkerSH3 sequence was obtained by PCR amplification of endophilin A3 N-BAR domain with primers containing AfI II and Eco R1 restriction sites.

- *BAR A2-SH3 A2*

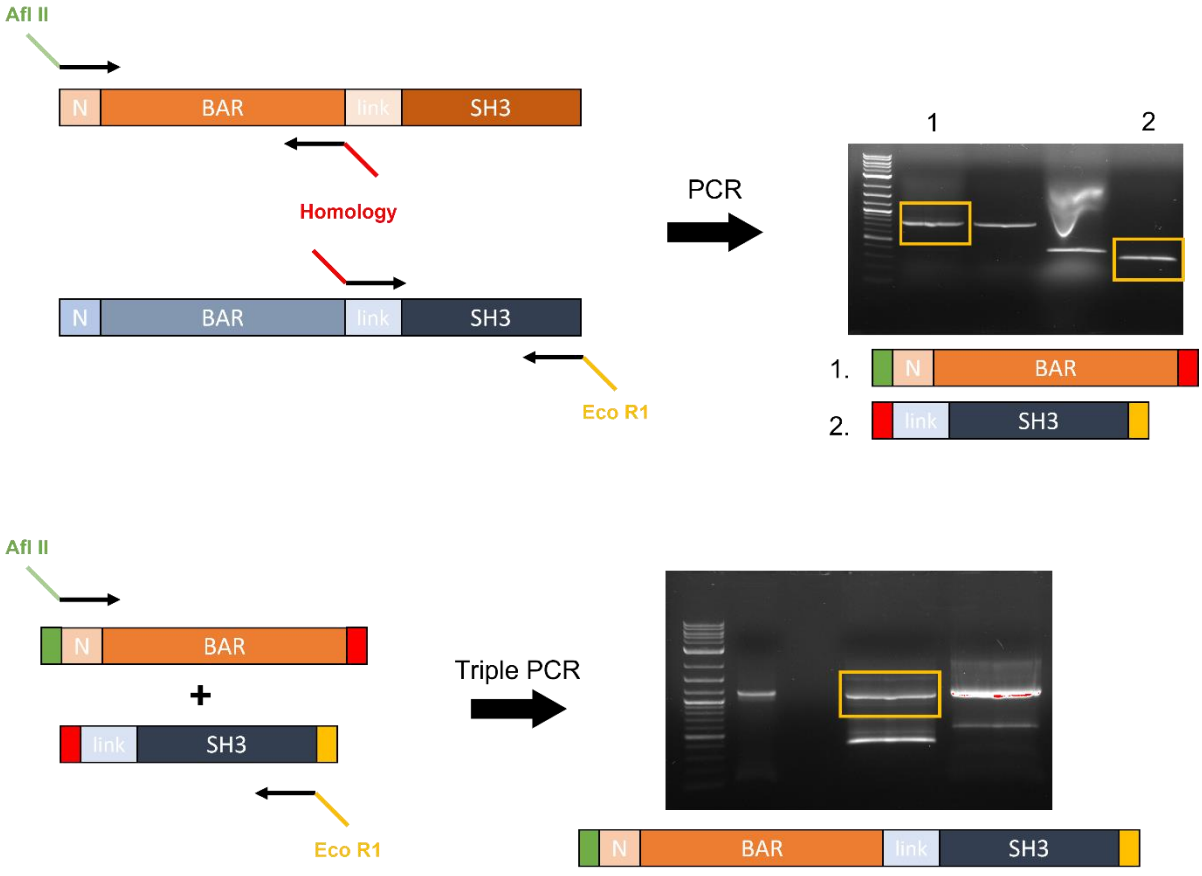


Figure S2. Obtention of BAR A2-SH3 A3 sequence. BAR A2 and linker-SH3 A3 sequences were obtained by PCR amplification of endoA2 N-BAR domain and endoA3 linker-SH3 domains respectively. Extremity primers contained AfI II and Eco R1 restriction sites while internal primers contained homology sequence for the other fragment. The two sequences were then assembled in a triple PCR .

- *BAR A3-SH3 A3*

The sequence for BAR A3-SH3 A2 was obtained similarly as BAR A2-SH3 A2 by swapping the sequence of endophilin A2 and A3 and adapting the primers (see section I.2.3 of material and methods)

1.1.2 Plasmid and inserts digestion by Afl II and Eco R1

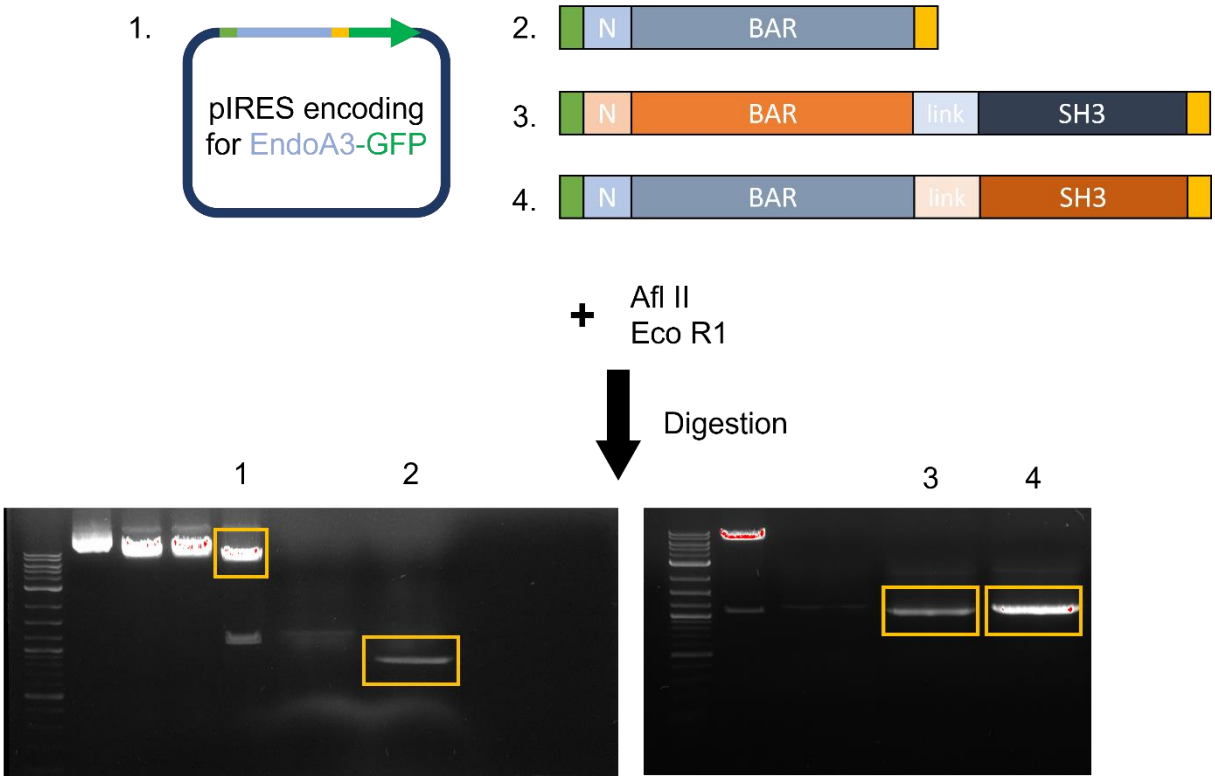


Figure S3. Digestion of pIRES encoding for endoA3-GFP, endoA3-ΔlinkerSH3, BAR A2-SH3 A3 and BAR A3-SH3 A2 by Afl II and Eco R1. Plasmid and mutant sequences were digested using Afl II and Eco R1 to create recombinant ends.

1.1.3 Ligation and verification of positive clones by restriction

After plasmid and inserts were digested, they were mixed together with a 5x molar excess of insert and T4 DNA ligase for ligation. Mixes were then transformed into *E. coli* (see section I.1.1 of material and methods) and positive clones for the two chimeric constructs BAR A2-SH3 A3 and BAR A3-SH3 A2 were verified by restriction with Pst I and Sma I respectively (**Figure S4**).

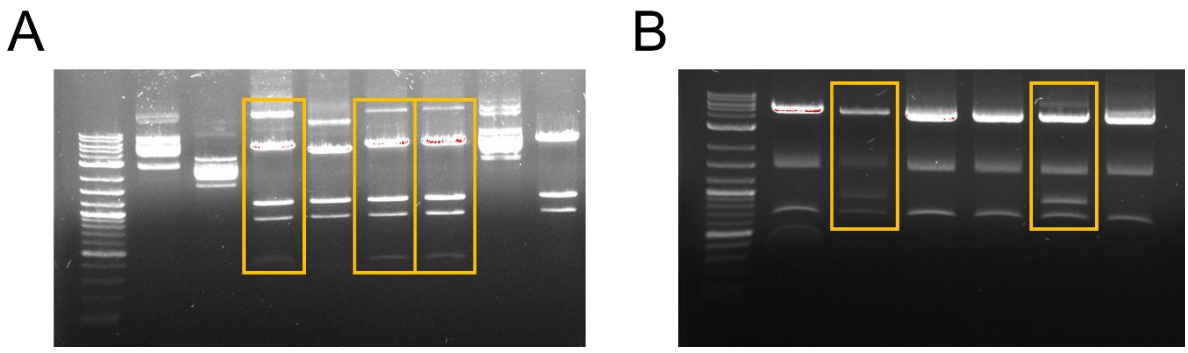


Figure S4. Verification by restriction of BAR A2-SH3 A3 and BAR A3-SH3 A2 constructs with Pst I and Sma I digestion respectively. The clones presenting the predicted pattern are boxed in yellow.

Positive clones were finally sent for sequencing (see section 1.3.3 of material and methods) and all of them matched the expected sequence.

1.2 Obtention and cloning of endoA2 and BAR A2-linker A3-SH3 A2 by the Gibson method

1.2.1 Fragments obtention

- *EndoA2*

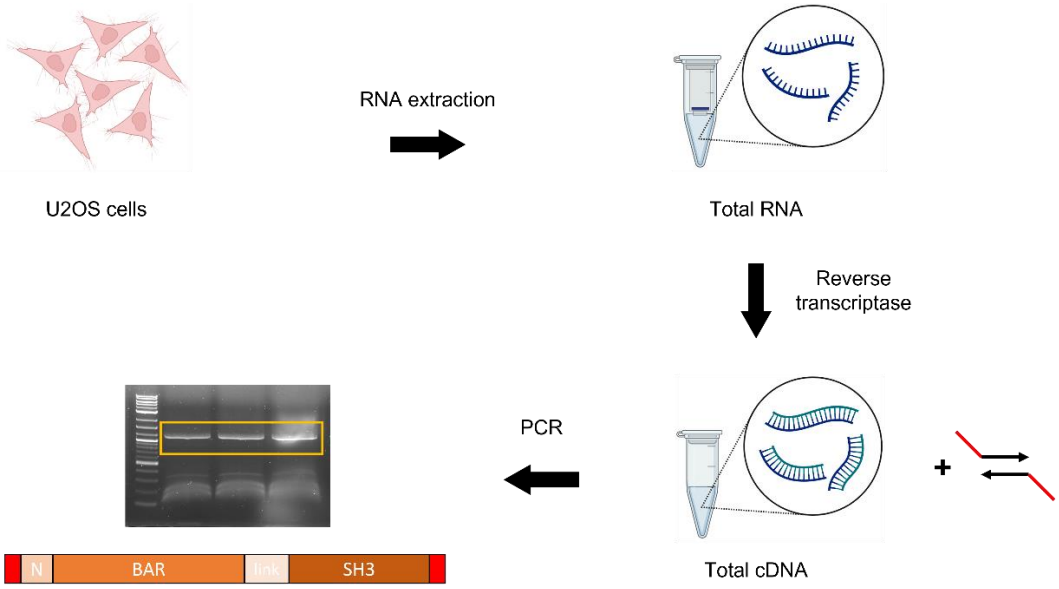


Figure S5. Obtention of endoA2 sequence. Total RNA was extracted from U2OS cells. cDNA was synthesized by reverse transcriptase. EndoA2 sequence was amplified using primers containing an homology sequence for the insertion site of the plasmid.

BAR A2-linker A3-SH3 A2

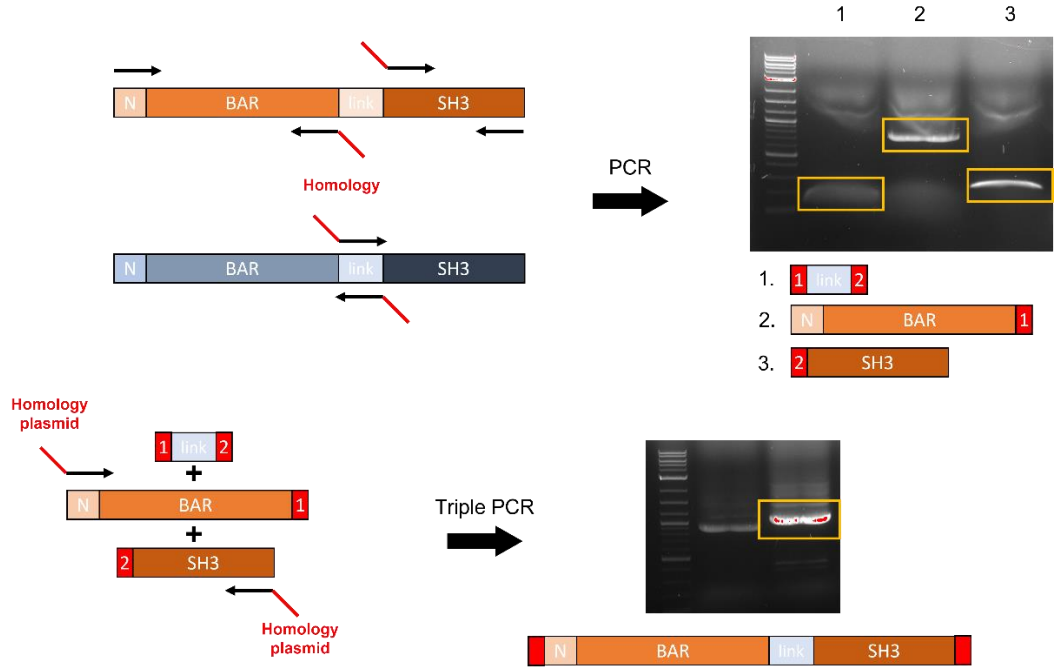


Figure S6. Obtention of BAR A2-linker A3-SH3 A2 sequence. BAR A2, linker A3 and SH3 A2 sequences were obtained by PCR amplification using full length endoA2 or A3 sequences as template. Internal primers contained homology sequence for the adjacent fragment in the final protein. Both sequences were then assembled in a triple PCR with extremity primers containing homology sequences for the insertion site of the plasmid.

- *pIRES*

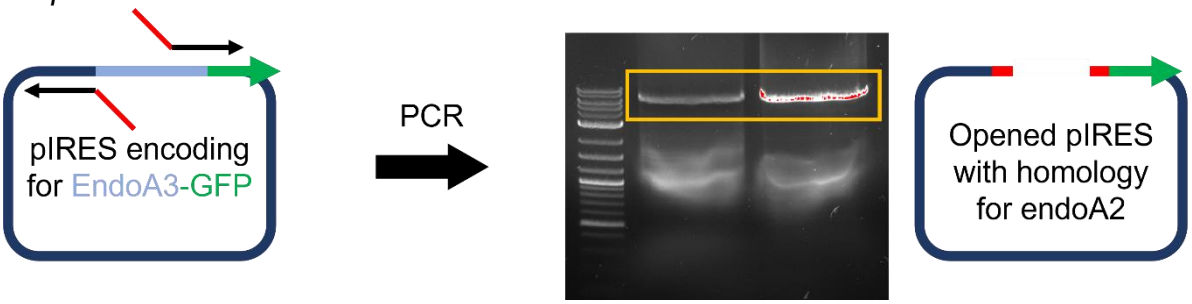


Figure S7. Obtention of pIRES for endoA2 and BAR A2-linker A3-SH3 A2 cloning by the Gibson method. The plasmid pIRES was amplified by PCR using as template pIRES encoding for endoA3. The primers contained homology sequences for endoA2 extremities.

1.2.2 Plasmid and inserts assembly by the Gibson method and verification of positive clones by restriction

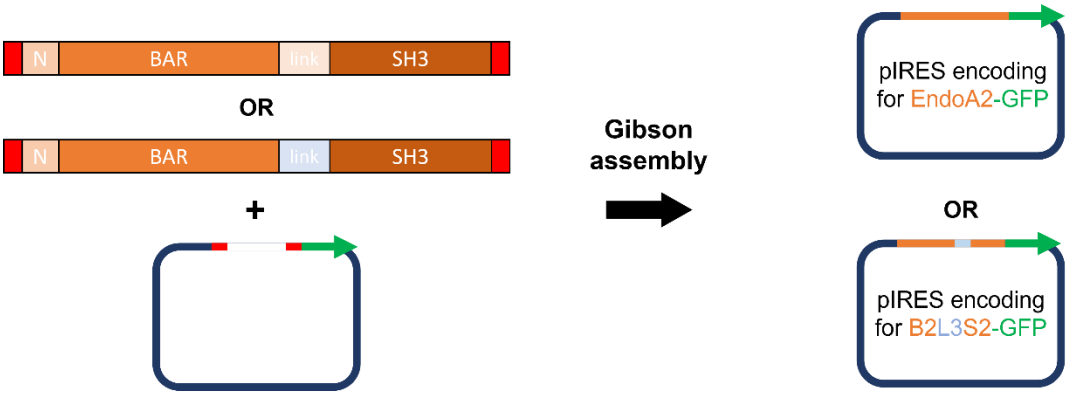


Figure S8. Cloning of endoA2 and BAR A2-linker A3-SH3 A2 by the Gibson method.

After 1h at 50°C, 5µL of each Gibson mix were transformed into *E. coli* (see section I.1.1 of material and methods). 10 positive clones were picked randomly and digested by Pst I for verification.

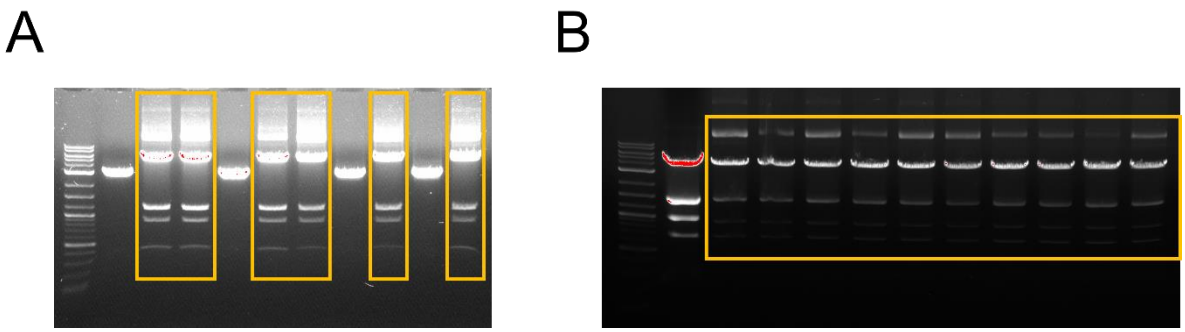


Figure S9. Verification by restriction of endoA2 and BAR A2-linker A3-SH3 A2 constructs with Pst I digestion. The clones presenting the predicted pattern are boxed in yellow.

3 positive clones per construct were finally sent for sequencing (see section 1.3.3 of material and methods) and all of them matched the expected sequence.

Annex 2. Colocalization assays

2.1 Endophilin A3 supplementary images

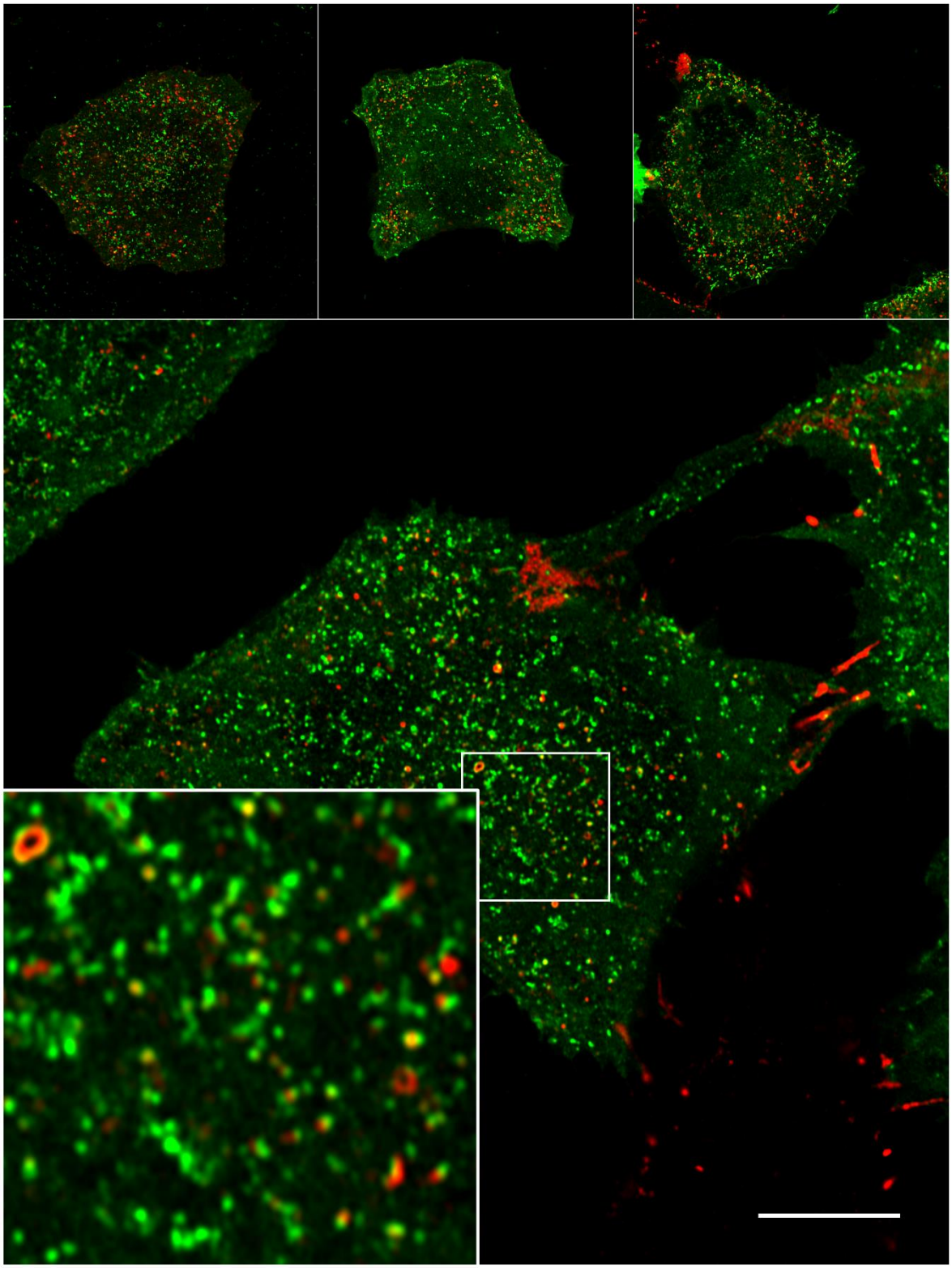


Figure S1. Supplementary image of colocalization assays between CD166 (red) and endoA3 (green). Scale bar = 10 µm.

2.2 Endophilin A2 supplementary images

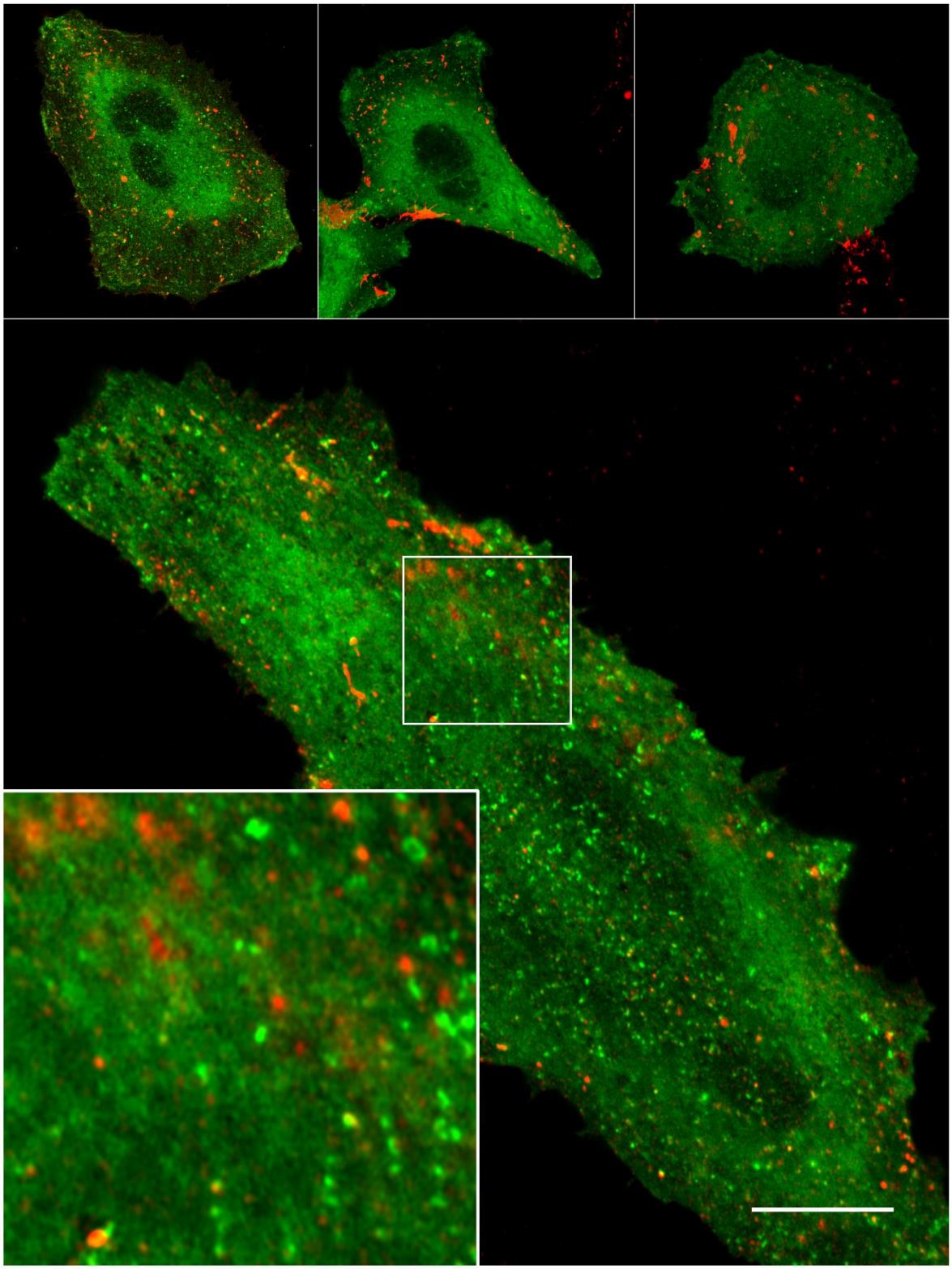


Figure S2. Supplementary image of colocalization assays between CD166 (red) and endoA2 (green). Scale bar = 10 µm.

2.3 Endo A3-ΔlinkerSH3 supplementary images

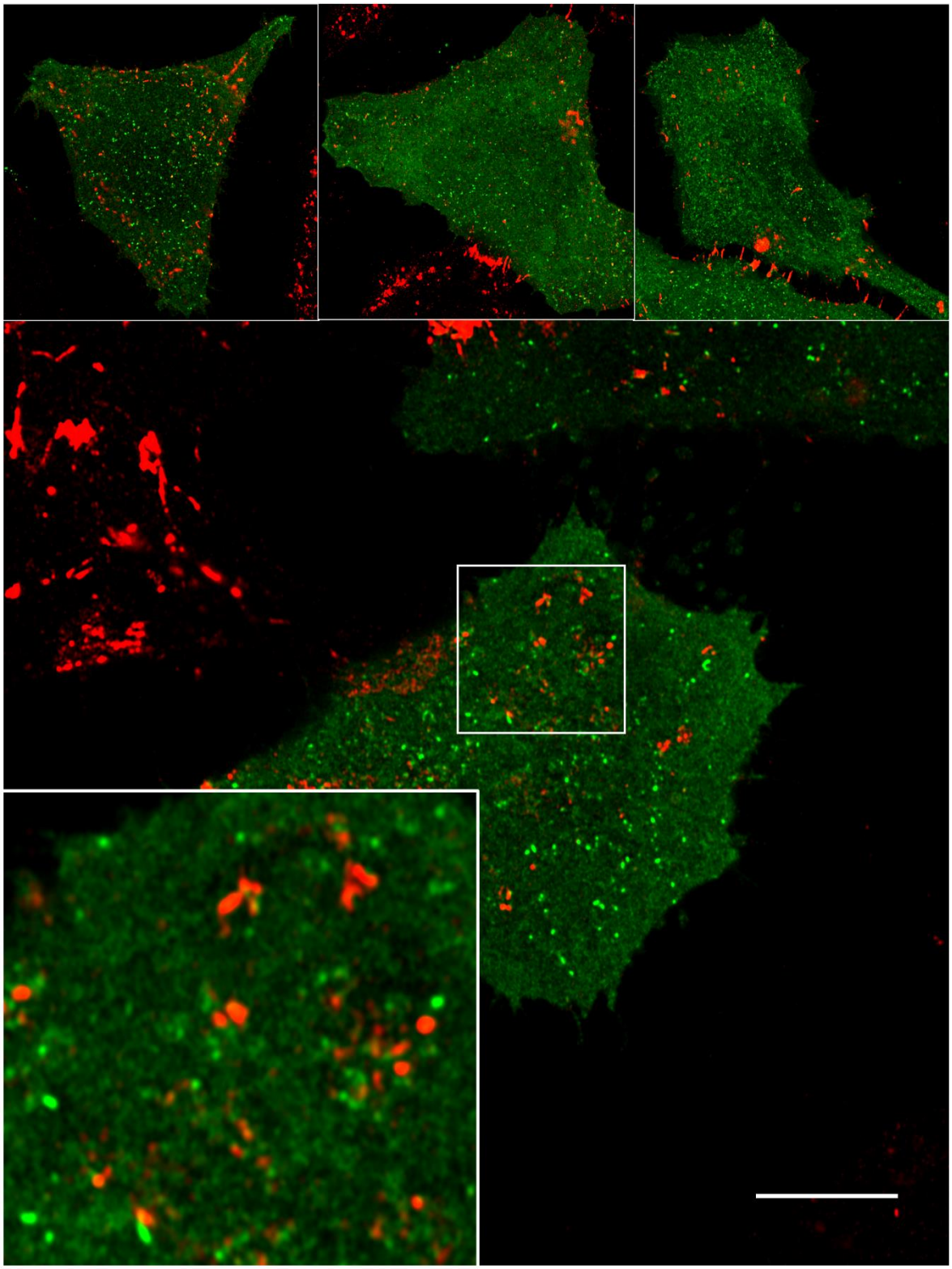


Figure S3. Supplementary image of colocalization assays between CD166 (red) and endoA3-ΔlinkerSH3 (green). Scale bar = 10 µm.

2.4 BAR A2-SH3 A3 supplementary images

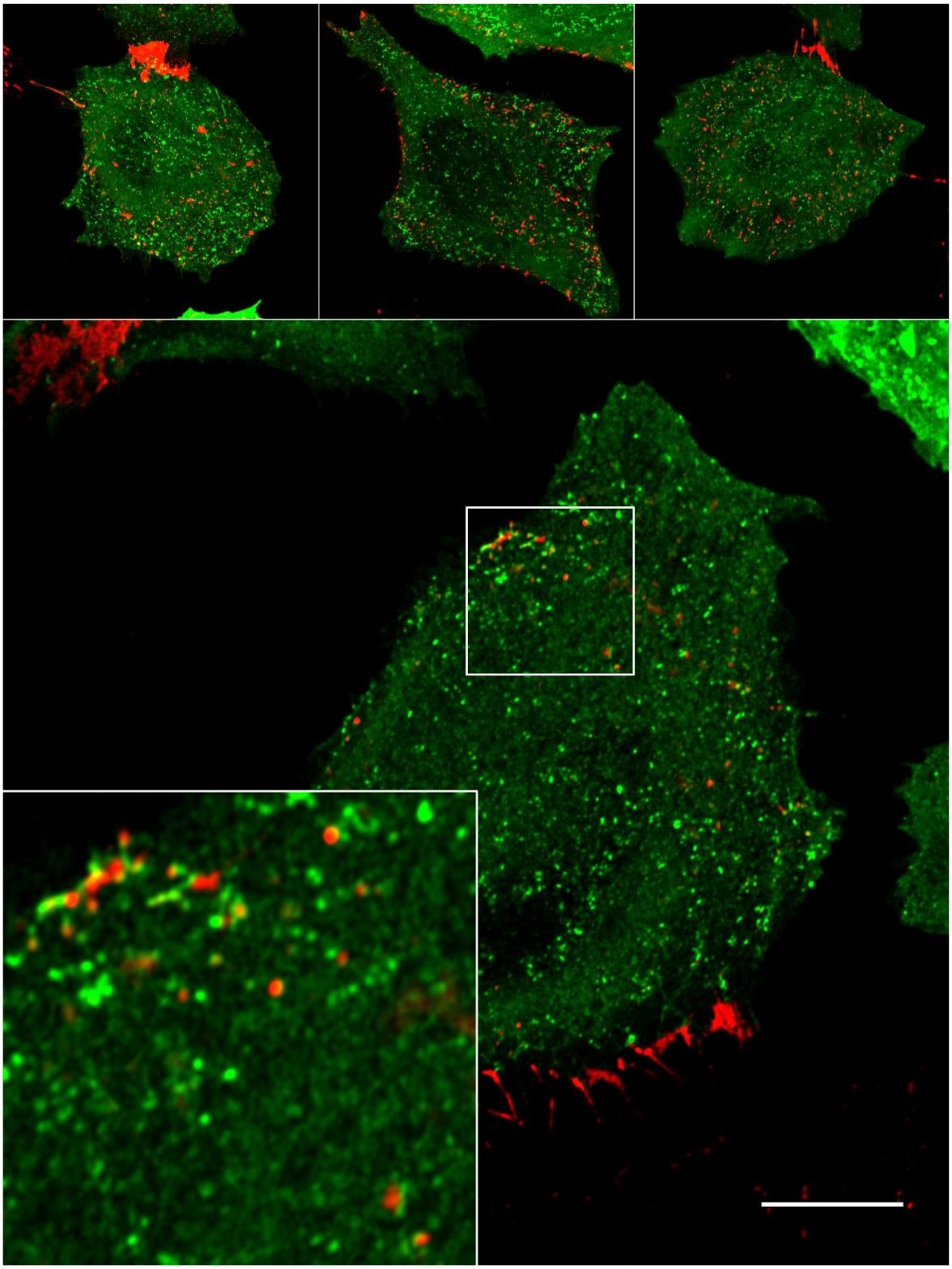


Figure S4. Supplementary image of colocalization assays between CD166 (red) and BAR A2-SH3 A3 (green). Scale bar = 10 µm.

2.5 BAR A3-SH3 A2 supplementary images

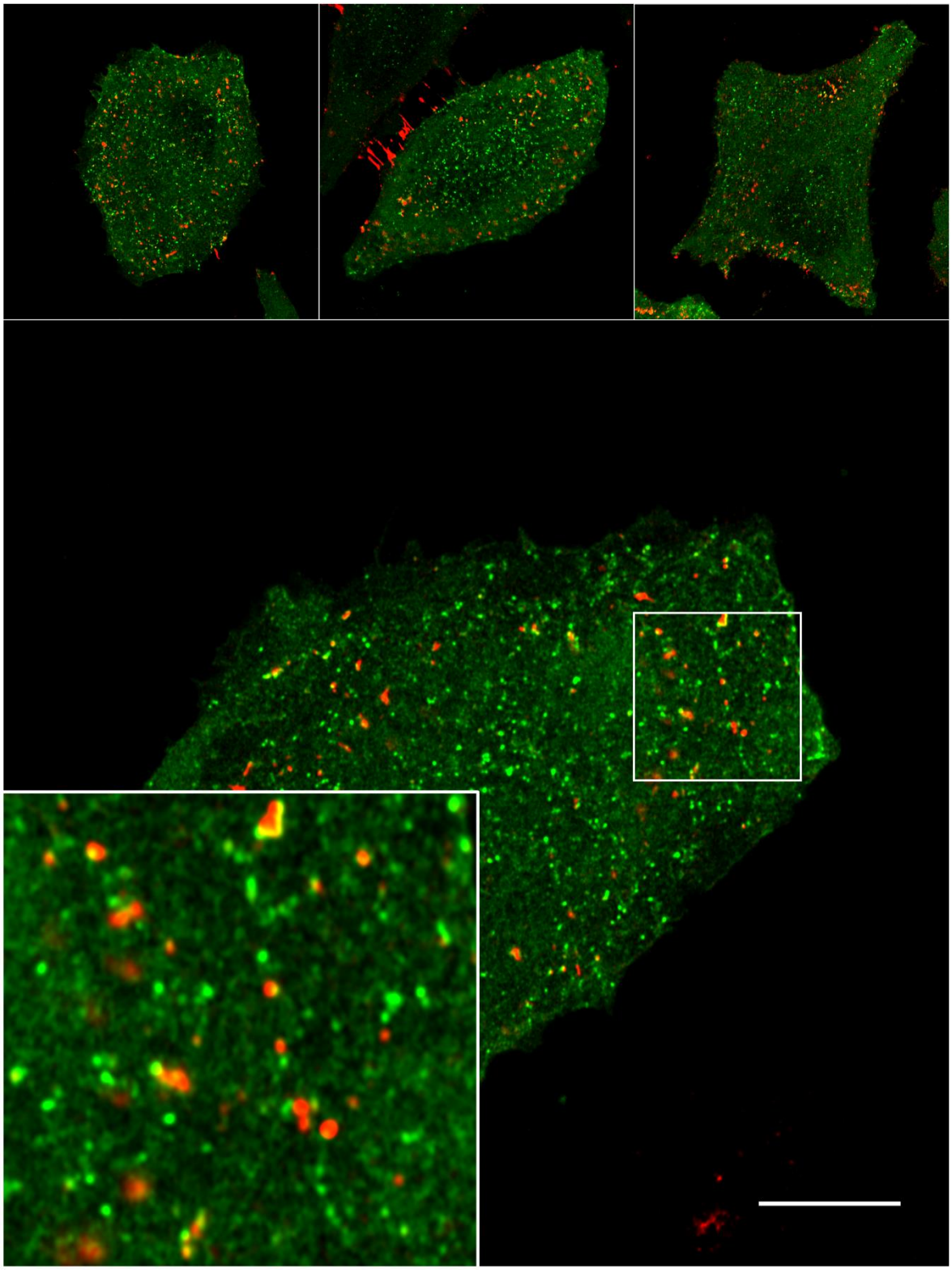


Figure S5. Supplementary image of colocalization assays between CD166 (red) and BAR A3-SH3 A2 (green). Scale bar = 10 µm.

Annex 3. Rescue assays

3.1 siCtrl and siEndoA3 supplementary images

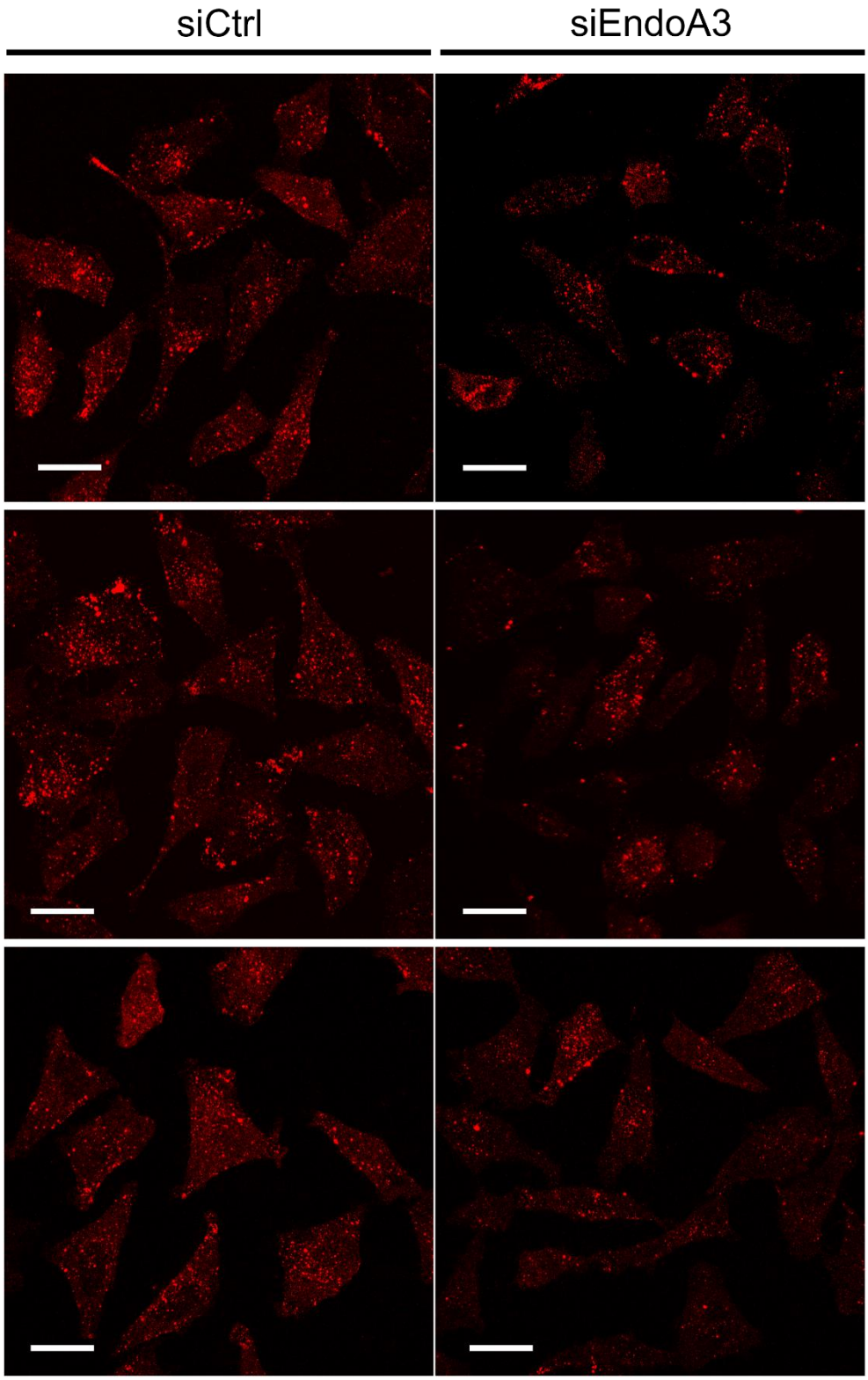


Figure S7. Supplementary images of rescue assays in control (siCtrl) and endoA3 depleted (siEndoA3) cells. Cells knocked-down for endophilin A3 harbor lower level of internalized CD166 (red signal), this effect is however variable between the cells. Scale bars = 25 µm.

3.2 siEndoA3 + endoA3 supplementary images

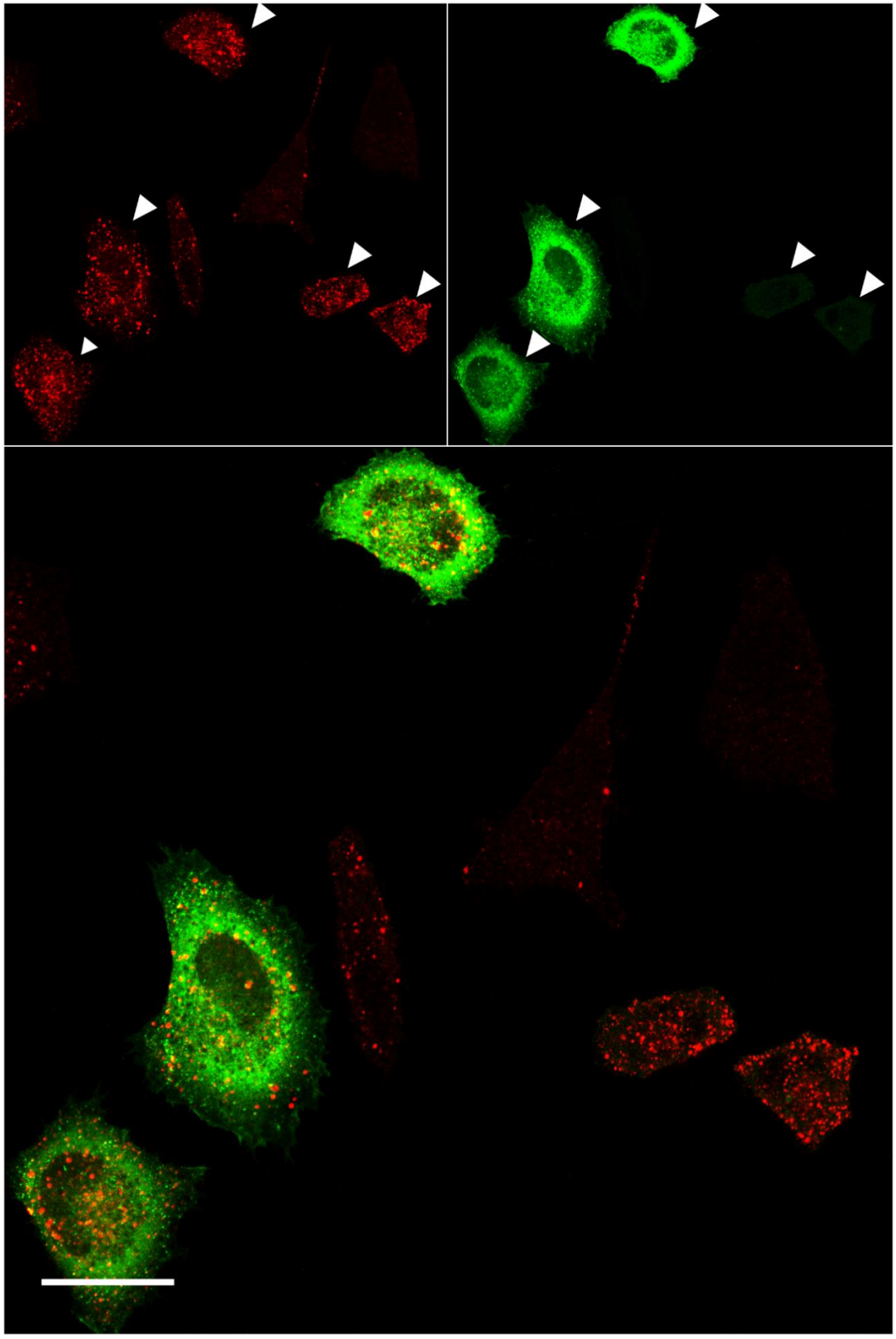


Figure S8. Supplementary images of rescue assays in endoA3 depleted cells transfected with a plasmid encoding for endoA3-GFP. Red signal: CD166-A546. **Green signal:** GFP-tagged endophilin A3. White triangles indicate cells effectively transfected with the construct. Scale bar = 25 µm.

3.3 siEndoA3 + endoA2 supplementary images

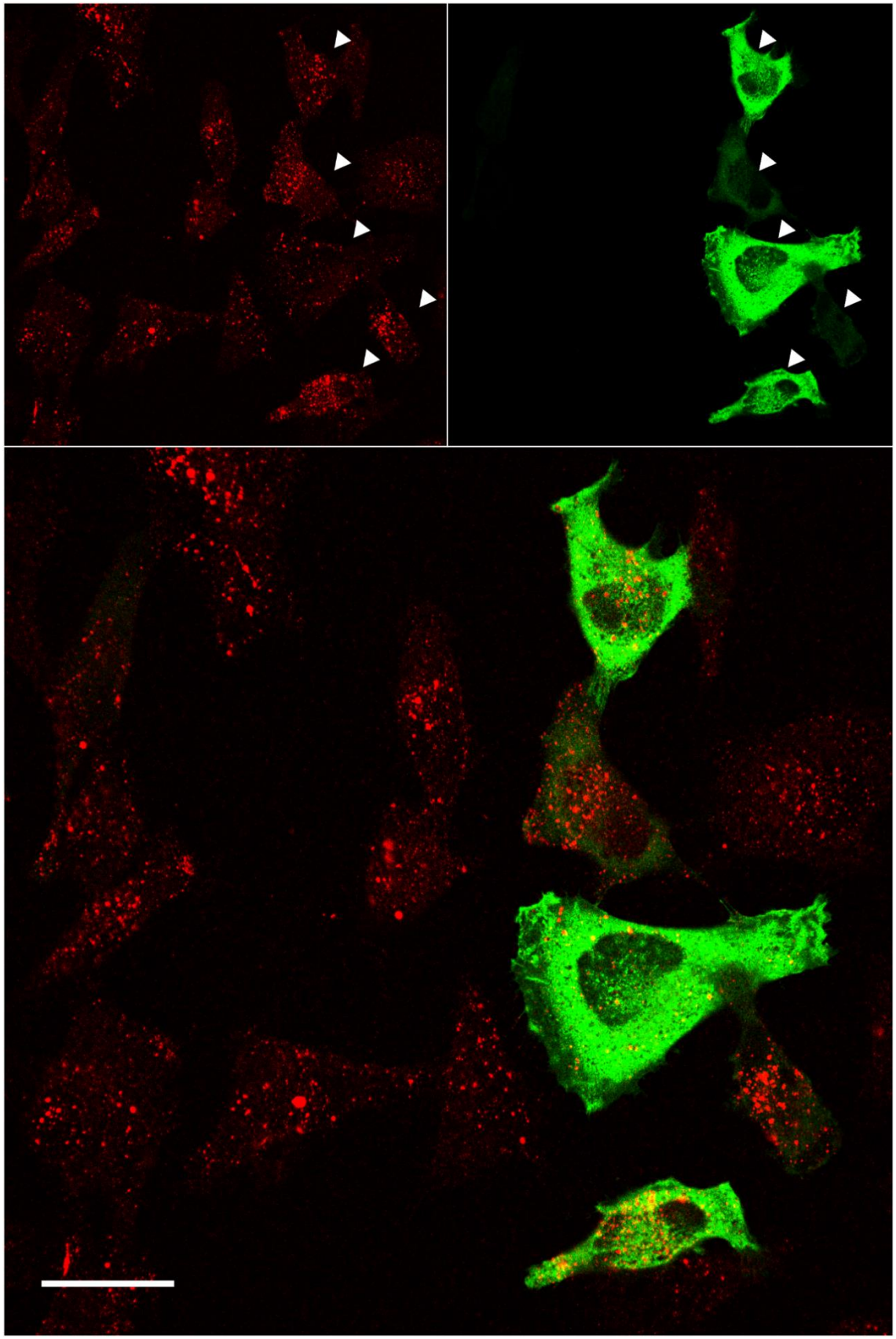


Figure S9. Supplementary images of rescue assays in endoA3 depleted cells transfected with a plasmid encoding for endoA2-GFP. Red signal: CD166-A546. **Green signal:** GFP-tagged endophilin A2. White triangles indicate cells effectively transfected with the construct. Scale bar = 25 µm.

3.4 siEndoA3 + endoA3-ΔlinkerSH3 supplementary images

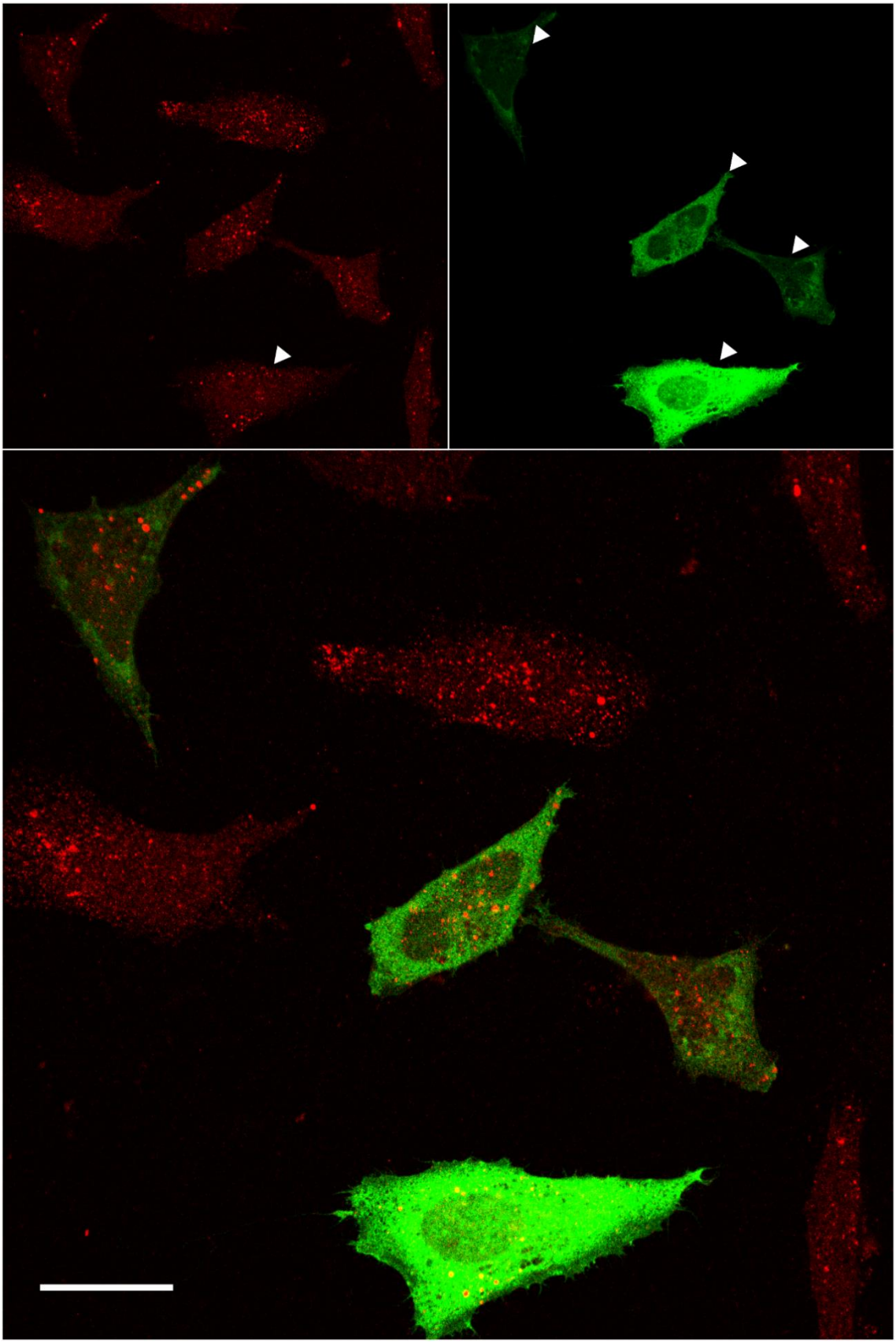


Figure S10. Supplementary images of rescue assays in endoA3 depleted cells transfected with a plasmid encoding for endoA3-ΔlinkerSH3-GFP. Red signal: CD166-A546. **Green signal:** GFP-tagged endoA3-ΔlinkerSH3. White triangles indicate cells effectively transfected with the construct. Scale bar = 25 µm.

3.5 siEndoA3 + BAR A2-SH3 A3 supplementary images

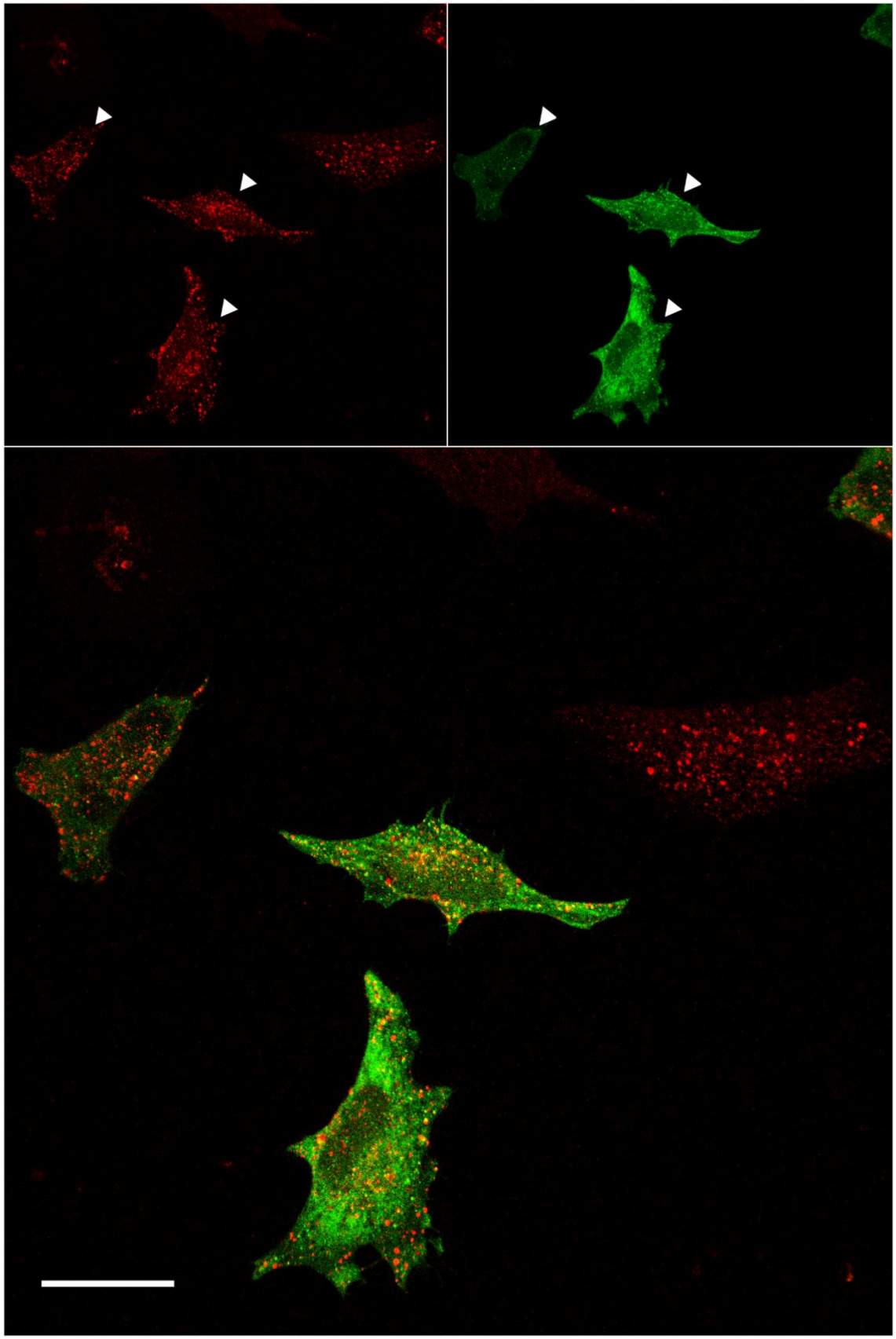


Figure S11. Supplementary images of rescue assays in endoA3 depleted cells transfected with a plasmid encoding for BAR A2-SH3 A3-GFP. Red signal: CD166-A546. **Green signal:** GFP-tagged BAR A2-SH3 A3. White triangles indicate cells effectively transfected with the construct. Scale bar = 25 µm.

3.6 siEndoA3 + BAR A3-SH3 A3 supplementary images

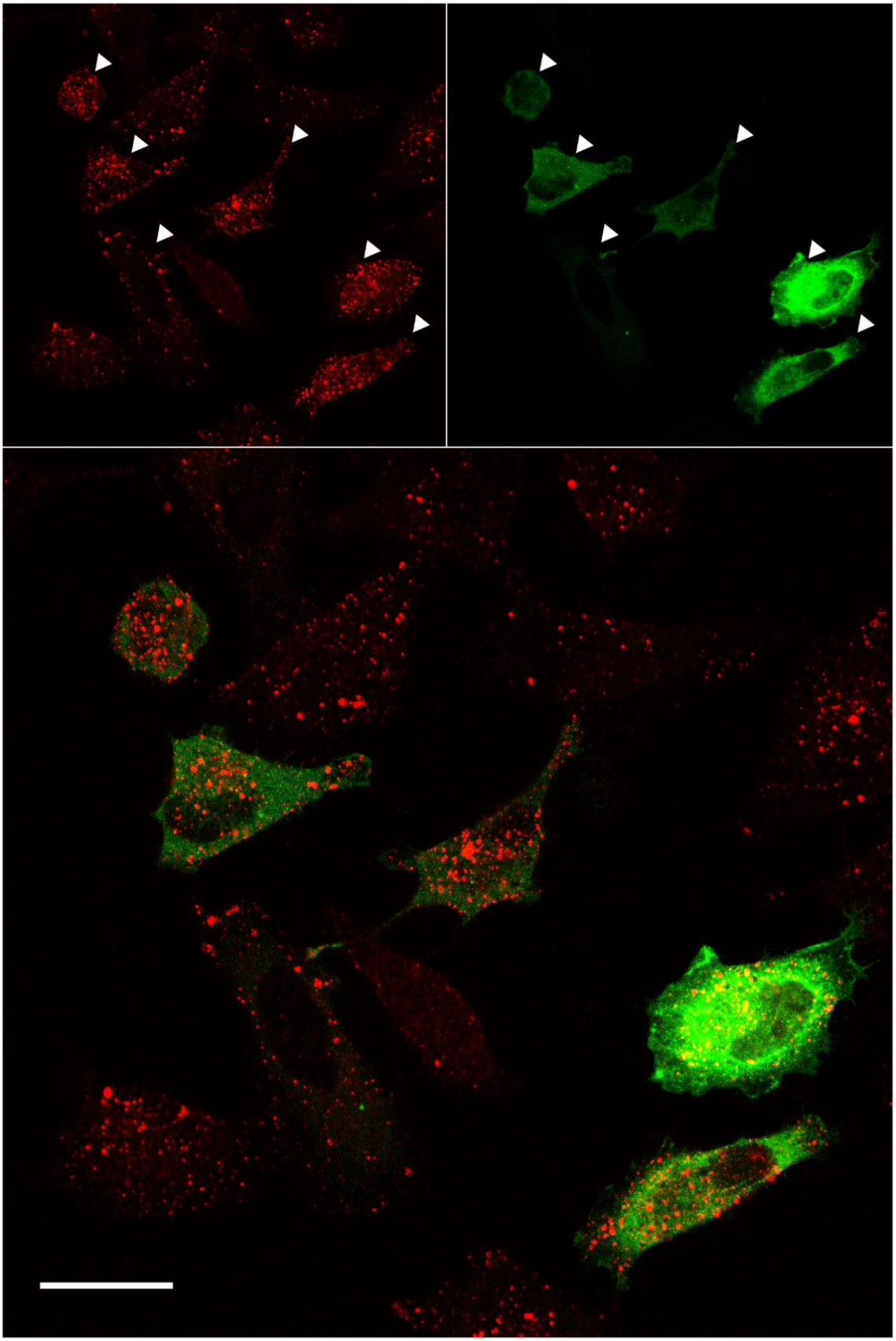


Figure S12. Supplementary images of rescue assays in endoA3 depleted cells transfected with a plasmid encoding for BAR A3-SH3 A2-GFP. Red signal: CD166-A546. **Green signal:** GFP-tagged BAR A3-SH3 A2. White triangles indicate cells effectively transfected with the construct. Scale bar = 25 µm.

3.7 siEndoA3 + BAR A2-linker A3-SH3 A2 supplementary images

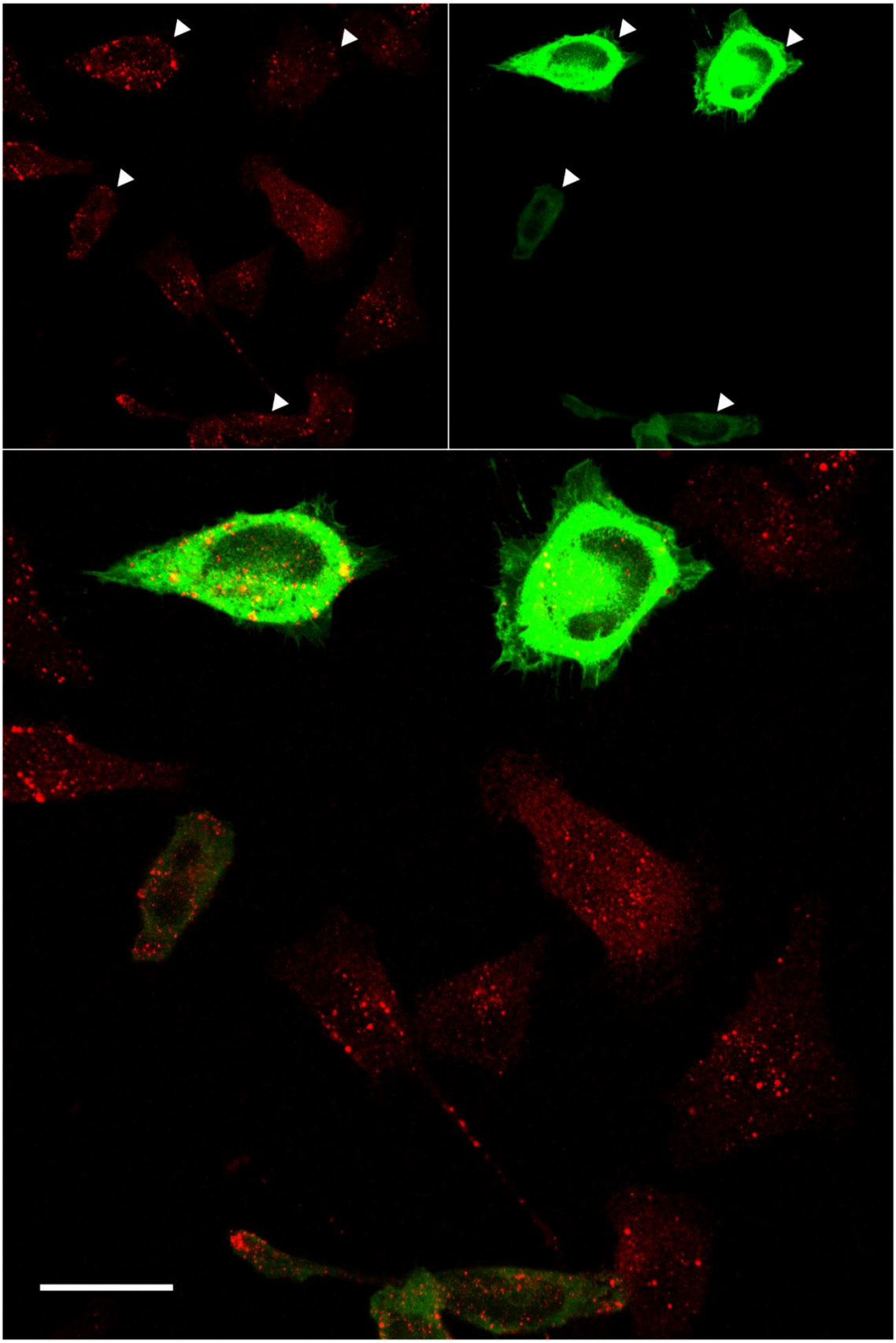


Figure S13. Supplementary images of rescue assays in endoA3 depleted cells transfected with a plasmid encoding for BAR A2-linker A3-SH3 A2-GFP. Red signal: CD166-A546. **Green signal:** GFP-tagged BAR A2-linker A3-SH3 A3. White triangles indicate cells effectively transfected with the construct. Scale bar = 25 µm.

Annex 4 : Endophilin triple knockdown (TKD)

One side objective of the master thesis was to confirm the previous observation that CD166 endocytosis is exclusively specific for endophilin A3. Results from colocalization and rescue assays already brought pieces of evidence in this direction, but we wanted to ensure that no redundancy could be observe with the two other isoforms endogenously expressed. We therefore compared CD166 internalization in cells knocked-down for endoA3 was (siEndoA3) or for the three endophilin isoforms (TKD).

Results obtained from two replicates (pulled together on **Figure S14**) clearly show that the additional silencing of endophilin A1 and A2 does not further decrease CD166 endocytosis, once again underlying the specificity of endoA3 for this mechanism.

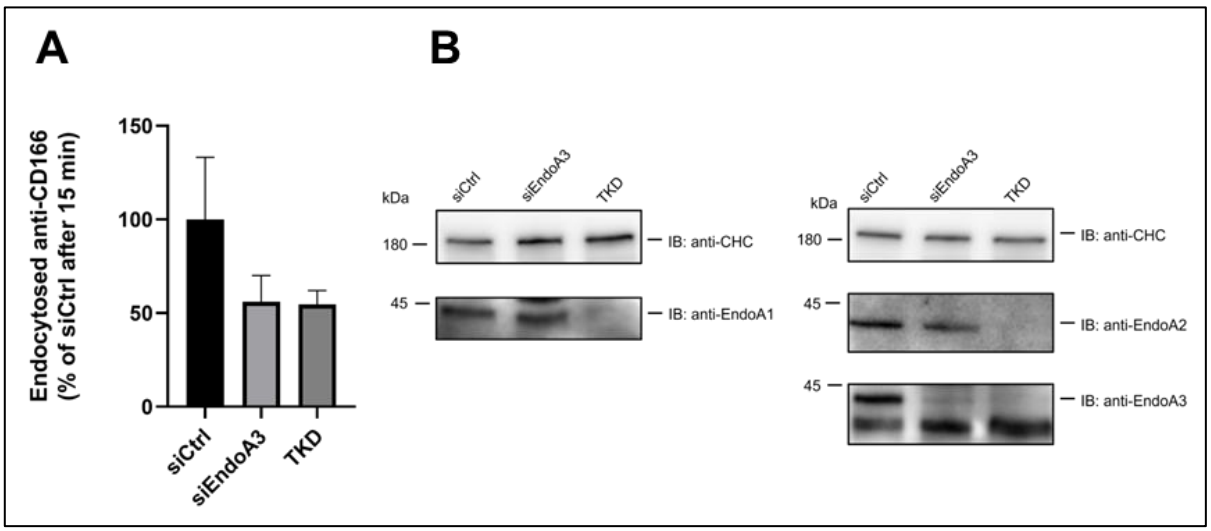


Figure S14. CD166 endocytosis in cells knocked-down for endophilin A3 alone (siEndoA3) and for the three endophilin isoforms combined (TKD). (A) The knockdown of three endophilin isoforms does not decrease further CD166 internalization compared to endoA3 knock-down alone. **(B)** Endophilins' knock down efficiency verified by western blot. Note the efficient depletion of each isoform.

Structural study of the endophilin A isoforms and their role in clathrin-independent endocytosis

Master's thesis presented by Ian Coupez

Among the wide variety of membrane bending proteins, the superfamily of BAR domain proteins has been one of the most extensively studied in the recent years. BAR stands for Bin/Amphiphysin/Rvs167 which are the first three described proteins of this superfamily. These proteins have been shown to function in many cellular processes requiring local membrane deformation. One of particular interest is endocytosis, the process by which cells internalize their nutrients but also recycle their plasma membrane components. Many BAR domain proteins were shown to act in different stages of endocytic processes. In particular, members of the endophilin A BAR subfamily held our attention for their strong involvement in these mechanisms.

The endophilin A subfamily is composed of three isoforms, endophilin A1, A2 and A3. They all share a common structure composed of four distinct domains: a N-terminal amphipathic helix, a BAR domain, a SH3 domain and a linker region connecting the last two. This global structure makes endophilins A remarkable proteins to build endocytic vesicles. Indeed, endophilin A were shown to be involved in well-described endocytic mechanisms such as clathrin-mediated endocytosis or fast endophilin-mediated endocytosis. In addition, recent research in my lab highlighted a new endocytic modality in which the A3 isoform plays a central role, but neither the A1 or A2 isoforms are involved. The first identified cargo of this new mechanism is the activated leukocyte cell adhesion molecule (ALCAM), also called CD166. The involvement of endophilin A3 in this mechanism but not of A1 and A2 raised the question of the origin of this isoform specificity.

To answer this question, we developed a strategy composed of four distinct steps. We started with the obtention of a library of plasmids encoding for mutant proteins fused with GFP and constructed using a domain swapping approach with endophilin A2 and A3. We then used these plasmids to assess the colocalization between CD166 and the mutants overexpressed in HeLa cells. We next asked whether the overexpression of these mutants could rescue CD166 endocytosis in HeLa cells previously knocked-down for endophilin A3. Finally, we performed pull-down experiment with GST fused with CD166 cytosolic tail as bait protein to highlight potential interaction between CD166 and the mutants.

The results from the three experiments all confirmed the specificity of endophilin A3 for CD166 endocytosis unlike endophilin A2. The linker-SH3 region appeared crucial for cargo recognition and the linker region's variability between endophilin isoforms appeared to be not sufficient to explain the specificity of endophilin A3 in CD166 endocytic mechanism. Unfortunately, further investigation are still required to conclude exactly which domain(s) of endophilin A3 provide(s) its specificity for this mechanism.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté des bioingénieurs
Croix du Sud, 2 bte L7.05.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/agro**