



OPEN ACCESS

EDITED BY

Anne Muesch,
Albert Einstein College of Medicine,
United States

REVIEWED BY

Giulio Preta,
Vilnius University, Lithuania

*CORRESPONDENCE

Bhuminder Singh,
✉ bhuminder.singh@vumc.org

RECEIVED 30 January 2024

ACCEPTED 21 February 2024

PUBLISHED 01 March 2024

CITATION

Harmych SJ, Tydings CW, Meiler J and Singh B (2024), Sequence and structural insights of monoleucine-based sorting motifs contained within the cytoplasmic domains of basolateral proteins.
Front. Cell Dev. Biol. 12:1379224.
doi: 10.3389/fcell.2024.1379224

COPYRIGHT

© 2024 Harmych, Tydings, Meiler and Singh. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Sequence and structural insights of monoleucine-based sorting motifs contained within the cytoplasmic domains of basolateral proteins

Sarah J. Harmych^{1,2}, Claiborne W. Tydings³, Jens Meiler ^{3,4} and Bhuminder Singh^{1,5*}

¹Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, United States,

²Department of Cell and Developmental Biology, Vanderbilt University, Nashville, TN, United States,

³Department of Chemistry and Center for Structural Biology, Vanderbilt University, Nashville, TN, United States, ⁴Institute for Drug Discovery, Leipzig University Medical School, Leipzig, Germany,

⁵Epithelial Biology Center, Vanderbilt University Medical Center, Nashville, TN, United States

Delivery to the correct membrane domain in polarized epithelial cells is a critical regulatory mechanism for transmembrane proteins. The trafficking of these proteins is directed by short amino acid sequences known as sorting motifs. In six basolaterally-localized proteins lacking the canonical tyrosine- and dileucine-based basolateral sorting motifs, a monoleucine-based sorting motif has been identified. This review will discuss these proteins with an identified monoleucine-based sorting motif, their conserved structural features, as well as the future directions of study for this non-canonical basolateral sorting motif.

KEYWORDS

epithelial polarity, polarized protein trafficking, basolateral sorting motif, clathrin adaptor proteins, monoleucine-based motif

1 Introduction

Polarized epithelial cells line body cavities and organs throughout the body, forming a barrier critical for organ function. The structure and function of polarized epithelial cells depends on maintaining biochemically distinct apical and basolateral membrane domains with asymmetric distributions of proteins and lipids (Apodaca et al., 2012; Mostov et al., 2003). Membrane lipid composition has been shown to play a key role in both the establishment and maintenance of epithelial polarity and protein trafficking (Bugda Gwilt and Thiagarajah, 2022). Epithelial cells achieve asymmetric protein distribution primarily through the selective targeting of newly synthesized proteins to either the apical or basolateral membrane (Matter, 2000). Newly synthesized transmembrane proteins are inserted into the endoplasmic reticulum and transit through the Golgi apparatus, eventually reaching the trans-Golgi network (TGN) (Mostov et al., 2003). At the TGN, transmembrane proteins are incorporated into vesicles that deliver them either directly to the plasma membrane or to endosomes and then to the cell surface (Ang et al., 2004) (Figure 1A). Alternatively, asymmetric distribution of proteins in polarized epithelial cells is achieved through selective retention at the plasma membrane. In this mechanism, newly synthesized proteins are delivered both to the apical and basolateral membrane, but endocytosed from

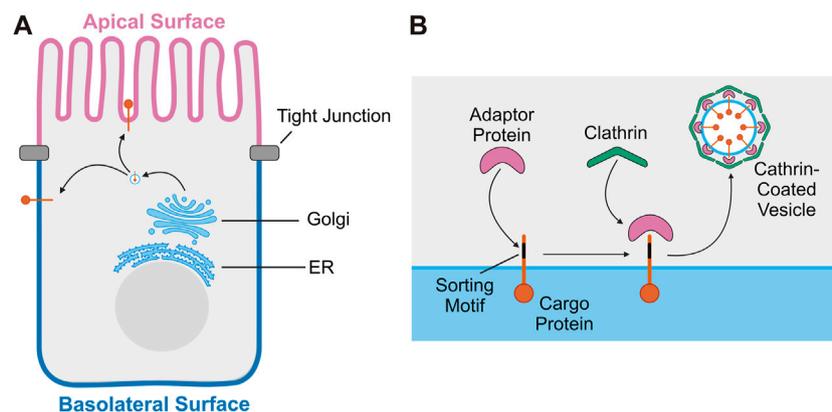


FIGURE 1

Trafficking of transmembrane proteins in polarized epithelial cells. (A) Overview of transport of transmembrane proteins from the Golgi to the plasma membrane in polarized epithelial cells. (B) Diagram of the sorting and loading of transmembrane proteins into vesicles at the Golgi. Sorting motif of cargo protein is recognized by adaptor protein. The adaptor protein then recruits clathrin, facilitating the formation of a clathrin-coated vesicle which will then transport the cargo protein to the plasma membrane.

or recycled back to a specific cell surface, leading to a steady state polarized distribution (Matter, 2000).

Short peptide motifs or post-translational modifications serve as signals that facilitate the selective targeting of proteins to a specific membrane domain in polarized epithelial cells (Cao et al., 2012; Bonifacino, 2014). Basolateral sorting motifs are well defined and are usually a short amino acid sequence in the cytoplasmic domain of the protein (Cao et al., 2012). Like basolateral sorting motifs, apical sorting motifs are also identified within short amino acid sequences, or post-translational modifications, including a glycosylphosphatidylinositol (GPI) anchor or N- and O-linked glycosylation (Cao et al., 2012; Weisz and Fölsch, 2020). Apical sorting motifs have been identified in the extracellular, transmembrane, or cytoplasmic region of apical proteins and lack a unifying mechanism to direct apical localization (Weisz and Fölsch, 2020). Basolateral localization is often directed by adaptor proteins, which recognize sorting motifs within cargo proteins to concentrate them into a small region of the TGN membrane (Cao et al., 2012). Adaptor proteins then recruit clathrin to the membrane, facilitating the generation of a clathrin-coated vesicle containing the cargo proteins to be brought to the plasma membrane (McMahon and Boucrot, 2011; Park and Guo, 2014) (Figure 1B). Using this mechanism, apical and basolateral proteins are sorted into separate vesicles and transported to their appropriate destination. Disruption of this process can impact epithelial function and cause disease (Stein et al., 2002).

2 Basolateral sorting in polarized epithelial cells

Basolateral sorting motifs most often reside in the cytoplasmic domain of the protein and are usually recognized by members of the clathrin adaptor protein (AP) family (Fölsch, 2005). These heterotetrameric adaptor complexes are responsible for recognizing and binding to cargo molecules and recruiting clathrin to generate a clathrin-coated vesicle for transport

(McMahon and Boucrot, 2011). Basolateral proteins in polarized epithelial cells are most often recognized by the epithelial cell-specific variant of AP-1, AP-1B, but there are known AP-1B independent pathways for basolateral targeting (Fölsch, 2005). AP complexes specifically mediate the trafficking of proteins within one compartment of the cell, such as AP-1B mediates trafficking from the TGN to the basolateral membrane. Selective recruitment of AP complexes to the membrane is directed by lipid composition of the membrane and GTPases present on the membrane (Park and Guo, 2014). In the case of AP-1, phosphatidylinositol 4 phosphate (PI(4)P) serves as the lipid directing its localization and ADP-ribosylation factor 1 (Arf1) serves as the GTPase (Ren et al., 2013; Park and Guo, 2014). Basolateral sorting motifs have been well characterized and often fall into two families, tyrosine-based and dileucine-based. Both motifs are recognized by AP complexes, specifically AP-1, which facilitates their incorporation into clathrin-coated vesicles destined for the basolateral surface (Robinson, 2004).

Tyrosine-based sorting motifs are characterized by the presence of a critical tyrosine residue and follow the general form YXXΦ (Y is tyrosine, X is any amino acid, and Φ is a bulky hydrophobic amino acid) (Bonifacino and Dell'angelica, 1999). While most well-known for targeting transmembrane proteins to the basolateral membrane, tyrosine-based sorting motifs also mediate rapid internalization of proteins from the cell surface and targeting proteins to other cellular compartments, such as lysosomes (Marks et al., 1997). The ability of tyrosine-based sorting motifs to differentially direct protein localization is mediated by their ability to interact with several members of the AP family, including AP-1B, which facilitates its role in directing proteins to the basolateral membrane (Bonifacino and Dell'angelica, 1999). Tyrosine-based sorting motifs were first determined to interact with the μ2 subunit of AP-2. The binding pocket for tyrosine-based sorting motifs on the μ2 domain has a banana-shaped structure with binding pockets for both the Y and Φ residues of the sorting motif (Owen and Evans, 1998). The X residues or residues adjacent to the motif may provide other points of contact, but the Y and Φ residues are critical (Marks

et al., 1997). Based on sequence alignment, the $\mu 2$ subunit residues important for this interaction are conserved across the μ subunits of the other AP family members, indicating that they interact with tyrosine-based sorting motifs in the same way. Interacting with different complexes in the AP family allows tyrosine-based sorting motifs to facilitate transport to multiple compartments within the cell (Bonifacino and Dell'Angelica, 1999).

The other major group of basolateral sorting motifs is dileucine-based. These motifs are characterized by the presence of critical leucine-leucine or leucine-isoleucine residues and come in two forms defined by the consensus motifs DXXLL and [DE]XXXL[LI]. (D is aspartic acid, E is glutamic acid, X is any amino acid, L is lysine, and I is isoleucine) (Janvier et al., 2003; Bonifacino, 2014). DXXLL motifs mediate intracellular sorting between the TGN and endosomes while [DE]XXXL[LI] motifs target proteins to endosomes, lysosomes, and the basolateral membrane. The leucine-leucine or leucine-isoleucine pairs are critical elements for the motifs while the upstream aspartic acid or glutamic acid is often important but not essential for function (Janvier et al., 2003). Even though the AP subunit interacting with tyrosine-based sorting motifs is well defined, the identity of the AP subunit that recognizes dileucine motifs is controversial. Mutational analysis has demonstrated that the binding site for dileucine motifs is on the β subunit of AP-1 (Rapoport et al., 1998). More recently, yeast two-hybrid analysis was unable to detect a strong interaction between the dileucine motifs and any of the AP subunits, which suggests a requirement for more than one subunit. In agreement with this, yeast-three hybrid experiments showed recognition by a combination of the γ and $\sigma 1$ subunits of AP-1 and the δ and $\sigma 3$ subunits of AP-3. This study was unable to determine whether the interaction requires binding sites on both the subunits or that the interaction between the subunits and the motif is more complex (Janvier et al., 2003). Co-crystallization of eight different dileucine-based sorting motif peptides with AP-2 further supports the idea of more than one subunit being involved in the interaction with dileucine-based sorting motifs, finding the binding site to be on the α - $\sigma 2$ heterodimer (Kelly et al., 2008).

Outside of the AP complexes, PDZ (PSD95, DlgA, ZO-1) proteins have also been shown to play a role in the polarized distribution of proteins in epithelial cells. PDZ proteins contain an 80–90 amino acid PDZ domain which bind to PDZ target motifs on the C-terminal end of target proteins (Muth and Caplan, 2003). The most common PDZ target motif follows the sequence X-S/T-X- Φ (X is any amino acid, Φ is a hydrophobic residue, typically I, L, V, or F). Unlike AP complexes, which bind internal sorting motifs, PDZ proteins typically bind C-terminal motifs, engaging with the carboxylate moiety during peptide binding (Amacher et al., 2020). While AP complexes are highly specific for certain sorting motifs, PDZ binding interactions are considered promiscuous. Due to the shallow nature of the PDZ binding cleft, there are a limited number of stereochemical restrictions on peptide recognition and binding, resulting in a number of PDZ domains being able to bind the same target (Amacher et al., 2020). PDZ domains have been shown to play a role in the selective retention of basolateral proteins in polarized epithelial cells, contributing to these proteins' basolateral distribution. For example, the PDZ protein LIN-7 has been implicated in the selective retention of the GABA transporter BGT-1 (Perego et al., 1999), the proteoglycan syndecan-1 (Maday

et al., 2008), and the receptor tyrosine kinase, ErbB2 (Shelly et al., 2003). PDZ proteins have also been implicated in the polarized distribution of apical proteins (e.g., CFTR; cystic fibrosis transmembrane conductance regulator), while AP complexes have only been implicated in the trafficking of basolateral proteins (Swiatecka-Urban et al., 2002).

Other less common sorting motifs have important implications for the localization of basolateral proteins. One of these motifs is a leucine-valine di-hydrophobic pair. This sorting motif was first identified in the cytoplasmic domain of CD44 and was shown to mediate its localization to the basolateral membrane in polarized epithelial cells (Sheikh and Isacke, 1996). A di-hydrophobic valine-leucine pair has also been shown to direct the basolateral localization of the MHC class I chain-related MICA molecule (Suemizu et al., 2001) and the basolateral localization of CD1d (Rodionov et al., 2000). CD1d is also an example of a basolateral protein whose cytoplasmic tail contains multiple sorting motifs. The cytoplasmic tail of CD1d contains the di-hydrophobic valine-leucine pair and a tyrosine-based sorting motif, both of which contribute to the basolateral localization of the protein (Rodionov et al., 2000). The cytoplasmic domain of epidermal growth factor receptor (EGFR) contains two functionally redundant basolateral sorting motifs, a dileucine-based sorting motif and a polyproline core-based sorting motif (He et al., 2002). EGFR ligands transforming growth factor α (TGF α) (Dempsey and Coffey, 1994) and betacellulin (Singh et al., 2015) also contain two basolateral sorting determinants. In the case of another EGFR ligand, epiregulin (EREG), when its basolateral motif is removed, EREG shifts to predominantly apical localization, suggesting the presence of a competing, recessive apical sorting motif in EREG (Singh et al., 2013).

While many basolateral proteins contain aforementioned sorting motifs, some lack standard identifying factors for basolateral sorting. Studies involving sequential cytoplasmic domain truncation mutants and site directed mutagenesis have revealed six proteins with a monoleucine-based basolateral sorting motif. These sorting motifs are characterized by a single leucine residue and an N-terminal acidic amino acid cluster with 3–5 residues between them. The remainder of this review will focus on the monoleucine-based sorting motif.

3 Proteins with identified monoleucine-based sorting motif

3.1 Stem cell factor (EDDXXXXXL)

A basolateral monoleucine-based sorting motif was first identified in stem cell factor (SCF) (Wehrle-Haller and Imhof, 2001). SCF is the ligand for the receptor tyrosine kinase c-Kit. Binding of SCF to c-Kit leads to the activation of signaling pathways involved in cell survival, migration, and proliferation. SCF belongs to a family of transmembrane growth factors with highly conserved cytoplasmic domains (Lennartsson and Rönstrand, 2012). Studies of an SCF mutant with an altered cytoplasmic domain suggested localization and secretion from the basolateral membrane is critical for SCF function (Wehrle-Haller and Weston, 1999). The cytoplasmic domain of SCF lacks the standard tyrosine- or dileucine-based basolateral sorting

motifs, but site-directed mutagenesis revealed a critical monoleucine residue and N-terminal cluster of acidic amino acids (Wehrle-Haller and Imhof, 2001). Substitution of the leucine residue with an alanine led to the apical localization of SCF in Madine-Darby Canine Kidney (MDCK) cells. Mutation of the acidic amino acid cluster N-terminal to the monoleucine led to a reduced efficiency of SCF basolateral targeting but did not recapitulate the complete apical localization caused by mutation of the leucine residue (Wehrle-Haller and Imhof, 2001).

3.2 CD147 (EDDXXXXXL)

Retinal pigment epithelial cells (RPE) display several plasma membrane proteins at the apical surface that typically localize to the basolateral surface in other epithelial cells, one of which is CD147 (Deora et al., 2004). In RPEs, CD147 shifts from a primarily basolateral localization in newborn rats, to a primarily apical distribution in adult rats. This shift suggests that the basolateral sorting motif in CD147 is suppressed or no longer recognized by RPE cells in mature rats (Marmorstein et al., 1998). The cytoplasmic domain of CD147, like SCF, lacked the standard tyrosine- or dileucine-based basolateral sorting motifs, which led to site-directed mutagenesis of the cytoplasmic domain to identify the sorting motif. Analysis of CD147 mutants revealed a monoleucine residue critical for basolateral targeting (Deora et al., 2004). When mutated, loss of this monoleucine led to almost complete apical localization of CD147. Unlike the sorting motif previously identified in SCF, basolateral targeting of CD147 did not require an acidic N-terminal amino acid cluster (Deora et al., 2004). Removal of a cluster of acidic amino acids N-terminal to the monoleucine residue resulted in apical mistrafficking, but replacement of these residues with alanine residues rescued normal localization. This suggests that these residues are not directly interacting with sorting machinery but instead are necessary for proper spacing within the cytoplasmic domain (Deora et al., 2004). Since the monoleucine-based sorting motif has only recently been identified, an adaptor protein recognizing this motif and facilitating its incorporation into clathrin-coated vesicles has not been determined. Basolateral localization of CD147 is not altered in cells which do not express AP-1B, the canonical basolateral adaptor protein, indicating that AP-1B may not mediate the trafficking of CD147 in polarized epithelial cells (Deora et al., 2004).

3.3 Amphiregulin (EEXXXL)

Amphiregulin (AREG) is a member of the epidermal growth factor (EGF)-like ligand family (Singh and Coffey, 2014). These ligands are produced as single-pass transmembrane proteins that, following delivery to the plasma membrane, are cleaved to produce a soluble signaling molecule. These soluble ligands then bind and activate members of the ErbB receptor family, including EGFR (Singh et al., 2016). These receptors activate signaling pathways that regulate processes such as cell growth, proliferation, and migration and are critical for epithelial tissue homeostasis (Wieduwilt and Moasser, 2008). Studies of both EGFR and some of the EGF-like ligands has demonstrated the importance of basolateral localization for their proper signaling (Kuwada et al., 1998; Singh and Coffey, 2014). AREG was shown to localize to the basolateral membrane in polarized epithelial cells in a cytoplasmic domain-dependent manner

(Brown et al., 2001). Further investigation of the cytoplasmic domain revealed a monoleucine-based basolateral sorting motif which when removed lead to a nonpolarized localization pattern (Gephart et al., 2011). While mutation of both the monoleucine and N-terminal cluster of amino acids led to altered localization, substitution of the three intervening amino acids did not alter localization. These intervening amino acids present a key difference between the previously identified monoleucine-based sorting motifs of SCF and CD147, which both had five amino acids between the monoleucine and acidic amino acids while the sorting motif of AREG has only three residues (Gephart et al., 2011). Knockdown of AP-1B in polarized epithelial cells led to an altered steady state localization of AREG but did not impact polarized delivery of newly synthesized molecules (Gephart et al., 2011). This suggests a role for AP-1B in the recycling of AREG to the basolateral surface but not in the trafficking of AREG from the TGN to the plasma membrane, although it has not been determined if recycling by AP-1B is also mediated by the monoleucine-based sorting motif or by another yet unidentified motif.

3.4 Betacellulin (EEXXXL)

Betacellulin (BTC) is also an EGF-like ligand that binds and activates EGFR and ErbB4 (Singh et al., 2016). Like previously studied EGF-like ligands, the steady state localization of BTC is also basolateral. Removal of the cytoplasmic domain led to a nonpolarized steady state distribution of BTC and the formation of lateral lumens (Singh et al., 2015). The cytoplasmic domain of BTC contained a monoleucine-based sorting motif like AREG, with a critical monoleucine residue, N-terminal cluster of acidic amino acids, with three amino acids in between. Analysis of sequential cytoplasmic domain truncation mutants of BTC showed that loss of a cysteine-rich motif at the beginning of the cytoplasmic domain further disrupted the basolateral-sorting specificity of BTC (Singh et al., 2015). While substitution of this cysteine-rich motif alone did not alter the localization of full length BTC, combining this mutation with mutation of the monoleucine-based sorting motif led to further loss of basolateral BTC compared to full length BTC with only the monoleucine-based motif mutated. This suggests that both motifs play a role in the basolateral localization of BTC, which was not seen in the other three proteins with a monoleucine-based sorting motif. Like other EGF-like ligands, loss of polarized distribution of BTC had significant impacts on cell growth and development (Singh et al., 2015). Cells expressing an apical mutant of BTC developed lateral lumens, indicating that mistrafficked BTC impacted the maturation of polarized epithelial cell layers. A mutation causing mistrafficking of BTC has been observed in human cancers and confers a growth advantage to the cells, suggesting mistrafficking of EGF-like ligands can have clinical implications (Singh et al., 2015). Mistrafficking of another EGFR ligand, epiregulin, led to epithelial transformation *in vivo* and loss of polarity *in vitro* (Singh et al., 2013; Singh et al., 2021).

Sequence analysis of heparin-binding EGF-like growth factor (HB-EGF), another EGF-like ligand, has revealed the presence of a putative monoleucine-based sorting motif. Given the important regulatory role trafficking plays for other EGF-like ligands, the localization of HB-EGF and the potential role of this motif in directing its trafficking warrant further study.

3.5 Na⁺/I⁻ symporter (EEXXXL)

The sodium/iodide symporter (NIS) is a protein expressed at the basolateral membrane of thyroid follicular cells and mediates the concentration of iodide for the generation of thyroid hormones. Uptake of radioiodine is used to treat thyroid cancer and relies on both the proper function and localization of NIS (Ravera et al., 2017). Thyroid tumors often become resistant to this treatment method through loss of proper targeting of NIS to the plasma membrane rather than loss of NIS function, making the mechanism of NIS trafficking a potential target for cancer therapy (Martín et al., 2019). Removal of the cytoplasmic domain of NIS led to a significant reduction in the amount of NIS at the plasma membrane compared to wild-type NIS. A monoleucine-based sorting motif was identified in the cytoplasmic domain that, when altered, led to NIS being localized primarily at the apical membrane. This altered localization reduced the accumulation of iodide in the cells (Martín et al., 2019). The trafficking of NIS to the basolateral membrane follows a clathrin-dependent mechanism in which both the AP-1A and AP-1B adaptor complexes can mediate its sorting in the TGN and recycling endosomes (Martín et al., 2019; Koumariou et al., 2022).

3.6 Matriptase (EEXXXXL)

Matriptase is an epithelial serine-protease whose expression has been seen in a variety of human tissues. Translated as a zymogen, matriptase requires a complicated series of steps to become active, including proteolytic cleavage and interaction with its cognate inhibitor (List et al., 2006). The basolateral localization of matriptase is important for both its proper activation and function, where these events take place (Tseng et al., 2020). Like most basolateral proteins, removal of the cytoplasmic domain of matriptase leads to a nonpolarized localization in MDCK cells (Murai et al., 2009). When a matriptase mutant lacking the cytoplasmic domain was expressed in keratinocytes, it accumulated in the nucleus and perinuclear area rather than at the plasma membrane like WT matriptase (Tseng et al., 2020). While complete removal of the cytoplasmic domain led to decreased plasma membrane delivery, a mutant with an altered monoleucine-based sorting motif was still delivered to the plasma membrane but showed a nonpolarized distribution. This suggests that while the entire cytoplasmic domain of matriptase plays a role in its delivery to the plasma membrane, the monoleucine-based sorting motif alone dictates preferential localization to the basolateral membrane (Tseng et al., 2020). Time-lapse video microscopy demonstrated that alterations to this motif impaired the movement of matriptase into the correct transport vesicles. The identification of the monoleucine-based sorting motif of matriptase introduced another variation of this motif. Previously identified motifs contained either three or five residues between the monoleucine and the cluster of acidic amino acids while the matriptase motif has four (Tseng et al., 2020).

4 Discussion

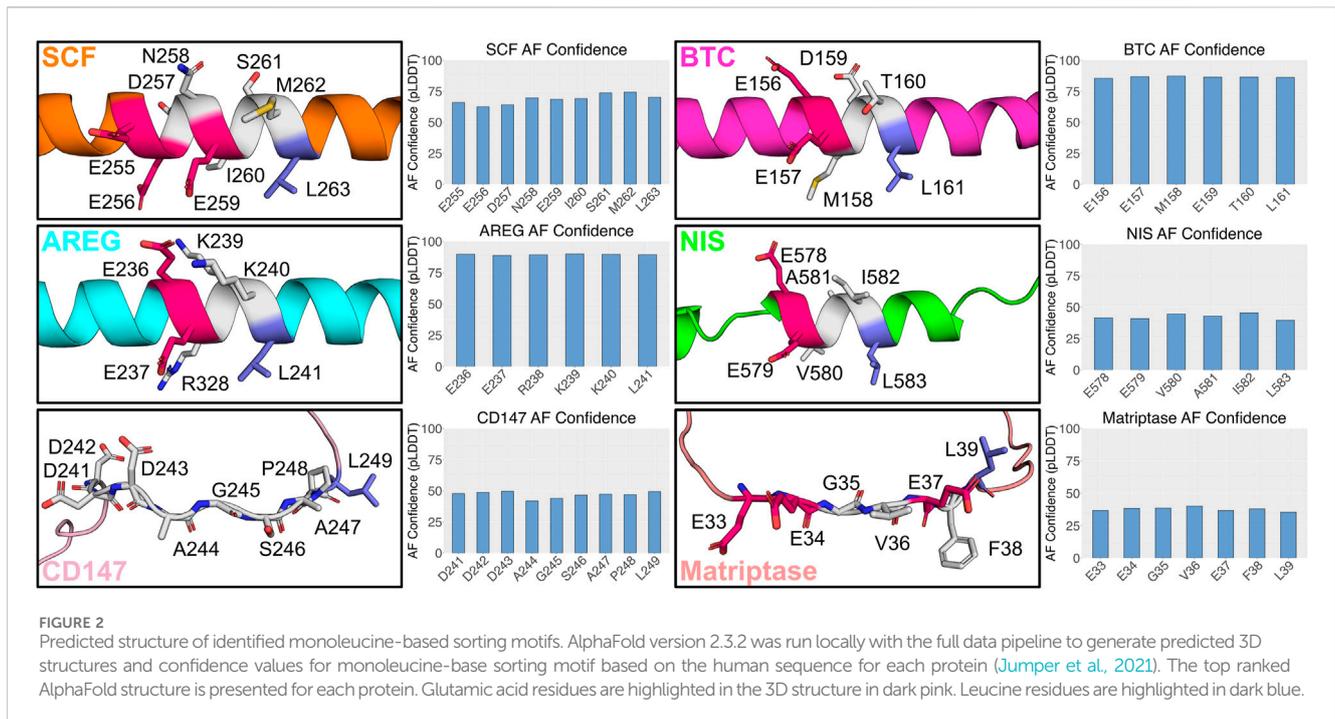
Disruption of normal protein localization has significant impacts on function, making their trafficking a critical regulatory mechanism to study. While the monoleucine-based sorting motif is a more recent

discovery than other basolateral sorting motifs, it warrants the same detailed study as tyrosine- and dileucine-based sorting motifs. Altered trafficking of proteins directed by a monoleucine-based sorting motif, such as NIS and BTC, has been shown to play a role in disease (Singh et al., 2015; Martín et al., 2019).

Recent studies of dominant nephrogenic diabetes insipidus (NDI) have demonstrated a role in the development of disease for a monoleucine-based sorting motif (Kamsteeg et al., 2003). In NDI, water reabsorption in the kidneys is diminished due to mutations of the aquaporin-2 (AQP2) water channel. In the case of recessive NDI, AQP2 mutants are misfolded and retained in the ER (Gao et al., 2020). But in dominant NDI, a functional AQP2 channel is expressed on the plasma membrane and is localized to the basolateral membrane instead of the apical membrane. A frameshift mutation in the AQP2 gene alters the cytoplasmic domain and introduces a leucine residue that when mutated restores the apical localization of AQP2. Unlike the other monoleucine-based sorting motifs identified, this motif lacks glutamic acids upstream of the leucine (Kamsteeg et al., 2003). This suggests, that like dileucine based sorting motifs, the critical residue is the leucine, and the glutamic acid residues are not essential to proper localization (Janvier et al., 2003). The identification of this motif in mutant AQP2 demonstrates that the monoleucine-based sorting motif is dominant since it overrides any apical sorting motif present in WT AQP2. It also provides a different role for the motif: instead of being altered in disease, introducing the motif can lead to the development of disease.

There is also potential role for monoleucine-based sorting motifs in the trafficking of neuronal proteins. Glycoprotein M6a (Gpm6a) is expressed in the neurons of the central nervous system, participating in filopodia formation, neurite expression, and synaptogenesis (Alfonso et al., 2005). Analysis post-removal of the N- and C- termini individually indicated that the C-terminus is essential for the role of Gpm6a in filopodia formation (Rosas et al., 2018). One explanation for this change in filopodia formation is that the C-terminus is needed to interact with cytosolic proteins needed for this process to occur. The other is that loss of the C-terminus impacts trafficking of Gpm6a in neurons. Alanine-scanning mutagenesis revealed three residues, K250, E258, and E259, that were necessary for filopodium formation (Rosas et al., 2018). E258/E259 are potentially apart of the monoleucine-based sorting motif EEQEL in the cytoplasmic C-terminus. Colocalization studies showed that deletion of the C-terminus diminished the association of Gpm6a with clathrin. This further supports the possibility that the C-terminus, and the monoleucine-based sorting motif within it, are involved in the trafficking of Gpm6a in neurons, and that alterations to this trafficking impact the function of Gpm6a (Rosas et al., 2018). More research needs to be done to confirm the role of the potential monoleucine-based sorting motif in Gpm6a trafficking, but this study suggests future work on monoleucine-based sorting motifs should look outside of polarized epithelial cells for other proteins relying on this motif to properly function.

Further study of monoleucine-based sorting motifs will likely involve the identification of more proteins containing this motif. This will require sequence analysis to identify putative monoleucine-based sorting motifs followed by analysis of the impact of mutations on the localization of these proteins to confirm the motif's role in trafficking. One major remaining question is what the role of the amino acids lying between the monoleucine and acidic amino acid cluster is. The number of residues occupying this space ranges from



three to five. Mutation of these residues in AREG does not impact localization (Gephart et al., 2011), suggesting they are not directly involved in the interaction between this motif and an adaptor protein. Future studies of the monoleucine-based sorting motif also need to include both *in vivo* and 3D studies of its impact on protein trafficking. The studies covered in this review have primarily conducted experiments on polarized epithelial cells grown on a permeable support, or in 2D. While 2D studies can provide valuable insights into the trafficking of proteins in polarized epithelial cells, 3D studies where epithelial cells are cultured on extracellular matrix components (e.g., Matrigel), or *in vivo* studies are more physiologically relevant (Singh et al., 2013; Singh et al., 2015).

We conducted a structural analysis for all the proteins described above containing the monoleucine-based sorting motif using AlphaFold. Among those, AREG, BTC, SCF, and NIS are predicted to have a helical structure in the region containing the monoleucine-based sorting motif, with varying degrees of confidence (Figure 2). AREG, BTC, and NIS all have three intervening residues between the monoleucine and acidic amino acid cluster. While SCF has five intervening residues, it does have a glutamic acid residue four residues N-terminal to the leucine. This matches the spacing between the leucine and glutamic acid in AREG, BTC, and NIS and suggests it may play a role in shaping the motif. CD147, with five intervening amino acids, and Matriptase, with four intervening amino acids, have a predicted disordered structure in the region of their monoleucine-based sorting motif (Figure 2). These analyses highlight structural heterogeneity within the monoleucine-based basolateral sorting motif, and a potential for differential binding to basolateral sorting adaptors. A better understanding of the impact of the intervening amino acids on the structure of the monoleucine-based sorting motif will better inform the function and interactions of these proteins, including the role of AP-1B in their trafficking.

Author contributions

SH: Conceptualization, Investigation, Visualization, Writing—original draft, Writing—review and editing. CT: Investigation, Visualization, Writing—review and editing. JM: Supervision, Writing—review and editing. BS: Conceptualization, Supervision, Writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. We acknowledge funding support to the Singh laboratory by NCI R01CA248505, American Cancer Society—Research Scholar Grant RSG-20-130-01-DDC, and CDMRP W81XWH-21-1-0694. CT is supported by the training grant T32 DK101003. JM is supported by a Humboldt Professorship of the Alexander von Humboldt Foundation. Work in the Meiler laboratory is further supported through the NIH (R01 CA227833, R01 LM013434, S10 OD016216, S10 OD020154, and S10 OD032234).

Acknowledgments

We thank Sarah E. Glass for editing this manuscript. Figure 1 was made using Biorender.com.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Alfonso, J., Fernández, M. E., Cooper, B., Flugge, G., and Frasnch, A. C. (2005). The stress-regulated protein M6a is a key modulator for neurite outgrowth and filopodium/spine formation. *Proc. Natl. Acad. Sci. U. S. A.* 102, 17196–17201. doi:10.1073/pnas.0504262102
- Amacher, J. F., Brooks, L., Hampton, T. H., and Madden, D. R. (2020). Specificity in PDZ-peptide interaction networks: computational analysis and review. *J. Struct. Biol. X* 4, 100022. doi:10.1016/j.jsbx.2020.100022
- Ang, A. L., Taguchi, T., Francis, S., Fölsch, H., Murrells, L. J., Pypaert, M., et al. (2004). Recycling endosomes can serve as intermediates during transport from the Golgi to the plasma membrane of MDCK cells. *J. Cell Biol.* 167, 531–543. doi:10.1083/jcb.200408165
- Apodaca, G., Gallo, L. I., and Bryant, D. M. (2012). Role of membrane traffic in the generation of epithelial cell asymmetry. *Nat. Cell Biol.* 14, 1235–1243. doi:10.1038/ncb2635
- Bonifacio, J. S. (2014). Adaptor proteins involved in polarized sorting. *J. Cell Biol.* 204, 7–17. doi:10.1083/jcb.201310021
- Bonifacio, J. S., and Dell'angelica, E. C. (1999). Molecular bases for the recognition of tyrosine-based sorting signals. *J. Cell Biol.* 145, 923–926. doi:10.1083/jcb.145.5.923
- Brown, C. L., Coffey, R. J., and Dempsey, P. J. (2001). The proamphiregulin cytoplasmic domain is required for basolateral sorting, but is not essential for constitutive or stimulus-induced processing in polarized madin-darby canine kidney cells. *J. Biol. Chem.* 276, 29538–29549. doi:10.1074/jbc.M102114200
- Bugda Gwilt, K., and Thiagarajah, J. R. (2022). Membrane lipids in epithelial polarity: sorting out the PIPs. *Front. Cell Dev. Biol.* 10, 893960. doi:10.3389/fcell.2022.893960
- Cao, X., Surma, M. A., and Simons, K. (2012). Polarized sorting and trafficking in epithelial cells. *Cell Res.* 22, 793–805. doi:10.1038/cr.2012.64
- Dempsey, P. J., and Coffey, R. J. (1994). Basolateral targeting and efficient consumption of transforming growth factor- α when expressed in Madin-Darby canine kidney cells. *J. Biol. Chem.* 269, 16878–16889. doi:10.1016/s0021-9258(19)89472-3
- Deora, A. A., Gravotta, D., Kreitzer, G., Hu, J., Bok, D., Rodriguez-Boulan, E., et al. (2004). The basolateral targeting signal of CD147 (EMMPRIN) consists of a single leucine and is not recognized by retinal pigment epithelium. *Mol. Biol. Cell* 15, 4148–4165. doi:10.1091/mbc.e04-01-0058
- Fölsch, H. (2005). The building blocks for basolateral vesicles in polarized epithelial cells. *Trends Cell Biol.* 15, 222–228. doi:10.1016/j.tcb.2005.02.006
- Gao, C., Higgins, P. J., and Zhang, W. (2020). AQP2: mutations associated with congenital nephrogenic diabetes insipidus and regulation by post-translational modifications and protein-protein interactions. *Cells* 9, 2172. doi:10.3390/cells9102172
- Gephart, J. D., Singh, B., Higginbotham, J. N., Franklin, J. L., Gonzalez, A., Fölsch, H., et al. (2011). Identification of a novel mono-leucine basolateral sorting motif within the cytoplasmic domain of amphiregulin. *Traffic* 12, 1793–1804. doi:10.1111/j.1600-0854.2011.01282.x
- He, C., Hobert, M., Friend, L., and Carlin, C. (2002). The epidermal growth factor receptor juxtamembrane domain has multiple basolateral plasma membrane localization determinants, including a dominant signal with a polyproline core. *J. Biol. Chem.* 277, 38284–38293. doi:10.1074/jbc.M104646200
- Janvier, K., Kato, Y., Boehm, M., Rose, J. R., Martina, J. A., Kim, B. Y., et al. (2003). Recognition of dileucine-based sorting signals from HIV-1 Nef and LIMP-II by the AP-1 gamma-sigma and AP-3 delta-sigma3 hemicomplexes. *J. Cell Biol.* 163, 1281–1290. doi:10.1083/jcb.200307157
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. doi:10.1038/s41586-021-03819-2
- Kamsteeg, E. J., Bichet, D. G., Konings, I. B. M., Nivet, H., Lonergan, M., Arthus, M. F., et al. (2003). Reversed polarized delivery of an aquaporin-2 mutant causes dominant nephrogenic diabetes insipidus. *J. Cell Biol.* 163, 1099–1109. doi:10.1083/jcb.200309017
- Kelly, B. T., McCoy, A. J., Späte, K., Miller, S. E., Evans, P. R., Höning, S., et al. (2008). A structural explanation for the binding of endocytic dileucine motifs by the AP2 complex. *Nature* 456, 976–979. doi:10.1038/nature07422
- Koumariou, P., Fernández-Méndez, C., Fajardo-Delgado, D., Mielu, L. M., Santisteban, P., and De La Vieja, A. (2022). Basolateral sorting of the sodium/iodide symporter is mediated by adaptor protein 1 clathrin adaptor complexes. *Thyroid* 32, 1259–1270. doi:10.1089/thy.2022.0163
- Kuwada, S. K., Lund, K. A., Li, X. F., Cliften, P., Amsler, K., Opresko, L. K., et al. (1998). Differential signaling and regulation of apical vs basolateral EGFR in polarized epithelial cells. *Am. J. Physiology* 275, 1419–1428. doi:10.1152/ajpcell.1998.275.6.C1419
- Lennartsson, J., and Rönstrand, L. (2012). Stem cell factor receptor/c-Kit: from basic science to clinical implications. *Physiol. Rev.* 92, 1619–1649. doi:10.1152/physrev.00046.2011
- List, K., Bugge, T. H., and Szabo, R. (2006). Matriptase: potent proteolysis on the cell surface. *Mol. Med.* 12, 1–7. doi:10.2119/2006-00022.list
- Maday, S., Anderson, E., Chang, H. C., Shorter, J., Satoh, A., Sfakianos, J., et al. (2008). A PDZ-Binding motif controls basolateral targeting of syndecan-1 along the biosynthetic pathway in polarized epithelial cells. *Traffic* 9, 1915–1924. doi:10.1111/j.1600-0854.2008.00805.x
- Marks, M., Ohno, H., Kirchnausen, T., and Bonracino, J. (1997). Protein sorting by tyrosine-based signals: adapting to the Ys and wherefores. *Trends Cell Biol.* 7, 124–128. doi:10.1016/S0962-8924(96)10057-X
- Marmorstein, A. D., Gan, Y. C., Bonilha, V. L., Finnemann, S. C., Csaky, K. G., and Rodriguez-Boulan, E. (1998). Apical polarity of N-cadherin and EMMPRIN in retinal pigment epithelium resulting from suppression of basolateral signal recognition. *J. Cell Biol.* 142, 697–710. doi:10.1083/jcb.142.3.697
- Martin, M., Modenutti, C. P., Peyret, V., Geysels, R. C., Darrouzet, E., Pourcher, T., et al. (2019). A carboxy-terminal monoleucine-based motif participates in the basolateral targeting of the Na⁺/I⁻ symporter. *Endocrinology* 160, 156–168. doi:10.1210/en.2018-00603
- Matter, K. (2000). Epithelial polarity: sorting out the sorters. *Curr. Biol.* 10, 39–42. doi:10.1016/s0960-9822(99)00256-0
- McMahon, H. T., and Boucrot, E. (2011). Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nat. Rev. Mol. Cell Biol.* 12, 517–533. doi:10.1038/nrm3151
- Mostov, K., Su, T., and Ter Beest, M. (2003). Polarized epithelial membrane traffic: conservation and plasticity. *Nat. Cell Biol.* 5, 287–293. doi:10.1038/ncb0403-287
- Murai, N., Miyake, Y., Tsuzuki, S., Inouye, K., and Fushiki, T. (2009). Involvement of the cytoplasmic juxtamembrane region of matriptase in its exclusive localization to the basolateral membrane domain of Madin-Darby canine kidney epithelial cells. *Cytotechnology* 59, 169–176. doi:10.1007/s10616-009-9205-0
- Muth, T. R., and Caplan, M. J. (2003). Transport protein trafficking in polarized cells. *Annu. Rev. Cell Dev. Biol.* 19, 333–366. doi:10.1146/annurev.cellbio.19.110701.161425
- Owen, D., and Evans, P. (1998). A structural explanation for the recognition of tyrosine-based endocytotic signals. *Science* 282, 1327–1332. doi:10.1126/science.282.5392.1327
- Park, S. Y., and Guo, X. (2014). Adaptor protein complexes and intracellular transport. *Biosci. Rep.* 34, e00123–e00390. doi:10.1042/BSR20140069
- Perego, C., Vanoni, C., Villa, A., Longhi, R., Kaech, S. M., Frö, E., et al. (1999). PDZ-mediated interactions retain the epithelial GABA transporter on the basolateral surface of polarized epithelial cells. *EMBO J.* 18, 2384–2393. doi:10.1093/emboj/18.9.2384
- Rapoport, I., Chen, Y. C., Cupers, P., Shoelson, S. E., and Kirchhausen, T. (1998). Dileucine-based sorting signals bind to the beta chain of AP-1 at a site distinct and regulated differently from the tyrosine-based motif-binding site. *EMBO J.* 17, 2148–2155. doi:10.1093/emboj/17.8.2148
- Ravera, S., Reyna-Neyra, A., Ferrandino, G., Amzel, L. M., and Carrasco, N. (2017). The sodium/iodide symporter (NIS): molecular physiology and preclinical and clinical applications. *Annu. Rev. Physiol.* 79, 261–289. doi:10.1146/annurev-physiol-022516-034125
- Ren, X., Farias, G. G., Canagarajah, B. J., Bonifacio, J. S., and Hurley, J. H. (2013). Structural basis for recruitment and activation of the AP-1 clathrin adaptor complex by Arf1. *Cell* 152, 755–767. doi:10.1016/j.cell.2012.12.042
- Robinson, M. S. (2004). Adaptable adaptors for coated vesicles. *Trends Cell Biol.* 14, 167–174. doi:10.1016/j.tcb.2004.02.002
- Rodionov, D. G., Nordeng, T. W., Kongsvik, T. L., and Bakke, O. (2000). The cytoplasmic tail of CD1d contains two overlapping basolateral sorting signals. *J. Biol. Chem.* 275, 8279–8282. doi:10.1074/jbc.275.12.8279
- Rosas, N. M., Alvarez Juliá, A., Alzuri, S. E., Frasnch, A. C., and Fuchsova, B. (2018). Alanine scanning mutagenesis of the C-terminal cytosolic end of Gpm6a identifies key residues essential for the formation of filopodia. *Front. Mol. Neurosci.* 11, 314. doi:10.3389/fnmol.2018.00314

- Sheikh, H., and Isacke, C. M. (1996). A di-hydrophobic leu-val motif regulates the basolateral localization of CD44 in polarized madin-darby canine kidney epithelial cells. *J. Biol. Chem.* 271, 12185–12190. doi:10.1074/jbc.271.21.12185
- Shelly, M., Mosesson, Y., Citri, A., Lavi, S., Zwang, Y., Melamed-Book, N., et al. (2003). Polar expression of ErbB-2/HER2 in epithelia: bimodal regulation by lin-7. *Dev. Cell* 5, 475–486. doi:10.1016/j.devcel.2003.08.001
- Singh, B., Bogatcheva, G., Krystofiak, E., McKinley, E. T., Hill, S., Rose, K. L., et al. (2021). Induction of apically mistrafficked epieregulin disrupts epithelial polarity via aberrant EGFR signaling. *J. Cell Sci.* 134, jcs255927. doi:10.1242/jcs.255927
- Singh, B., Bogatcheva, G., Starchenko, A., Sinnaeve, J., Lapierre, L. A., Williams, J. A., et al. (2015). Induction of lateral lumens through disruption of a monoleucinebased basolateral-sorting motif in betacellulin. *J. Cell Sci.* 128, 3444–3455. doi:10.1242/jcs.170852
- Singh, B., Bogatcheva, G., Washington, M. K., and Coffey, R. J. (2013). Transformation of polarized epithelial cells by apical mistrafficking of epieregulin. *Proc. Natl. Acad. Sci. U. S. A.* 110, 8960–8965. doi:10.1073/pnas.1305508110
- Singh, B., Carpenter, G., and Coffey, R. (2016). EGF receptor ligands recent advances. *F1000Res* 5, F1000 Faculty Rev-2270. doi:10.12688/f1000research.9025.1
- Singh, B., and Coffey, R. J. (2014). Trafficking of epidermal growth factor receptor ligands in polarized epithelial cells. *Annu. Rev. Physiol.* 76, 275–300. doi:10.1146/annurev-physiol-021113-170406
- Stein, M., Wandinger-Ness, A., and Roitbak, T. (2002). Altered trafficking and epithelial cell polarity in disease. *Trends Cell Biol.* 12, 374–381. doi:10.1016/s0962-8924(02)02331-0
- Suemizu, H., Radosavljevic, M., Kimura, M., Sadahiro, S., Yoshimura, S., Bahram, S., et al. (2001). A basolateral sorting motif in the MICA cytoplasmic tail. *Proc. Natl. Acad. Sci. U. S. A.* 99, 2971–2976. doi:10.1073/pnas.052701099
- Swiatecka-Urban, A., Duhaime, M., Coutermarsh, B., Karlson, K. H., Collawn, J., Milewski, M., et al. (2002). PDZ domain interaction controls the endocytic recycling of the cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem.* 277, 40099–40105. doi:10.1074/jbc.M206964200
- Tseng, C. C., Jia, B., Barndt, R. B., Dai, Y. H., Chen, Y. H., Du, P. W. A., et al. (2020). The intracellular seven amino acid motif EEGEVFL is required for matriptase vesicle sorting and translocation to the basolateral plasma membrane. *PLoS One* 15, e0228874. doi:10.1371/journal.pone.0228874
- Wehrle-Haller, B., and Imhof, B. A. (2001). Stem cell factor presentation to c-Kit. Identification of a basolateral targeting domain. *J. Biol. Chem.* 276, 12667–12674. doi:10.1074/jbc.M008357200
- Wehrle-Haller, B., and Weston, J. A. (1999). Altered cell-surface targeting of stem cell factor causes loss of melanocyte precursors in steel 17H mutant mice. *Dev. Biol.* 210, 71–86. doi:10.1006/dbio.1999.9260
- Weisz, O. A., and Folsch, H. (2020). “Molecular mechanisms of apical and basolateral sorting in polarized epithelial cells,” in *Basic epithelial ion transport principles and function: ion channels and transporters of epithelia in health and disease*. Editors K. L. Hamilton, and D. C. Devor (Springer Nature), 135–158. doi:10.1007/978-3-030-52780-8
- Wieduwilt, M. J., and Moasser, M. M. (2008). The epidermal growth factor receptor family: Biology driving targeted therapeutics. *Cell. Mol. Life Sci.* 65, 1566–1584. doi:10.1007/s00018-008-7440-8