

Compartmentalization of Aquaporins in the Human Intestine

Hari H. P. Cohly^{1*}, Raphael Isokpehi¹ and Rajendram V. Rajnarayanan²

¹Department of Biology, Jackson State University, Jackson, Mississippi 39217, USA

²Department of Chemistry, Tougaloo College, Jackson, Mississippi 39174, USA

*Correspondence to Dr. Hari H. P. Cohly. E-mail: hari.cohly@jsums.edu

Received: 25 October 2007 / Accepted: 30 April 2008 / Published: 30 June 2008

Abstract: Improper localization of water channel proteins called aquaporins (AQP) induce mucosal injury which is implicated in Crohn's disease and ulcerative colitis. The amino acid sequences of AQP3 and AQP10 are 79% similar and belong to the mammalian aquaglyceroporin subfamily. AQP10 is localized on the apical compartment of the intestinal epithelium called the glycocalyx while AQP3 is selectively targeted to the basolateral membrane. Despite the high sequence similarity and evolutionary relatedness, the molecular mechanism involved in the polarity, selective targeting and function of AQP3 and AQP10 in the intestine is largely unknown. Our hypothesis is that the differential polarity and selective targeting of AQP3 and AQP10 in the intestinal epithelial cells is influenced by amino acid signal motifs. We performed sequence and structural alignments to determine differences in signals for localization and post-translational glycosylation. The basolateral sorting motif "YRLL" is present in AQP3 but absent in AQP10; while N-glycosylation signals are present in AQP10 but absent in AQP3. Furthermore, the C-terminal region of AQP3 is longer compared to AQP10. The sequence and structural differences between AQP3 and AQP10 provide insights into the differential compartmentalization and function of these two aquaporins commonly expressed in human intestines.

Keywords: Aquaporin, colitis, compartmentalization, protein sorting, sequence analysis, water channel proteins

Introduction

Aquaporins (AQPs) are cell surface proteins that provide selective and rapid transport of water across the cell membrane [1]. There are 13 known mammalian aquaporins (AQP0-AQP12). Improper localization of AQPs has been reported to affect fluid permeability [2-5]. Alterations in fluid flux can result in disease conditions such as ulcerative colitis and Crohn's disease as a consequence of mucosal injury [3,6]. In the epithelial cells, AQPs are localized either in the apical or basolateral membrane [3]. Most cell surface proteins contain signals within their cytoplasmic termini that permit their recruitment into endocytic vesicles, which in turn facilitates their selective compartmentalization in the apical or basolateral membranes selectively [7, 8].

There are four known classes (I-IV) of endocytosis recruitment/localization signals [9]. Class I: Tyrosine-based signals within YXXØ sequence (where X can be any amino acid and Ø is an amino acid with a bulky hydrophobic group) [10]. Many of these signals are

constitutively active and are found to be independent of ligand binding [10]. Class II: The di-leucine motif (LL) is another well known endocytosis motif that is present within many transmembrane cell surface proteins [11]. Class III: Ligand-induced phosphorylation of serine residues in G protein-coupled receptors (GPCR) serves as a signal for receptor endocytosis. Class IV: A phenylalanine-isoleucine motif together with the acidic EEDE cluster has been reported to regulate intracellular traffic of the endoprotease furin [12]. Class V: Ubiquitination has been recently identified as another regulator of the endocytosis of several membrane receptors. Ubiquitination is a post translational modification (PTM) that likely affects all proteins at some point in their life cycle. The most common role ubiquitination plays is in protein degradation [13].

In many cases, more than one endocytosis signal act in tandem, as in the case of the T-cell receptor CD3 gamma where the di-leucine motif acts in cooperation with a phosphorylated serine residue to mediate endocytosis. The low-density lipoprotein receptor-related proteins consist of NPXY motif in addition to the YXXL

motif. However, recent reports suggest that only YXXL motif is sufficient for endocytic trafficking [14]. Other than phosphorylation, glycosylation is another PTM that can facilitate compartmentalization of cell surface proteins. Almost all cell surface proteins are glycosylated. Proper folding of the translated protein is facilitated by glycosylation events, and thus protein function is often dependent on or refined by the carbohydrate moieties attached to the polypeptide. Heterogeneity often exists in the multiple oligosaccharide chains attached to a single protein. The structures of these oligosaccharides are regulated by very specific glycosyltransferases and glycosidases [15].

The amino acid sequences of AQP3 and AQP10 are 79% similar and belong to the mammalian aquaglyceroporin subfamily. AQP10 is localized on the apical compartment of the intestinal epithelium called the glycocalyx while AQP3 is selectively targeted to the basolateral membrane [16]. Despite the high sequence similarity and evolutionary relatedness, the molecular mechanism involved in the polarity, selective targeting and function of AQP3 and AQP10 in the intestine is largely unknown. Our hypothesis is that the differential polarity and selective targeting of AQP3 and AQP10 in the intestinal epithelial cells is influenced by amino acid signal motifs. We performed sequence and structural alignments to determine differences in signals for localization and post-translational glycosylation. We observed that the YRLL motif is present in AQP3 but absent in AQP10 while the N-glycosylation signal is present in AQP10 but absent in AQP3. Finally, the C-terminal region of AQP3 is longer compared to AQP10. It is therefore of interest to decipher the significance in the structural and sequence differences between AQP3 and AQP10.

Methods

Examination of AQP3 and AQP10 for localization and glycosylation signals: The SwissProt Protein Sequences for human aquaporins AQP3 (Q92482) and AQP10 (Q96PS8) were retrieved and aligned using ClustalX [17]. The Prediction of N-Glycosylation sites for AQP3 and AQP10 was done with NetNGlyc 1.0 Server (www.cbs.dtu.dk/services/NetNGlyc/). Structural models of AQP3 and AQP10 were built with MODELLER (<http://www.salilab.org/modeller/>) using the high-resolution X-ray crystal structure of *E. coli* glycerol facilitator protein (1LDF) structure as a template (Protein Data Bank; <http://www.rcsb.org/>). The resultant models were subjected to Molecular Dynamics simulations and energy minimization using DISCOVER module of Insight II (Accelrys, Inc). Molecular dynamics simulations consisted of an initial equilibration of 5 pico seconds (ps) and followed by 100 ps dynamics at 300 K followed by a sequence of 10,000 steps of steepest descent and conjugate gradient energy minimization procedure. For all the above calculations, a distance-dependent dielectric constant and non-bonded distance cutoff of 20 Å were used. Molecular graphics images

were produced using SYBYL7.0 and the UCSF Chimera package from the Computer Graphics Laboratory, University of California, San Francisco.

Results and Discussion

Membrane trafficking proteins especially aquaporins found in the intestine are sorted proteins that show variations in functions and specificity. In order to understand the function of aquaporins it is essential to understand their modular three-dimensional architecture. In cellular membranes, four AQP monomers associate into a tetramer. To date only AQP1 has been crystallized and its biological unit (tetramer) has been reported [18]. In addition to providing four pores for bidirectional water transport, an additional pore formed by the biological unit transports ions (Fig. 1).

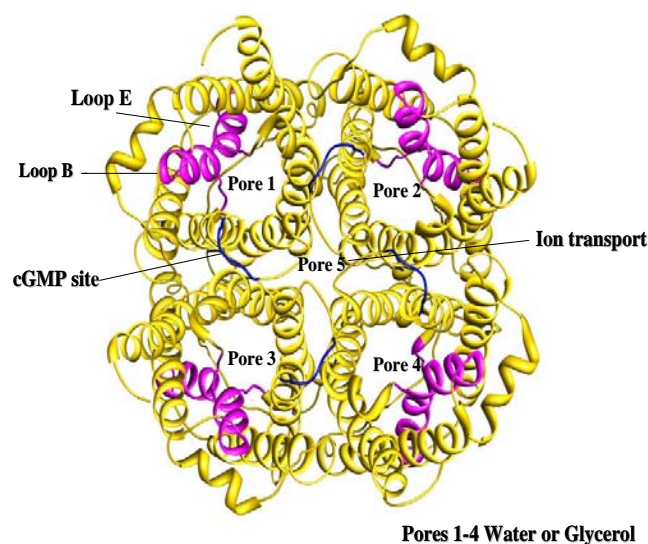


Figure 1: Illustration of aquaporin homotetramer. Water transport through pores 1-4 and an ion transport is predicted to occur via the central pore 5.

Due to high homology of AQPs, structure of water specific aquaporins can be predicted using AQP1 crystal structure as a model system. Several theoretical models of AQP3 and AQP10 were generated using the X-ray crystal structure of *E. coli* glycerol facilitator protein (1LDF) as a template. The final homology model was aligned with mammalian AQP1 crystal structures to predict/compare residue interactions and the overall structure-function relationship. The AQP1, AQP3 and AQP10 monomers contain 269, 292 and 301 amino acid residues respectively. Each AQP is comprised of two tandem repeats that fold in a similar manner, with three transmembrane (TM)-helices and one short connecting loop helix. Both the N- and C-terminus of AQP are to be found on the cytoplasmic side of the membrane.

There are six TM helices, plus two half-helices that form the overall helical bundle in the membrane. A single aqueous pathway is formed through the centre of the protein. There is a site at R41 of AQP1, within loop A

(extracellular), for attachment of a polylectosaminoglycan moiety; loop D, the cytoplasmic counterpart of loop A, does not have such a site. Loops B and E, each containing an asparagine-proline-alanine (NPA) motif and short helices B and E, bend into the six-helix bundle to form the channel. Sequence analysis of all human AQPs shows that the inter-motif distances of functional water only AQPs is 113 residues and that of aquaglyceroporins is 129 residues. The analysis also revealed a conserved [KR]xxxY motif at a distance of 14 residues from the first NPA .

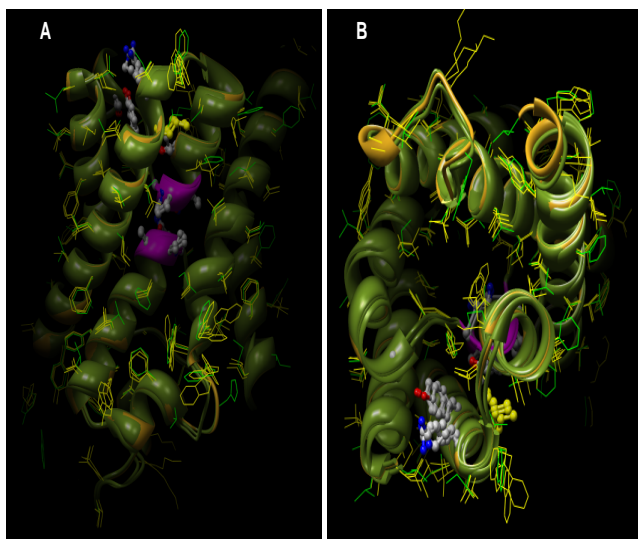


Figure 2 A & B: Hydrophobic residues are displayed in stick representation (AQP3 in green and AQP10 in yellow)

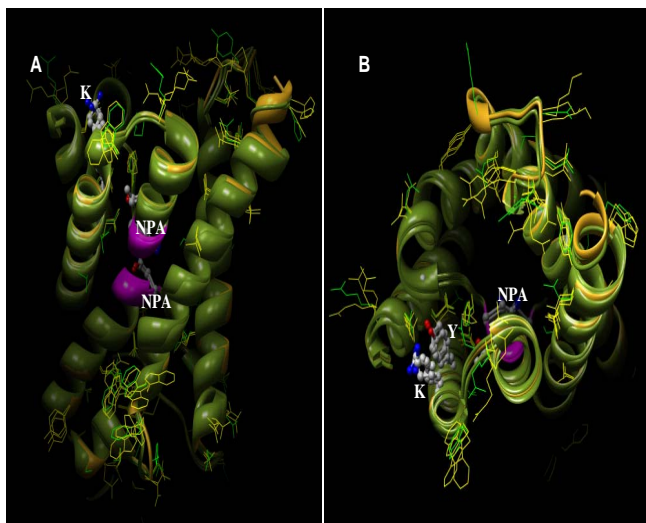


Figure 3 A & B: Polar residues are displayed in stick representation (AQP3 in green and AQP10 in yellow)

Differential display of hydrophobic residues from the side (Fig. 2A) and the top view (Fig. 2B) post structural superposition of AQP3 and AQP10 theoretical models reveal a strikingly similar orientation of corresponding conserved amino acids and motifs (Fig. 2) and a similar display of polar residues is illustrated in Fig 3. Three

dimensional molecular surface of the aquaglyceroporins from the side (Fig. 4A [solid surface] and Fig. 4B [transparent surface]) and the top view (Fig. 4C [solid surface] and Fig. 4D [transparent surface]) reveal the pore architecture and volume available for fluid transport. A residue level analysis shows that the AQP loops are held together by interactions between P77 and P193 in AQP1; P84 and P216 in AQP3; and P83 and P215 in AQP10. Also the positively charged arginine residues cap the N-terminal ends of helix-B and helix-E via hydrogen-bonding interactions. The positions of helices B and E are further stabilized by the ion-pair H74 with E17 in AQP1; H81-E28 in AQP3 and H80-E27 in AQP10, the salt bridge between R195-E142 in AQP1 and the corresponding arginine residue in AQP3 and AQP10 form hydrogen bonds with glutamine Q164 and Q163 respectively. In addition to providing comparative structural analyses of residue level interactions, the theoretical models have also provided an opportunity to understand the volume available for fluid transport and the critical residues on the periphery for future protein-protein or protein-small molecular interaction studies.

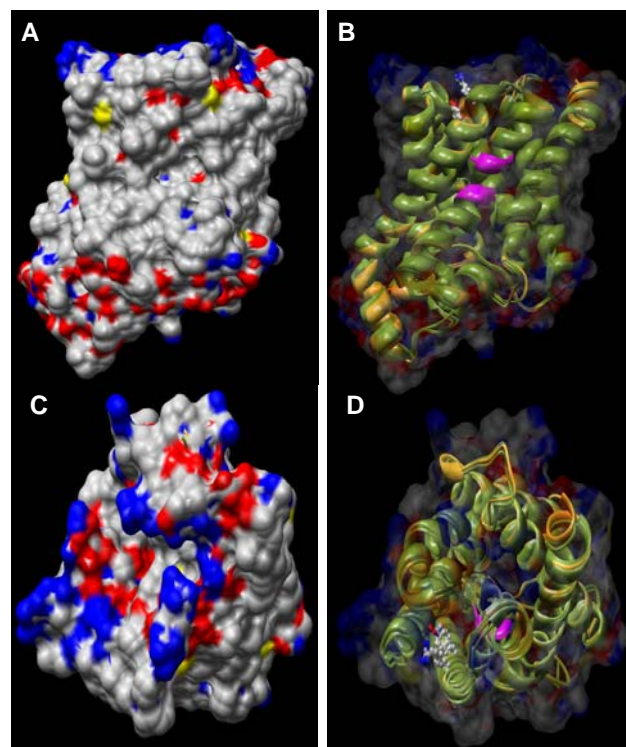


Figure 4 A-D: Surface rendering clearly shows polar areas and regions responsive to polar small molecular transport.

Sorting Signals

The N-terminus of AQP3 contains a four amino acid motif YRLL which includes the crucial tyrosine (Y) residue and the dileucine (LL) motif typical of basolateral targeting proteins. Alignment of AQP3 with that of AQP10 shows that AQP10 lacks the basolateral sorting motif (Fig. 5).

```

AQP03 MGRQKELVSRGEMLHIRYRLLRQALAECLGTLILVMFGCGSVAQVVLRSRGTHTGGFLTIN 60
AQP10 MVFTQAPAEIMG-HLRIRSLRLARQCLAEFLGVFVLMMLLTQGAQAQAVTSGETKGNFFTMF 59
* : .. * *:* * **.* **.:*::: *:*:* * * *:*:*:*:
AQP03 LAFGFVAVTLGILIAGQVSGAHLNPAVTFAMCFLAREPWIKLPIYTLAQTLAGAFLGAGIVF 120
AQP10 LAGSLAVTIAIYVGGNVSGAHLNPAFSLAMCIVGRLPWVKLPIYILVQLLSAFCASGATY 119
** .:*:*:* * :.*:*:*:*:*:*:*:*:*:*:*:*:*:*:* * * *.* * * *.* * : * :
AQP03 GLYDAIWHFADNQLFVSGPNGTAGIFATYPSGHLDMINGFFDQFIGTASLIVCVLAIVD 180
AQP10 VLYHDALQNYTGGNLTVTGPKETASIFATYPAPYLSLNNGLDQVLGTGMLIVGLLAILD 179
**:*:*: :*:..* *:*:*: *.*****: :*.: *:*:*:*:*:*: * * *:*:*:*:
AQP03 PYNPNVPRGLEAFTVGLVVLVIGTSMGFNSGYAVNPARDFGPRFLTALAGWGSVAVFTTQ 240
AQP10 RRNKGVPAGLEPVVVGMILALGLSMGANCGIPLNPARDLGPRFLTIVAGWGPEVFSAGN 239
* : * * * * .. * * : * : * * * * * . * * * * * : * * * * * : * * * * * :
AQP03 HWWWVPIVSPLLGSIAGVFVYQLMIGCHLEQPPP-----SNEEENVKLAHVKHKEQI-- 292
AQP10 GWWWVPPVAPLVGATYQLLVALHHPEGPEPAQDLVSAQHKASELETPASAQMLEC 299
*****:*:*:*:*: *. . *****: . * : * * : : : * : :
AQP03 --
AQP10 KL 301
    
```

Figure 5: Sequence alignment of human AQP3 and AQP10.

<i>Danio rerio</i>	gi 47087333 ref NP_998633.1	-----MGWQKSVLDKLAQTFQIRNKLLRQGLAECLGTLILVMFGCG 41
<i>Tribolodon hakonensis</i>	gi 18147584 dbj BAB83082.1	-----MGWQKAMLDKLAQTFRIIRNKLLRQGLAECLGTLILVMFGCG 41
<i>Anguilla anguilla</i>	gi 28207812 emb CAC85286.1	-----MGKQKPYLDKLAQTFQIRNRLRQGLAECLGTLILVMFGCG 41
<i>Dicentrarchus labrax</i>	gi 109639282 gb ABG36519.1	-----MGRHKVYLDKLSRFPQIRHLLRQALAECLGTLILVMFGCG 41
<i>Xenopus laevis</i>	gi 4090337 emb CAA10517.1	-----MGSQKDFVNKCQMLRLRNKLLRQALAECLGTLILVMFGCG 41
<i>Xenopus tropicalis</i>	gi 89273397 emb CAJ82459.1	-----MGRQKDFINKNQLLRLRNKLLRQALAECLGTLILVMFGCG 41
<i>Homo sapiens</i>	gi 4826645 ref NP_004916.1	-----MGRQKELVSRGEMLHIRYRLLRQALAECLGTLILVMFGCG 41
<i>Macaca mulatta</i>	gi 109111220 ref XP_001099318	-----MGRQKELMSRGEMLHIRHRLRQALAECLGTLILVMFGCG 41
<i>Bos taurus</i>	gi 119331228 ref NP_001073262	-----MGRQKELVNRGEMLHIRYRLLRQALAECLGTLILVMFGCG 41
<i>Canis familiaris</i>	gi 73971745 ref XP_854596.1	-----MGRQKELVSRGEMLHIRYRLLRQALAECLGTLILVMFGCG 41
<i>Mus musculus</i>	gi 121949813 ref NP_057898.2	-----MGRQKELMNRGEMLHIRYRLLRQALAECLGTLILVMFGCG 41
<i>Rattus norvegicus</i>	gi 13928980 ref NP_113891.1	-----MGRQKELMNRGEMLHIRYRLLRQALAECLGTLILVMFGCG 41
<i>Notomys alexis</i>	gi 22095319 gb AAM92149.1	-----EMLHIRYRLLRQALAECLGTLILVMFGCG 29
<i>Pan troglodytes</i>	gi 114624138 ref XP_001159733	-----MGRQKELVSRGEMLHIRYRLLRQALAECLGTLILVMFGCG 41
<i>Leishmania major</i>	gi 68128057 emb CAJ08141.1	-----EDQHESKRNFMSQNRWPLYRYRWRLREYVAEFPFGTFFLVTPGTG 82
<i>Trypanosoma brucei</i>	gi 70801809 gb AAZ11715.1	ENELVGDNADNEAHDADVNVYWAPRQLRLRYRNYMGEFLCTFVLLFMGNG 100
<i>Streptomyces coelicolor A3</i>	gi 21222460 ref NP_628239.1	-----MAERIVHRRGYYRKGGLGCEMCAEFLCTFVLLLPVGV 35
<i>Aedes aegypti</i>	gi 94469108 gb ABF18403.1	-----MKKSTLDTTISILLAEVGTGLLVLLGCM 28

Figure 6: Multiple sequence alignment of unique AQP3 obtained from NCBI GenBank

In order to check if this motif is conserved across the species, we examined all the AQP3 sequences deposited in GenBank. Of the mammalian AQP3 sequences (Fig. 6) only rhesus monkey (*Macaca mulatta*) lack the YRLL motif denoting the significance of the sorting signal (highlighted in Fig. 6), the sorting motif was absent in all the non-mammalian sequences examined. The absence of

the YRLL motif in AQP10 may be responsible for the differential compartmentalization. *Glycocalyx and Glycosylated Proteins*

Glycocalyx is the site for absorption of water, minerals, simple sugars and amino acids while the tight junctions regulate permeability of water and electrolytes

across intestinal epithelium and prevents leakage of macromolecules from the gut lumen. AQP3 has only one glycosylation site 141-NGT-143 in loop C while the AQP10 has three glycosylation sites 75-NVS-77; 128-NYT-130; and 133-NLT-135 in loops B & C. The additional glycosylation sites in AQP10 provide the necessary signal to sort the aquaporin in the apical region.

Conclusion

The fluid balance in the intestinal epithelium is controlled by the sorting of AQP3 and AQP10. Misrouting of aquaporins result in fluid imbalance, which may induce Crohn's disease and ulcerative colitis. Sorting signals YRLL in AQP3 and additional N-glycosylation sites in AQP10 lead to differential sorting. Our analysis provides valuable insight into membrane sorting of these functionally related AQPs and structural basis for drug discovery.

Acknowledgements

Mississippi NSF-EPSCoR "Innovations through Computational Sciences" Award (EPS-0556308); Research Centers in Minority Institutions (RCMI) – Center for Environmental Health at Jackson State University (NIH-NCRR G12RR13459-09); and NIH RIMI Grant 1P20MD002725-01 to Tougaloo College.

References

1. Agre, P.: The aquaporin water channels. *Proc. Am. Thorac. Soc.* **2006**, *3*, 5-13.
2. Asai, T.; Kuwahara, M.; Kurihara, H.; Sakai, T.; Terada, Y.; Marumo, F.; Sasaki S.: Pathogenesis of nephrogenic diabetes insipidus by aquaporin-2 C-terminus mutations. *Kidney Int.* **2003**, *64*, 2-10.
3. Rai, T.; Sasaki, S.; Uchida, S.: Polarized trafficking of the aquaporin-3 water channel is mediated by an NH2-terminal sorting signal. *Am. J. Physiol Cell Physiol* **2006**, *290*, C298-C304.
4. Wellner, R. B.; Cotrim, A. P.; Hong, S.; Swaim, W. D.; Baum B. J.: Localization of AQP5/AQP8 chimeras in MDCK-II cells: exchange of the N- and C-termini. *Biochem. Biophys. Res. Commun.* **2005**, *330*, 172-177.
5. Wellner, R. B.; Hong, S. A.; Cotrim, P.; Swaim, W. D.; Baum, B. J.: Modifying the NH2 and COOH termini of aquaporin-5: effects on localization in polarized epithelial cells. *Tissue Eng.* **2005**, *11*, 1449-1458.
6. Hardin, J. A.; Wallace, L. E.; Wong, J. F.; O'Loughlin, E. V.; Urbanski, S. J.; Gall, D. G.; MacNaughton, W. K.; Beck, P. L.: Aquaporin expression is downregulated in a murine model of colitis and in patients with ulcerative colitis, Crohn's disease and infectious colitis. *Cell Tissue Res.* **2004**, *318*, 313-323.
7. Dempsey, P. J.; Meise, K. S.; Coffey, R. J.: Basolateral sorting of transforming growth factor-alpha precursor in polarized epithelial cells: characterization of cytoplasmic domain determinants. *Exp. Cell Res.* **2003**, *285*, 159-174.
8. Le Mount, S.; Welling, P. A.; Brejon, M.; Olsen, O.; Merot, J.: Basolateral membrane expression of a K⁺ channel, Kir 2.3, is directed by a cytoplasmic COOH-terminal domain. *Proc. Natl. Acad. Sci. U.S.A* **2001**, *98*, 10475-10480.
9. Bonifacino, J. S.; Traub, L. M.: Signals for sorting of transmembrane proteins to endosomes and lysosomes. *Annu. Rev. Biochem.* **2003**, *72*, 395-447.
10. Marks, M. S.; Woodruff, L.; Ohno, H.; Bonifacino, J. S.: Protein targeting by tyrosine- and di-leucine-based signals: evidence for distinct saturable components. *J. Cell Biol.* **1996**, *135*, 341-354.
11. Letourneur, F.; Klausner R. D.: A novel di-leucine motif and a tyrosine-based motif independently mediate lysosomal targeting and endocytosis of CD3 chains. *Cell.* **1992**, *69*, 1143-1157.
12. Simmenm, T.; Nobile, M.; Bonifacino J. S.; Hunziker, W.: Basolateral sorting of furin in MDCK cells requires a phenylalanine-isoleucine motif together with an acidic amino acid cluster. *Mol Cell Biol.* **1999**, *19*, 3136-3144.
13. Chen, C.; Matesic, L. E: The Nedd4-like family of E3 ubiquitin ligases and cancer. *Cancer Metastasis Rev.* **2007**, *26*, 587-604.
14. Fuller, G. L.; Williams, J. A.; Tomlinson, M. G; Eble, J. A.; Hanna, S. L.; Pöhlmann, S.; Suzuki-Inoue, K.; Ozaki, Y.; Watson, S. P.; Pearce, A. C.: The C-type lectin receptors CLEC-2 and Dectin-1, but not DC-SIGN, signal via a novel YXXL-dependent signaling cascade. *J Biol Chem.* **2007**, *282*, 12397-12409.
15. Potter, B. A; Hughey, R. P.; Weisz, O. A.: Role of N- and O-glycans in polarized biosynthetic sorting. *Am J Physiol Cell Physiol.* **2006**, *290*, C1-C10.
16. Mobasher, A.; Shakibaei, M.; and Marples, D.: Immunohistochemical localization of aquaporin 10 in the apical membranes of the human ileum: a potential pathway for luminal water and small solute absorption. *Histochem. Cell Biol.* **2004**, *121*, 463-471.
17. Chenna, R.; Sugawara, H.; Koike, T.; Lopez, R.; Gibson, T. J.; Higgins, D. G.; Thompson, J. D.: Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* **2003**, *31*, 3497-3500.
18. Ren, G.; Cheng, A.; Reddy, V.; Melnyk, P; Mitra, A.K.: Three-dimensional fold of the human AQP1 water channel determined at 4 Å resolution by electron crystallography of two-dimensional crystals embedded in ice. *J. Mol Biol.* **2000**, *301*, 369-387.