Available online at www.sciencedirect.com
ScienceDirect
Biochimica et Biophysica Acta 1758 (2006) 989–993
BIOCHIMICA ET BIOPHYSICA ACTA
BBA
www.elsevier.com/locate/bbamem

Review

Aquaporin subfamily with unusual NPA boxes

Kenichi Ishibashi

Molecular Biology, Clinical research Center, Chiba-east Hospital, 673 Nitona, Chuuou-ku, Chiba 260-8712, Japan

Received 7 October 2005; received in revised form 18 February 2006; accepted 21 February 2006
Available online 20 March 2006

Abstract

Aquaporins have been identified based on highly conserved two asparagine–proline–alanine (NPA) boxes that are important for the formation of a water-permeating pore. Some aquaporin-like sequences, however, have less conserved NPA boxes. Although they have lower homology with conventional aquaporins, they should be included in aquaporin family based on their conserved six transmembrane domains and hydrophobic NPA box-like repeats. They are widely distributed in multicellular organisms. Only SIPs from plants and AQP11/12 from mammals were examined previously and found to be localized inside the cell. Intracellular localization will be a common feature of these aquaporin-like proteins since most of them have positively charged amino acid clusters at the carboxy-termini similar to di-lysine motif (–KKXX) for an endoplasmic reticulum retention signal. Accordingly, they are tentatively named subcellular-aquaporins in this review. Currently, studies on their functions and biological roles are limited. SIPs were shown to function as water channels and the disruption of AQP11 produced neonatally fatal polycystic kidneys. Further works on subcellular-aquaporins will reveal new insights into the roles of aquaporins.

Keywords: NPA box; Aquaporin; Subcellular; SIP; Polycystic kidney

Contents

1. Introduction

Although overall primary sequences are not well conserved (~30% identity with each other), all AQPs have two highly conserved hydrophobic asparagine–proline–alanine (NPA) repeats (NPA boxes), that form a pore for water and/or small molecules such as glycerol and urea. Remarkably, the upstream of the first NPA box and the downstream of the second NPA box are particularly conserved: SG(A/G)HXNPA and NPAR(S/D/A) as shown in the upper half in Table 1. In database, AQP-like sequences with less conserved NPA boxes are also present as shown in the lower half in Table 1. The first NPA box of these AQP-like sequences are highly deviated and diversified, while the second NPA boxes are relatively conserved, where the arginine (R) just after the second NPA is changed to valine (V) or alanine (A). NPA(V/A)A in the second NPA box seems be a clue to identify this new aquaporin subfamily as well as deviated NPA

E-mail address: kishiba@nitona.hosp.go.jp.

0005-2736/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbamem.2006.02.024

Table 1
Sequence alignments of aquaporins at the first and second NPA boxes and the carboxy-termini

	First NPA boxes	Second NPA boxes	Carboxy-termini
AQPZ	–VGHISGGHFNPAVTIGLWAG–	–SIPVTNTSVNPARSTAVAIFQG–	–GGLIYRTLLEKRD
GlpF	–TAGVSGAHLNPAVTIALWLF–	–MGPLTGFAMNPARDFGPKVFAW–	–KETTTPSEQKASL
Chlam	–VGHISGGKLNPAVSIGVLIG–	–GGGLTGAAMNPARAFGPALVSG–	–MLKDLTTLSTPSV
Ustil	–CATTSGGQFHPAFTIAQVVF–	–CFSSSNVVANSARDIGARLVCS–	–QKQRTNLGVKNF
A.nid	–YLATPSPACNPAISIIMALI–	–LGWQTGYAINPARDFGPRLFSA–	–DKAADRNGELRLD
A.fumig	–FYRVTGGLFNPVVSPTHELQ–	–GDYYTGGSLNPARSLGPDVINR–	–NRPVSGAEQV
Tryp1	–FGYISGGHFNPAVTMAVFLV–	–VGRISGGAFNPAAATGLQLALC–	–SAANGVAPVQ
Tryp2	–FGYISGAHFNPAITFATFIN–	–VGGFTGGAFNPAVATGTQLVGC–	–DRVAPIELSGQVF
Leish	–FGYISSSHFNPAVSIAVFLV–	–AGRISGGAFNPAAASGLQVAMC–	–ATTSWEGPTFNRR
TIP1.1	–GANISGGHVNPAVTFGAFIG–	–GGAFSGASMNPAVAFGPAVVSW–	–INTTHEQLPTTDY
PIP2.6	–TAGISGGHINPAVTFGLFLA–	–TIPITGTGINPARSFGAAVIYN–	–YGSVRSQLHELHA
NIP1.2	–LGHISGAHFNPAVTIAFASC–	–AGPVSGASMNPGRSLGPAMVYS–	–GSFLKTVRNGSSR
AQP1	–VGHISGAHLNPAVTLGLLLS–	–AIDYTGCGINPARSFGSAVLTR–	–NSRVEMKPK
AQP3	–AGQVSGAHLNPAVTFAMCFL–	–MGFNSGYAVNPARDFGPRLFTA–	–VKLAHMKHKEQI
AQP8	–LGNISGGHFNPAVSLAVTVI–	–GGSISGACMNPARAFGPAVMAG–	–DEKTRLILKSR
SIP1.1	–TVIFGSASFNPTGSAAFYVA–	–GSKYTGPAMNPAIAFGWAYMYS–	–KKQKKA
SIP1.2	–GNVLGGASFNPCGNAAFYTA–	–GSKFTRPFMNPAIAFGWAYIYK–	–KKQKKA
SIP2.1	–QQATKGGLYNPLTALAAGVS–	–GSDLTGGCMNPAAVMGWAYARG–	–PKAKSE
CeAQP9	–IEFQRDAVAHPCPLVTNCYR–	–GINYTGMYANPIVAWACTFNCL–	–EESEEQEKDTKKKE
CeAQP10	–NIFNRGAMTNCAPIFEQFVF–	–LYVVGVPGLNPIVATARLYGCR–	–KAEKKAKAAAKKSD
CeAQP11	–ALCNRTAFCSPLAPIEQYLF–	–VTFVGDQALDPLVASTLFFGCR–	–KEKKAEKKRAKKNE
Brigg9	–CYFQRDAVAHPCPLVTNCYR–	–GINYTGMYANPIVAWACTFNCL–	–EEQTESQKESKKTD
Brigg10	–GIFNRDAWTNCAPIFEQFIF–	–LYIVGVPGLNPIVATARLYGCQ–	–KKEKKAAAAAKKSD
Brigg11	–ALCNRTAFCSPLAPIEHYLF–	–VTFVGDQALDPLVASTLFYGCR–	–KEKKAEKKRAKKSE
Dros.me	–GRVWGDASACPYTHMEDVVE–	–AFNFSGGYFNPVLATALKWGCR–	–EGAASKSKQE
Dros.ps	–GKVWGDASACPYTHMEDVLE–	–AFNFSGGYFNPVLATALKWGCR–	–KLPIVRRFLLGE
Anoph	–GRNWGSATACPYTYMEQIVE–	–AFNYSGGYFNPVLATALKWGCA–	–FLTDTKTKSE
Urch1	–LTFDGDSTANTCMIWQSMLK–	–GLEWTGMMFNPALAAGITLNCG–	–NRPTPAVPPTKED
Urch2	–NEELSNAGDAPLGQAVQVQP–	–GLEYTGAPMNPILGFASGWGCK–	–NEVNGWVTSV
ZF1	–GFSFRGAICNPTGALELLSR–	–GGRLTGAVFNPALAFSIQFPCP–	–QQLNSNGLKKKKMK
ZF2	–TAVMQDVSGNPAVTLLRLLQ–	–ANNYTSGYVNPALAYAVTLTCP–	–RLPKGKTNDEKSS
Xeno	–GFTFNKASGNSAVSLQDFLL–	–AGSYTGAFFNPTLAAALTFQCS–	–RKAATLPAQKRSS
Chic11	–GLTLPGSTCNPCGTLQPLWG–	–GGNLTGAIFNPALAFSLHPHCF–	–KSFLGHQKTLKSQL
Chic12	–AACANGAASNPTVSLQEFLL–	–AAPATGAFFNPALATASTFLCA–	–QKGKGEKSNPAPRA
AQP11	–GLTLVGTSSNPCGVMMQMML–	–GGSLTGAVFNPALALSLHFMCF–	–WLHNNQMTNKKE
AQP12	–GVTLDGASANPTVSLQEFLM–	–AGPFTSAFFNPALAASVTFACS–	–SPGSVDAKMHKGE

The sequences above the line are conventional aquaporins (water selective aquaporins and aquaglyceroporins); the sequences below the line are subcellular-aquaporins. Highly conserved NPAs (asparagine–proline–alanine) are underlined. Except for SIPs, the second NPA boxes are conserved in subcellular-aquaporins: NPA(L/V/A/I)AXXXXXXC. Most of the carboxy-termini of subcellular-aquaporins are lysine (K)-rich, especially in nematodes, *C. elegans* and *C. briggsae*.
AQPZ and GlpF are from bacteria, *Escherichia coli* (NP_415396, NP_418362). Chlam is from a chlamydia (*parachlamydia* sp. UWE25; CAF23520). The next three are from fungi, Ustil (*Ustilago maydis* 521; XP_758316.1), A. nid (*Aspergillus nidulans* FGSC A4; XP_658434.1), and A. fumig (*Aspergillus fumigatus* Af293; EAL84488). The next three are from the protozoa, Tryp1/2 (*Trypanosoma cruzi*; XP_815990, AF31269.1) and Leish (*Leishmania major*; CAJ08765.1). TIP1.1 (tonoplast intrinsic protein), PIP2.6 (plasma membrane intrinsic protein), NIP1.2 (NOD26-like intrinsic proteins) are from a plant (*Arabidopsos thaliana*; P25818, Q9ZV07, Q8LFP7). AQP1/3/8 are from mice (NP_031498, NP_057898, NP_031500). SIP1.1, 1.2, 2.1 are from a plant (*Arabidopsos thaliana*; Q9M8W5, Q9FK43, Q9M1K3). CeAQPs and Briggs are from nematodes (*C. elegans*; NP_001021552.1, NP_496105.1, NP_499821.2, *C. briggsae*; CAE57975.1, CAE60164.1, CAE71341.1). The next three are from insects, Dros. me (*Drosophila melanogaster*; AAF58409.2), Dros. ps (*Drosophila pseudoodscura*; EAL25342) and Anoph (*Anopheles gambiae*; Xp_309823.2). Urch1/2 are from purple sea urchin (*Strongylocentrotus purpuratus*; XP_780933.1, XP_787329.1). ZF1/2 are from zebrafish (*Danio rerio*; AAH95775.1, AAH95564.1); Xeno from a frog (*Xenopus laevis*; AAH82904.1). Chic11/12 are from chickens (*Gallus gallus*; XP_424343.1, NP_001030011.1). AQP11/12 are from mice (NP_780314, NP_808255).

boxes. Furthermore, cysteine (C) at the downstream of the second NPA is also highly conserved except for plant SIPs (Small basic Intrinsic Proteins) [1]. In the literature, only SIPs and mammalian AQP11/12 have been reported. SIPs function as water channels and are localized subcellularly to the endoplasmic reticulum (ER) [2]. AQP11/12 are also localized inside the cell [3,4]. Accordingly, these aquaporin-like proteins with less conserved NPA boxes are tentatively named 'subcellular-aquaporins' in this review.

This review will be focused on the classification of a new aquaporin-like subfamily: subcellular-aquaporin family. The criteria for subcellular-aquaporins is the presence of less conserved sequences around 'both' NPA boxes. Thus, an aquaporin-like sequence with only one deviated NPA box should not be included and will be classified on the basis of the conserved NPA box. Although such a sequence is very rare, it may be an intermediate or transitional form between conventional AQPs and subcellular-aquaporins and merits future classification. Currently, this classification solely depends on the primary sequences. However, in the near future subcellular-aquaporins will be reclassified on the basis of their functions and/or cellular localizations.