Correspondence

Structural and functional features of the intracellular amino terminus of DEG/ENaC ion channels

Nektarios Tavernarakis*, John K. Everett*, Nikos C. Kyrpides† and Monica Driscoll*

The degenerin/epithelial sodium channel (DEG/ENaC) protein family includes related ion channel subunits from organisms ranging from the simple nematode *Caenorhabditis elegans* to humans. Members of this protein family have been implicated in functions as diverse as touch transduction and proprioception [1–4], pain sensation and maintenance of sodium balance [5].

Several blocks of sequence are conserved in DEG/ENaC subunits, but understanding structure/function relations in this channel class is in its infancy. There is only one conserved region in the intracellular amino termini of all DEG/ENaC family members, and this region is clearly critical for channel function (Figure 1). Several channel-inactivating substitutions in *C. elegans* degenerins affect the conserved domain [6,7] and a substitution in this motif in human βENaC causes the salt-wasting disorder pseudohypoaldosteronism, type I [8]. This region can influence ENaC gating properties [9,10] as well as ion permeation and selectivity properties [11] and has been implicated in endocytosis [12]. Still, how the conserved amino-terminal domain actually influences channel function remains a mystery.

We have analyzed all novel degenerin family members revealed by the *C. elegans* genome sequencing project [13] and found that, with no exception, they contain the intracellular amino-terminal conserved motif (Figure 1b). Based on the genetic properties of mutations altering key amino acids within this motif in two nematode degenerins, UNC-8 and MEC-4, we propose that this region might serve as an interaction domain that associates with other proteins that form the channel complex.

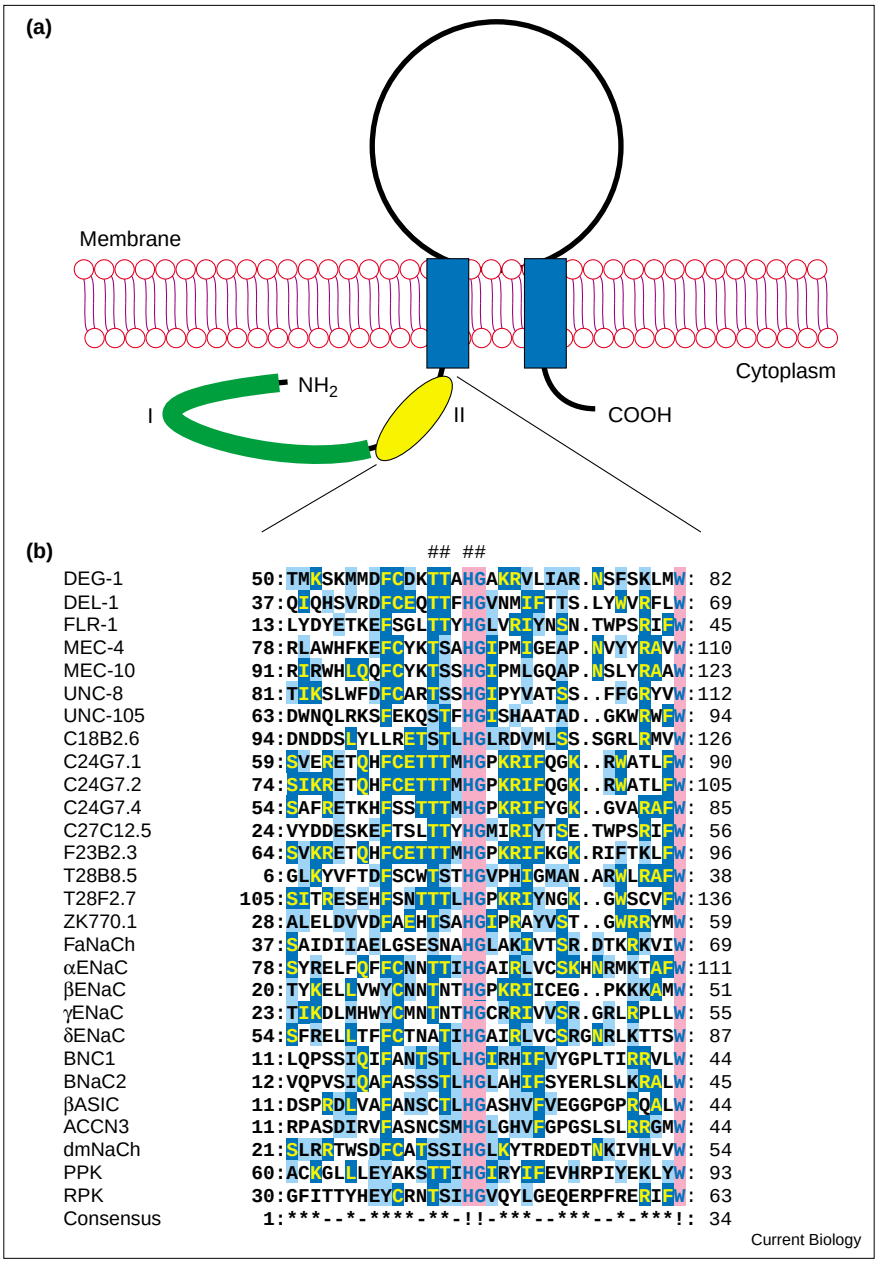
unc-8 is expressed in motor neurons and command interneurons of the nematode nervous system and has been implicated in proprioception and regulation of locomotion [6]. Dominant, gain-of-function mutations in the *unc-8* gene cause transient neuronal swelling and dysfunction and render the canonical sinusoidal movement of the worm severely uncoordinated. Absence of the UNC-8 protein in *unc-8* loss-of-function mutant strains results in a pronounced reduction of the amplitude and wavelength of the worm's normal sinusoidal movement [6].

Interestingly, the effects of a dominant mutation in the UNC-8 protein can be completely blocked by mutating the absolutely conserved histidine residue, within the conserved amino-terminal motif, to tyrosine, highlighting the functional importance of this motif. Suppression is observed both when the histidine substitution resides *in cis*, on the same protein molecule as the dominant mutation or *in trans*, on different co-expressed molecules, as observed in heterozygote animals carrying a dominant allele on one chromosome and a histidine-substituted allele on the other [14] (see also Supplementary material). Such a pattern of genetic suppression suggests that UNC-8 proteins interact as a dimeric or multimeric complex to form a channel. Mutations in the amino-terminal conserved region could disrupt critical interactions required for the formation of the channel complex.

More evidence which supports this hypothesis comes from studies of the related degenerin MEC-4. Overexpression of a MEC-4 amino-terminal fragment encompassing the conserved intracellular motif, disrupts endogenous channel function. However, these interfering properties depend on the integrity of the conserved motif. Substitution of any of three conserved amino acid residues within this sequence (T91, S92 and G95; see Figure 1b) results in loss of interference [7]. These observations are consistent with the hypothesis that this region could participate in critical interactions in the channel complex. Excess amino-terminal fragments might compete for interaction surfaces with endogenous MEC-4 molecules thereby disrupting the function of the channel. Such fragments, mutant at highly conserved positions possibly critical for interactions, would not be able to interfere when overproduced.

MEC-4 is required for normal body touch sensation in the nematode and was identified as one of the first candidate metazoan, mechanically gated ion channels [1]. Mechanosensitive channels are postulated to be tightly tethered to rigid structures both inside (cytoskeleton) and outside (extracellular matrix) the cell [15]. Such attachments are required to relay gating tension to the core channel subunits. The topology of MEC-4 is such that both the amino terminus and the carboxyl terminus are intracellular with a large extracellular loop between two transmembrane domains [16]. The intracellular amino terminus is 110 amino acids long while the intracellular carboxy-terminal part has a length of only 15 amino acids. The amino-terminal part contains highly conserved motifs whereas the carboxyl terminus has diverged within DEG/ENaC proteins. The extended amino terminus is

Figure 1



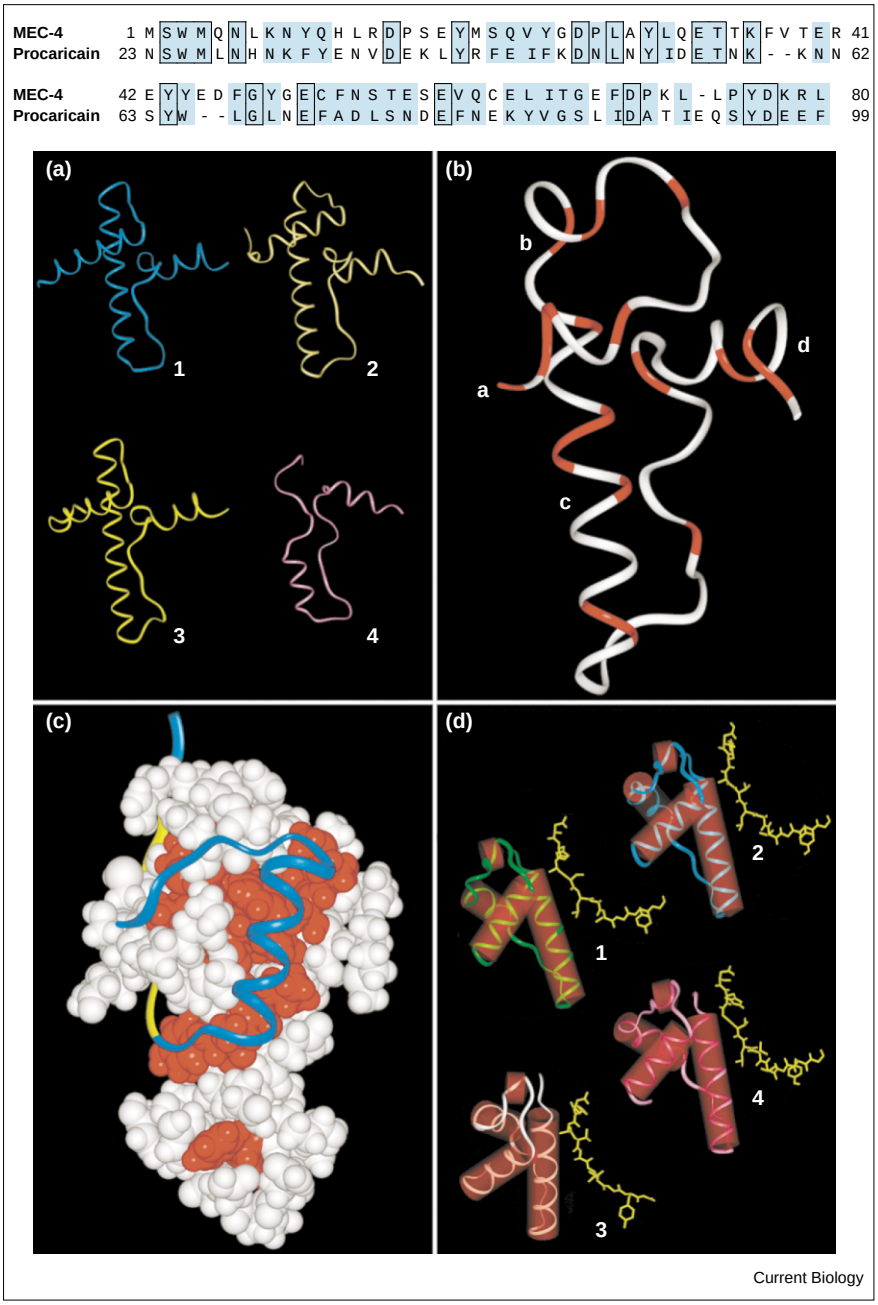
The unique conserved region in the amino-terminal intracellular domain of DEG/ENaC family channels. (a) Transmembrane topology of DEG/ENaC family members. Amino termini are always intracellular, with the most amino-terminal region (sub-domain I) variable in sequence and in length. Sub-domain II includes a highly conserved region of approximately 33 amino acids positioned near the first transmembrane domain, which is thought to start at the conserved tryptophan residue. Relative lengths of intracellular amino and carboxyl termini are as in MEC-4 degenerin; the extracellular domain is not drawn to scale. (b) The conserved intracellular amino-terminal domain of DEG/ENaC family members. Listed are DEG/ENaC family members from nematodes, snails, flies, and mammals. Nine new uncharacterized family members from the *C. elegans* genome are also included in the alignment. # symbols on top of the alignment denote amino acid positions affected in mutant proteins from humans and nematodes. Amino acids substituted in mutant proteins from nematodes (MEC-4, T91I, S92F, G95E; UNC-8 H114Y) and humans (βENaC G37S) are underlined for clarity. Pink boxes indicate 100% amino acid identity, dark blue boxes indicate >50% amino acid sequence identity and light blue shading indicates conservative amino acid substitutions. In the consensus line, ! indicates 100%, * >50% and - <50% sequence similarity. See Supplementary material for detailed methods and sequence accession numbers.

therefore the domain likely to mediate critical intracellular interactions for the assembly and function of the MEC-4 channel.

To gain insight into the structure and function of the MEC-4 amino terminus we constructed a three-dimensional model of this domain based on amino acid sequence similarity to proteins of solved structure. Our model is primarily based upon sequence similarity with the pro-domain of procaricain, a eukaryotic thiol protease (25% identity, 62% similarity over 80 residues; Figure 2, top). The procaricain pro-domain is structurally homologous to the pro-domains of members of the cathepsin and other protease families. (Figure 2a). Our model includes four α-helices: a long central helix (helix c) flanked by two smaller helices (helix a and d) with a very small helical segment (helix b) residing between helix a and helix c (Figure 2b). One extended chain packs against helix c to form a hydrophobic core (Figure 2b). An interesting feature of the model is an exposed patch of three aromatic residues — F37, Y43, and Y44 — which may form a critical domain for protein–protein association (Figure 2c).

The remaining 30 residues of the MEC-4 amino terminus (residues

Figure 2



A three-dimensional model of the MEC-4 amino terminus. Top. MEC-4 sub-domain I (amino acids 1–80) is similar to procaricain prodomain; identical residues are boxed, similar residues are shaded. **(a)** The pro-domain of protease procaricain (1: 1PCI) is structurally similar to several pro-domains in cathepsin family members including procathepsin L (2: 1CS8), procathepsin K (3: 7PCK), procathepsin B (4: 3PBH). **(b)** Ribbon diagram of the proposed model of MEC-4 amino acids 1–80. Red are hydrophobic residues, white are hydrophilic residues; letters a–d indicate the four predicted helices that are characteristic of protease pro-domains. **(c)** CPK rendering for MEC-4 sub-domain I (amino acids 1–80) also showing the hypothesized association with sub-domain II (amino acids 81–111). Sub-domain II is depicted as blue ribbon with the conserved motif highlighted in yellow. An amphipathic helix of sub-domain II may pack with the large hydrophobic patch (red) situated on sub-domain I. **(d)** Comparison of packing of 4 solved structures of protease prodomains (colored ribbon) relative to their histidine active sites (yellow sticks). 1, cathepsin K (1BY8); 2, cathepsin H (8PCH); 3, cathepsin L (1CS8); 4, papain (1PEG). (See Supplementary material for methods and atom coordinates for the model.)

81–111, sub-domain II) that comprise the conserved intracellular motif present in all DEG/ENaC family members, are not homologous to procaricain nor any other protease pro-domains of known structure. However, it is intriguing that by searching motif databases, we have detected the histidine active site signature of a cathepsin thiol protease within this conserved region (see Supplementary material for alignment and description of the search and statistical analyses). The conserved region in DEG/ENaC channels corresponds to one domain of a tripartite active site in thiol proteases, which in the cathepsin family folds together with two additional domains to create the functional catalytic center [17]. The motif is intact in *C. elegans* MEC-4 and MEC-10 and is highly conserved in several additional degenerins. The histidine residue corresponding to the critical histidine active site core and the flanking small sidechain amino acid (glycine) are 100% conserved in all DEG/ENaC family members. Interestingly, although the

motif is more divergent in the ENaC subfamily, we have detected a Kunitz-type protease inhibitor motif preceding the conserved amino-terminal region in αENaC. In this context, it is intriguing that two nematode proteins, MEC-2 and UNC-1, that have been proposed to interact with and modulate degenerin channels, contain a domain implicated in regulating proteolysis of membrane-associated proteins [18].

We predicted the folding of sub-domain II based on secondary structure prediction algorithms and shared sequence similarity with a napthalene dioxygenase and a histone acetyltransferase domain (see Supplementary material). Available data suggest that sub-domain II is most likely to exist primarily as random coil with the exception of a 3-turn amphipathic α-helix preceding the protease histidine active site motif. The large hydrophobic patch in sub-domain I might readily accept the amphipathic α-helix found in sub-domain II (Figure 2c). Intriguingly, the helices that comprise our model are of comparable lengths and orientations to those comprising the core of several thiol proteases, though threaded in a different order. The proposed homology model places the histidine active site motif in an orientation analogous to that in thiol proteases, facing helix c and roughly perpendicular to helix a and helix d (Figure 2d). However, since sub-domain II is positioned only a few residues away from a transmembrane region, it may be somewhat extended from the folded amino terminus, a position that could enable it to associate with a complementary domain supplied by another protein.

Structural studies are clearly required to test the three-dimensional MEC-4 amino terminus homology model. Such studies are not trivial however, because of inherent difficulties in efficiently expressing degenerins in a heterologous expression system (our unpublished observations). An additional difficulty could arise from the transmembrane nature of these proteins. Our model provides a first approximation to the structure of one of the most functionally important domains in a candidate mechpanosensory channel subunit. Moreover, it holds interesting implications that can be experimentally tested for the function of the MEC-4 amino terminus and the role of the intracellular motif that is strikingly conserved in all DEG/ENaC proteins.

Supplementary material

Supplementary material including methodological details and accesssion numbers is available at http://current-biology.com/supmat/supmatin.htm.

References

1. Tavernarakis N, Driscoll M: Molecular modeling of mechanotransduction in the nematode ***Caenorhabditis elegans***. *Annu Rev Physiol* 1997, 59:659-689.
2. Adams CM, Anderson MG, Motto DG, Price MP, Johnson WA, Welsh MJ: Ripped pocket and pickpocket, novel ***Drosophila*** DEG/ENaC subunits expressed in early development and in mechanosensory neurons. *J Cell Biol* 1998, 140:143-152.
3. Drummond HA, Price MP, Welsh MJ, Abboud FM: A molecular component of the arterial baroreceptor mechanotransducer. *Neuron* 1998, 21:1435-1441.
4. Price MP, Lewin GR, McIlwrath SL, Cheng C, Xie J, Heppenstall PA, Stucky CL, Mannsfeldt AG, Brennan TJ, Drummond HA, *et al.*: The mammalian sodium channel BNC1 is required for normal touch sensation. *Nature* 2000, 407:1007-1011.
5. Alvarez de la Rosa D, Canessa CM, Fyfe GK, Zhang P: Structure and regulation of amiloride-sensitive sodium channels. *Annu Rev Physiol* 2000, 62:573-594.
6. Tavernarakis N, Shreffler W, Wang S, Driscoll M: ***unc-8***, a DEG/ENaC family member, encodes a subunit of a candidate mechanically gated channel that modulates ***C. elegans*** locomotion. *Neuron* 1997, 18:107-119.
7. Hong K, Mano I, Driscoll M: ***In vivo*** structure-function analyses of ***Caenorhabditis elegans*** MEC-4, a candidate mechanosensory ion channel subunit. *J Neurosci* 2000, 20:2575-2588.
8. Chang SS, Grunder S, Hanukoglu A, Rosler A, Mathew PM, Hanukoglu I, Schild L, Lu Y, Shimkets RA, Nelson-Williams C, *et al.*: Mutations in subunits of the epithelial sodium channel cause salt wasting with hyperkalaemic acidosis, pseudohypoaldosteronism type 1. *Nat Genet* 1996, 12:248-253.
9. Grunder S, Firsov D, Chang SS, Jaeger NF, Gautschi I, Schild L, Lifton RP, Rossier BC: A mutation causing pseudohypoaldosteronism type 1 identifies a conserved glycine that is involved in the gating of the epithelial sodium channel. *EMBO J* 1997, 16:899-907.
10. Grunder S, Jaeger NF, Gautschi I, Schild L, Rossier BC: Identification of a highly conserved sequence at the N-terminus of the epithelial Na^+ channel α subunit involved in gating. *Pflugers Arch* 1999, 438:709-715.
11. Coscoy S, de Weille JR, Lingueglia E, Lazdunski M: The pre-transmembrane 1 domain of acid-sensing ion channels participates in the ion pore. *J Biol Chem* 1999, 274:10129-10132.
12. Chalfant ML, Denton JS, Langloh AL, Karlson KH, Loffing J, Benos DJ, Stanton BA: The NH(2) terminus of the epithelial sodium channel contains an endocytic motif. *J Biol Chem* 1999, 274:32889-32896.
13. The *C. elegans* Sequencing Consortium: Genome sequence of the nematode ***C. elegans***: a platform for investigating biology. *Science* 1998, 282:2012-2018.
14. Shreffler W, Magardino T, Shekdar K, Wolinsky E: The unc-8 and sup-40 genes regulate ion channel function in ***Caenorhabditis elegans*** motorneurons. *Genetics* 1995, 139:1261-1272.
15. Sackin H: Mechanosensitive channels. *Annu Rev Physiol* 1995, 57:333-353.
16. Lai CC, Hong K, Kinnell M, Chalfie M, Driscoll M: Sequence and transmembrane topology of MEC-4, an ion channel subunit required for mechanotransduction in ***Caenorhabditis elegans***. *J Cell Biol* 1996, 133:1071-1081.
17. Rawlings ND, Barrett AJ: Families of cysteine peptidases. *Methods Enzymol* 1994, 244:461-486.
18. Tavernarakis N, Driscoll M, Kyrpides NC: The SPFH domain: implicated in regulating targeted protein turnover in stomatins and other membrane-associated proteins. *Trends Biochem Sci* 1999, 24:425-427.

Acknowledgements

We thank Peter Kahn for providing resources essential for this work. Research was supported by grants from the National Institutes of Health (NS34435, NS37955) and partly by Integrated Genomics Inc. (Chicago, Illinois). N.T. was supported by a Research Fellowship from Human Frontiers in Science Program Organization.

Addresses: *Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey, Nelson Biological Laboratories A220, 604 Allison Road, Piscataway, New Jersey 08855, USA. †Integrated Genomics Inc., 2201 W. Campbell Park Drive, Chicago, Illinois 60612, USA.

Correspondence: Nektarios Tavernarakis
E-mail: tavernarakis@mbcl.rutgers.edu