

Supplemental Information:

Characterization of the honeybee AmNav_v1 channel and tools to assess the toxicity of insecticides

Pascal Gosselin-Badaroudine¹, Adrien Moreau¹, Lucie Delemotte², Thierry Cens⁴, Claude Collet⁵, Mathieu Rousset⁴, Pierre Charnet⁴, Michael L. Klein², and Mohamed Chahine^{1,3}

¹Centre de recherche, Institut universitaire en santé mentale de Québec, Quebec City, QC, Canada G1J 2G3

²Institute of Computational Molecular Science, Temple University, Philadelphia, PA 19122 USA

³Department of Medicine, Université Laval, Quebec City, QC, Canada G1K 7P4

⁴Centre de Recherche de Biochimie Macromoléculaire (CRBM), CNRS UMR 5237, 1919 Montpellier, France.

⁵INRA UR 406, Abeilles et Environnement, 228 Route de l'Aérodrome, Domaine Saint Paul - Site Agroparc, CS40509, 84914 Avignon Cedex 9, France

Short title:

Functional expression and characterization of the honeybee Nav_v1 channel

Key words: Apis mellifera, voltage gated ion channel, voltage gated sodium channel, insecticide toxicity, Nav_v1

Correspondence to:

Mohamed Chahine, Ph.D.

Centre de recherche

Institut universitaire en santé mentale de Québec

2601 chemin de la Canardière

Quebec City, QC, Canada G1J 2G3

Telephone: 1-418-663-5747, ext.4723

Fax: 1-418-663-8756

Email: mohamed.chahine@phc.ulaval.ca

Supplemental results: sequence and features of the honeybee's Nav1 channel regulatory subunits

Five regulatory subunits for the Nav1 channel have been identified previously in *Drosophila Melanogaster*¹. Interestingly, these subunits were reported to be similar to the regulatory subunits of mammalian calcium-activated potassium channels (K_{Ca}) which are distant cousins of Nav channels in the VGL-kanome². Screening a HMM protein profile inferred from K_{Ca}β subunits against the bee proteome resulted in the successful identification of the TipE protein and its homologs (e-values from 3,5E-66 to 5,8E-68). Unlike the Nav1 sequence characterization, the TipE and TEH sequence characterizations were performed with the jackhammer algorithm. This algorithm consists in iterative hmmsearch-like queries. Each new query uses a protein profile inferred from the selected proteins associated with the significant matches of the previous query. This procedure was performed in order to make sure that all TipE homolog proteins were identified from the bee proteome. Using a similar approach with a protein profile inferred from mammalian regulatory subunits of Nav channels did not highlight any homologous protein in the bee proteome.

Comparison of the identified subunits with the known K_{Ca}β (*KCNMB*) subunits shows that, these proteins belong to a single protein family (**Figure 1D**). Moreover, based on sequence homology, the TEH family is divided into two subfamilies. TEH3 and TEH4 seem to feature sequence motifs slightly different from TipE, TEH1 and TEH2

All the regulatory subunits of AmNav1 are located within a span of 2 686 kB in the LG1 chromosome (**Figure S2C**). The TipE, TEH2, TEH3 and TEH4 genes are all located in closer proximity (within 74 kB). All the TEH subunits are oriented in the same direction. Similar arrangements have been reported for the TipE and TEH subunits of the *Drosophila Melanogaster*, *Drosophila Pseudoobscura*, *Drosophila Yakuba* and *Anopheles gambiae*¹. Such an arrangement either points to early gene duplication and inversion events in the phylogeny or to common regulation of the genes.

Using the SOSUI web server³, we identified two putative transmembrane segments in all regulatory subunit but TEH1. Given that all the putative transmembrane segments identified are in the same regions and that drosophila's TEH1 subunit is thought to feature

two transmembrane segments¹, it is probable that the honeybee's TEH1 also has two transmembrane segments in the same regions (**Figure S3**).

As noted by Derst and collaborators for the drosophila's TipE and TEH subunits¹, some cysteines thought to be involved in the formation of disulphide bridges in hK_{Ca}β2 are conserved in the honeybee's Nav₁ regulatory subunits¹. Other cysteines not found in hK_{Ca}β2 seem to be highly conserved in both drosophila's and honeybee's regulatory subunits (**Figure S3A-B**). Those residues may be involved in the formation of disulphide bridges as well. Furthermore, TEH3 and TEH4 both feature two EGF-like domains. Those domains are typically associated with protein-protein interactions and calcium binding. TEH3's first EGF-like domain seems to be involved in calcium binding as it features the D/N-x-D/N-E/Q-x₁-D/N*-x₂-Y/F consensus sequence (where x₁ and x₂ are of variable length and * represents β-hydroxylated residue). TEH4's first EGF-like domain, however, only features some residues of the consensus sequence.

No alternative sequences were found for the honeybee's TEH1, TEH2 and TEH3 subunits. Nevertheless, alternative sequences for TipE and TEH4 were found as a result of genomic variations and RNA editing (**Figure S2B**).

References:

- 1 Derst, C., Walther, C., Veh, R. W., Wicher, D. & Heinemann, S. H. Four novel sequences in *Drosophila melanogaster* homologous to the auxiliary Para sodium channel subunit TipE. *Biochem Biophys Res Commun* **339**, 939-948, doi:10.1016/j.bbrc.2005.11.096 (2006).
- 2 Yu, F. H. & Catterall, W. A. The VGL-chnome: a protein superfamily specialized for electrical signaling and ionic homeostasis. *Sci Signal* **2004**, re15, doi:10.1126/stke.2532004re15 (2004).
- 3 Hirokawa, T., Boon-Chieng, S. & Mitaku, S. SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14**, 378-379 (1998).

Supplemental methods: cloning AmNav₁ and its regulatory subunits

Total RNA from honeybee heads was extracted using Trizol kits (Sigma). cDNA was produced using Transcriptor first strand cDNA synthesis kits (Roche). The cDNA corresponding to two overlapping fragments of the Nav₁, the TipE, and TEH1-4 proteins was obtained by PCR amplification. Prior to insertion in a plasmid with a standard

digestion and ligation method, restriction sites were introduced at the beginning and end of each fragment cloned using PCR amplification with the following primers:

TipE :

Forward primer with **BamH1** restriction site:

GAGAG**GGATCC**GCCGCCACCAT**ATG**GCGGAGGAAAAGGAGAAG

Reverse primer with **HindIII** restriction site:

GAGAA**AAGCTT**GCTTCAGACTGCCGCCGT

TEH1

Forward primer with **BamH1** restriction site:

GAGAG**GGATCC**GCCGCCACCAT**ATG**CGCGGGAGCAGCTCTGA

Reverse primer with **EcoR1** restriction site:

GAGAGA**AATTC**TACCTGAGATCGTTGCCCTG

TEH2

Forward primer with **BamH1** restriction site:

GAGAG**GGATCC**GCCGCCACCAT**ATG**GGCGCCCAGGAAGAG

Reverse primer with **HindIII** restriction site:

GAGAA**AAGCTT**TTACTCTTCCTTCCTTGCGTACTTGT

TEH3

Forward primer with **BamH1** restriction site:

GAGAGGGATCCGCCGCCACCATGCCTAAGCAAGTGCCAGTG

Reverse primer with EcoR1 restriction site:

GAGAGGAATTCTCAGAGCGCCATAGCGGAGT

TEH4

Forward primer with EcoR1 restriction site:

GAGAGGAATTCGCCGCCACCATGGGTCGAAAGCATAAACG

Reverse primer with HindIII restriction site:

GAGAAAGCTTCTGCTAAAGCGCCATGCT

Nav1 fragment #1 (from ATG to the 4131th nucleotide, starting with A from ATG)

Forward primer with Kpn1 restriction site:

GAGAGGGTACCGCCGCCACCATGTCCGAAGATTCTGACTC

Reverse primer:

GCCAGCGCCACAGAGCGACGCTACGAAGTTAATGAGTGAC

Nav1 fragment #2 (from 4092th nucleotide to the stop codon)

Forward primer:

GTCACTCATTA~~ACTTCGTAGCGTCGCTCTGTGGCGCTGGC~~

Reverse primer with Xma1 restriction site:

GAGACCCGGGTCAGACGTCCGCGGTTCTCG

The cDNA of all the regulatory subunits were inserted into the pPol_Not1 vector (a generous gift from Dr. Paul Isenring, Université Laval), an oocyte expression vector

containing the T7 promoter (5' to 3'), the *Xenopus laevis* β -globin 5'-untranslated region, a multiple cloning site, the *Xenopus laevis* β -globin 3'-untranslated region, a linearizing site, and polyA and polyC tracts. The cDNA corresponding to the two Nav1 fragments were inserted in the TOPO vector.

The vectors containing the regulatory subunits were amplified in *Escherichia coli* XL2 Blue (Agilent) and were purified using GenElute HP Plasmid Maxiprep kits (Sigma).

The constructs were linearized with Not1, and T7 RNA polymerase was used to make sense RNA using mMESSAGE mMACHINE T7 kits (Ambion). mRNA was also generated using mMESSAGE mMACHINE kits with a custom DNA featuring a T7 promoter, the AmNav1 channel sequence, the *Xenopus laevis* β -globin 3'-untranslated region, and polyA and polyC tracts.

To identify sequence variations, genomic DNA from whole bees was extracted and was sequenced using QIAamp DNA mini kits (Qiagen).

Supplemental methods: oligonucleotide primer sequences for tissue expression analysis

The following primer sequences were used for RT-PCR amplifications in order to visualize the tissue expression pattern of the proteins under investigation:

For AmNav1

5'-GATCCGATGCTCGAACAAGGGC-3' (forward primer)

5'-GGGCATGATCATGAGTATGCAGGC-3' (reverse primer)

For TipE

5'-AAGTGGACCCCGGGATAGTCATCAGC-3' (forward primer)

5'-CGACCTTCAGCTGATGCCCAAAGC-3' (reverse primer)

For TEH1

5'-CATCTGCACCACCTCCAGACGC-3' (forward primer)

5'-ATTTCGCGGGTGAAGTTCTCGGC-3' (reverse primer)

For TEH2

5'-TAAATTCTACACCTCCCTCTGCCTGC-3' (forward primer)

5'-CACTGTATGGAGCCCAACGGTTTCGC-3' (reverse primer)

For TEH3

5'-TCTACCTGACCGTTGCCATCTACAGC-3' (forward primer)

5'-TTCGACACTCGTCCGCTATGCGC-3' (reverse primer)

For TEH4

5'-GAACAAAGCGAATGGAAATCTGCCGC-3' (forward primer)

5'-TAGAGGCGAGAACTGTTGTAGATGC-3' (reverse primer)

Supplemental methods: Equations

The fraction of channels modified by pyrethroid insecticides was calculated using the following equation:

$$M(\%) = 100 \times \frac{I_{tail}/(E_{tail} - E_{rev})}{I_{dep}/(E_{dep} - E_{rev})}$$

where I_{tail} is the maximal amplitude of the tail current measured in the 170s following the last conditioning pulse (the first 5ms are excluded to avoid the inclusion of a capacitive component in the current measured), E_{tail} is the membrane potential at which the tail current was measured, E_{rev} is the reversal potential of sodium ions as calculated using an I-V curve, and I_{dep} is the peak current observed in response to a depolarization at the E_{dep} potential.

The dose-response curves obtained from these calculations were fitted to a Hill curve using the following equation:

$$M = \frac{M_{Max}}{1 + (EC_{50}/dose)^h}$$

where h is the Hill coefficient, M_{max} is the maximal fraction of channels affected, EC_{50} is the half-maximal effective concentration, and $dose$ is the concentration at which the observation was made.

The voltage-dependence of activation was fitted to the following Boltzmann equation:

$$\frac{G(V)}{G_{Max}} = \frac{1}{1 + \exp\left(\frac{V-V_{1/2}}{k}\right)}$$

while the voltage-dependence of inactivation was fitted to the following equation:

$$\frac{I(V)}{I_{Max}} = \frac{1}{1 + \exp\left(\frac{V-V_{1/2}}{k}\right)}$$

where $V_{1/2}$ is the half-maximal voltage of activation (or inactivation), G is the conductance, I is the current measured at a given voltage (V), and k is the slope factor (in mV).

Current decay and tail current kinetics were fitted to the following equation when a single exponential was used:

$$y = y_0 + A \left(1 - \exp\left(\frac{-x}{\tau}\right)\right)$$

and to the following equation when double exponential was used:

$$y = y_0 + A_1 \left(1 - \exp\left(\frac{-x}{\tau_1}\right)\right) + A_2 \left(1 - \exp\left(\frac{-x}{\tau_2}\right)\right)$$

where y_0 is the initial value, A is the weight of the exponential, and τ is the time constant of the exponential.

All fits were performed using the Levenberg-Marquardt algorithm.

Supplemental tables:

Table S1: Correspondence of residues in the honeybee Nav1 sequence with known kdr-linked mutations

Table S2: Residues implicated in the formation of the pyrethroid binding site in the honeybee's Nav1 channel according to the participating residues in *Aedes aegypti*.

Supplemental figure legends:

Figure S1: Sequence alignment of the homologous domains of the honeybee Nav1 channel. **(A)** The homologous domains of AmNav1 are aligned with the sequences of the homologous domains from various mammalian sodium channels. The crystallized bacterial sodium channel NavAb was used as reference to identify the 6 transmembrane segments comprised in each domain. Conserved residues are annotated with “*”, “:” or “.” according to the ClustalW consensus representation. The highly conserved amino acids which form the GCTC, the selectivity filter (SF) and the conserved positive charges on S4 are highlighted. Conserved negative, positive, aromatic and uncharged residues are shown in red, blue, green and grey respectively. **(B)** Sequence alignment of the putative hydrophobic plug for AmNav1. The IFMT motif and its MFMT counterpart are shown in bold. The putative phosphorylation site identified with pBlast is highlighted in orange. **(C)** 2-D representation of the motifs found in the honeybee Nav1 channel.

Figure S2: Sequence variants of the investigated proteins and chromosomal organization of the AmNav1 regulatory subunits. **(A-B)** Sequence variants found for the AmNav1 channel and its regulatory subunits. Green annotations correspond to insertions whereas the red annotation represents a deletion. Red and green circles represent alterations which are the result of genomic variations and RNA edition respectively.

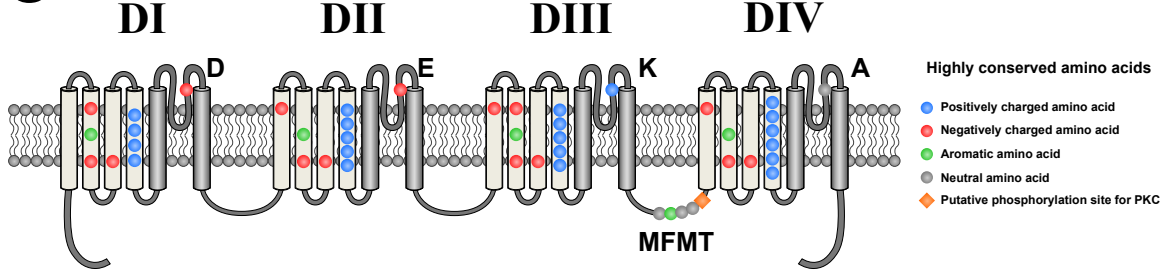
Figure S3: Sequence alignment the TipE and TEH regulatory subunits. **(A)** The regulatory subunits are aligned with the hK_{Ca}β2 protein which modulates the biophysical properties of the mammalian K_{Ca} channels. Conserved residues are annotated with “*”, “:” or “.” according to the ClustalW consensus representation. The transmembrane segments identified with SOSUI are shown in bold. The regions where most transmembrane segments are found are framed. The cysteines thought to be important for the formation of disulphide bridges in hK_{Ca}β2 are highlighted in grey. Highly conserved cysteines found in the drosophila’s and honeybee’s Nav1 regulatory subunits are highlighted in orange. **(B)** Alignment of the two EGF-like domains found in TEH3 and TEH4. Cysteines involved in disulphide bonds are highlighted in grey. Residues in the first EGF-like domain potentially involved in calcium binding are shown in bold. **(C)** 2-D representation of the regulatory subunits.

Figure S4: AmNav_{v1} single-channel conductance. **(A)** Single-channel traces from cell-attached oocytes patch are illustrated for -10, -20, -30 and -40 mV. Arrows indicate the onset of patch depolarization from -100 mV to the indicated voltage for 150 ms. **(B)** All-points histogram of the single-channel current amplitudes at -30 mV. The smooth curve is a Gaussian fit with amplitude peaks at 1.77 pA. **(C)** Plot of the current–voltage relationship ($n \geq 7$). The straight line is a linear regression yielding a unitary conductance of 24.5 ± 0.2 pS. Data are presented as mean \pm SEM.

Figure S5: Effect of TTX on AmNav_{v1}. **(A)** The channel exhibited a reversible, dose-dependent block by TTX. Using oocytes expressing AmNav_{v1} and TipE, the fraction of channels blocked at different concentrations of TTX were recorded ($n = 5$). The data was then fitted with a Hill curve in order to extract the Hill coefficient ($h = 1,6$) and the half maximal inhibitory concentration ($IC_{50} = 0,7$ nM).

Original Amino Acid	Numbering <i>Musca Domestica</i>	Mutated Amino acid	Associated mutations	Species	<i>Apis Mellifera</i>	
					Amino Acid	Numbering
I	254	N	None		I	264
V	410	A, G, L, M	None E435K+C785R	<i>Drosophila melanogaster</i>	V	419
E	435	K	V410M+C785R, C785R+L1014F	<i>Helicoverpa zea</i> , <i>Cimex luctularis</i>	E	444
* E	485	K	C785R + L1014F	<i>Blattella germanica</i>	D	489
* C	785	R	E485K+ L1014F, L1014F	<i>Blattella germanica</i>	C	789
M	827	I, T	None, T929I, L932F, L1014F	<i>Pediculus humanus capitis</i> , <i>Pediculus humanus corporis</i>	M	831
* M	918	I, L, T, V	None, M918L+L925I, M918I+L1014F, M918T+L1014F	at least 11 different	M	922
L	925	I	None	<i>Bemisia tabaci</i> , <i>Cimex luctularis</i> , <i>Trialeurodes vaporariorum</i> , <i>Rhipicephalus microplus</i>	L	929
* T	929	C, I, N, V	None, M827I, L932F, L1014F, L1014F+A1060T+P1879S	at least 13 different	T	933
L	932	F	M827I, M827I+T929I	<i>Pediculus humanus capitis</i> , <i>Pediculus humanus corporis</i>	L	936
G, C	933	V, A	None	<i>Rhipicephalus microplus</i>	C	937
I	936	V	None	<i>Helicoverpa zea</i>	I	940
Q	945	R	None	<i>Lepeophtheirus salmonis</i>	Q	949
* F	979	S	L1014F	<i>Myzus persicae</i>	F	983
S	989	P	V1016G	<i>Aedes aegypti</i>	S	993
* V	1010	L	L1014S	<i>Anopheles culicifacies</i> , <i>Anopheles sinensis</i>	V	1014
I	1011	M, V	None	<i>Aedes aegypti</i> , <i>Anopheles sinensis</i>	I	1015
N	1013	S	None	<i>Anopheles sinensis</i>	N	1017
L	1014	C, F, H, S, W	Identified with *	At least 20 different	L	1018
V	1016	G, I	None, D1763Y	<i>Aedes aegypti</i>	V	1020
F	1020	C	None	<i>Blattella germanica</i> , <i>Plutella xylostella</i>	F	1024
L	1024	V	None	<i>Tetranychus urticae</i>	L	1028
* A	1060	T	T929I+L1014F+P1879S, L1014F+N1575Y, L1014F+P1879S, P1879S	<i>Plutella xylostella</i>	A	1099
A	1215	D	None	<i>Tetranychus urticae</i>		Poor alignment
A	1410	V	None	<i>Drosophila melanogaster</i>	A	1409
A	1494	V	None	<i>Drosophila melanogaster</i>	A	1493
M	1524	I	None	<i>Drosophila melanogaster</i>	M	1523
F	1528	L	L1596P+I1752V+M1823I	<i>Varroa destructor</i>	F	1527
F	1534	C	None	<i>Aedes aegypti</i> , <i>Aedes albopictus</i>	F	1533
F	1538	I	None	<i>Rhipicephalus microplus</i> , <i>Tetranychus cinnabarinus</i> , <i>Tetranychus urticae</i>	F	1538
D	1549	V	E1553G	<i>Helicoverpa armigera</i> , <i>Heliothis virescens</i>	D	1548
E	1553	G	D1549V	<i>Helicoverpa armigera</i> , <i>Heliothis virescens</i>	E	1552
* N	1575	Y	L1014F	<i>Anopheles gambiae</i>	N	1574
L	1596	P	None, F1528L+I1752V+M1823I	<i>Varroa mite</i>	P	1595
I	1752	V	F1528L+L1596P+M1823I	<i>Varroa mite</i>	I	1751
D	1763	Y	V1016G	<i>Aedes aegypti</i>	D	1762
M	1823	I	F1528L+L1596P+I1752V	<i>Varroa destructor</i>	I	1822
* P	1879	S	T929I+L1014F+A1060T, A1060T	<i>Plutella xylostella</i>	P	1878

Aedes aegypti				Apis Mellifera		
Residue label	Original amino acid	Numbering	Mutated amino acid	Numbering	Original amino acid	Conserved
1k7	I	259	A	258	I	Yes
1k11	V	263	A	262	V	Yes
1o6	L	270	I	269	L	Yes
1o10	I	274	C	273	I	Yes
1o13	T	277	W	266	T	Yes
1i18	L	281	G	280	L	Yes
2i13	I	990	M	1015	I	Yes
2i16	L	993	F	1018	L	Yes

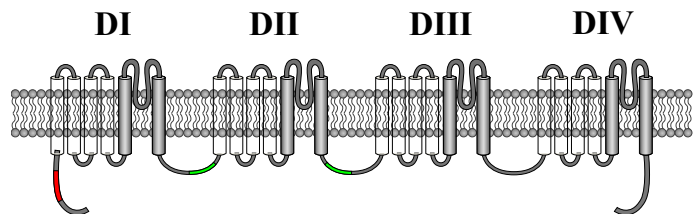
a**DI****DII****DIII****DIV****DI****DII****DIII****DIV****b****c****S1 S2 S3 S4 S5**

Amla.1	133	MLDPPN-PIRVAIVYLIYVHLSFLPIITLLNCLIMMPTF-PI	ESVYVYVGLTISAVVNVARGPILPPYLRD	QWVIVVAVLAVVYVGI	DLGN	SALRFRVLRALKTVISVPLKRTIVGAVIESKRVLEIVLIMFSLVFMALMQLTIVGLVQRKIKNFPEDES
Dmla.1	134	WMLDPPN-PIRVAIVYLIYVHLSFLPIITLLNCLIMMPTF-PI	ESVYVYVGLTISAVVNVARGPILPPYLRD	QWVIVVAVLAVVYVGI	DLGN	SALRFRVLRALKTVISVPLKRTIVGAVIESKRVLEIVLIMFSLVFMALMQLTIVGLVQRKIKNFPEDES
hNa.1.1	110	ILIFPN-PLRKAIAIKILVHLSFMSLIMCTLLNCFVMSNPNF-DW	TKNVTYVGLTIFESLILKILARGCLEDLFFLRD	NWMLDITVYFAVYVFFV	DLGN	SALRFRVLRALKTVISVPLKRTIVGAVIESKRVLEIVLIMFSLVFMALMQLTIVGLVQRKIKNFPEDES
hNa.1.2	111	ILIFPN-PLRKAIAIKILVHLSFMSLIMCTLLNCFVMSNPNF-DW	TKNVTYVGLTIFESLILKILARGCLEDLFFLRD	NWMLDITVYFAVYVFFV	DLGN	SALRFRVLRALKTVISVPLKRTIVGAVIESKRVLEIVLIMFSLVFMALMQLTIVGLVQRKIKNFPEDES
hNa.1.4	128	ILHLSFMSLIMCTLLNCFVMSNPNF-DW	SKNVTYVGLTIFESLILKILARGCLEDLFFLRD	NWMLDITVYFAVYVFFV	DLGN	SALRFRVLRALKTVISVPLKRTIVGAVIESKRVLEIVLIMFSLVFMALMQLTIVGLVQRKIKNFPEDES
hNa.1.6	114	ILISFPN-LIRRAIAIKILVHLSFMSLIMCTLLNCFVMSNPNF-DW	SKNVTYVGLTIFESLILKILARGCLEDLFFLRD	NWMLDITVYFAVYVFFV	NLGN	SALRFRVLRALKTVISVPLKRTIVGAVIESKRVLEIVLIMFSLVFMALMQLTIVGLVQRKIKNFPEDES
hNa.1.7	121	-----KILVHLSFMSLIMCTLLNCFVMSNPNF-DW	TKNVTYVGLTIFESLILKILARGCLEDLFFLRD	NWMLDITVYFAVYVFFV	NLGN	SALRFRVLRALKTVISVPLKRTIVGAVIESKRVLEIVLIMFSLVFMALMQLTIVGLVQRKIKNFPEDES
Amla.1	798	DCCHMLKEPKQVVALVDFVFLPTITLCTVNVVLPALDHD	MD-KMSEVYVGLTISAVVNVARGPILPPYLRD	MD-KMSEVYVGLTISAVVNVARGPILPPYLRD	MD-KMSEVYVGLTISAVVNVARGPILPPYLRD	MD-KMSEVYVGLTISAVVNVARGPILPPYLRD
Dmla.1	799	DCCHMLKEPKQVVALVDFVFLPTITLCTVNVVLPALDHD	MN-KEMSEVYVGLTISAVVNVARGPILPPYLRD	MN-KEMSEVYVGLTISAVVNVARGPILPPYLRD	MN-KEMSEVYVGLTISAVVNVARGPILPPYLRD	MN-KEMSEVYVGLTISAVVNVARGPILPPYLRD
hNa.1.1	750	CSYPLKRVKVVNVLVMDVFLDATTICVNLTPMANSHP-MT	DHFNVYVGLTISAVVNVARGPILPPYLRD	DHFNVYVGLTISAVVNVARGPILPPYLRD	DHFNVYVGLTISAVVNVARGPILPPYLRD	DHFNVYVGLTISAVVNVARGPILPPYLRD
hNa.1.2	741	CCPKLWVYVHVLVMDVFLDATTICVNLTPMANSHP-MT	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD
hNa.1.4	971	-----IHLVMDVFLDATTICVNLTPMANSHP-MT	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD
hNa.1.6	735	CHPYVYVYVHVLVMDVFLDATTICVNLTPMANSHP-MT	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD
hNa.1.7	738	-----YFVMDVFLDATTICVNLTPMANSHP-MT	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD
Amla.1	1271	-----RQGNMLRKTFLQENLVEFETAVIMLMSLSALALD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD
Dmla.1	1284	-----RQGNMLRKTFLQENLVEFETAVIMLMSLSALALD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD
hNa.1.1	1200	-----RQGNMLRKTFLQENLVEFETAVIMLMSLSALALD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD
hNa.1.2	1190	-----RQGNMLRKTFLQENLVEFETAVIMLMSLSALALD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD
hNa.1.4	1024	-----CFVYVYVHVLVMDVFLDATTICVNLTPMANSHP-MT	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD
hNa.1.6	1180	-----LQKSMWLRKTFLQENLVEFETAVIMLMSLSALALD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD	HLQRPILQDLYVQRIYVYVGLTISAVVNVARGPILPPYLRD
hNa.1.7	1187	-----KIVHLSFMSLIMCTLLNCFVMSNPNF-DW	TKNVTYVGLTIFESLILKILARGCLEDLFFLRD	NWMLDITVYFAVYVFFV	NLGN	SALRFRVLRALKTVISVPLKRTIVGAVIESKRVLEIVLIMFSLVFMALMQLTIVGLVQRKIKNFPEDES
Amla.1	1588	IFRPFN-RQAIVFVIVDQVSDIMLIMFLGMLMFLDLD	QD-QTFRVYVGLTISAVVNVARGPILPPYLRD	QD-QTFRVYVGLTISAVVNVARGPILPPYLRD	QD-QTFRVYVGLTISAVVNVARGPILPPYLRD	QD-QTFRVYVGLTISAVVNVARGPILPPYLRD
Dmla.1	1601	IFRPFN-RQAIVFVIVDQVSDIMLIMFLGMLMFLDLD	AS-DTVAIVYVGLTISAVVNVARGPILPPYLRD	AS-DTVAIVYVGLTISAVVNVARGPILPPYLRD	AS-DTVAIVYVGLTISAVVNVARGPILPPYLRD	AS-DTVAIVYVGLTISAVVNVARGPILPPYLRD
hNa.1.1	1523	IFRPFN-RQAIVFVIVDQVSDIMLIMFLGMLMFLDLD	QS-EYVYVGLTISAVVNVARGPILPPYLRD	QS-EYVYVGLTISAVVNVARGPILPPYLRD	QS-EYVYVGLTISAVVNVARGPILPPYLRD	QS-EYVYVGLTISAVVNVARGPILPPYLRD
hNa.1.2	1513	IFRPFN-RQAIVFVIVDQVSDIMLIMFLGMLMFLDLD	QS-QMNVYVGLTISAVVNVARGPILPPYLRD	QS-QMNVYVGLTISAVVNVARGPILPPYLRD	QS-QMNVYVGLTISAVVNVARGPILPPYLRD	QS-QMNVYVGLTISAVVNVARGPILPPYLRD
hNa.1.4	1024	-----CFVYVYVHVLVMDVFLDATTICVNLTPMANSHP-MT	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD	EPQSYVYVGLTISAVVNVARGPILPPYLRD
hNa.1.6	1504	IFRPFN-RQAIVFVIVDQVSDIMLIMFLGMLMFLDLD	QS-QMNVYVGLTISAVVNVARGPILPPYLRD	QS-QMNVYVGLTISAVVNVARGPILPPYLRD	QS-QMNVYVGLTISAVVNVARGPILPPYLRD	QS-QMNVYVGLTISAVVNVARGPILPPYLRD
hNa.1.7	1510	-----IFRPFN-RQAIVFVIVDQVSDIMLIMFLGMLMFLDLD	QD-QTFRVYVGLTISAVVNVARGPILPPYLRD	QD-QTFRVYVGLTISAVVNVARGPILPPYLRD	QD-QTFRVYVGLTISAVVNVARGPILPPYLRD	QD-QTFRVYVGLTISAVVNVARGPILPPYLRD
Na.Ab	1	-----WLR-----ITNIVES	EFYKFLIVLVNGLTGL	ETFLNIVYVITL	ELIILAY	-----RR-ISFQD

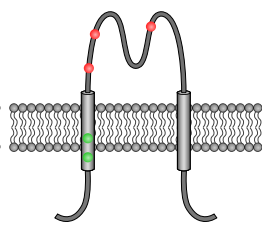
SF**S6****DI****DII****DIII****DIV****Highly conserved amino acids**

- Positively charged amino acid
- Negatively charged amino acid
- Aromatic amino acid
- Neutral amino acid
- Putative phosphorylation site for PKC

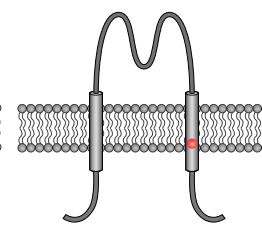
MFMT

a**AmNa_v1**

■ Δ49-59: EGGFGRKKKKK
■ ins 761: VSSTYYFPT
■ ins 1088: GEGPSNSWKE

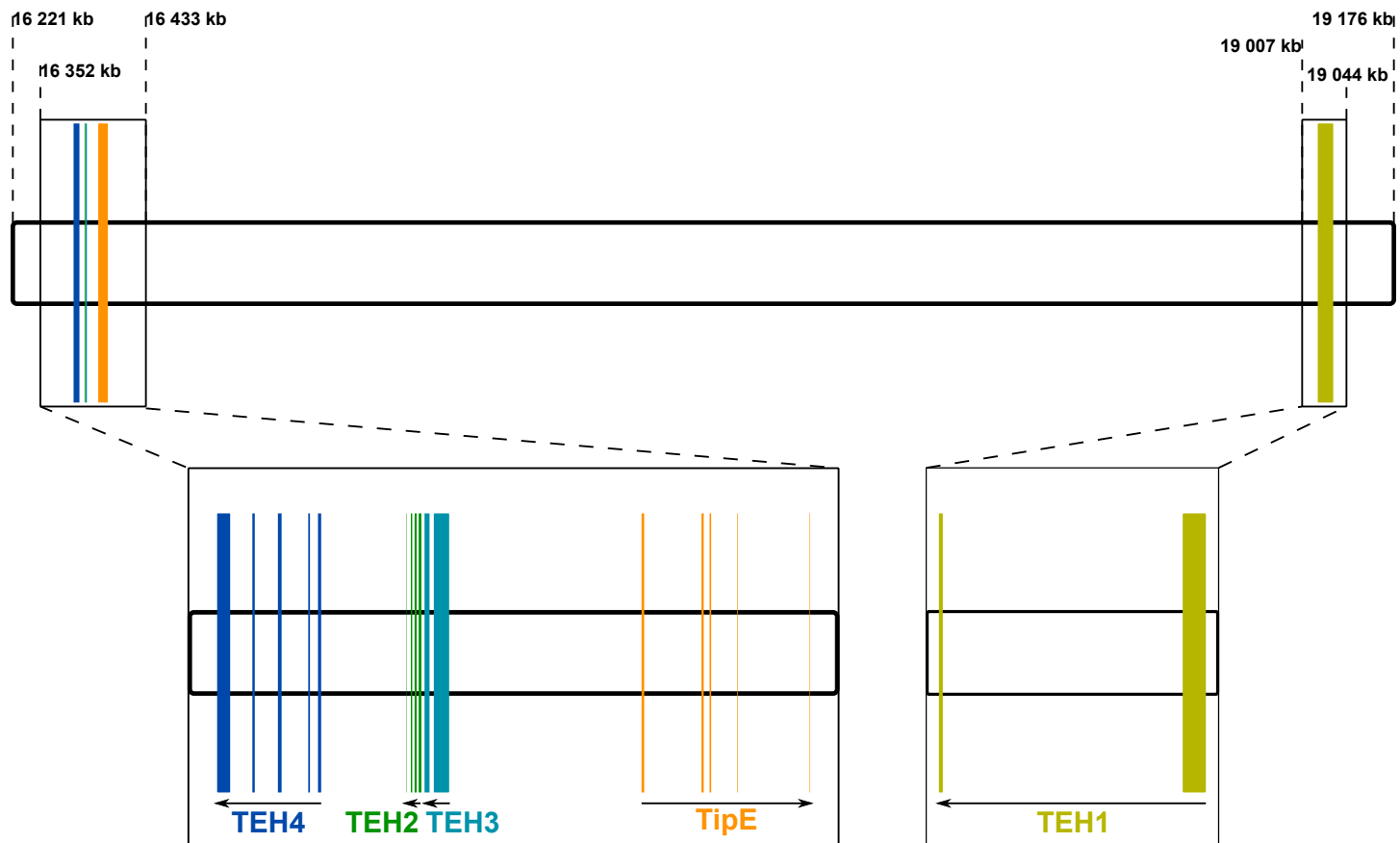
b**TipE**

● 23: R→K
● 26: S→G
● 70-72: TSN→EQS
● 115: A→V
● Δ165: G

TEH4

● 407: I→V

- Splice variant: insertion
- Splice variant: deletion
- RNA edition
- Genomic variation

c

a

```

Am_TipE 1 -----MAEEKEKQTLFQKLLFYTAFAFILLSTFLFAFLFLVFPVFDPAFTIFMFQFDTRPAEITTYVESRRGTNSCSWTSREGCTK-ELYDCTQIRVNYKLPNTSEDS-----DGQG-----
Am_TeH1 55 LGFGASRRRPRRTCRQRFNFYATSA-LAFVATVSGGAALLFLVPLVYDPAISTLAADFSPDPVICTTSRREELAGLFNCTWSSCREGCTS-DVYRCTHIYVYTPWSNASMKN-----D-----
Am_TeH2 10 STIGGLVPEVVEVVLVEKAKFYSLC-LGITALLAVFAFLFLIFVIEPAVITLADFSHPHAAVVTVDHYVAEGLKNCSWASCREGNS--AALRCHQIRVNYTRLSFEFFVA-----K-----
Am_TeH3 1 -M-PKQVPVENLVIPQDGRICGTTICICQMTLVSSVALVYLTVAIYMPSTRAFQSGISEVPMVCTTIRAV---NADNCWEGSCGEMCLSKTSGPCQIHVNLRRNSGRILLANCTNTNTKCYGIDQENAKKSKCIA-DECRNLTGTFNCSGG
Am_TeH4 1 -M-GR--KHKRRVLEPQDRICRAITCICFCQFTIVSICALVYLVSVAIYMPSTRAFAGIDDPDPMQTVNIT---LNNNCWASCGEMCLTKTSGPCQIHTVRRNGTDIVFENCTKFNISCPQVNTASLKKVYCNNGSECVLSGLFNCSLGG
Dm_TipE 1 -----MGDQKDRRTGKELLFYTAFAFILLSTFLFAFLFLVFPVFDPAFTIFMFQFEVPALEITTYEYGAKNCSWSSREGCTK-DLYTCTQIRVNYKLPNTSEDS-----FNFTHEYHINKLEAERIL-----
Dm_TeH1 24 LAPKKNKGNRRFRSRRERARFYVST-LAFPSVTAGASLLFLVPLVYDPAITLSDHDFIEKPTLTTTRREDLVGFINCSWSSCREGCTS-DLYRCVHIYVYTFEQNITIPEN-----M-----
Dm_TeH2 65 DAIKAKREEIEMDTLEKAKFYVSVCLGTTALLSVTFFLFLIFVYDPAITTIADYDPVPTQIVIDHIIYAEGKNCSWSSCREGCTS-SLTKEHQLFVNYTRIPFSEWER-----N-----
Dm_TeH3 1 MG-KKVPVENLVIPQDGRICGTTICICQMTLVSSVALVYLTVAIYMPSTRAFQSGIDPTFVMTTIRAV---NADNCWEGSCGEMCLSKTSGACIQIYVNLRSNGSLIYVNLRRNSGRILLANCTNTNTKCYGIDQDRADKARCIN-DECKNLTGTFNCTAG
Dm_TeH4 1 MG-RR--KDKPRVLEPQDARICRAITCICQMTLVSSVALVYLVSVAIYMPSTRAFQSGIDELPDMQTVDRQ---MPNNCWASCGEMCLTKTSGPCQIHSIVRRNGTDIQLNCTRVNTSCAMIDLSRLNKNFCNNGTACNIRGVFNCSNG
hKc_62 32 ---KRRKTVTALKAGEDRAIL-----LGLAMVCSIMMYFLL-GI--LTLISYMQSVWTEESQVITLLNAS-ITETFNCSF-----SCGPDCKWLSQYPCQLVYVNLTSSEGKLLLYHTEETI-----KINQKQSYIPKCGKN-----FEESMS

```

```

Am_TipE 111 -----EGGAVGVEDEDEDSTTMGKPRYDRSLREYDYVEDLD-ED-----FAED--DE-----A--
Am_TeH1 168 -----GGRGNSTGI-----A-----
Am_TeH2 123 -----LGS-----
Am_TeH3 149 VGINITDAFECIFHDTDPMPKCSGRRGKITCIDIDGLFNCRNGTCTERIRTPYNCDRRCVDPITRNKNMILLTGDKVYLSQCEAIDVQ-----TNREIWH-EDRGDVMMSACYGIFNSTLGVAVDCINGSVLEKDLLTDL-TNFT-
Am_TeH4 148 HCVNISELMLCHYKADGIVVDS-----EKDNMKNLNGYFSCHNSRCKTIKSPFSCDRYCPDVTSDVNVFLMQDDNIVTVKERCGLALNANGNLPGVRLTPHFQWE-DRNGSIIIVSCLAVDKKM-NDVRTQDCVNGTLKEIPLPQPTINFTS
Dm_TipE 124 -----PPVKRTDRY-----ERALRSDYEYDN-----LGGGT--GLDIDLGAGRMEQLNFGDADGNSNGYLIEDS-EDTRG-LSASGTLISDERP-----F--
Dm_TeH1 137 -----P-----D-----
Dm_TeH2 178 -----RLDLD-----
Dm_TeH3 150 QCLNITDAFECIFHNSDAPVKCSGRRGKINCMDISGLYSRGTCTCRKIRTPYNCDRRCVDPITRNKNVVVLSGDKVYLSQCAINAIE-----TLEEVWN-ESSENAVMSICYFIRHTSDQVDAVDCINGSTLETNMLSDL-TNFT-
Dm_TeH4 149 HCKNMSEFFLCHHKADGLTVNS-----QKDNKLNKGFECGHVHCTIKKPFSCDRYCSKITTNNVNTLIMHEDNLIADCCENAVAFNQARGSEHGVRV-EPFEFWK-EDDGNLLTNCATVTRSDNRITATDCINGTLLHEDTLPAPFMNFTQ
hKc_62 158 LVNVVY-----MENFRKQHFSCYSDPE-----GNQKSVLTL-LYSS-----NVLFHSLFWNPT-----CMMAGGVA-----IVAMVKLTQ

```

```

Am_TipE 157 -----GLPKPFPPTGLMGNDSSEWYFIGAKLFPNVKGCYPPMLNCSIFYRQYAI-----GQNFSCYYSKVDP-GIVI-----SDDLMOVYVNLVYAMAIPSPFIISVIYLTIAYFKIYNEDEVVL-----V--G--
Am_TeH1 179 -----HTSTTSVPTFGDVEAVLLVNIKGCYPPIVDCENFTREMGEV-----GAKFPCHYSRVNG-SIVM-----ANYNREAVQTTIIHFFAAPPVVLATSVLVCVMHCDRCSPFPRHSSRG-----I--R--
Am_TeH2 127 -----IQWVSDTKFVFNTEKGCYPPRNVCSDFAKNYGYS-----NMKGIFFCYYSRTHP-ETVAVXRYNXYRYSWDENLRLHVLALVPTVTVFVSGVSLGVLVYCPFMGKTCGGPGRNLDKYARKEE-----
Am_TeH3 288 YLSYLNIFATKPLDETRMVAPEQDLIIANESRLINLEGCVNTLRECKEFLHEYGKDGSDHNARARFPFCYAESNT-GIVV-----SRFLENITYKEFMIALLLPSILFVSSGLTILFCQKTVVVDGDAKMRFKGTGVALSMEKSASGN-
Am_TeH4 295 FLNIEKSLQYVDPDPTNIYVPAQRSLTIYNSRLINIFEGCQVNTLRGCKEDFLATHGRDGNQTAQRSRYCYNKNSS-LLVV-----ARFDLNTRETLIAIIVPSGLFVVISLTLVITRISYQVGDADKMRCRYVDKQVEGEDEGLVE-
Dm_TipE 200 -----DEISELNEGLMGRSMYYVGARLFPNVKGCYPPMLNCRILWLRKYTKI-----GMKFFCYYSKVDP-SLVI-----SDLDYQWNTLNLVYVYAMAIPSPFIISVIYLTIAYFKIYNEDEETA-----P--L--
Dm_TeH1 139 -----YSNFTSDMEQSGEATLVNIIKGCYPPSVTCKNFNGYVYIE-----GAIFPFCYYSRKNK-TVVL-----TSYNHDDQVAMIHFFAVFPVITVISSIALCIMIHCDCRCKKDRSRRNR-----PQCR-
Dm_TeH2 184 -----VNWVSYTFLINSEGGYPTTNCISIFARQYGFSS-----HGEPPFCYYSRAYP-EVVI-----GRYSWENLNYHLRSLIIPNVLFAISIGVLSVYWCPCQCEKACNKSRSRYAEKFPTEKDKILLCH-
Dm_TeH3 289 YLSHKLHVSATFPV-----EIAAPPDVLTIISNESKLMINLEGCVNTLMDCKEFLKDFGRDGSDDHNARARFPFCYSPGKK-DVVV-----ARFDLEVTYRQYFVFAVPSVSLFVVSGLIMCQTTIVVVDGDAKMRFKGCVDTETVLNKNVYVAP
Dm_TeH4 296 FWAIYENST-RSVDPEQRYPNQNANLTIYSWKKLFINLEGCVNTLRGECKDFVARYGNDGDNNTAQRSRYCYNKNDSNVEFVV-----ARYDLDRYRELVSLVIVLVISSISLICITKRSYKVGDDAKMRVCAGDSDNDGPFPG-
hKc_62 222 YLSLLCERIQRINR-----

```

b

```

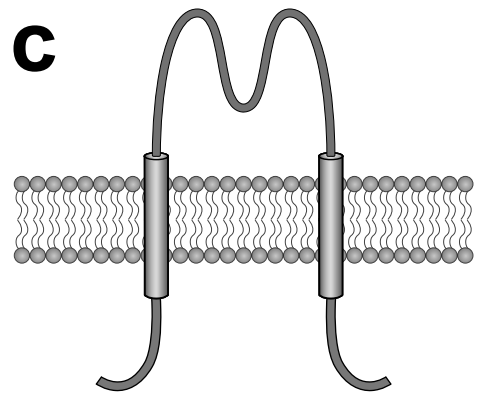
Am_TeH3 111 TNTNTKCYGIDQENAKKSKCIA-DECRNLTGTFNCSGG-----VGINITDAFECIFHDTDPP
Am_TeH4 109 TKFNISICPQVNTASLKKVYCNNGSECVLSGLFNCSLGG-----HCVNISELMLCHYKADGIV
Dm_TeH3 112 TNSANKTCYGIDQDRADKARCIN-DECRNLTGTFNCTAG-----OCLNITDAFECIFHNSDAP
Dm_TeH4 110 TRVNTSCAMIDL SRLNKNFCNNGTACNIRGVFNCSNG-----HCKNMSEFFLCHHKADGLT
hLFB1 1195 DD--NKTQCDINECEHPG-LCGPQCELETEGFSFHCVCCQ--GFSISADGRTCEIDIECVNNT-VCDSHGCPDNTAGSFRCLCYQGFQ
hFBN2 1359 KK-GTTGCTDVECEI GAHNCDMHASCLNIPGSFKSCREG--WIGNIKCIDLDECSNGTHQCSINAQCVNTPGYSRACACSEG---
fFBLN1 472 SDVDGVTCEIDICALPTGGHICSYRCLNIPGSFQSCPPSSGYRLAPNGRNQDIDECVTGHNCSINETCFNIGGFRGLAFECPEN

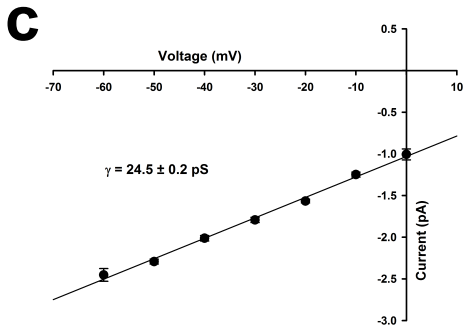
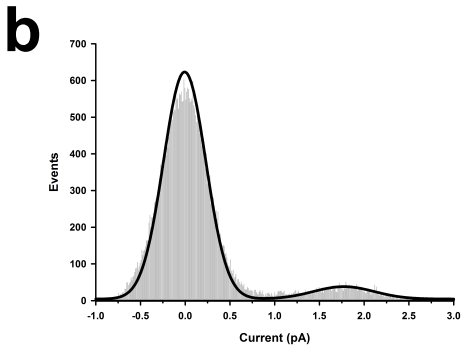
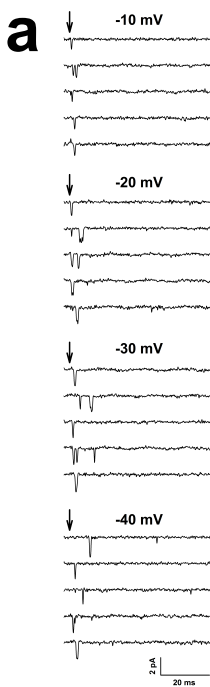
```

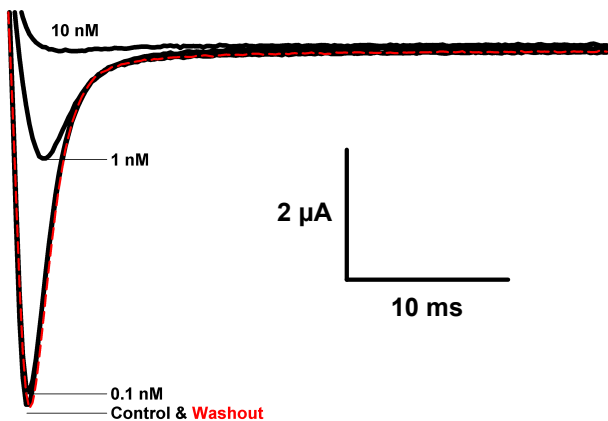
```

Am_TeH3 190 RGTCTERIRTPYNCDRRCVDPITRNKNMILLTGDKVYLSQCEAIDVQ-----TNREIWHEDRG
Am_TeH4 184 NSRQTKIKSPFSCDRYCPDVTSDVNVFLMQDDNIVTVKERCGLALNANGNLPGVRLTPHFQWEDRNG
Dm_TeH3 191 RCTCRKIRTPYNCDRRCVDPITRNKNVVVLSGDKVYLSQCAINAIE-----TLEEVWNESSE
Dm_TeH4 185 GVHCTKIKKPPSFCDRYCSKITTNNVNTLIMHEDNLIADCCENAVAFNQARGSEHGVRV-EPFEFWKEDDG

```

c



a**b**