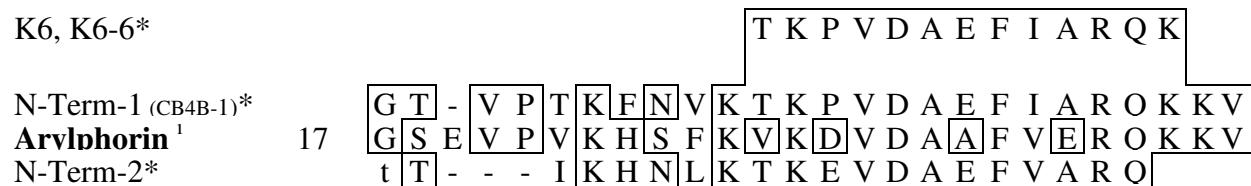


Table 2. Peptide sequences identified by Edman sequence analysis or mass spectrometry and their alignment to the most similar database entry.

Entry names und Datenbank accession numbers: ¹Arylphorin β; ARYB_MANSE, P14297 (*Manduca sexta*), ²Arylphorin, Q9U5K4 (*Spodoptera litura*), ³Q9VTW1, *Drosophila melanogaster* genomic sequence (SAM-methyltransferase); ⁴Arylphorin α, ARYA_MANSE, P14296; ⁵ABC transporter, Q8U6S8, *Agrobacterium tumefaciens*; ⁶Arylphorin, ARY_ANTP, Q7Z1F8, *Antheraea pernyi*, ⁷Arylphorin, O17478, *Hyalophora cecropia*; ⁸Arylphorin, SSP2_BOMMO, P20613, *Bombyx mori*; ⁹Arylphorin, Q86QQ8, *Heliorithis virescens*; ¹⁰Arylphorin, ARY_AHTP, Q7Z1F8, *Antheraea pernyi*. Small letters denote amino acids, which were not identified unequivocally. X, no clear signal detected. #, this peptide showed heterogeneities at certain positions indicating that it was a mixture of two very similar peptides. Modified N-glycosylation sequons are shown as black boxes. K, CB, T (ESI-MS/MS only!) or E (ESI-MS/MS only!) denotes peptides (or partial peptide sequences) derived from cleavage with lysyl endopeptidase, CNBr, trypsin or endoproteinase Glu-C, respectively. Numbers and other letters in peptide designations for K and CB peptides denote chromatographic pools. *, sequences from Edman sequence analysis; +, peptides identified by mass spectrometry (MS/MS). In MS-derived sequences, the amino acid L stands for either L or I.



K7, K6-5c*

— — — — — [e v d a e f v a r q k]

K6-8b*

Arylphorin⁷

61 i a t t y s v e e n i n n y t n k
I G K D Y D I E A N I D N Y T N K

63 E Y N I E A N I D N Y S N K
E Y S I E A N I E N Y T N K

K7-4*

T4⁺

Arylphorin¹

78 A V E D F L Q L Y R
A V E E F L Q L Y R

T6⁺

Arylphorin⁹

98 Y F F L P T A L D F Y Q
Y A F M P T A L D F Y Q

CB3,4A-3*

Arylporin⁴

156

g I [V L P A P Y E V] F [P Q Y F] V [N] X X L
G L [V L P A P Y E V] Y [P Q Y F] T [N] M E V

T11⁺

Arylporin⁴

210

[V F Y] T [N Y S] N S L T Y P N D E Q R
[V F Y] A N Y S N S L [S] Y P N E E E R
[Y A] [N Y S] D [S L T Y P N E E] S R

T13⁺

CB7A-1*

CB5A*

Q9VTW1³

226

D E L V K Q
D E L V K Q t G [N K Y L] Y K I
D E L V [M] Q T C [N K Y L] N V I

T2 ⁺	T G Y N P L L T T Y A Y P Y A Q R
K2,3-1*	T G Y N P L L T T Y A Y P Y A Q R S N N Y D I H N E Q N Y e
Arylphorin⁷	300 T G Y Y P L M T S Y Y Y P F A Q R P E N Y N L H S V K N Y E
Arylphorin²	316 S N D Y N L H N E K N Y E Y I R
T1 ⁺	S N N Y E L H N E q n y e - - R
T7 ⁺	S N N Y D L H N E Q N Y E N L R

K8-3*	T F F Q Y L Q K
Arylphorin²	339 T F F Q F L Q K
Arylphorin⁴	342 T F L Q Y L Q K G H F K
K7-5*	T F F Q Y L Q S G H F K

T3 ⁺	A L N F V G N Y W Q N N A D L Y S E E V T K D Y Q
Arylphorin⁸	363 A I N F V G N Y W Q D N A D L Y G E E V T K D Y Q

Sequence alignment diagram showing the C-terminal region of various proteins. The alignment is rooted at position 366. Amino acids are shown as boxes, with identical residues sharing the same color. The sequence starts with 'a' at position 366.

Protein	Position	Sequence
K5*	366	a X n f v g n y w q r n
Arylphorin ⁴	366	A V N F V G N Y W Q A N A D L Y N E E V T K
K6-2*		N F V G N Y W Q N N A D L Y s e e v t k
K4-4#,*		A L N F V G N Y W Q N N A D L Y S E E V T K
Arylphorin ¹⁰	372	Y W Q D N A D L Y G E E V T K D Y Q R S Y E I V A R
T8 ⁺		W Q N N A D L Y S E E V T K D Y Q H S Y E V V S R
T9 ⁺		Y W Q T N A D L Y S E D V T K D Y Q H S Y E V V S R
K4-1*		D Y Q H S Y E I T A R R V L G A G P E Q T N K
Arylphorin ¹	385	D Y Q R S Y E I I A R Q V L G A A P K P F D K
K4-3*		N Y Q H S Y E I T A R R V L G A A P K
K7-1*		L I Q R I Y G V K
Q8U6S8 ⁵	398	K L I E R I Y G V F
K7-2*		K L I Q R I Y G V K

T5⁺

Arylporin⁴

411

T5 ⁺	Y S Y L P S S L D F Y L T
Arylporin ⁴	Y S F I P S A L D F Y Q T

K4-2*

Arylporin⁶

469

K4-5*

T14⁺

K4-2*	L V T Y F D Y Y D F X A T N K
Arylporin ⁶	L V T [F] F [E] Y Y D F [D] A T N T
K4-5*	L V T Y F D Y F D F X I T N K
T14 ⁺	L V T Y F D Y F D F N I T N K

K8-2*

Arylporin²

510

K8-2*	S E [V] S T D A V F K
Arylporin ²	S D [V] A S D A V F K

K8-5*	F Y E L D W F I Q K
K8-4*	M Y E L D W F V N K
CB2A-1*	Y E L D W F V N K L V P G E N K I E R K
Arylphorin²	542 F Y E L D W F V Q K L V P G E N K I E R K S S E F V
CB2A-2*	F V N K L V P G E N K I E R K S S N F A
K6-4*	L V P G E N K I E R K
E1 ⁺	L D W F V N K L V P G E N K L E

CB6A-1*	S V L P D N I P R R
Arylphorin²	593 S V Q P D N M P R R
T12 ⁺	N V T L Y H E G E P F A W E Y N V P A Y F S K
Arylphorin⁶	675 D V S V Y H E G E Y H A Y E Y N V P A Y F S H

CB1A-1*

Arylphorin⁴

K6-1*

615	L P K G T K G G Y P F Q L F V F I Y P Y N G V T K
	L P R G T K D G Y P F Q L F V F V Y P Y Q A V P K
	G T N G G Y P F Q L F V F V Y P Y K

K4-6*

Arylphorin⁶

675 D V T I Y H E G E P F A W E Y N V P A Y F S K
D V S V Y H E G E Y H A Y E Y N V P A Y F S H