The Journal of Experimental Biology 209, 3183-3198
Published by The Company of Biologists 2006
doi:10.1242/jeb.02374

Molecular characterization of the first aromatic nutrient transporter from the sodium neurotransmitter symporter family

Ella A. Meleshkevitch[1], Poincyane Assis-Nascimento[1,2], Lyudmila B. Popova[1,3], Melissa M. Miller[1], Andrea B. Kohn[1], Elizabeth N. Phung[1], Anita Mandal[4], William R. Harvey[1,5] and Dmitri Y. Boudko[1,*]

[1]The Whitney Laboratory for Marine Bioscience, University of Florida, 9505 Ocean Shore Blvd., St Augustine, FL 32080, USA, [2]Barry University, FL 33161, USA, [3]A. N. Belozersky Institute, Moscow State University, Russia, [4]Department of Biology, University of North Florida, FL 32224, USA and [5]Department of Physiology and Functional Genomics, College of Medicine, University of Florida, FL 32610, USA

*Author for correspondence (e-mail: boudko@whitney.ufl.edu)

Accepted 5 June 2006

Summary

Nutrient amino acid transporters (NATs, subfamily of sodium neurotransmitter symporter family SNF, a.k.a. SLC6) represent a set of phylogenetically and functionally related transport proteins, which perform intracellular absorption of neutral, predominantly essential amino acids. Functions of NATs appear to be critical for the development and survival in organisms. However, mechanisms of specific and synergetic action of various NAT members in the amino acid transport network are virtually unexplored. A new transporter, agNAT8, was cloned from the malaria vector mosquito *Anopheles gambiae* (SS). Upon heterologous expression in *Xenopus* oocytes it performs high-capacity, sodium-coupled (2:1) uptake of nutrients with a strong preference for aromatic catechol-branched substrates, especially phenylalanine and its derivatives tyrosine and L-DOPA, but not catecholamines. It represents a previously unknown SNF phenotype, and also appears to be the first sodium-dependent B^0 type transporter with a narrow selectivity for essential precursors of catecholamine synthesis pathways. It is strongly and specifically transcribed in absorptive and secretory parts of the larval alimentary canal and specific populations of central and peripheral neurons of visual-, chemo- and mechano-sensory afferents. We have identified a new SNF transporter with previously unknown phenotype and showed its important role in the accumulation and redistribution of aromatic substrates. Our results strongly suggest that agNAT8 is an important, if not the major, provider of an essential catechol group in the synthesis of catecholamines for neurochemical signaling as well as ecdysozoan melanization and sclerotization pathways, which may include cuticle hardening/coloring, wound curing, oogenesis, immune responses and melanization of pathogens.

Supplementary material available online at
http://jeb.biologists.org/cgi/content/full/209/16/3183/DC1

Key words: malaria, mosquito, SLC6, phenylalanine, tyrosine, dopamine, catecholamine, sensory neurons, neurotransmitter synthesis, essential amino acids, brain evolution.

Introduction

The Sodium neurotransmitter symporter family (SNF, a.k.a. SLC6) is one of the largest, most ancient, and most diversified families of secondary transporters. Its members mediate absorption of neurotransmitters: dopamine (DA), norepinephrine (NE), epinephrine, octopamine (OA), serotonin (5-HT), γ-aminobutyric acid (GABA); neuromodulators: glycine and proline; intracellular osmolites: taurine and betaine; intracellular energy substrates: creatine and proline; and a number of 'orphan' proteins with as yet unidentified phenotypes and physiological roles (Chen et al., 2004). This family also includes transporters for Na^+- or K^+-dependent absorption of neutral and cationic amino acids (Castagna et al., 1997; Chen et al., 2004). Their physiological activities were determined in mouse blastocysts (Van Winkle et al., 1985) and brush border membrane vesicles (BBMVs) isolated from vertebrates (Broer, 2002; Palacin et al., 1998; Van Winkle et al., 1985). Earlier studies revealed two distinct transport mechanisms: broad substrate spectra neutral (B^0) and broad substrate spectra neutral plus cationic ($B^{0,+}$) amino acid transport systems (Broer, 2002; Kilberg et al., 1993; Stevens, 1992). Molecular cloning and heterologous characterization showed that the $B^{0,+}$ system corresponds to a unique mammalian SNF member, SLC6A14, the human (Sloan and

Mager, 1999) and mouse (Ugawa et al., 2001) orthologs of which transport neutral and basic substrates. The molecular identities of the mammalian B^0 system presently comprise two transporters: SLC6A19 [a.k.a. B^0AT1 (Bohmer et al., 2005; Broer et al., 2004)] and SLC6A15 [a.k.a. B^0AT2 (Broer et al., 2006b; Takanaga et al., 2005a)]. SLC6A19 accepts a broad spectrum of neutral amino acids. SLC6A15 accepts a more specific subset of large side-chain amino acids. Mammalian B^0 and $B^{0,+}$ transporters form orthologous clusters, suggesting conservative evolution of their phenotypes (Boudko et al., 2005a). The preferred substrates for both systems are essential amino acids, which metazoan cells cannot synthesize and must absorb from their environs. It was proposed that B systems represent diversified variants of an ancestral phenotype (Kilberg et al., 1993). The molecular counterparts of these essential mammalian systems are expected to be present in other organisms but remain to be identified.

B type-like systems for the absorption of neutral and cationic amino acids have been reported in insects (Giordana et al., 1998; Giordana and Parenti, 1994; Hennigan et al., 1993a; Hennigan et al., 1993b; Sacchi et al., 1993) and (for reviews, see Boudko et al., 2005c; Giordana et al., 1989; Sacchi et al., 2001). Functional and morphological similarities as well as prominent physiological differences of mammalian and insect B systems have been examined at the molecular level. For example, the first 'B system-like' insect transporter, msKAAT1, cloned from the caterpillar *Manduca sexta* (ms) midgut (Castagna et al., 1998), and the closely related msCAATCH1 (Feldman et al., 2000), both preferentially utilize inverted potassium *vs* sodium ion electrochemical gradients. However, a sodium-preferring B^0-like transporter was cloned from larval midgut of the yellow fever mosquito, *Aedes aegypti* aeAAT1 (Boudko et al., 2005a). In addition, aeAAT1 showed substrate binding differences. Concomitantly emerging genomic data allowed us to identify an insect-specific expansion of nutrient amino acid transporters (NATs), which includes all three cloned insect NATs and a number of orphan members, seven in *An. gambiae* and six in *D. melanogaster* (Boudko et al., 2005a; Boudko et al., 2005c). Based on this data the best deduction was that paralogous groups of metazoan NATs constitute multi-unit, lineage-specific, B^0-like transport systems with a role in the active absorption and redistribution of essential amino acids that is similar to that in the mammalian B^0 system. Molecular cloning and comparative analysis of NAT phenotypes in a few distant metazoan models is essential to validate such a hypothesis as well as to establish the roles of particular NATs in the amino acid transport network.

In the present study we report the cloning and characterization of the first NAT member from the African malaria vector mosquito, *An. gambiae*, which is emerging as the predominant genomic model for the biology of tropical disease vector mosquitoes and vector–pathogen interactions during the transmission of malaria. The available genomes, explicit pattern of functional specialization, and simple cellular organization of the alimentary canal in mosquito larvae, provide unique opportunities to study the molecular and integrative mechanisms of membrane transport phenomena (Boudko et al., 2001a; Boudko et al., 2001b; Assis et al., 2004; Boudko et al., 2005a; Boudko et al., 2005b; Boudko et al., 2005c). In the presented work we have identified a new NAT phenotype, which presumably plays essential roles in the intestinal absorption and somatic redistribution of catechol-branched phenylalanine, tryptophan and the metabolic intermediate L-DOPA. This transporter appears to be an important if not a unique substrate provider in melanization and sclerotization pathways as well as in the synthesis of catecholamine neurotransmitters. The biomedical relevance of mammalian NATs to several metabolic disorders has attracted broad attention during recent years (for a review, see Broer et al., 2006a). A role for NATs in neurotransmitter synthesis and corresponding neuronal disorders has also emerged (Broer et al., 2006b; Takanaga et al., 2005a), and is extended by the data presented here. Insect NATs are relevant to tropical medicine and agriculture since they are potential targets for developing lineage-specific and environmentally safe agents for the control of disease vector mosquitoes and agricultural pests.

Materials and methods

Model organism

Anopheles gambiae SS G3 strain larvae were hatched from eggs supplied by MR4 (The Malaria Research and Reference Reagents Resource Center) at the Centers for Disease Control and Prevention in Atlanta, GA, USA (http://www.malaria.atcc.org) and raised to the 3rd–4th instar as described in the supplier manual (www2.ncid.cdc.gov/vector/vector.html) under a 12 h:12 h dark:light interval and 2 days feeding schedule.

Bioinformatics

Sequence analysis and assembly were performed using SeqMAN II (DNASTAR, Inc, Madison, WI, USA). Phosphorylation sites were predicted using PepTools 2.0 software (BioTools, Inc., Edmonton, Alberta, Canada). Homologous protein sequences were derived as best reciprocal matches to agNAT8 and other SNF members, which were acquired using NCBI BLASTp (gap BLAST) by screening non-redundant protein databases (NCBI and EMBL).

The phylogenetic analysis included SNF populations from two dipterous arthropods with concomitant annotation of genomes and proteomes. Protein alignments were generated and fine-tuned with different gap penalties, using the ClustalX platform (Thompson et al., 1997) and MEGA 3.0 interface (Kumar et al., 2004). This alignment was exported into MEGA 3.0 and trimmed by deleting weakly conserved N and C termini along with unique inserts. The alignment was converted into the NEXUS format and a tree was reconstructed using Bayesian inference of posterior probability for 2 million generations with mixed models of amino acid substitution given an equal priority [MrBayes v3.1 (Huelsenbeck and Ronquist, 2001)]. The selected sequence alignment was also exported for a representative graph of protein alignment, along with protein secondary structures and specific functional

Table 1. *Primers used for amplification and cloning of agNAT8*

Oligo name	Description	Nucleotide sequence
Lu4St11	5′ RACE adapter	CGACGTGGACTATCCATGAACGCAACTCTCCGACCTCTCACCGAGTACG
Lu4Trsa	3′ RACE adapter	CGACGTGGACTATCCATGAACGCA CGCAGTCGGTACTTTTTTTTTTTTT
NsLu4	nested 5′′and 3′ RACE adapter	TCGAGCGGCCGCCCGGGCAGGT CGACGTGGACTATCCATGAACGCA
agNAT8_338-57_1>	internal primer	AGGCGGCGCGTTCGTGATAC
agNAT8_1115-34_2>	internal primer	GTCGCTGCTGGCCGGGTGTA
agNAT8_551-70_1<	internal primer	GCCCACGGGAGCGGGTACAT
agNAT8_1412-31_2<	internal primer	CCGCCCGGCGTCACGTAGAT
agNAT8_3′	agNAT8 3′ end	TCAGTCCTCGTCGATGAACTTTTTGTACT
agNAT8_5′	agNAT8 5′ start	ATGGAAGGACGGGATAATAATGGGTTCA
5′agNAT8*Bam*HI	5′ primer for pXOOM cloning	cgggatcccgATGGAAGGACGGGATAATAATGGGTTCA*
3′agNAT8*Hin*dIII	3′ primer for pXOOM cloning	cccaagcttgggTCAGTCCTCGTCGATGAACTTGTACTCGT*
agNAT8-7F	qPCR primer	GGACGGGATAATAATGGGTTCA
agNAT8-64R	qPCR primer	GGTACTGGCCCGCTATCGAT
ag18SrRNA-403F	qPCR primer	CGGCGCGAGAGGTGAA
ag18SrRNA-466R	qPCR primer	CCTTGGCAAACGCTTTCG

**Bam*HI and *Hin*dIII restriction sites with restriction enhancers are shown in lower case with restriction sites underlined.

motifs. The alignment was visualized and exported into Microsoft Word (Microsoft®) using RTF export of the GeneDoc software (www.psc.edu/biomed/genedoc). Hydropathy analysis and prediction of secondary and tertiary structure were performed at remote servers (bioweb.pasteur.fr/seqanal/interfaces/toppred.html) and (protinfo.compbio.washington.edu/protinfo_abcmfr), respectively. To confirm 2D and 3D structures, the sequence and structure of the *Aquifex aeolicus* VF5 SNF member, $LeuT_{Aa}$ (Yamashita et al., 2005) were used as templates.

5′, 3′ RACE and TA cloning

Fragments of agNAT8 were amplified from a cDNA collection prepared from isolated midgut tissues of 4th instar *An. gambiae* (G3 strain) larvae using techniques described earlier (Matz, 2002). Exact primers agNAT8_338-57_1> and agNAT8_1412-31_2< (Table 1), based on NAT-specific conserved regions of predicted sequences (NCBI Accession no. EAA05568), were used to obtain and sequence an internal fragment of agNAT8. 5′ and 3′ RACE (Rapid Amplification of cDNA Ends) was performed using adapter primers and exact internal primers agNAT8_1115-34_2> and agNAT8_551-70_1< to verify the terminal parts of the predicted NAT sequence. Primers were designed using an OLIGO tool of Jellyfish software (LabVelocity Inc., Los Angeles, CA, USA) against the verified agNAT8 transcript sequence (Table 1). Exact ORF-flanking primers agNAT8_3′ and agNAT8_5′ were used to re-amplify transcripts from the midgut specific cDNA collections. For amplification we used TaKaRa LA Taq™ polymerase with 3′→5′ exonuclease activity (proof reading activity). Polymerase chain reactions (PCR) included 25 cycles with denaturation at 95°C, 30 s, annealing at 60°C (adjusted for each primer set), 1 min, and extension at 72°C, 3 min. The resulting PCR product was isolated (QIAquick; QIAgen Inc., Valencia, CA, USA) and ligated into the pGEM-T vector according to a TA-cloning procedure (Promega), and transformed into TOP10 cells (Invitrogen Corp., Carlsbad, CA, USA). Positive colonies were confirmed by sequencing of extracted plasmids using the BigDye Terminator Cycle Sequencing Kit (ABI, Columbia, ML, USA).

pXOOM cloning

A low copy-number pXOOM plasmid was used to obtain an expression construct in which agNAT8 was flanked with *Xenopus* globin 5′ and 3′ UTRs, which mediate robust *in vivo* expression, presumably by increasing RNA stability in the heterologous system (Jespersen et al., 2002). Cloning protocols were similar to those for TA cloning except that ORF primers were flanked with two non-redundant restriction enzyme sites, *Bam*HI and *Hin*dIII, for ligation into pXOOM (Table 1).

Heterologous expression

cRNA for oocyte injections was obtained by *in vitro* transcription of *Pme*I-linearized pXOOM-agNAT8 plasmids using mMessage mMachine®, a high yield, capped RNA transcription kit (Ambion Inc., Carlsbad, CA, USA). The integrity and quantity of the agNAT8 transcript were confirmed by agarose gel electrophoresis. Surgically isolated and collagenase-treated stage V–VI *Xenopus laevis* oocytes (Nasco, Fort Atkinson, WI, USA) were injected with ~40 ng of agNAT8 cRNA and incubated for 2–6 days at 17°C in sterile N98 oocyte medium (98.0 mmol l^{-1} NaCl, 2.0 mmol l^{-1} KCl, 1.0 mmol l^{-1} $MgCl_2$, 10 mmol l^{-1} Hepes, 1.8 mmol l^{-1} $CaCl_2$, adjusted to pH 7.4 with NaOH) that was supplemented with 2.5 mmol l^{-1} sodium pyruvate, 100 units ml^{-1} penicillin, 100 mg ml^{-1} streptomycin, and 5% horse serum. Oocytes were conditioned in the N98 serum-free media for 2–3 h prior to recording.

Electrophysiological characterization

Freshly isolated oocytes were randomly evaluated for endogenous responses to a selected set of amino acids; only

groups consistently lacking endogenous responses were used in the heterologous expression experiments. agNAT8-injected oocytes were placed in a small volume (50–100 μl) of bathing medium in a micro-machined plastic chamber with a capillary channel that allows rapid laminar perfusion. The chamber solution was connected through two 3 mol l^{-1} KCl 1% agar-bridged, Ag/AgCl reference electrodes to a virtual ground of a current monitor head stage (VG-2A-x100, Molecular Devices, San Francisco, CA, USA). An eight-channel electromechanical valve (Warner Instruments; Salt Lake City, UT, USA) was used for switching of specified solutions. Amino acid-induced currents were measured with a two-electrode voltage clamp (GeneClamp 500, Molecular Devices). Microelectrodes (0.5–1 MΩ) were prepared with a Sutter-2000 puller from 1.2 mm borosilicate glass capillaries and filled with 1 mol l^{-1} KCl/1 mol l^{-1} (saturated) K_2SO4 electrolyte, resulting in 1–5 MΩ tip resistance. Current/voltage (*I*/*V*) signals were amplified/filtered with CyberAmp380 (Molecular Devices) and recorded using two parallel acquisition systems: MP100 (WPI, Sarasota, FL, USA) for slow pharmacological responses and DigiData1200 (Molecular Devices) for snapshots of *I*/*V* plots (1 Hz and 2 kHz low-pass filter cut-off setting, respectively). Data were analyzed using ClampFit (Molecular Devices) and SigmaPlot 9.0 software (SYSTAT Software Inc., Point Richmond, CA, USA). The composition of solutions used for ion substitution assays has been described earlier (Boudko et al., 2005a).

Data analysis

Values depicted in graphs represent the mean ± s.e.m. or s.d. from at least three independent experiments involving at least three different oocytes. Since expression analysis includes subjects with distinct expression intervals between 72–96 h after injection, the response amplitudes were normalized relative to the maximum cationic inward current in each experimental set and were fit using a nonlinear regression tool of Sigma Plot 9.0 software (SYSTAT Software Inc.). Kinetic profiles and constants were derived by curve fitting of normalized data sets with a three-parameter sigmoidal Hill function: $y=a*x^b/(c^b+x^b)$; 200 iterative steps were used, where a=max(*y*)=derived normalized mean maximum current; b=η=order of the transport process; $c=E_{0.5}$, (*x*,*y*), substrate concentration at 50% of the apparent transport velocity.

Isotope uptake assay

agNAT8 transcripts were injected into oocytes as described above. Random electrophysiological tests of a few oocytes from each set were performed to confirm agNAT8 expression and absence of background levels of amino acid uptake in the water injected oocytes. L-[methyl-^{3}H] methionine (specific activity: 84 Ci mmol l^{-1}; 1 Ci=3.7×10^{10} Bq) and L-[4-^{3}H]phenylalanine (specific activity: 27.0 Ci mmol l^{-1}) were obtained from Amersham Biosciences. Uptake assays were performed on Day 4 post-injection using 1 mmol l^{-1} concentrations of methionine or phenylalanine in the presence of 100 mmol l^{-1} NaCl. Briefly, oocytes were first washed in Na^+-free medium (100 mmol l^{-1} choline chloride, 2 mmol l^{-1} KCl, 1 mmol l^{-1} $CaCl_2$, 1 mmol l^{-1} $MgCl_2$, 10 mmol l^{-1} Hepes/Tris, pH 7.5). To initiate the uptake, this medium was replaced with 200 μl of Na^+-containing medium supplemented with radioactive methionine or phenylalanine (10 μCi ml^{-1}), and brought to a concentration of 1 mmol l^{-1} with the corresponding unlabeled amino acids. Uptake assays were performed at room temperature for different incubation times (1–10 min) and stopped by removing the uptake solution and washing the oocytes five times with 5 ml of ice-cold Na^+-free medium. Individual oocytes were placed in separate scintillation vials, dissolved in 200 μl of 10% SDS, to which 4 ml of scintillation fluid was added, and the radioactivity was counted in a Beckman-Coulter LS 6500 scintillation counter.

Whole-mount in situ hybridization

A purified pGEM-T agNAT8 plasmid was linearized with *Nco*I or *Not*I restriction enzymes to obtain full-length, run-off transcripts using SP6 and T7 promoters for anti-sense and sense (control) probes, respectively. DIG-labeled probes were transcribed *in vitro* using a DIG RNA labeling kit (Roche, Basel, Switzerland). 4th instar *An. gambiae* larvae were immobilized in ice-cold 0.1 mol l^{-1} phosphate buffered saline (PBS), opened by a lateral incision, and fixed in 4% paraformaldehyde/PBS overnight. Preparations were dehydrated/rehydrated by passing through a PBS/methanol gradient set (100% PBS-3:1–1:1–1:3–100% methanol, then in reverse order), 10 min for each mixture, pre-treated with Proteinase K/detergent solution (0.1% Tween-20 in PBS supplemented with 10 μg ml^{-1} Proteinase K) for 30–40 min and post-fixed in 4% PFA/PBS at 4°C for 40 min. After fixation preparations were washed in 0.1% Tween-20 in PBS supplemented with 2 mg ml^{-1} L-glycine followed by 0.1 mol l^{-1} triethanolamine hydrochloride, pH 8.0. Acetic anhydride (2.5 μl ml^{-1}) was added to the last wash for 5 min. Following prehybridization for 6–8 h at 50°C in hybridization solution (HS; 50% formamide, 5 mmol l^{-1} EDTA, 5× SSC, 1× Denhardt's solution, 0.1% Tween-20, 0.5 mg ml^{-1} yeast tRNA), hybridization was performed by incubation of preparations with approximately 1 μg DIG-labeled RNA probe ml^{-1} HS at 50°C for 1–3 days. A series of stringent washes were performed after hybridization: 50% formamide, 5× SSC and 1% SDS at 60°C for 30 min; 50% formamide, 2× SSC and 1% SDS at 60°C for 30 min; 0.2× SSC at 55°C for 30 min twice. Hybridized preparations were labeled with alkaline phosphatase-conjugated, anti-DIG antibodies (Boehringer Ingelheim Pharmaceuticals Inc., Point Richmond, CA, USA) according to the manufacturer's protocol. Hybridization patterns were visualized in a NBT/BCIP alkaline buffer solution (Boehringer). Labeled preparations were fixed in 4% paraformaldehyde in methanol. Following fixation, preparations were embedded in 3:1 glycerol:PBS on glass slides and photographed using an SMZ-1500 stereo microscope (Nikon) equipped with a Pro 600ES CCD imaging device (Pixera, Corp., Los Gatos, CA, USA). Alternatively, an inverted DIAPHOT 300 (Nikon) equipped with Hoffman

contrast optics and Fuji 2S Pro digital SLR camera was used. ImageJ software was used for image contrast and image stack reconstruction.

Quantitative real-time PCR (qPCR)

RNA was isolated from 70% methanol-fixed tissues of *An. gambiae* 4th instar larvae, pupae and adults using an RNAqueous-Micro kit (Ambion). Purified DNase-treated RNA was converted to cDNA using the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen) with random hexamers. agNAT8 expression levels were determined using SYBR Green dye technology on an ABI Prism 7000 Sequence Detection System (Applied Biosystems, Columbia, MD, USA). *An. gambiae* 18S ribosomal RNA (NCBI Accession no. AF417778) was selected as an evenly expressed endogenous control transcript to normalize the calculations of target mRNA relative to differences in the amount of cDNA added to each reaction. qPCR primers were designed using Primer Express v. 2.0 software (Applied Biosystems). Sequences of qPCR primers are present in Table 1. qPCR reactions were performed in triplicate in a total volume of 25 μl containing 12.5 μl SYBR Green PCR Master Mix (Applied Biosystems), 200 nmol l^{-1} (agNAT8) or 300 nmol l^{-1} (18S rRNA) of each primer, and approximately 1–10 ng cDNA template. Control runs included a subset of PCR components lacking the cDNA template. The PCR program was run as follows: 50°C for 2 min, 95°C for 10 min, and 40 cycles of 95°C for 15 s, 60°C for 1 min followed by a dissociation stage of 95°C for 15 s, 60°C for 20 s, and 95°C for 15 s. Dissociation curve analysis was performed to confirm the specificity of the reaction products. Relative expression values were calculated according to the equation:

$$\text{Ratio} = (E_{18S})^{\text{Ct,sample}} / (E_{\text{agNAT8}})^{\text{Ct,sample}} .$$

The efficiency (E) of each primer set was determined by plotting template dilutions against Ct values and is equal to $10^{[-1/\text{slope}]}$. Ct refers to the cycle number at which the fluorescence rises above a set threshold. Expression of agNAT8 in each tissue was normalized to that found in whole larvae, which was set to a value of one. Data represent three averaged replicates of two independent experiments. EXCEL software (Microsoft®) was used to analyze data and SigmaPlot 8.0 (Systat Software Inc., IL, USA) was used to generate the final graphs.

Results

Molecular characteristics of agNAT8 gene, transcript and protein

An amino acid transporter, agNAT8, was cloned from a midgut tissue-specific cDNA collection of 4th instar *An. gambiae* larvae (NCBI Accession no. AAN40409). The identified functional agNAT8 transcript represents a 1911 nucleotide-long open reading frame that encodes a protein with 636 amino acid residues and an estimated molecular mass of 70 247 Da. The agNAT8 protein has an estimated isoelectric point=7.8108 and a formal net charge of 6.0 at neutral pH. The agNAT8 transcript is a splice product of *An. gambiae* EMBL gene ID ENSANGG00000015692, which is located on the 3L chromosome (Fig. 1). The splicing scheme for the coding region of the agNAT8 transcript includes six introns and seven exons (Table S1 in supplementary material). The predicted versions of agNAT8, XP_309840.2 (NCBI) and ENSANGP00000018181 (EMBL) have nine non-synonymous SNPs *vs* the cloned agNAT8, which result in amino acid differences in positions 149 K/E; 181 L/P; 189 S/A; 192 N/Y; 230 S/W; 294 T/A; 328 Y/C; 486 Y/C; 599 Q/H (N>3 clones for each position). In addition, 26 synonymous SNPs were identified in the pool of cloned transcripts, which may be the result of a normal polymorphism of the G3 *An. gambiae* strain as well as usage of multiple specimens for cDNA preparation (Table S1 in supplementary material).

The NATs population in the *An. gambiae* genome represents a condensed cluster of seven genes with all of its members located on the

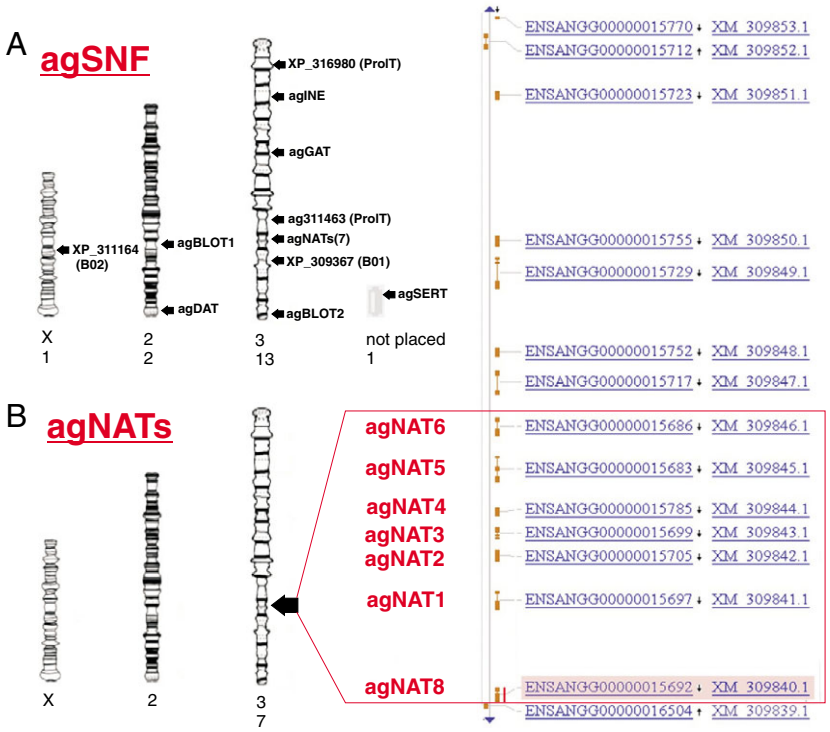

Fig. 1. Physical map of SNF and NAT members in the *An. gambiae* genome. agSNF (A) represents the relative chromosomal distribution of the entire SNF population. agNATs (B) show the exact location of all agNATs on chromosome 3, which is expanded by a virtual map of particular genes on the right. Values beneath chromosomes represent chromosome numbers and the number of genes belonging to the agSNF or agNAT cluster on the corresponding chromosomes. agSERT is a serotonin transporter, the location of which is presently unknown.

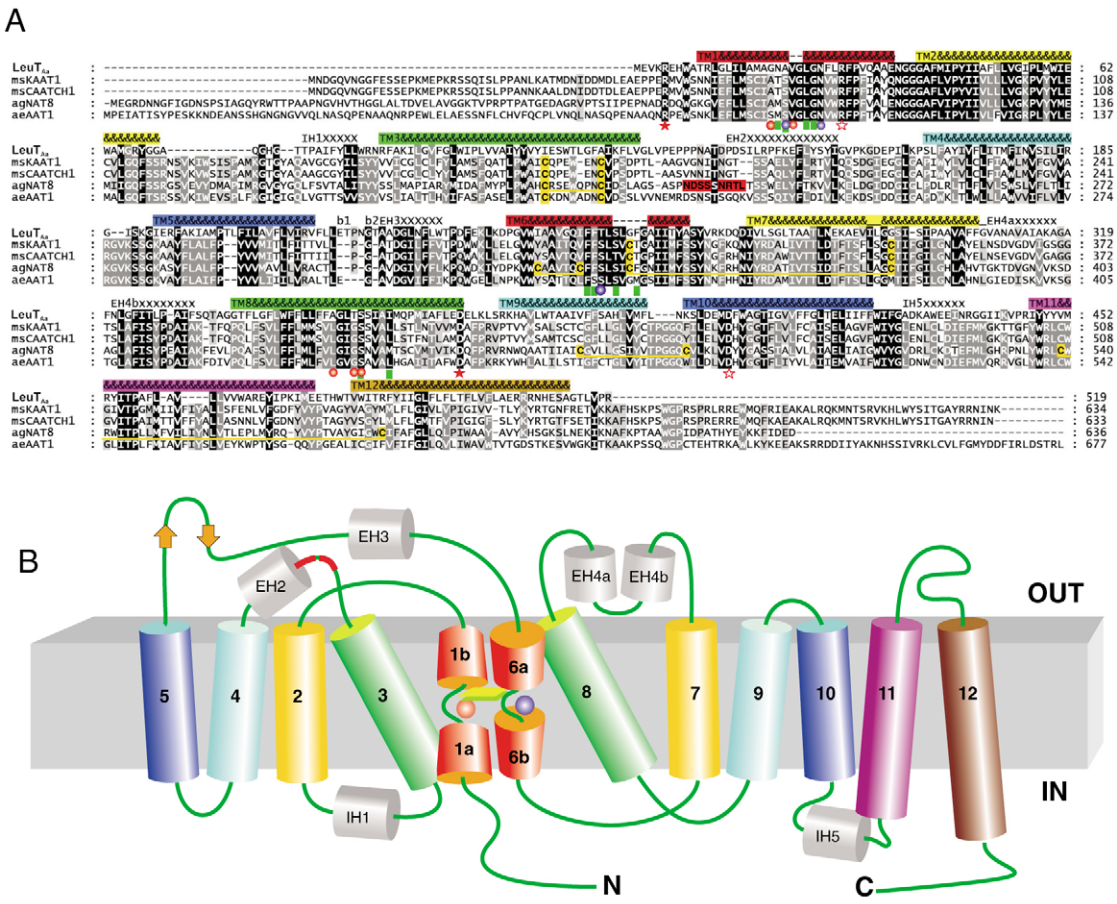

Fig. 2. Alignment and reconstruction of agNAT8 structure. (A) Sequence/structure alignment of characterized insect transporters relative to the first crystallized bacterial NAT from *A. aeolicus*, $LeuT_{Aa}$ (Yamashita et al., 2005). (B) 2D structure of agNAT8 based on structural homology with the $LeuT_{Aa}$ protein sequence. NCBI Accession no.: $LeuT_{Aa}$, NP_214423 (PDB no., 2A65); msKAAT1, AAC24190; msCAATCH1, AAF18560; aeAAT1, AAR08269; agNAT8, AAN40409. Filled and open stars represent putative cationic gates at extra- and intracellular interfaces, respectively. Squares indicate putative substrate binding sites; red and blue spheres outline sites, which interact with the first and second sodium ion respectively. Red and yellow boxes show putative glycosylation motifs and disulfide bridges, respectively. The 12 transmembrane domains are numbered 1–12.

positive strand without apparent interruption by other known genes (Fig. 1). The agNAT8 gene represents the last downstream unit in this gene cluster. The agNATs cluster is flanked downstream by a negative strand-encoded glucosamine/galactosamine-6-phosphate isomerase gene (ENSANGG00000016504). Five genes are also present upstream of the NAT cluster on the positive strand and appear to be structurally related to membrane transporters (data not shown). These upstream genes represent SNF-unrelated, uncharacterized transporters, which share sequence homology with a membrane-associated transporter protein, AIM-1 (a.k.a. Melanoma antigen AIM1) and more distant homology to proton-driven sugar symporters of plants and animals.

The agNAT8 peptide includes a Sodium Neurotransmitter symporter Family domain, PFAM ID=00209 (79–614 amino acid position, Expect=2×10^{-119}), overlapping with a Na^+-dependent transporter of the SNF family domain, COG0733, (79-556, E=2×10^{-50}), which represent the structural fingerprints of SNF members (Clark and Amara, 1993; Kanner, 1994; Reizer et al., 1994). Despite moderate sequence identities of 42%, the hydropathy profile and putative secondary structure of the agNAT8 protein are virtually identical to the earlier cloned NATs, msKAAT1 and msCAATCH1 from the

Fig. 3. A comparative Bayesian evolutionary dendrogram of selected SNF members. This tree includes SNF members from three prokaryote and five eukaryote genomes and the most characterized SNF members from various insects; species abbreviations and color codes are given in the lower right corner insert. Branch lengths correspond to an evolutionary distance scale of 0.1 mutation per site (mps). The numbers at the nodes are posterior probabilities. Node values in the range 0.95–0.99 are shown, whereas nodes with a probability less then 0.95 were collapsed, and probability values of 1 are not shown. The red arrow indicates the position of cloned agNAT8. White and black triangles correspond with apparent gene expansion and conservation nodes. metNTT, metazoan neurotransmitter transporters; miNAT, mammalian-insect nutrient amino acid transporters (NAT); iNAT, insect NATs; ceNAT, *Caenorhabditis elegans* NATs, archNAT, archaean NATs; bacNAT, bacterial, *Bacillus cereus* NATs.

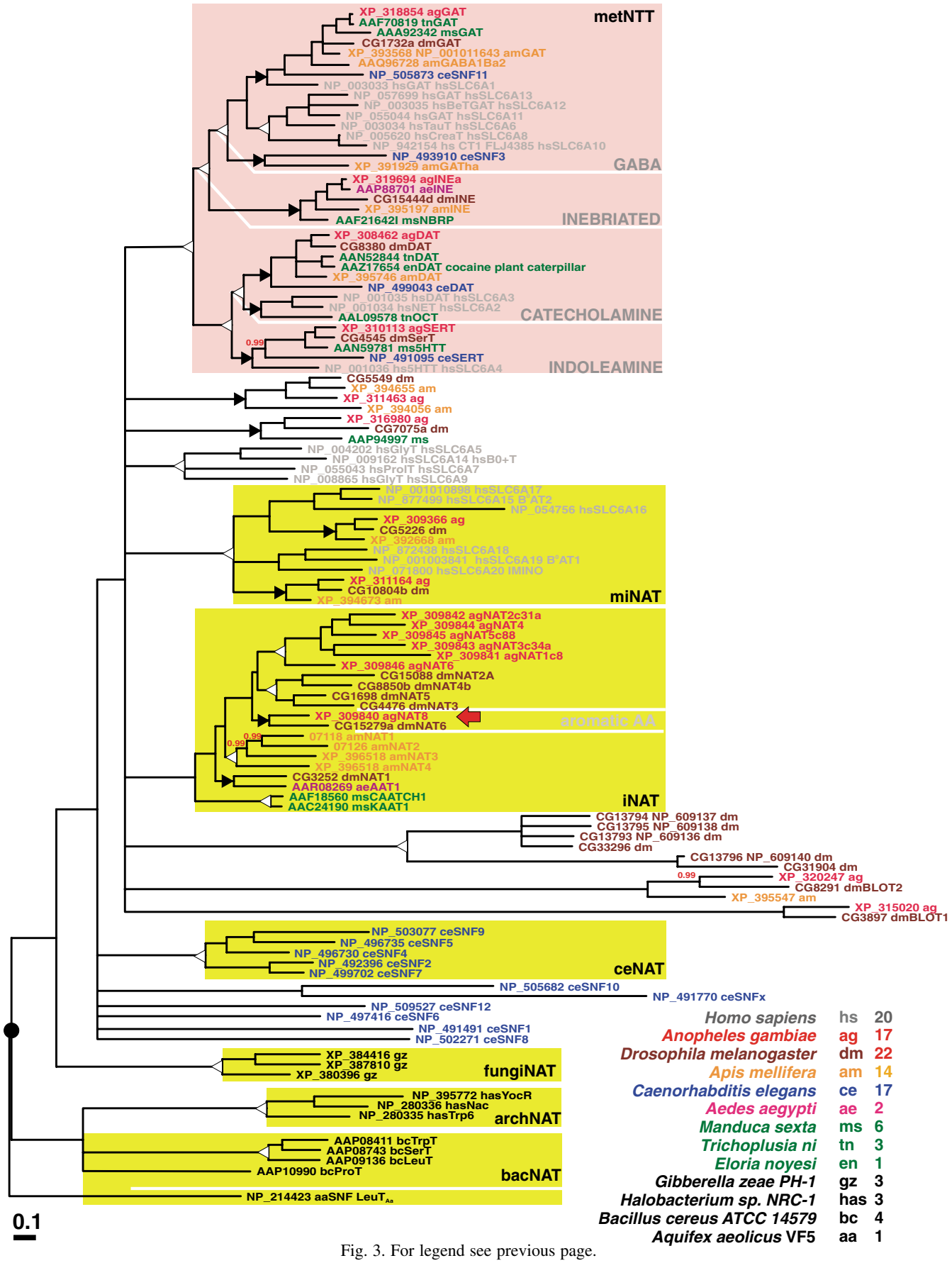

Fig. 3. For legend see previous page.

caterpillar midgut, and aeAAT1 from the *Ae. aegypti* larval midgut (Fig. 2A) (Boudko et al., 2005a). The predicted secondary structure includes 12 transmembrane domains (Fig. 2A,B). 33 enzyme-specific phosphorylation sites are predicted, six of which could be phosphorylated by more than one kinase (Table S2 in supplementary material). Several putative phosphorylation sites and other post-translational modification motifs are recognizable in external loops and transmembrane domains (Table S2 in supplementary material). Putative glycosylation sites are identified at positions 218–221 (NDSS) and 223–226 (NRTL), which reside in the extracellular loop between the 3rd and 4th TMDs (Fig. 2). This loop has been proposed to bear the motifs contributing to substrate selectivity mechanisms (Nelson and Lill, 1994).

Phylogenetic position of agNAT8 in the insect NAT subfamily

A Bayesian inference evolutionary tree of the selected SNF members (Fig. 3) confirms that the cloned agNAT8 is a novel member of the earlier identified insect-specific NAT cluster (iNAT), a subfamily of the SNF (Boudko et al., 2005a), which in insects splits into paralogous groups. For example, most insect NATs: agNAT1-6, msKAAT1 and msCAATCH1, dmNAT2-5 and amNAT1-4, represent a paralogous diversification pattern that is consistent with rapid gene duplication and paralogous modification of transporter structures and phenotypes (represented by white expansion nodes in Fig. 3). On the other hand, agNAT8 and dmNAT6 as well as dmNAT1 and aeAAT1 lie outside of this paralogous diversification scheme, suggesting some degree of conservation of their phenotypes. *Ae. aegypti* and *M. sexta* NATs represent only a portion of the NAT populations existing in these species, since their genomes and protein annotation are not yet available and analyzed. agNAT8 together with dmNAT6 represent the basal ancestral unit in the large expansion of dipteran NATs (Fig. 3). In addition, agNAT8 is a basal unit in a group of *Anopheles* NATs, which is different from other groups including earlier cloned *M. sexta* and *Ae. aegypti* transporters. At present, the insect-specific NAT clusters of genomic model insects include seven members from *An. gambiae*, six from *D. melanogaster*, and four from *Ap. mellifera*; moreover, they lack SNF members from any other selected species. However, there are several mammalian and nematode SNF members in groups that are proximal to the insect NAT cluster and likely include counterparts of the iNAT phenotypes (Fig. 3).

Electrochemical characteristics of agNAT8 transport mechanism

Steady-state heterologous expression of agNAT8 in *Xenopus* oocytes was observed by measurements of amino acid induced currents on the 2nd–6th days after transcript injection. There was a dramatic increase in phenylalanine-induced sodium currents and in the ratio of phenylalanine uptake of experimental versus control, un-injected or deionized water-injected oocytes (Figs 4–8), along with a moderate increase of leak currents (data not shown). Screening of putative substrates including 20 proteinogenic amino acids, selected amino acid derivatives and several neurotransmitters, revealed a unique substrate selectivity profile that has not previously been reported (Fig. 4). L-phenylalanine induced the largest currents, associated with the highest transport velocities, whereas other

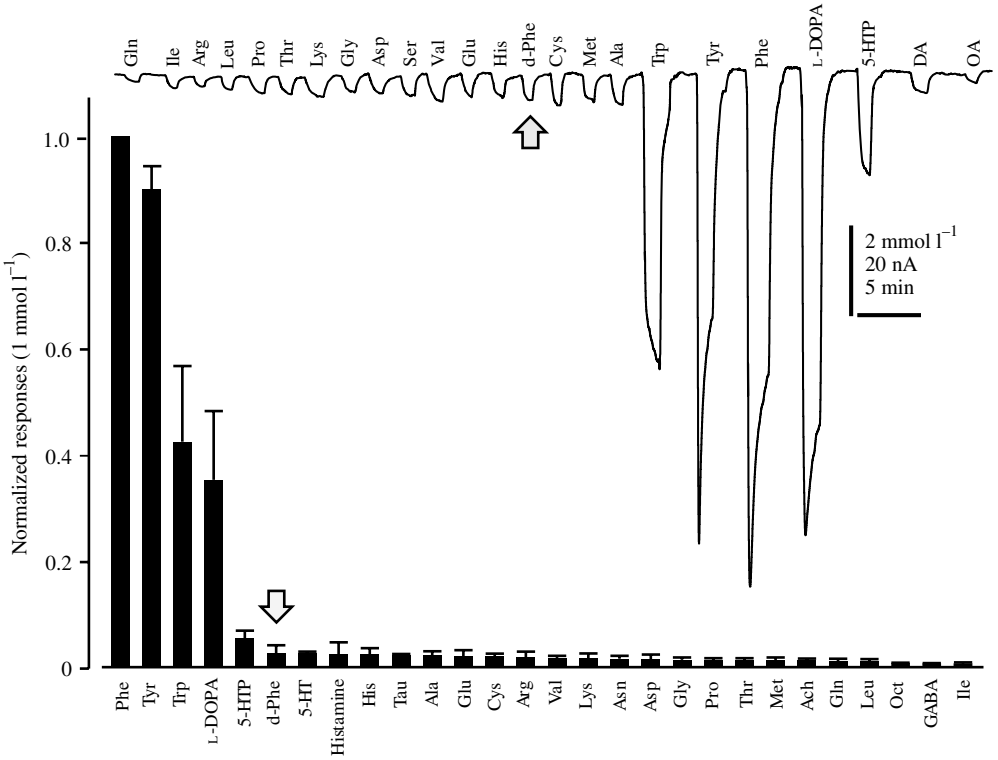

Fig. 4. Profile of agNAT8-mediated amino acid-induced currents. The upper line combines typical currents of agNAT8-expressing oocytes upon a trial application of selected substrates at 2 mmol l^{-1} concentrations. The lower histogram incorporates mean values of induced responses upon application of 1 mmol l^{-1} concentration of selected substrates normalized to Phe induced currents (values are mean + s.d.; $N \geqslant 3$ recordings for at least three different oocytes for each data point). Holding potentials for all recordings were –50 mV. Standard abbreviations of amino acids and neurotransmitters are used. Arrows highlight the D-isomer of phenylalanine.

catechol-branched substrates, L-tyrosine and L-DOPA, induced the second and third highest current magnitudes (Fig. 4). Indole-branched L-tryptophan and its immediate metabolic derivative, 5-hydroxytryptophan (5-HTP) induced the 4th and 5th highest currents. Some currents were induced in agNAT8-injected oocytes by other neutral amino acids, but with much smaller magnitudes (Fig. 4). There were responses to cationic and anionic amino acids (Fig. 4). These responses were correlated with agNAT8 expression, but they were relatively small and not significantly different within that group (based on paired *t*-tests, data not shown). The rank order of agNAT8-mediated currents was unchanged at substrate concentrations of 0.25, 0.5, 1 and 2 mmol l^{-1}. At 1 mmol l^{-1}, very small or no responses were observed upon applications of classical neurotransmitters, including GABA, 5-HT, dopamine, norepinephrine, octopamine and histamine (Fig. 4). The substrate selectivity of agNAT8 was stereo-specific with strong preferences for L-amino acids (e.g. D-Phe *vs* L-Phe induced currents; Fig. 4, arrow).

Substitution of Na^+ with K^+ and Li^+ dramatically reduced or completely abolished substrate-induced currents at –50 mV holding transmembrane voltage (Fig. 5). Significant decay of currents was also determined upon Cl^- substitution with gluconate$^-$ (15.2±4.6% of normalized responses, paired *t*-test, $P<0.0001$, $N\geqslant3$). Typical examples of a voltage-induced family of currents in agNAT8 expressing oocytes are presented in Fig. 6A. The voltage step reconstruction revealed sub-linear *I/V* characteristics of agNAT8 at physiological values of transmembrane voltage (Fig. 6B,C). Moreover, the *I/V* plot suggests that an agNAT8 mechanism may utilize an inward K^+ electrochemical gradient (Fig. 6C, red line in –60 to –120 mV range) and appears to be rectified (irreversible). Therefore, agNAT8 represents a Na^+- and/or K^+-driven and Cl^- facilitated aromatic amino acid transporter with preferences for catechol-branched amino acids and products of their metabolism.

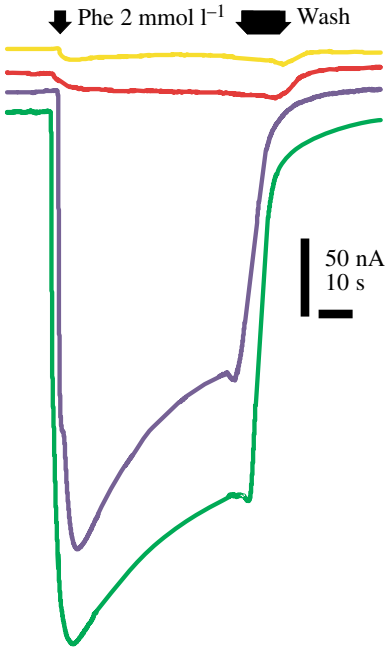

Fig. 5. Ion dependency of agNAT8. Representative data from a typical ion substitution experiment. Different coloured lines represent the response upon application of 2 mmol l^{-1} phenylalanine in oocyte medium with different ion concentrations: yellow, 100 mmol l^{-1} LiCl; red, 98 mmol l^{-1} KCl; blue, 98 mmol l^{-1} sodium gluconate; green, 98 mmol l^{-1} NaCl. Holding potentials were –50 mV.

Comprehensive saturable kinetic analysis of amino acid-induced currents was performed for aromatic substrates that induced the largest sodium-dependent currents. Fig. 7A,B are examples of kinetic graphs for phenylalanine induced currents. Similar analyses were used to derive $E_{0.5}$ and Hill coefficients for different substrates, which are summarized in Fig. 7C,D. $E_{0.5}$ values for selected substrates were within the 0.1–4 mmol l^{-1} range, with an order of L-Tyr<L-Phe<L-Tr$P\ll$L-DOPA<5-HTP (Fig. 7C), while the apparent Hill coefficients for those substrates were close to 1 (Fig. 7D). Accumulative agNAT8-mediated transport was confirmed by uptake of isotope-labeled phenylalanine and methionine as representative of substrates with the highest and moderate apparent transport velocity and $E_{0.5}$, respectively (Fig. 8A,B). Finally, substrate-induced currents in agNAT8-expressing oocytes were pH dependent, increasing near-linearly from pH 6.2 to pH 9.3 (Fig. 8C).

Spatial and temporal profiles of agNAT8 expression pattern

In situ hybridization of whole-mount preparations with antisense probes revealed a site-specific expression pattern of the agNAT8 transcript in *An. gambiae* larvae (Fig. 9). The control preparations exposed to sense probes produced small non-specific background signals (data not shown). Strong and specific hybridization patterns were observed in the posterior midgut, in the proximal portion of the Malpighian tubules, and specific cellular populations of the rectal gland (Fig. 9A,B). All these sites correspond with primary apical absorption of nutrient amino acids from the alimentary canal lumen. Very strong hybridization was also detected in the salivary glands (Fig. 9C), which mediate synthesis and secretions of salivary and digestive enzymes into the alimentary canal lumen. Slightly weaker signals were observed in the gastric caeca, a deposit site for nutrients that are utilized in earlier stages of larval development; and in the cardia, a structure that aids in the synthesis of the peritrophic membranes (Fig. 9D). agNAT8 expression was also detected in specific neuronal populations of the ventral nerve cord (Fig. 9E) and peripheral neuronal populations that are associated with chemo-, visual- and mechano-sensory afferents of mosquito larvae (Fig. 9F–J). Specific hybridization signals were also observed in the tracheal plexus and tissues attached to the integument, including muscles and fat body (data not shown). Variations in the labeling were detected in the posterior midgut, where signals were relatively weaker than in the proximal portion of the Malpighian tubules (data not shown).

The results of a quantitative real-time PCR (qPCR)

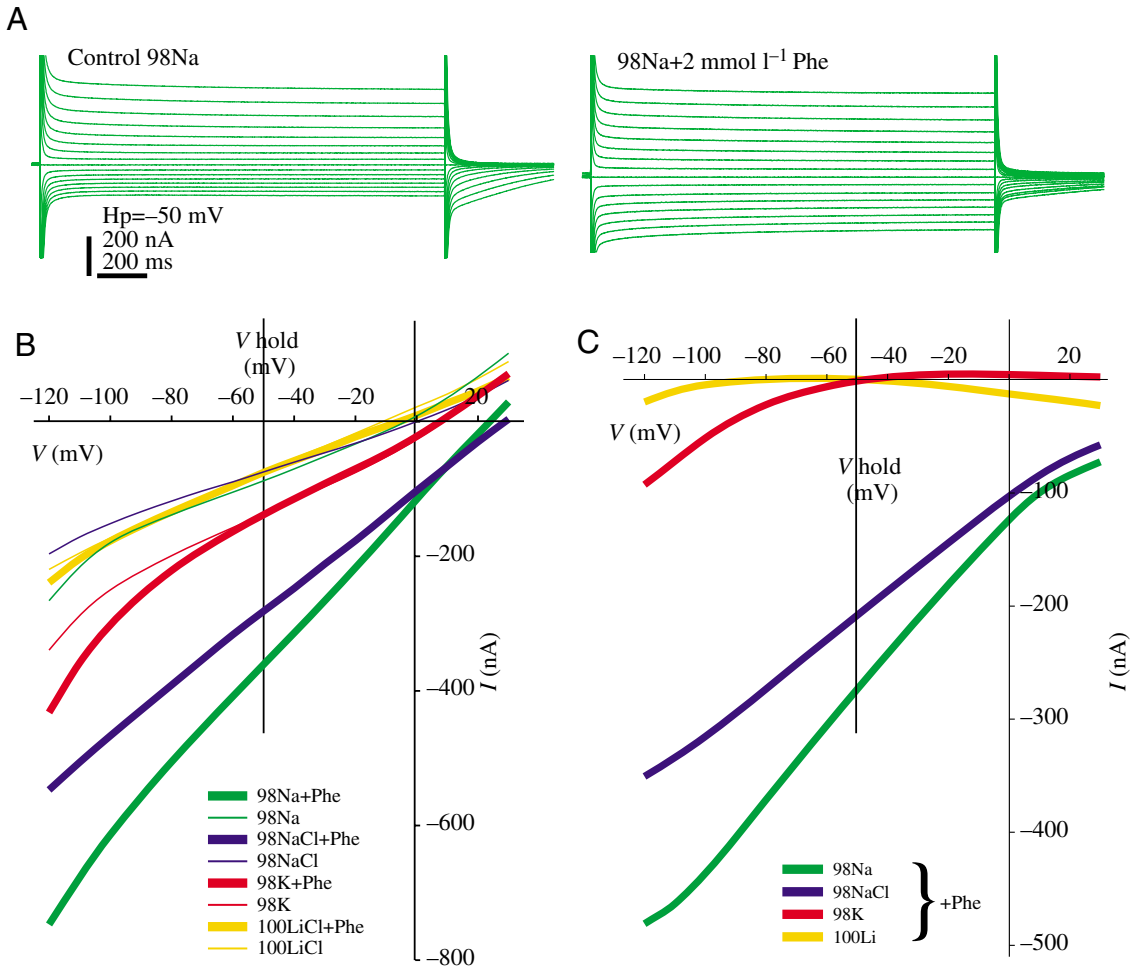

Fig. 6. Electrical properties of agNAT8 transporter. (A) Families of current traces induced by square pulse polarization of oocyte membranes in a 98 mmol l^{-1} NaCl solution (left) and after application of 2 mmol l^{-1} phenylalanine (right) (V_{hold}, holding voltage=–50 mV, first pulse=+30 mV, last pulse=–120 mV; increment=–10 mV). (B) Family of *I*/*V* plots upon ion substitution in the bathing media. Line colors and thickness represent different ion and substrate profiles, respectively (see inserts). (C) *I*/*V* profiles with subtracted 'control' leak currents in the absence of substrate molecules. Phe concentration in B and C = 1 mmol l^{-1}.

experiments confirmed the presence of the agNAT8-specific transcript in the areas of elevated expression, which were identified by *in situ* hybridization (Fig. 10). In addition, it allows a quantitative comparison of relative expression values, in which the transcript expression is strongest in the rectal glands and decreases in the set of tissues: salivary gland/cardia>Malpighian tubules>posterior midgut>CNS>gastric caeca>anterior midgut (Fig. 10A). Intriguingly high expression levels were detected in the preparation of whole larvae, which may suggest that some tissues have a very high expression of agNAT8 transcripts. This qPCR data may correspond with high *in situ* hybridization signals in the epithelia tracheal plexus, which was included in whole larvae but absent in isolated integument samples. Additional analysis conducted for selected tissue samples from post-larval developmental stages showed very high expression of agNAT8 transcripts in whole pupae and adult gut as well as extremely high expression in the adult head (Fig. 10B).

Discussion

Arguably, all families of secondary transporters that accept amino acids as substrates have now been identified [see HUGO SLC classification and recent reviews (Boudko et al., 2005c; Broer, 2002)]. Remarkable progress in the characterization of novel phenotypes and the functional and phylogenetic classification of membrane transporters has been made. A few transporters involved in the transport of neutral amino acids have been the subject of intensive study in biomedically relevant mammalian models; however, the principal molecular mechanisms that mediate transmembrane accumulation of these essential substrates remains incomplete in mammals and is virtually unknown in the rest of the Animal Kingdom. We have cloned the first aromatic substrate-selective transporter from the NAT subfamily of SNF whose functional characterization provides new insights into unraveling the molecular and evolutionary basis of the essential amino acid transport network.

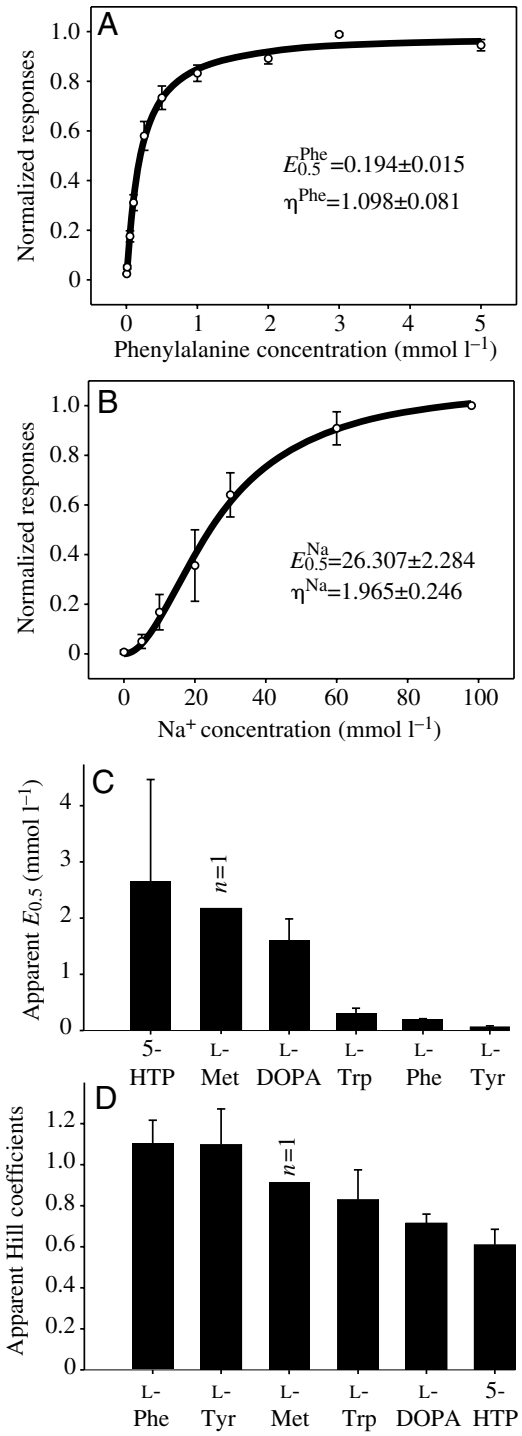

Fig. 7. Kinetic properties of agNAT8 transport mechanism. (A) The effects of increasing phenylalanine concentration on agNAT8-mediated currents; values are means ± s.e.m.; $N \geqslant 3$. (B) The effects of increasing sodium ion concentration on phenylalanine-elicited, agNAT8-mediated currents; values are means ± s.e.m.; $N \geqslant 3$, except for L-Met, $N=1$. (C) $E_{0.5}$ for substrates with highest apparent transport velocity; values are means ± s.d.; $N \geqslant 3$. (D) Estimated Hill coefficient for selected substrates; values are means ± s.d.; $N \geqslant 3$. Lines in A and B represent iterative nonlinear regression of the data with three-parameter Hill function ($E_{0.5}^{[S]}$ and $\eta^{[S]}$ are incorporated in the graph panels; see inserts).

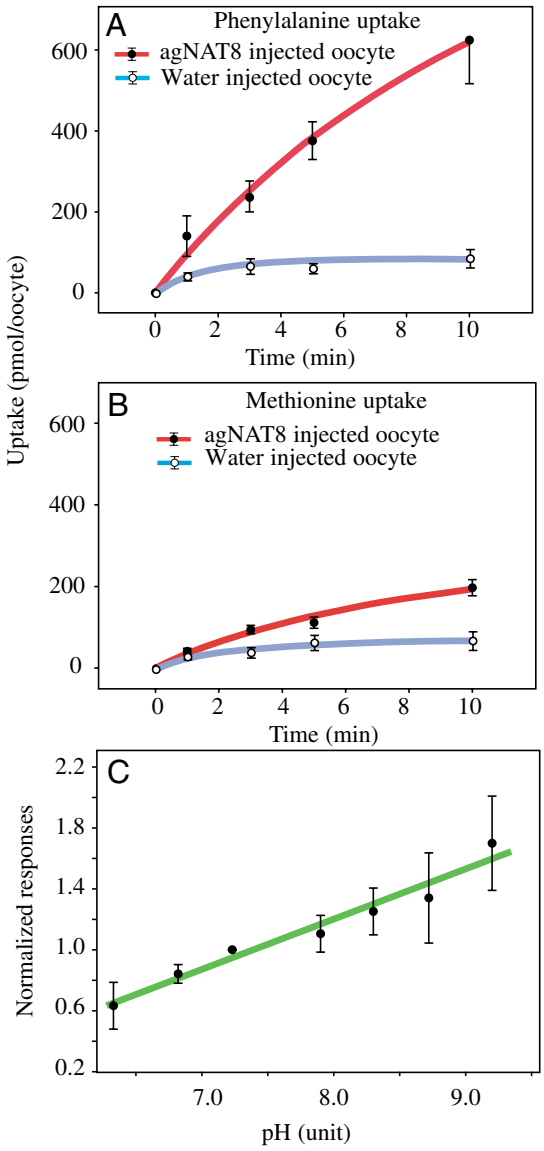

Fig. 8. Accumulation of isotope-labeled phenylalanine (A) and methionine (B) in deionized water-injected (empty circles and blue lines) and agNAT8 transcript-injected (black filled circles and red lines) oocytes. Values are means ± s.e.m. for at least three oocytes per data point. Lines are nonlinear regression fits of the data sets by a single rectangular function, $f=a*x/(b+x)$. (C) pH dependency of phenylalanine-induced currents mediated by agNAT8. Values are means ± s.e.m. for at least three oocytes/experiment per data point. Data from different oocytes were normalized in each group relative to values at pH 7.3. The straight line represents a linear regression of the entire data set, $f=y_0+ax$.

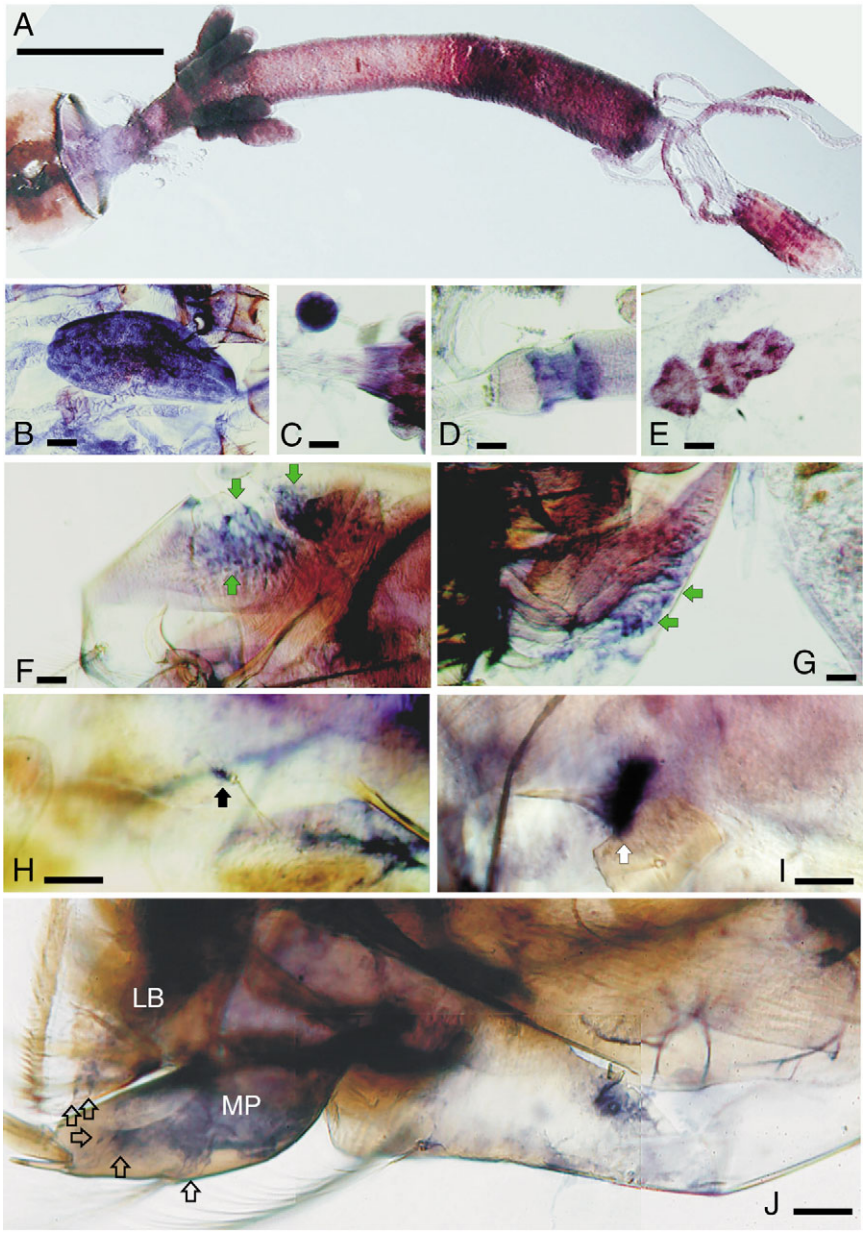

Fig. 9. Spatial expression of agNAT8 transcript in the alimentary canal, CNS and PNS of *An. gambiae* larvae. (A) *In situ* hybridization of agNAT8-specific probes in the whole-mount midgut from 4th instar *An. gambiae* larvae. (B) Magnified image of the rectal gland. (C) Intensive hybridization in the salivary gland. (D) Stretched-opened esophageal epithelial infolding, cardia; this epithelial region is associated with secretion of the tubular peritrophic membrane. (E) Area-specific labeling in the isolated neuronal ventral cord of CNS; three thoracic ganglia are shown. (F,G) Dorsal and lateral views of agNAT8 hybridization pattern in the larval stemma (eye); this labeling is associated with primary photosensitive neurons (green arrows). (H) An example of hybridization of sensory neurons associated with vibration-sensitive hair sensilla on the larval head (black arrow). (I) Labeling in the basal part of the peg sensillum at the head capsule (white arrow); this structure has an unknown sensory role in larvae, but it is morphologically similar to temperature-sensitive structures in adults. (J) agNAT8-pecific hybridization of specific populations of chemosensory neurons in maxillary pulp (MP) and labium afferents (LB) (open arrows). AMG, anterior midgut; CA, cardia, a sub-esophageal invagination; GC, gastric caeca; MT, Malpighian tubules; PMG, posterior midgut. All images were acquired using a Hoffman contrast microscope except for A and E, which were acquired using a stereo microscope. Bars, 1000 μm (A), 200 μm (B–E), 100 μm (F–H,J), 50 μm (I).

agNAT8 resides in the NAT cluster, which is the largest subfamily of the SNF. More specifically, agNAT8 is a member of an insect-specific paralogous gene expansion that we have designated as 'insect NATs'. To date four insect NATs have been characterized including msKAAT1 (Castagna et al., 1998), msCAATCH1 (Feldman et al., 2000), aeAAT1 (Boudko et al., 2005a) and now, agNAT8. Partial properties of a few other novel NATs have been reported at scientific meetings (Assis et al., 2004; Boudko et al., 2005b). All characterized insect NATs share the principal function of transporting neutral amino acid substrates. However, individual insect NATs differ from one another in their substrate spectra. Specifically, agNAT8 has a high transport velocity and apparent affinity for large aromatic amino acids, especially for catechol-branched substrates including phenylalanine and its metabolites e.g. tyrosine and L-DOPA. Both the narrow preferences and aromatic substrate spectra represent unique properties of agNAT8 among known NATs.

In addition to substantial intra- and cross-specific variations in substrate preferences, insect NATs can utilize alternative electrochemical gradients for Na^+, K^+ or both ions. For example, caterpillar msKAAT1 exhibits an electrical gradient-driven K^+:amino acid symport in contrast to Na^+ motive force-driven mammalian NATs (Castagna et al., 1998). Like aeAAT1 (Boudko et al., 2005a), agNAT8 is Na^+-dependent under physiological conditions, with a coupling stoichiometry of 2 sodium ions to 1 amino acid substrate ion. However, both mosquito NATs showed non-linear increases in inward potassium current at negative transmembrane potentials relative to the K^+ equilibrium potential for oocyte membranes

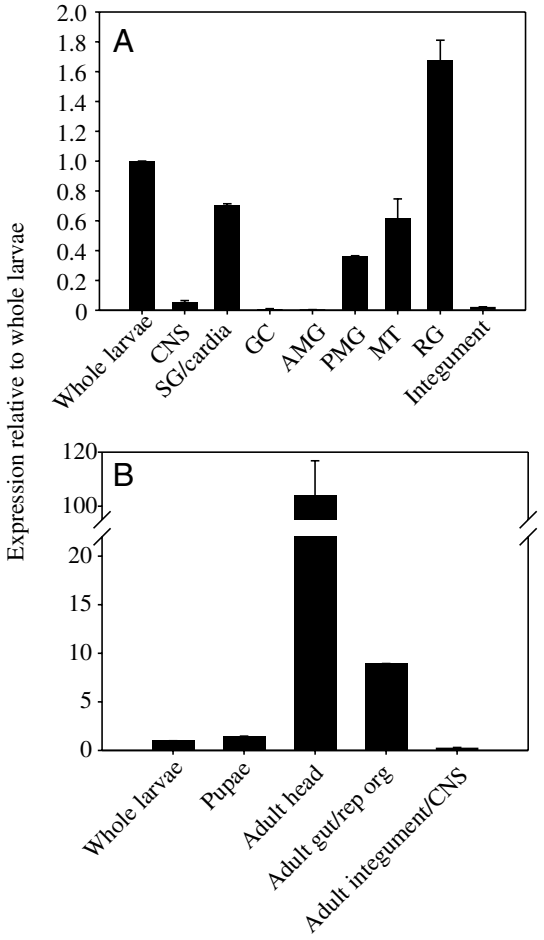

Fig. 10. Quantitative real-time PCR analysis of agNAT8 expression in distinct larval tissues and developmental stages of *An. gambiae*. (A) Relative expression levels of agNAT8 in tissues of 4th instar larvae. (B) Relative expression levels of agNAT8 at various developmental stages and in adult tissues. Note the different scales on the *x* axis for A and B. Data are represented as the mean of three averaged replicates of two independent experiments plus standard error. *An. gambiae* 18S ribosomal RNA was used as a reference gene. Results were normalized to values for whole larvae, which was set to a value of 1.0. CNS, central nervous system; SG, salivary glands; GC, gastric caeca; AM, anterior midgut; PM, posterior midgut; MT, Malpighian tubules; RG, rectal gland; rep org, reproductive organs.

(Fig. 6C). This observation is consistent with a substrate-induced inward K^+ current. Although actual K^+-coupled uptake was not confirmed due to the instability of agNAT8-expressing oocytes in Na^+-free K^+ medium, the inward and substrate-dependent K^+ current strongly suggest a coupled substrate uptake mechanism, which can utilize both direct Na^+ and/or inward K^+ motive forces. The ability to utilize K^+ gradients for intestinal absorption of amino acids is generalized in herbivorous insects, which have but traces of Na^+ ions in their diets (Giordana et al., 1998). Freshwater mosquito larvae have to deal with trace quantities of both Na^+ and K^+ ions. Presumably, sodium ions are recycled to the midgut lumen by the putative H^+ V-ATPase-coupled cation exchanger pathways that alkalinize anterior midgut lumen (Boudko et al., 2001a). It may satisfy the requirement for sodium ion-dependent amino acid uptake by recycling the pool of coupling ions in the posterior midgut of mosquito larvae. Therefore, the ability to switch between sodium and potassium ions appears to be a facultative property of mosquito NATs. This property may benefit amino acid absorption in mosquito larvae, which experience a large shift in intestinal cation profiles upon exposure to different mineral environments.

The number of NATs in insects varies, e.g. seven in *Anopheles* and six in *Drosophila*. The emerging *Apis mellifera* genome includes at least four NATs with a comprehensive structure although only two NAT genes have been annotated in the databases. In mosquito (Fig. 1) and bee genomes (data not shown), all NATs are condensed in tight proximity. They are also condensed in the phylogenetic tree, representing multiple and rapidly duplicated groups of genes (Fig. 3). Extensive gene duplications reduce stabilizing pressure on the entire group and facilitate the acquisition of new phenotypes, which enables a high plasticity of an integrated NAT mechanism. Allocation of NATs within the genome is also useful for the temporal and spatial coordination of the expression patterns.

Insect NATs share upstream proximity with a gene cluster that includes several characterized mammalian amino acid transporters, designated here as 'mammalian + insect NATs' (miNATs in Fig. 3) because it also includes two orthologous groups of orphan insect transporters (Boudko et al., 2005a). Partial data are available regarding expression of *Drosophila* NATs from that cluster (http://www.fruitfly.org/cgi-bin/ex/basic.pl): one of them, CG10804, is so meagerly expressed that it is undetectable in whole mounts by *in situ* hybridization of *Drosophila* embryos whereas the other one, CG5226, is expressed in the embryonic CNS, including both neuronal and glial cells.

A principal role in the absorption of nutrient amino acids is clear for the majority of mammalian NATs (three of five in human; Fig. 3). The mammalian B^0 system presently comprises two NATs. SLC6A19 is the broad spectrum neutral amino acid transporter, which also represents a molecular basis of Hartnup disorder, a disease characterized by severe neutral aminoaciduria due to deficiency of neutral amino acid resorption (Bohmer et al., 2005; Broer et al., 2006a; Broer et al., 2004; Broer et al., 2005; Romeo et al., 2005; Seow et al., 2004). The cloning of a transporter (SLC6A20), which appears to be the closest relative to B^0 (SLC6A19), revealed a phenotype of the IMINO system, characterized by sodium and chloride ion dependent proline transport (Kowalczuk et al., 2005; Takanaga et al., 2005b). IMINO system transporters mediate absorption of the nutrient proline. In insects proline is involved in osmotic regulation and extensively utilized for cell proliferation and tissue repair (Tapiero et al., 2002) as well as a flight jumping fuel (Gade and Auerswald, 2002). However, the molecular identity of an insect transporter with properties of the IMINO system at present is unknown. Another mammalian NAT, SLC6A15, was designated earlier as an

orphan v7-3 transporter since it does not transport neurotransmitters (Farmer et al., 2000; Sakata et al., 1999; Uhl et al., 1992). However, recent heterologous expression and analysis of human (Takanaga et al., 2005a) and rodent (Broer et al., 2006b; Takanaga et al., 2005a) orthologs have shown that mammalian SLC6A15 transporters also accept B^0 substrates but differ in substrate preferences *vs* earlier characterized SLC6A19 homologs. In addition, the analysis showed interesting expression profiles, restricted to human brain tissues in man (Takanaga et al., 2005a) but more broadly expressed in mouse, including CNS, intestine and kidney (Broer et al., 2006b). Despite orthologous structures and a cumulative role in nutrient amino acid transport, mammalian NATs are distinct with regard to their electrochemical signature including substrate affinity profiles, apparent kinetics and expression patterns. The data presented here suggest that even more prominent differences in substrate selectivity are present between mammalian and insect NATs. It appears that both populations participate in the B^0 systems; however the insect system operates with a larger number of NATs (Fig. 3). There are also substantial differences in electrochemical mechanisms that may utilize a K^+ gradient. In addition, narrow-spectra and aromatic substrate-specific B^0 transporters are unknown in mammals.

Elevated agNAT8 transcript expression was identified in the rectal glands, posterior midgut, proximal portion of the Malpighian tubules, cardia, salivary glands, CNS and sensory afferents of mosquito larvae. The expression pattern of agNAT8 in the alimentary canal is similar but not identical to that observed for aeAAT1. agNAT8 has a wider spatial expression band in the posterior midgut and proximal portion of the Malpighian tubules and more uniform expression in the gastric caeca *vs* aeAAT1(Boudko et al., 2005a). Fig. 11 is a simplified diagram of the agNAT8 expression pattern relative to functional specializations of midgut epithelia and putative roles of the transporter *in vivo*. Absorptive roles of iNATs in the posterior alimentary canal of mosquito larvae are consistent with roles of homologous mammalian NATs expressed in apical membranes of intestinal and kidney epithelia (Romeo et al., 2006). In contrast, the expression of agNAT8 in salivary glands, gastric caeca and cardia appear to correspond with an inverse, basal polarity, because these cells mediate extensive protein synthesis and subsequent secretion of amino acid biopolymers through the apical membrane that require substrate absorption from the hemocoel *via* the basal cell membrane (Fig. 11).

Hybridization signals in the CNS ganglia of larvae were limited to specific neuronal populations, and were especially intense in the cerebral sensory afferents (Fig. 9E–J). These are the first evidence of neuron-specific NAT expression in the CNS and PNS. Such a pattern suggests an elevated role of NATs in particular neurons and neuronal populations. The

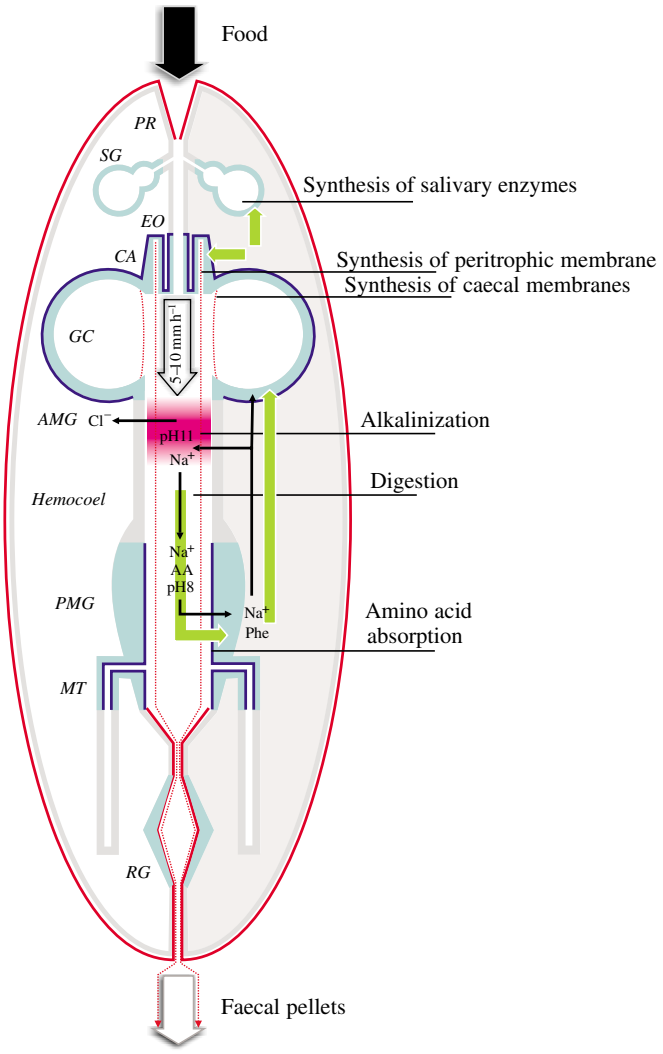

Fig. 11. Model of agNAT8 functions superimposed with patterns of its spatial expression and putative membrane localizations, along with specific physiological roles of particular epithelial areas and with a simplified Na^+ ion recycling scheme. Light blue fill corresponds to epithelial areas with areas of hybridization and qPCR detection of agNAT8 transcripts. Dark blue outlines correspond with theoretically possible membrane localization of agNAT8. The posterior midgut is associated with apical Na^+-dependent absorption of free amino acids from the midgut lumen into epithelial cells (green arrow). A complementary transporter for emitting amino acids *via* the basal membrane of the posterior midgut into the hemocoel is expected but presently unknown; also, it may not require an active transport mechanism since the intracellular concentration of free amino acids in this area is high *vs* the concentration in the hemolymph. In contrast, the cardia, salivary glands and gastric caeca clearly are secretory parts of the alimentary canal, which mediates the synthesis and apical secretion of various amino acids in the form of peptides and protein polymers. A basal accumulative transport mechanism is necessary here (green arrow), which is complemented by an apical secretion process. The disposable pool of sodium ions in the midgut lumen, which is involved in the transport process, is recycled *via* the area of anterior alkalinization (pH 11), shown as a pink gradient band in the anterior midgut lumen. PR, pharynx; SG, salivary glands; EO, oesophagus; CA, cardia; GC, gastric caeca; AMG, anterior midgut; PMG, posterior midgut; MT, Malpighian tubule; RG, rectal gland.

agNAT8 mechanism mediates intracellular accumulation of essential substrates, which supply the synthesis of dopamine and octopamine, the insect equivalent of norepinephrine in vertebrates. We assume that neuronal agNAT8 is an essential substrate provider for the synthesis of catecholamine neurotransmitters. The role of agNAT8 in supplying tryptophan for 5-HT synthesis is also emerging; however, it may not be very efficient in the presence of catechol-branched substrates and other more efficient and specific mechanisms may be needed.

The elevated agNAT8 expression in adult mosquitoes and especially in the adult head is intriguing (Fig. 10B). It is unlikely that such over-expression corresponds to the formation of new cuticle sheaths because adult ecdysis is final. In contrast, it correlates with elevated neuronal and sensory development of adult *vs* larvae because adult sensory and neuronal structures, including eyes and chemosensory antennas, are far more massive and perhaps more active. Future morphological analysis with better spatial resolution techniques may reveal the exact roles of agNAT8 in the adult mosquitoes.

Roles of mammalian NATs in the synthesis of non-monoamine neurotransmitters have been proposed for the mammalian NAT SLC6A15 from mouse [B^0AT2 (Broer et al., 2006b; Romeo et al., 2006)] and human [SBAT1 (Takanaga et al., 2005a)]. Both are strongly expressed in the brain and may supply anaplerotic intermediates in the TCA cycle for synthesis of excitatory and inhibitory neurotransmitters e.g. glutamic acid and GABA (Yudkoff et al., 1996).

In summary, agNAT8 is a novel transporter from the insect-specific NATs cluster. It is the first NAT with a narrow specialization in providing essential aromatic amino groups in various tissues, including neuronal. We propose that the variety of NAT phenotypes is driven by rapid gene duplication and reflects an adaptive plasticity of the NAT populations. The primary factor in the structural and functional diversification of NATs appears to be the balance between the demand and availability of essential amino acids (Boudko et al., 2005a; Broer, 2002). Secondary factors of the diversification could be adaptation to different electrochemical motive forces, coupling stoichiometry, as well as cell-specific expression and membrane docking mechanisms. New data from studies of insect and mammalian transporters support an earlier hypothesis that NATs comprise an integrated transport system with a high plasticity of individual phenotypes but a conserved integrative role in the active absorption of essential amino acids (Boudko et al., 2005a).

List of abbreviations

B^0	broad substrate spectrum neutral amino acid transport system
$B^{0,+}$	broad substrate spectrum cationic amino acid transport system
BBMV	brush border membrane vesicle
CNS	central nervous system
DA	dopamine
GABA	gamma-aminobutyric acid
HS	hybridization solution
5HTP	5-hydroxytryptophan
L-DOPA	L-3,4-dihydroxyphenylalanine
NAT	nutrient amino acid transporter
NE	norepinephrine
OA	octopamine
PBS	phosphate buffered saline
PNS	peripheral nervous system
SLC6	solute carrier family 6
SNF	sodium neurotransmitter symporter family
TBS	Tris buffered saline
TMD	transmembrane domain

We thank Philip Tran (U.N.F.) for assistance with uptake assays. P.A.-N. and E.N.P. were supported by a NSF summer research fellowship, Research Experience for Undergraduates (R.E.U.). This work was supported by a R01 5R01AI030464 NIH-NIAID research grant (D.Y.B.).

References

Assis, P., Boudko, D. Y., Meleshkevitch, E. A. and Phung, E. (2004). Molecular expression and electrochemical analysis of phenylalanine-tyrosine transporter from *Anopheles gambiae* larvae. *FASEB J.* **18**, A1269.

Bohmer, C., Broer, A., Munzinger, M., Kowalczuk, S., Rasko, J. E. J., Lang, F. and Broer, S. (2005). Characterization of mouse amino acid transporter B(0)AT1 (slc6a19). *Biochem. J.* **389**, 745-751.

Boudko, D. Y., Moroz, L. L., Harvey, W. R. and Linser, P. J. (2001a). Alkalinization by chloride/bicarbonate pathway in larval mosquito midgut. *Proc. Natl. Acad. Sci. USA* **98**, 15354-15359.

Boudko, D. Y., Moroz, L. L., Linser, P. J., Trimarchi, J. R., Smith, P. J. S. and Harvey, W. R. (2001b). *In situ* analysis of pH gradients in mosquito larvae using noninvasive, self-referencing, pH-sensitive microelectrodes. *J. Exp. Biol.* **204**, 691-699.

Boudko, D. Y., Kohn, A. B., Meleshkevitch, E. A., Dasher, M. K., Seron, T. J., Stevens, B. R. and Harvey, W. R. (2005a). Ancestry and progeny of nutrient amino acid transporters. *Proc. Natl. Acad. Sci. USA* **102**, 1360-1365.

Boudko, D. Y., Meleshkevitch, E. A. and Harvey, W. R. (2005b). Novel transport phenotypes in the sodium neurotransmitter symporter family. *FASEB J.* **19**, A748.

Boudko, D. Y., Stevens, B. R., Donly, B. C. and Harvey, W. R. (2005c). Nutrient amino acid and neurotransmitter transporters. In *Comprehensive Molecular Insect Science*. Vol. 5 (ed. L. I. Gilbert, K. Iatrou and S. S. Gill), pp. 255-307. Amsterdam: Elsevier.

Broer, A., Klingel, K., Kowalczuk, S., Rasko, J. E. J., Cavanaugh, J. and Broer, S. (2004). Molecular cloning of mouse amino acid transport system B^0, a neutral amino acid transporter related to Hartnup disorder. *J. Biol. Chem.* **279**, 24467-24476.

Broer, A., Cavanaugh, J. A., Rasko, J. E. J. and Broer, S. (2006a). The molecular basis of neutral aminoacidurias. *Pflugers Arch.* **451**, 511-517.

Broer, A., Tietze, N., Kowalczuk, S., Chubb, S., Munzinger, M., Bak, L. K. and Broer, S. (2006b). The orphan transporter v7-3 (slc6a15) is a Na^+-dependent neutral amino acid transporter (B^0AT2). *Biochem. J.* **393**, 421-430.

Broer, S. (2002). Adaptation of plasma membrane amino acid transport mechanisms to physiological demands. *Pflugers Arch.* **444**, 457-466.

Broer, S., Cavanaugh, J. A. and Rasko, J. E. J. (2005). Neutral amino acid transport in epithelial cells and its malfunction in Hartnup disorder. *Biochem. Soc. Trans.* **33**, 233-236.

Castagna, M., Shayakul, C., Trotti, D., Sacchi, V., Harvey, W. and Hediger, M. (1997). Molecular characteristics of mammalian and insect amino acid transporters: implications for amino acid homeostasis. *J. Exp. Biol.* **200**, 269-286.

Castagna, M., Shayakul, C., Trotti, D., Sacchi, V. F., Harvey, W. R. and

Hediger, M. A. (1998). Cloning and characterization of a potassium-coupled amino acid transporter. *Proc. Natl. Acad. Sci. USA* **95**, 5395-5400.

Chen, N. H., Reith, M. E. and Quick, M. W. (2004). Synaptic uptake and beyond: the sodium- and chloride-dependent neurotransmitter transporter family SLC6. *Pflugers Arch.* **447**, 519-531.

Clark, J. A. and Amara, S. G. (1993). Amino acid neurotransmitter transporters: structure, function, and molecular diversity. *BioEssays* **15**, 323-332.

Farmer, M. K., Robbins, M. J., Medhurst, A. D., Campbell, D. A., Ellington, K., Duckworth, M., Brown, A. M., Middlemiss, D. N., Price, G. W. and Pangalos, M. N. (2000). Cloning and characterization of human NTT5 and v7-3: two orphan transporters of the Na^+/Cl^--dependent neurotransmitter transporter gene family. *Genomics* **70**, 241-252.

Feldman, D. H., Harvey, W. R. and Stevens, B. R. (2000). A novel electrogenic amino acid transporter is activated by K^+ or Na^+, is alkaline pH-dependent, and is Cl^- independent. *J. Biol. Chem.* **275**, 24518-24526.

Gade, G. and Auerswald, L. (2002). Beetles' choice-proline for energy output: control by AKHs. *Comp. Biochem. Physiol.* **132B**, 117-129.

Giordana, B. and Parenti, P. (1994). Determinants for the activity of the neutral amino acid/K^+ symport in lepidopteran larval midgut. *J. Exp. Biol.* **196**, 145-155.

Giordana, B., Sacchi, V. F., Parenti, P. and Hanozet, G. M. (1989). Amino acid transport systems in intestinal brush-border membranes from lepidopteran larvae. *Am. J. Physiol.* **257**, R494-R500.

Giordana, B., Leonardi, M. G., Casartelli, M., Consonni, P. and Parenti, P. (1998). K(+)-neutral amino acid symport of *Bombyx mori* larval midgut: a system operative in extreme conditions. *Am. J. Physiol.* **274**, R1361-R1371.

Hennigan, B. B., Wolfersberger, M. G. and Harvey, W. R. (1993a). Neutral amino acid symport in larval *Manduca sexta* midgut brush-border membrane vesicles deduced from cation-dependent uptake of leucine, alanine, and phenylalanine. *Biochim. Biophys. Acta* **1148**, 216-222.

Hennigan, B. B., Wolfersberger, M. G., Parthasarathy, R. and Harvey, W. R. (1993b). Cation-dependent leucine, alanine, and phenylalanine uptake at pH 10 in brush-border membrane vesicles from larval *Manduca sexta* midgut. *Biochim. Biophys. Acta* **1148**, 209-215.

Huelsenbeck, J. P. and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754-755.

Jespersen, T., Grunnet, M., Angelo, K., Klaerke, D. A. and Olesen, S. P. (2002). Dual-function vector for protein expression in both mammalian cells and *Xenopus laevis* oocytes. *Biotechniques* **32**, 536-538, 540.

Kanner, B. I. (1994). Sodium-coupled neurotransmitter transport: structure, function and regulation. *J. Exp. Biol.* **196**, 237-249.

Kilberg, M. S., Stevens, B. R. and Novak, D. A. (1993). Recent advances in mammalian amino acid transport. *Annu. Rev. Nutr.* **13**, 137-165.

Kowalczuk, S., Broer, A., Munzinger, M., Tietze, N., Klingel, K. and Broer, S. (2005). Molecular cloning of the mouse IMINO system: an Na^+- and Cl^- dependent proline transporter. *Biochem. J.* **386**, 417-422.

Kumar, S., Tamura, K. and Nei, M. (2004). MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief. Bioinformatics* **5**, 150-163.

Matz, M. V. (2002). Amplification of representative cDNA samples from microscopic amounts of invertebrate tissue to search for new genes. *Methods Mol. Biol.* **183**, 3-18.

Nelson, N. and Lill, H. (1994). Porters and neurotransmitter transporters. *J. Exp. Biol.* **196**, 213-228.

Palacin, M., Estevez, R., Bertran, J. and Zorzano, A. (1998). Molecular biology of mammalian plasma membrane amino acid transporters. *Physiol. Rev.* **78**, 969-1054.

Reizer, J., Reizer, A. and Saier, M. H., Jr (1994). A functional superfamily of sodium/solute symporters. *Biochim. Biophys. Acta* **1197**, 133-166.

Romeo, E., Dave, M. H., Kleta, R., Ristic, Z., Camargo, S. M. R., Makrides, V., Wagner, C. A. and Verrey, F. (2005). B(0)AT1 (SLC6A19) and other SLC6 orphan transporters: Hartnup disorder and beyond. *FASEB J.* **19**, A747-A747.

Romeo, E., Dave, M. H., Bacic, D., Ristic, Z., Camargo, S. M., Loffing, J., Wagner, C. A. and Verrey, F. (2006). Luminal kidney and intestine SLC6 amino acid transporters of B^0AT-cluster and their tissue distribution in *Mus musculus*. *Am. J. Physiol.* **290**, F376-F383.

Sacchi, V. F., Martino, P., Dassi, M., Perego, C., Parenti, P. and Giordana, B. (1993). Sodium and potassium interaction on the leucine cotransporter. *Pflugers Arch.* **423**, R9-R9.

Sacchi, V. F., Castagna, M., Trotti, D., Shayakul, C. and Hediger, M. A. (2001). Neutral amino acid absorption in the midgut of lepidopteran larvae. *Adv. Insect Physiol.* **28**, 168-184.

Sakata, K., Shimada, S., Yamashita, T., Inoue, K. and Tohyama, M. (1999). Cloning of a bovine orphan transporter and its short splicing variant. *FEBS Lett.* **443**, 267-270.

Seow, H. F., Broer, S., Broer, A., Bailey, C. G., Potter, S. J., Cavanaugh, J. A. and Rasko, J. E. J. (2004). Hartnup disorder is caused by mutations in the gene encoding the neutral amino acid transporter SLC6A19. *Nat. Genet.* **36**, 1003-1007.

Sloan, J. L. and Mager, S. (1999). Cloning and functional expression of a human Na^+ and Cl^- dependent neutral and cationic amino acid transporter B^{0+}. *J. Biol. Chem.* **274**, 23740-23745.

Stevens, B. R. (1992). Vertebrate intestine apical membrane mechanisms of organic nutrient transport. *Am. J. Physiol.* **263**, R458-R463.

Takanaga, H., Mackenzie, B., Peng, J. B. and Hediger, M. A. (2005a). Characterization of a branched-chain amino-acid transporter SBAT1 (SLC6A15) that is expressed in human brain. *Biochem. Biophys. Res. Commun.* **337**, 892-900.

Takanaga, H., Mackenzie, B., Suzuki, Y. and Hediger, M. A. (2005b). Identification of mammalian proline transporter SIT1 (SLC6A20) with characteristics of classical system IMINO. *J. Biol. Chem.* **280**, 8974-8984.

Tapiero, H., Mathe, G., Couvreur, P. and Tew, K. D. (2002). II. Glutamine and glutamate. *Biomed. Pharmacother.* **56**, 446-457.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876-4882.

Ugawa, S., Sunouchi, Y., Ueda, T., Takahashi, E., Saishin, Y. and Shimada, S. (2001). Characterization of a mouse colonic system B0+ amino acid transporter related to amino acid absorption in colon. *Am. J. Physiol.* **281**, G365-G370.

Uhl, G. R., Kitayama, S., Gregor, P., Nanthakumar, E., Persico, A. and Shimada, S. (1992). Neurotransmitter transporter family cDNAs in a rat midbrain library: 'orphan transporters' suggest sizable structural variations. *Brain Res. Mol. Brain Res.* **16**, 353-359.

Van Winkle, L. J., Christensen, H. N. and Campione, A. L. (1985). Na^+-dependent transport of basic, zwitterionic, and bicyclic amino acids by a broad-scope system in mouse blastocysts. *J. Biol. Chem.* **260**, 12118-12123.

Yamashita, A., Singh, S. K., Kawate, T., Jin, Y. and Gouaux, E. (2005). Crystal structure of a bacterial homologue of Na^+/Cl^- dependent neurotransmitter transporters. *Nature* **437**, 215-223.

Yudkoff, M., Daikhin, Y., Nelson, D., Nissim, I. and Erecinska, M. (1996). Neuronal metabolism of branched-chain amino acids: flux through the aminotransferase pathway in synaptosomes. *J. Neurochem.* **66**, 2136-2145.

Table S1. *Splicing scheme and SNPs of agNAT8 (ENSANGG00000015692) gene*

#	Exon / Intron Chr 3	Start	End	Length	Sequence
	5' upstream sequence				.....ctgccgaaaagacctcgaaataaactttactatctgctttttccagata
1	ENSANGE00000062606	11089528	11090052	525 bp	aacaacttcttctcgcgcacttgccaactcctccaccacaccctacgaccATGGAAGGACG GGATAATAATGGGTTCATCGGGGACAACAGCCCATCGATAG CGGGCCAGTACCGTTGGACTACACCGGCAGCACCGAATGGA GTTCACGTCACCCACGGAGGACTGGCCCTTACTGATGTTGAG CTGGCTGTAGGAGGTAAAAC(a/t)GTCCCCAGACCAACTCCAG CTACAGG(g/a)GAAGATGCGGGACGAGTCCCTACC(c/t)C(c/t)AT AATACCCGAGCCCAATGCCGATCGGGACCAGTGGGGCAAGG GCGTCGAGTTCCTG(a/t)TGTCCTGCATTGCCATGTCGGTCGGG TTGGGCAACGTGTGGCGGTTTCCGTTCGTAGCGCTCGAGAAC GGAGGCGGCGC(a/g)TTCGTGATACCGTACAT(c/t)ATAGTGCT GCTGCTTGTCGGCAAACCCGTCTACTACATGGAGATGA(t/c)C ATCGGACAGTTCTCGAGCCGGGGCAGCGTC(a/g)AGGTGTACG ATATGGCACCGATCATGCGAG
	Intron 1-2	11090053	11090649	597 bp	gtaagtggaggcgtcctttggactt.....ccctctttctctctgacgcgccaag
2	ENSANGE00000062603	11090650	11090878	229 bp	GTGT(t/c)GGTTACGGGCAGCTGTTCTCGGTGACCGCTCTAAT CACGTACTACTCCTCCCTAATGGCAC(t/c)AATTGCACGGTAC ATGATCGAC(t/g)CGTTCATG(a/t)ACCCGCTCCCGTGGGCCCA CTGTCGCTCGGAATGGCAACCGAACTGCATCGATTCCGTGG CCGGCAGTGCAAGCCCTAATGACAGCTCATCCAATCGCACA CTGACCAGTAGTT(c/g)GGAGCTTTACTTTAC
	Intron 2-3	11090879	11091033	155 bp	gtaagttttggcgggttcggtgtta.....ctcatactgttctgtttacttacag
3	ENSANGE00000062605	11091034	11091149	116 bp	CAAGGT(a/g)GTCCTGAAGGAGCTGGATGGCATTGACGATGG AATCGGACTTCCAGATTTGAGACTAACGCTATTCCTGGTCC TATCTTGGTCGCTCGTGTTCCTCACCCTCAT(t/a)AAAG
	Intron 3-4	11091150	11091220	71 bp	gtaacgaaaactattctttcttcaa.....taaattcgtttcacacctttgcag
4	ENSANGE00000062608	11091221	11091846	626 bp	GTGTTAAAAGCTCCGGCAAAGCGTCCTACTTCCTAGCGCTC TTCCCGTACGTCGTCATG(a/g)CGGTGCTGCT(c/t)GTCCGCGC TTG(t/c)ACACT(c/g)CCGGG(c/t)GCAGTGGA(t/c)GGTATCGTGT ACTTCCTCAAGCC(g/c)CAGTGGGACAAAATCTACGACCC(a/g)AAGGTGTGGT(a/g)CGCCGCCGTCACGCAGTGCTTCTTCTCG CTGTCGATCTGCTTCGGCAACATCATCATGTACTCGTCGTAC AACAA(a/g)TTCCGGCACAATGTGTACCGCGATGCGACGATC GTGACGTCGATCGATACGTTCACGTCGCTGCTGGCCGGGTG TACAATCTTCGGCATACTGGGACATCTGGCACACGTGACGG GGAAGACGGACGTGGGAAATGTGGTGAAATCCGATGCGGG ACTTGCATTTATTTCGTACCCGGAAGCGATCGCAAAGTTCG AGGTGCTGCCGCAAGCATTTTCGGTGCTGTTCTTCCTGATGC TGTTTGTGCTCGGCATCGGCAGCAATGTGGCGATGACGTCG TGCGTGATGACGGTGATAAAGGATCAGTTCCCGCGGGTGCG CAACTGGCAGGC(c/g)GCCACGATTATTGCCATCTGTGGCGT GCTGCTCGGTAGCATCTACGTGACGCCG
	Intron 4-5	11091847	11091918	72 bp	gtaggttgtttctttagcttcttta.....atccttttctcctcgttcatttcag
5	ENSANGE00000062609	11091919	11092195	277 bp	GGCGGTCAGT(a/g)TGTGTTGAAG(C/t)TGGTGGATTATTACGG CGCCTCATCGATCGCGCTCGTGCTTGCCATTGCTGAGCTGAT CGCGATCGGGTGGGTTTA(t/c)GGCGTGGATCGACTGTGCAA GGACACGGAGTTTATGCTTGGCCATCGGCCCAACCTGTACT GGCGTCTGTGCTGGCGCTGGATCACGCCCCTGCTAATGTTT GTCATTCTCATCTACAACCTTGTCACGCTCGAGCCGCTCATG TACA(a/g)GCAGTACGTTTATCCTACGGTCGCTTACG
	Intron 5-6	11092196	11092279	84 bp	gtgagtatgagccattaagagctaa.....actgctgctgctttcatccccgcag
6	ENSANGE00000062604	11092280	11092374	95 bp	GCATTGGTTGGTGCATTTTTGCCTTTGGGCT(g/t)CTGCAGCT (c/t)CCGATTTGGGC(g/t)GCCTACGCCGTCTACAAGCAGAGC GGCAAATCGCT(c/g)AACGAG
	Intron 6-7	11092375	11092437	63 bp	gtacggcacccgaaactagctcgcc.....ttaactgtttcattattctttttag
7	ENSANGE00000062602	11092438	11092530	93 bp	AAAATTAA(a/g)AACGCCTTCAAACCGACCGCTGCCTGGGGT CCGATTGATCCTGCCACGCATTACGAGTACAAAAAGTTCAT CGACGAGGACTGA
	3′ downstream sequence				ctttagagggaaaccttaaccgggacggctcacagcagtttaattcttta.....

SNPs are shown in parentheses (EMBL prediction/vs./cloned agNAT8).

Table S2. *Putative post-translational modification sites in the agNAT8 protein along with predicted transmembrane domain (TMD) boundaries*

Putative motif definition	Start	Stop	Length
NH_2			
1. Proline dependent kinase phosphorylation site	13	16	4
2. Myosin 1 heavy chain kinase phosphorylation site	21	25	5
3. Proline dependent kinase phosphorylation site	24	27	4
4. Casein kinase II phosphorylation site (S or T PO4′D)	41	46	6
5. Myosin 1 heavy chain kinase phosphorylation site	54	58	5
6. Proline dependent kinase phosphorylation site	56	59	4
7. Casein kinase II phosphorylation site (S or T PO4′D)	59	64	6
TMD1	**84**	**111**	**27_TMD1**
TMD2	**115**	**144**	**29_TMD2**
8. Protein kinase C phosphorylation site	142	146	5
9. Glycogen synthase kinase 3 phosphorylation site (last S must be PO4 S)	142	148	7
10. cGMP dependent protein kinase phosphorylation site	143	146	4
TMD3	**176**	**212**	**36_TMD3**
11. Glycogen synthase kinase 3 phosphorylation site (last S must be PO4 S)	209	215	7
12. Proline dependent kinase phosphorylation site	215	218	4
13. Casein kinase II phosphorylation site (S or T PO4′D)	215	220	6
14. Glycogen synthase kinase 3 phosphorylation site (last S must be PO4 S)	215	221	7
1g. Tentative N-glycosylation site	218	221	NDSS
15. Protein kinase C phosphorylation site	221	225	5
2g. Tentative N-glycosylation site	223	226	NRTL
16. Casein kinase II phosphorylation site (S or T PO4′D)	227	232	6
TMD4	**252**	**272**	**20_TMD4**
17. Myosin 1 heavy chain kinase phosphorylation site	254	258	5
18 Protein kinase C phosphorylation site	277	281	5
TMD5	280	300	**20_TMD5**
19. Glycogen synthase kinase 3 phosphorylation site (last S must be PO4 S)	277	283	7
1c. Tentative chaperonin binding epitope	293	300	8
20. Tyrosine kinase phosphorylation site	314	323	10
TMD6	**324**	**349**	**25_TMD6**
21. Consensus sequence for phosphorylation of mannose	347	352	6
22. Consensus sequence for phosphorylation of mannose	350	355	6
23. Casein kinase II phosphorylation site (S or T PO4′D)	365	370	6
TMD7	**358**	**390**	**32_TMD7**
24. Protein kinase C phosphorylation site	390	394	5
25. Casein kinase II phosphorylation site (S or T PO4′D)	409	414	6
TMD8	422	454	32_TMD8
2c. Tentative chaperonin binding epitope	428	435	8
TMD9	**460**	**481**	**21_TMD9**
26. Proline dependent kinase phosphorylation site	480	483	4
TMD10	486	512	**26_TMD10**
1n. NADP binding consensus sequence	494	513	20
Tmd11	535	569	**34_TMD11**
27. Proline dependent kinase phosphorylation site	543	546	4
28. Consensus sequence for phosphorylation of mannose	563	568	6
TMD12	575	604	**29_TMD12**
29. Consensus sequence for phosphorylation of mannose	593	598	6
30. Protein kinase C phosphorylation site	599	603	5
31. Casein kinase II phosphorylation site (S or T PO4′D)	602	607	6
32. Casein kinase II phosphorylation site (S or T PO4′D)	624	629	6
33. Consensus sequence for phosphorylation of mannose	628	633	6
COOH			