PNAS RESEARCH ARTICLE | EVOLUTION

Co-option of a conserved host glutamine transporter facilitates aphid/*Buchnera* metabolic integration

Rebecca P. Duncan[a,1,2], Catriona M. H. Anderson[b], David T. Thwaites[c], Charles W. Luetje[d], and Alex C. C. Wilson[a,2]

Edited by Douglas Schemske, Michigan State University, East Lansing, MI; received May 20, 2023; accepted September 14, 2023

Organisms across the tree of life colonize novel environments by partnering with bacterial symbionts. These symbioses are characterized by intimate integration of host/endosymbiont biology at multiple levels, including metabolically. Metabolic integration is particularly important for sap-feeding insects and their symbionts, which supplement nutritionally unbalanced host diets. Many studies reveal parallel evolution of host/endosymbiont metabolic complementarity in amino acid biosynthesis, raising questions about how amino acid metabolism is regulated, how regulatory mechanisms evolve, and the extent to which similar mechanisms evolve in different systems. In the aphid/*Buchnera* symbiosis, the transporter ApGLNT1 (*Acyrthosiphon pisum* glutamine transporter 1) supplies glutamine, an amino donor in transamination reactions, to bacteriocytes (where *Buchnera* reside) and is competitively inhibited by *Buchnera*-supplied arginine—consistent with a role regulating amino acid metabolism given host demand for *Buchnera*-produced amino acids. We examined how ApGLNT1 evolved a regulatory role by functionally characterizing orthologs in insects with and without endosymbionts. ApGLNT1 orthologs are functionally similar, and orthology searches coupled with homology modeling revealed that GLNT1 is ancient and structurally conserved across insects. Our results indicate that the ApGLNT1 symbiotic regulatory role is derived from its ancestral role and, in aphids, is likely facilitated by loss of arginine biosynthesis through the urea cycle. Given consistent loss of host arginine biosynthesis and retention of endosymbiont arginine supply, we hypothesize that GLNT1 is a general mechanism regulating amino acid metabolism in sap-feeding insects. This work fills a gap, highlighting the broad importance of co-option of ancestral proteins to novel contexts in the evolution of host/symbiont systems.

symbiosis | *A. pisum* | *Buchnera* | amino acid metabolism | gene co-option

Significance

Symbiotic systems are integrated in many ways, including metabolically. Less well understood are the evolution and generalizability of mechanisms regulating host/symbiont integration. To address this gap, we examined how an aphid glutamine transporter, ApGLNT1 (*Acyrthosiphon pisum* glutamine transporter 1), evolved its proposed role regulating amino acid biosynthesis in response to demand for symbiont-supplied arginine. Functionally characterizing and modeling GLNT1 from insects with and without endosymbionts revealed that ApGLNT1 evolved its regulatory role from an ancient, conserved function, likely facilitated by loss of aphid-mediated arginine biosynthesis. Functional and structural conservation of GLNT1 suggests that it is a general mechanism regulating amino acid metabolism in the evolution of host/symbiont systems, supporting the role of constraint of ancestral gene function in the evolution of host/symbiont systems.

Bacterial endosymbionts, by providing access to their metabolic repertoires, enable the evolution of novel eukaryotic host phenotypes, thus facilitating host expansion into new ecological niches (1). Ancient endosymbioses are characterized by hosts and endosymbionts so tightly integrated that one cannot live without the other. There are countless examples of how hosts and endosymbionts are genetically, developmentally, and metabolically integrated (2–6), but the evolutionary and mechanistic process of integration is still not fully understood. A fundamental step in the establishment of an endosymbiont within its host is metabolic integration, i.e., the coordinated bidirectional flux of metabolites between endosymbiont and host. Elucidating the mechanisms of metabolic integration is key to understanding how hosts and endosymbionts become integrated.

Metabolic integration is a complex process that requires multiple complementary mechanisms working together to coordinate host/endosymbiont metabolism, regulate metabolism, and mediate metabolite exchange (4, 7–10). Despite its complexity, obvious targets for studying metabolic integration are the transporters that facilitate host/endosymbiont metabolic flux. Indeed, evidence across diverse systems suggests that host-derived proteins were co-opted to facilitate host/endosymbiont metabolite exchange (9, 11, 12). Here, we focus on a co-opted host glutamine transporter in the aphid/*Buchnera* system that has been implicated in facilitating host/endosymbiont metabolic integration (8).

The endosymbiont of aphids, *Buchnera aphidicola*, resides in specialized aphid cells called bacteriocytes (Fig. 1*A*) (13, 14). Bacteriocyte cells aggregate with sheath cells to form the bacteriome (Fig. 1*B*), an organ specialized in aphids and other phloem-feeding insects for amino acid biosynthesis (14–16). While phloem sap is rich in sugar and free amino acids, it contains limiting concentrations of most essential amino acids, so phloem-feeding insects typically host endosymbionts that provision essential amino acids (17). In most phloem-feeding insect systems, including that of aphids and *Buchnera*, some amino acid biosynthetic pathways are encoded in the host genome, and some are encoded by the symbiont genome. Remarkably, some amino acid biosynthetic pathways require contributions from the host and endosymbiont genomes—that is, they are collaboratively completed by

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]Present address: School of Biological Sciences and Center for Microbial Dynamics and Infection, Georgia Institute of Technology, Atlanta, GA 30332.

[2]To whom correspondence may be addressed. Email: rduncan41@gatech.edu or acwilson@miami.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2308448120/-/DCSupplemental.

Published October 16, 2023.

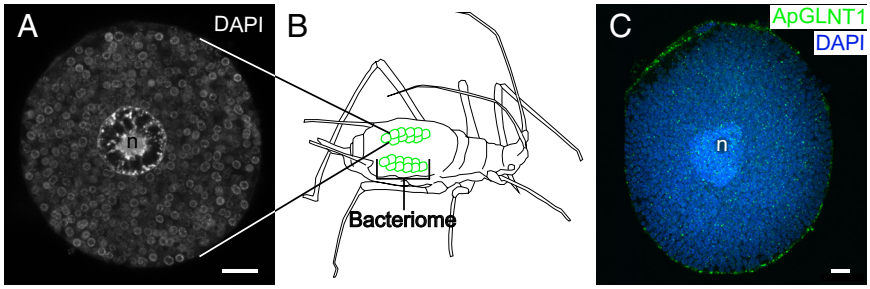

Fig. 1. Aphids house *Buchnera* in bacteriocyte cells and ApGLNT1 localizes to the bacteriocyte membrane. (*A*) Confocal microscope image of a bacteriocyte from a three-day-old asexual female aphid. DNA is stained in DAPI (white color). Bacteriocytes make up the symbiotic organ located in the aphid's abdomen. (*B*) Drawing of an aphid showing the bacteriome, the symbiotic organ made up of bacteriocytes (outlined in green). (*C*) Confocal microscope image of ApGLNT1 immunolocalization to the plasma membrane of a bacteriocyte from an asexual female adult aphid. ApGLNT1 is shown in green and DNA is stained with DAPI (blue). The nucleus in each bacteriocyte image is marked with an "n". (Scale bars are 10 μm.)

enzymes encoded in both host and endosymbiont genomes (18). A convergent genomic feature of such nutritional endosymbionts compared to their free-living relatives is loss of the ability to synthesize most nonessential amino acids (19, 20). Host genome evolution also features loss of amino acid biosynthetic capacity, a notable example being loss of the ability to synthesize arginine. Ancestrally, animals make arginine as a by-product of the urea cycle, yet aphids and related insects have lost almost all genes in the urea cycle and instead depend on their endosymbionts to supply arginine (21, 22). Previously, Price et al. (8) proposed that amino acid biosynthesis in the aphid bacteriome is regulated in the pea aphid, *Acyrthosiphon pisum*, by a glutamine transporter called *A. pisum* glutamine transporter 1 (ApGLNT1; gene ID: ACYPI001018), which localizes to the plasma membrane of bacteriocyte cells (Fig. 1*C*). Through heterologous expression in *Xenopus laevis* oocytes, Price et al. (8) demonstrated that ApGLNT1 takes up, with high specificity, the nonessential amino acid glutamine—a crucial amino acid precursor which provides amino groups directly to three essential amino acids and indirectly (through glutamate, serine, and aspartate) to the remaining seven (21, 23, 24). Intriguingly, glutamine transport by ApGLNT1 is competitively inhibited by arginine, an essential amino acid for aphids and other hemipterans that is synthesized by *Buchnera*. Together, these data are consistent with the hypothesis that ApGLNT1 is part of the machinery that mediates aphid/*Buchnera* metabolic integration by controlling supply of the precursor glutamine based on host demand for a symbiont-provisioned essential amino acid, arginine. Accumulating arginine in aphid hemolymph inhibits glutamine supply to bacteriocytes, thereby down-regulating all amino acid biosynthesis in aphid bacteriomes (8).

While the model proposed by Price et al. (8) fits well with what is known about amino acid metabolism in the aphid/*Buchnera* symbiosis (25), it leaves some important questions unanswered. First, ApGLNT1 has broad spatiotemporal expression (26, 27) and orthologs in insects lacking bacteriocytes (26), indicating that it must have nonsymbiotic roles in addition to its symbiotic role in aphids. The potential of ApGLNT1 to perform additional nonsymbiotic roles raises the question of whether its functional role in aphid bacteriocytes is derived from its ancestral role or if bacteriocyte function required evolution in transport function. For example, glutamine transport could be an essential functional role of ApGLNT1 orthologs in a nonsymbiotic context, while inhibition by arginine is not. Second, since only ApGLNT1 has been functionally characterized, it is unclear whether this model of regulation is particular to aphids or whether it could apply more broadly across related endosymbioses, as has been found for coordinated nutrient biosynthesis (4, 6, 18, 21–24). Both of these questions can be addressed within a phylogenetic framework, by inferring transport function of ApGLNT1 orthologs from insects with and without bacteriocytes.

We used a comparative framework to test whether co-option of ApGLNT1 to the aphid/*Buchnera* symbiotic interface involved evolution of transport function. To elucidate the functional evolutionary history of GLNT1, we functionally characterized ApGLNT1 orthologs from insects with and without endosymbionts (Fig. 2) by heterologous expression of GLNT1 orthologs in *Xenopus* oocytes and electrophysiological measurement of transporter function (Figs. 3–6). Our empirical observations of functional transport activity using ApGLNT1 orthologs demonstrate conserved core transport characteristics (ion coupling, substrate selectivity) across all five orthologs investigated. Such conservation of functional activity is most likely explained by conservation of the GLNT1 transporter binding pocket structure. Homology modeling, used here to predict the structural basis of substrate recognition, supports the hypothesis, based upon direct measurements of transporter

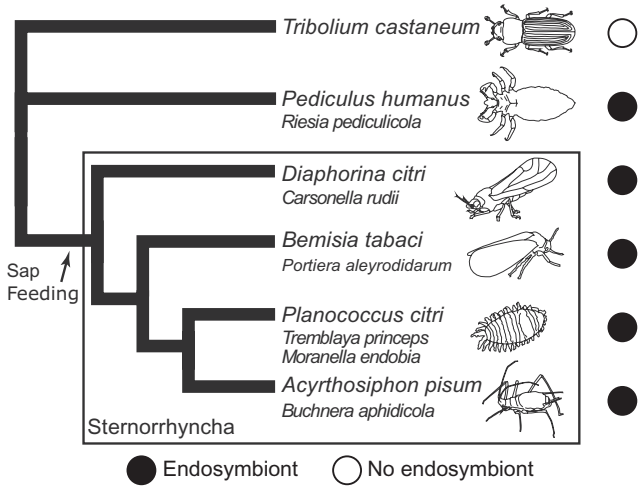

Fig. 2. Phylogenetic framework and taxon sampling. Phylogenetic relationships among insects sampled in this study for functional characterization and/or urea cycle gene annotation. In addition to *A. pisum*, we functionally characterized ApGLNT1 orthologs from two sap-feeding insects (*Planococcus citri* and *Diaphorina citri*), one blood-feeding insect (*Pediculus humanus*), and one insect without an intracellular symbiont (*Tribolium castaneum*). *Bemisia tabaci* was included only in our annotation of urea cycle genes. Insect names appear at the end of each branch, with the names of intracellular bacterial symbionts (if any) listed below. The sap-feeding taxa (*P. citri*, *B. tabaci*, and *D. citri*) represent superfamilies of the hemipteran suborder Sternorrhyncha (highlighted by the box), to which aphids belong. The evolution of sap-feeding is mapped onto the tree with an arrow. Relationships shown here are based on published phylogenetic studies (28–30).

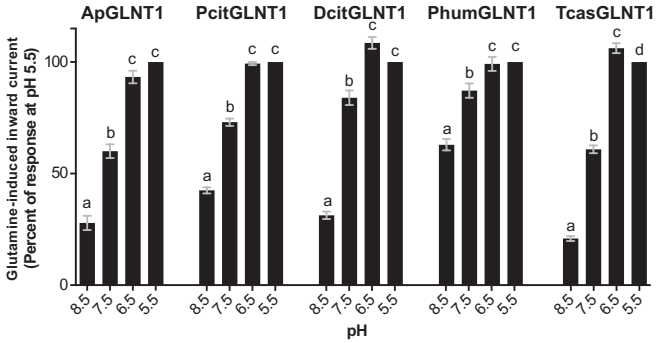

Fig. 3. ApGLNT1 orthologs induce an inward current in response to glutamine that is dependent upon extracellular pH. Graphs show mean ± SE of response amplitudes to 5 mM glutamine at pH 8.5, 7.5, 6.5, and 5.5 normalized to the mean response amplitude at pH 5.5. Letters above bars in each graph mark statistically significant differences based on a Friedman test [ApGLNT1: $\chi^2(2)$= 17, $P < 0.0001$; PcitGLNT1: $\chi^2(2) = 11.10$, $P = 0.0009$; DcitGLNT1: $\chi^2(2) = 18.77$, $P = 0.0003$; PhumGLNT1: $\chi^2(2) = 21.75$, $P < 0.0001$; TcasGLNT1: $\chi^2(2) = 33.30$, $P < 0.0001$]. Representative current recordings for pH dependence runs are shown in *SI Appendix*, Fig. S1, along with representative current recordings for control, water-injected oocytes. Sample sizes are given in *SI Appendix*, Table S2.

function, that ApGLNT1 orthologs across the insect phylogeny are structurally and, therefore, functionally conserved. Our results support the idea that highly selective glutamine transport and inhibition by arginine were characteristics of the ancestral protein from which ApGLNT1 and its orthologs evolved, prior to the evolution of endosymbiosis. Functional conservation of ApGLNT1 orthologs suggests that co-option of ancestral proteins could be a general mechanism of facilitating host/endosymbiont metabolic integration in insect nutritional endosymbioses.

Results

ApGLNT1 Orthologs Function as pH-Dependent, H^+-Coupled, Amino Acid Cotransporters with High Selectivity for Glutamine as a Substrate. To test our prediction that ApGLNT1 orthologs selectively transport glutamine, and to glean additional insight into transport function, we heterologously expressed the transporters in *Xenopus* oocytes for functional characterization. The orthologs we examined belonged to two sap-feeding insects with endosymbionts (*Planococcus citri* and *Diaphorina citri*), one non-sap-feeding hemipteran with an endosymbiont (*Pediculus humanus*), and one nonhemipteran with no endosymbiont (*Tribolium castaneum*) (Fig. 2). Price et al. (8, 31) previously showed that ApGLNT1 is electrogenic, cotransporting glutamine and H^+, and that the H^+/glutamine cotransport function, measured as inward current in the presence of extracellular glutamine, was dependent upon extracellular pH. We tested transport function of ApGLNT1 orthologs by measuring amino acid–induced inward current using two-electrode voltage clamp electrophysiology in extracellular bathing solutions ranging from pH 5.5 to 8.5. ApGLNT1 and ortholog transport function is detected as an inward positive current representing H^+/glutamine cotransport (Fig. 3 and *SI Appendix*, Fig. S1). Each transporter ortholog, when exposed to extracellular glutamine, showed a significant (based on a Friedman test) increase in transport as extracellular pH was lowered from pH 8.5 to 5.5, typical of H^+-coupled amino acid cotransport. Post hoc analysis with a corrected Dunn's test found a significantly higher response at pH 5.5 compared with pH 8.5 (ApGLNT1: Z = 3.801, corrected $P = 0.0009$; *P. citri*: Z = 3.012, corrected $P = 0.0155$; *D. citri*: Z = 2.898, corrected $P = 0.0225$; *P. humanus*: Z = 3.679, corrected $P = 0.0014$; *T. castaneum*: Z = 4.269, corrected $P = 0.0001$), consistent with H^+-coupled glutamine transport in ApGLNT1 orthologs (Fig. 3 and *SI Appendix*, Fig. S1). Based on evidence that the ApGLNT1 orthologs we examine here transport glutamine, we decided to name them PcitGLNT1 (*P. citri*), DcitGLNT1 (*D. citri*), PhumGLNT1 (*P. humanus*), and TcasGLNT1 (*T. castaneum*).

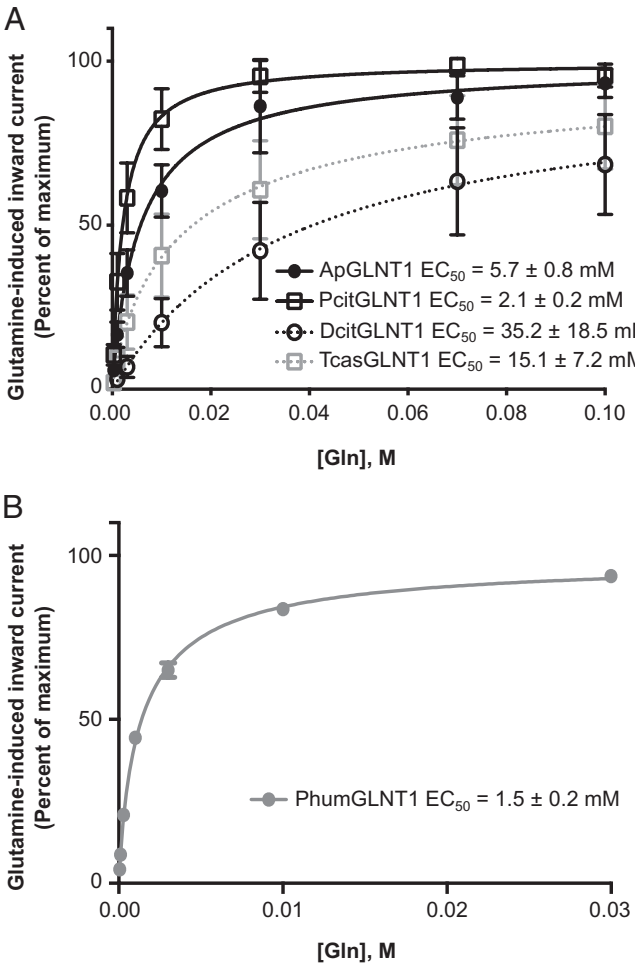

Fig. 4. GLNT1 orthologs are low-affinity glutamine transporters. Glutamine-induced inward current at different concentrations of glutamine and EC_{50} values for GLNT1 orthologs from (*A*) *A. pisum*, *P. citri*, *D. citri*, and *T. castaneum* and (*B*) *P. humanus*. *P. humanus* was plotted separately from the other species because the scale on the horizontal axis is different. Response amplitudes were normalized and concentration–response analysis performed as described in Methods. Points on each graph represent means ± SE. Sample sizes are given in *SI Appendix*, Table S2. Representative current recordings for glutamine response and control runs with water-injected oocytes are shown in *SI Appendix*, Fig. S2.

ApGLNT1 Orthologs Are Low-Affinity Glutamine Transporters with a Narrow Substrate Selectivity. Price et al. demonstrated that ApGLNT1 is a selective glutamine transporter (8). Given that the ApGLNT1 orthologs typically demonstrate an inward current during exposure to extracellular glutamine (*SI Appendix*, Fig. S1), we examined whether the orthologs showed saturation kinetics typical of ion-coupled amino acid symport systems. Consistent with earlier observations of ApGLNT1 only (8, 31), all ApGLNT1 orthologs investigated were characterized by saturable inward current responses in the presence of increasing concentrations of extracellular glutamine (Fig. 4 and *SI Appendix*, Fig. S2). Our finding that glutamine-induced inward current responses are saturable indicates that ApGLNT1 orthologs are amino acid transporters that carry glutamine. Estimated EC_{50} values for ApGLNT1 and its orthologs were: 1.5 ± 0.2 mM (*P. humanus*), 2.1 ± 0.2 mM (*P. citri*), 5.7 ± 0.8 mM (ApGLNT1),

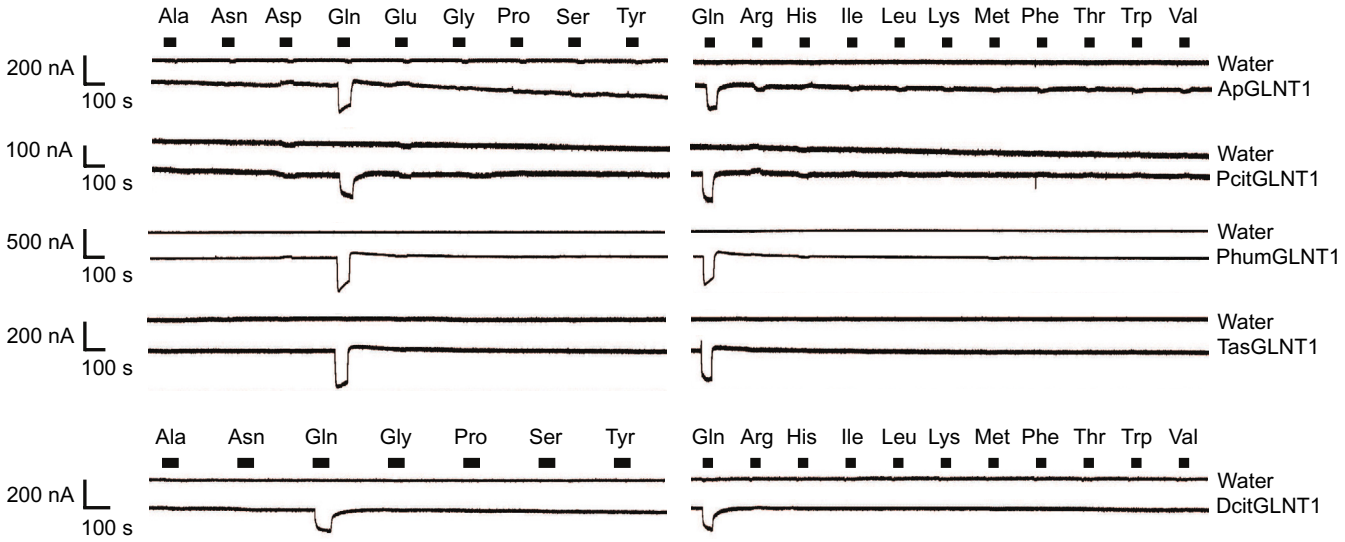

Fig. 5. GLNT1 orthologs have narrow substrate selectivity. Representative current recordings for amino acid screens in GLNT orthologs and water-injected oocytes (controls). Black bars above the water-injected oocyte traces for ApGLNT1 and DcitGLNT1 runs mark the timing and duration of amino acid applications. Results shown in representative current responses were qualitatively obtained in 100% of experimental runs with slight quantitative differences in response amplitudes. Amino acids were applied at 2 mM to oocytes expressing ApGLNT1, PcitGLNT1, PhumGLNT1, and TcasGLNT1 and at 10 mM to oocytes expressing DcitGLNT1 (because of low sensitivity), with the exception of tyrosine, which was applied at 2.5 mM because it is insoluble at higher concentrations. Sample sizes are given in *SI Appendix*, Table S2.

15.1 ± 7.2 mM (*T. castaneum*), and 35.2 ± 18.5 mM (*D. citri*), suggesting that ApGLNT1 orthologs are consistently low-affinity (with affinity values in the mM range) glutamine transporters with varying absolute affinity for glutamine. ApGLNT1 is related to the mammalian SLC36 solute carrier family (26, 32). The affinity for amino acid substrates of other, functionally characterized, vertebrate and invertebrate SLC36-related transporter proteins varies from those with relatively high-affinity (in the μM range) (12, 33–35) to those, more similar to the GLNT1 orthologs reported here, which have relatively low-affinity (in the mM range) (33, 34, 36–38). Our estimated EC_{50} for ApGLNT1 at an extracellular pH of 7.5 (5.7 ± 0.8 mM) was consistent with the estimated K_m values of 2.7 ± 0.4 mM (at pH 6.0) and 5.4 ± 0.5 mM (at pH 7.5) found by Price et al. (8, 31). Importantly, glutamine concentration in aphid hemolymph is high enough to nearly saturate ApGLNT1 (39), implying that this transporter functions continuously to import glutamine into bacteriocytes in vivo.

Apart from cotransporting glutamine and protons, another characteristic of ApGLNT1 is that it does not transport other amino acids well (8). To determine substrate selectivity in GLNT1 orthologs, we conducted amino acid screens where we sequentially applied essential amino acids (with glutamine as a positive control) or nonessential amino acids to oocytes expressing each ortholog, with transport function measured as H^+/amino acid inward current. The ApGLNT1 orthologs all failed to transport amino acids other than glutamine at the concentration used, showing that GLNT1 orthologs from other insects are also highly selective for glutamine (Fig. 5).

Glutamine Transport by GLNT1 is Inhibited by Arginine. A key feature of ApGLNT1, underlying its proposed ability to regulate bacteriocyte amino acid metabolism, is that glutamine transport is competitively inhibited by arginine, a *Buchnera*-produced essential amino acid (8). We found that the responses of all GLNT1 orthologs to glutamine at the appropriate EC_{50} concentration, determined individually for each ortholog (Fig. 4), are inhibited by the presence of arginine (Fig. 6 and *SI Appendix*, Fig. S3). We calculated IC_{50} values of arginine inhibition to be 2.5 ± 0.2 mM (*P. citri*), 2.8 ± 2.0 mM (*P. humanus*), 3.3 ± 0.1 mM (*T. castaneum*), 5.1 ± 0.9 mM (*A. pisum*), and 6.1 ± 0.8 mM (*D. citri*). Our IC_{50} value of arginine for ApGLNT1 is consistent with the 3.9 ± 1.1 mM IC_{50} value of arginine found by Price et al. (8). Price et al. (8) additionally found that while glutamine was the primary substrate of ApGLNT1, it could mediate some arginine uptake at high concentrations, though the level of arginine uptake at 10 mM was only ~12% of glutamine uptake at 10 mM. This result implies effective binding of arginine within the transporter binding pocket but severely limited translocation. Consistent with their finding, we observed a small arginine-induced inward current in oocytes expressing ApGLNT1, but we did not observe an inward current in oocytes expressing GLNT1 orthologs up to a concentration of 5 mM (*SI Appendix*, Fig. S4).

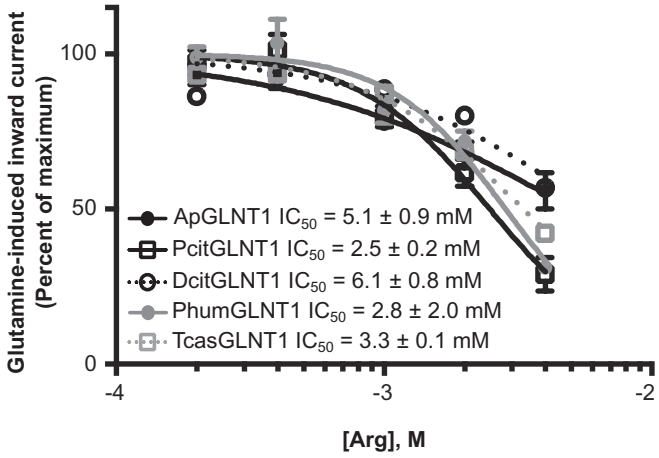

Fig. 6. Glutamine transport is inhibited by arginine in GLNT1 orthologs. Arginine inhibition curves and estimated IC_{50} values for each GLNT1 ortholog. Arginine inhibition of glutamine response was measured by applying a mixture of glutamine at EC_{50} plus increasing concentrations of arginine. Responses to glutamine + arginine were normalized to the preceding response to glutamine alone. Concentration inhibition analysis was performed as described in Methods. Points on each graph represent mean ± SE. N values for each transporter are given in *SI Appendix*, Table S2. Representative current recordings for arginine inhibition and control, water-injected oocytes are shown in *SI Appendix*, Fig. S3.

GLNT1 is Widespread in Insects. Given evidence for the strong functional conservation of GLNT1 orthologs in the taxa we sampled for functional characterization, including from a species without endosymbionts, (Figs. 3–6), we posit that GLNT1 fulfills an important, possibly housekeeping, role in insects. Given its broad spatio-temporal expression in *A. pisum* (26, 40), ApGLNT1 and orthologs likely function in general cellular maintenance in nonsymbiotic tissues. If GLNT1 indeed has an ancestral housekeeping role in a range of tissue types, then we hypothesize that this transporter is ancient and ubiquitous in insects. To gain insight into the age of GLNT1, we mined genomes and transcriptomes of representative taxa across the insect phylogeny (Dataset S1) for GLNT1 orthologs using a reciprocal BLAST approach and phylogenetic analysis. Orthology searches found evidence that all of the insects we sampled encode GLNT1, including the order Odonata (dragonflies and damselflies), which shared a common ancestor with the rest of the insects we sampled over 400 Mya (28) (*SI Appendix*, Figs. S5 and S6). These data indicate that GLNT1 is widespread among insects and minimally predates the diversification of most insects. Orthology searches additionally found that most insects we sampled encode a single GLNT1 ortholog, implying that ApGLNT1 was ancestrally a single-copy transporter. In contrast, members of the hemipteran suborder Auchenorrhyncha (including *Nilaparvata lugens, Philaenus spumarius*, *Laodelphax striatellus*, and *Diceroprocta semicincta*—see Datasets S1 and S2) encode between two and four loci that we assigned orthology to GLNT1 based on reciprocal BLAST searches. Notably, our phylogenetic analysis failed to support monophyly of GLNT1 orthologs when all auchenorrhynchan paralogs were included (*SI Appendix*, Fig. S5). Whereas, the GLNT1 clade had decent (85%) bootstrap support with the inclusion of a single GLNT1 paralog from each auchenorrhynchan species (*SI Appendix*, Fig. S6). These observations suggest that additional auchenorrhynchan paralogs are either not true GLNT1 orthologs and/or that their sequences are extremely divergent. Indeed, in the phylogeny with a well-supported GLNT1 clade (*SI Appendix*, Fig. S6), the included auchenorrhynchan paralogs shared percent identities with ApGLNT1 of 52.82 to 65.41%, while the other auchenorrhynchan GLNT1-like paralogs shared much lower percent identities with ApGLNT1 (33.73 to 47.72%) (Dataset S3). Interestingly, GLNT1 homology modeling (see below) found that only auchenorrhynchan paralogs that shared higher percent identities with ApGLNT1 (and were supported as belonging in the GLNT1 clade by phylogenetic analyses) are structurally conserved with GLNT1 proteins that we functionally characterized (see below). Phylogenetic analyses support the hypothesis that GLNT1 is an ancient, functionally important transporter in insects.

GLNT1 Binding Pocket Residues Are Conserved between Orthologs. To investigate the structural basis of GLNT1 functional conservation, we performed a comparative analysis of substrate-binding pocket models for each GLNT1 ortholog included in our functional analyses. ApGLNT1, like the mammalian SLC36 transporters, belongs to the Amino Acid/Auxin Permease (AAAP, 2.A.18) family within the Amino Acid-Polyamine-Organocation (APC) Superfamily (12, 41). We modeled ApGLNT1 on the resolved crystal structure of GkApcT (42), a bacterial H^+-coupled amino acid cotransporter from the APC Superfamily, using HHPred and Modeller (43, 44). This analysis showed that ApGLNT1 has the classic APC Superfamily 10 transmembrane (TM) core structural fold, consisting of a 5 TM structural repeat, with each repeat sitting inverse to the other in the membrane (45) (Fig. 7*A*). The binding pocket of APC Superfamily transporters is primarily formed by TM domains TM1, TM3, TM6, and TM8. Within the predicted binding pocket, we examined two key positions, one in each of TM3 and TM8, known to influence substrate specificity

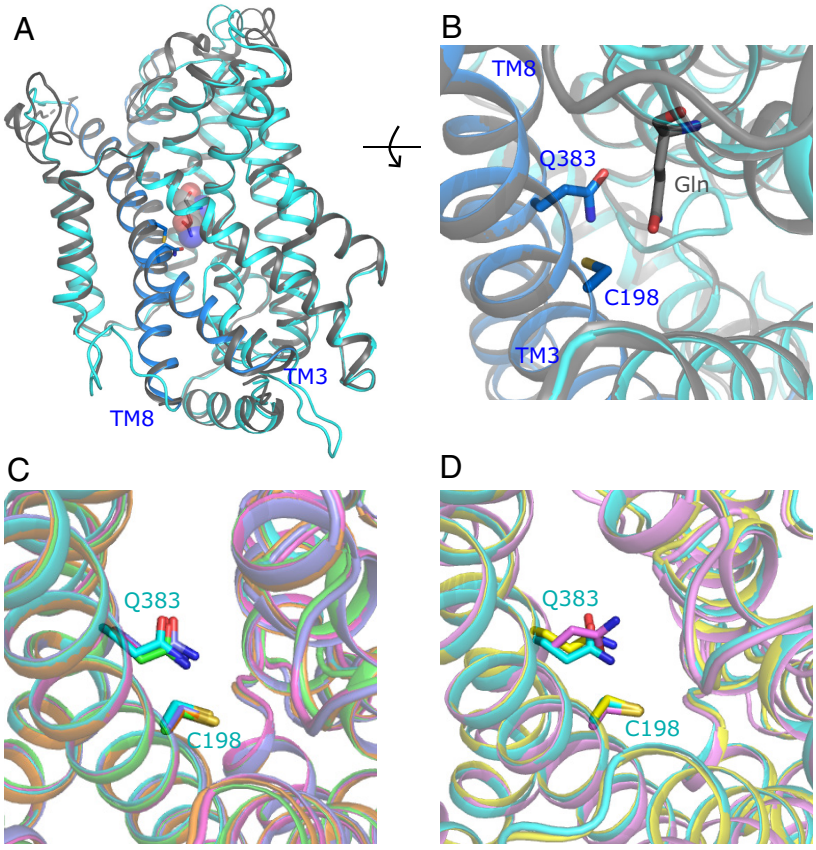

Fig. 7. Conservation of substrate binding pocket residues in GLNT1 orthologs. (*A* and *B*) Homology model of ApGLNT1 created using HHPred and Modeller superimposed on to the crystal structure of bacterial amino acid transporter GkApcT (PDB 5OQT, gray) from which it was derived. ApGLNT1 is shown as cyan with TM domains 3 and 8 shown in blue. Key residues C198 (TM3, blue) and Q383 (TM8, blue) are shown as sticks. Glutamine (gray sticks and spheres) is shown as a substrate in the binding pocket (positioned using crystal-bound alanine in GkApcT). C198 and Q383 line the deep areas of the substrate binding pocket and are adjacent to the substrate side chain. In (*A*), proteins are shown "side on", as they would reside in the membrane and in (*B*), "top down", above the plane of the membrane, looking down into the binding pocket. (*C* and *D*) Models of GLNT1 orthologs produced by AlphaFold (models are calculated without substrate). The overlaid models are from functionally characterized and more distant taxa as follows: ApGLNT1 (cyan), DcitGLNT1 (magenta), PcitGLNT1 (purple), PhumGLNT1 (orange), TcasGLNT1 (green); (*D*) ApGLNT1 (cyan), CsplGLNT1 (lilac) and Dmel (*D. melanogaster*) CG43693 (yellow). The models show conservation of a Cys residue, in the helix of TM3, in the equivalent position to C198 in ApGLNT1 and a Gln residue, in the helix of TM8, in the equivalent position to Q383 in ApGLNT1. In all figures, the intracellular N-terminal section has been omitted for clarity. Four-letter taxon IDs for sequences are listed in Dataset S1.

and inhibitor access in transporters of the APC Superfamily (Fig. 7 *B* and *C*) (12, 37, 42, 46). The amino acid residue occupying the TM3 position dictates the space available within the binding pocket for the amino acid substrate chain in the mammalian H^+-coupled amino acid transporter slc36a2 and related mammalian and insect transporters (12, 37). In ApGLNT1, this position was identified as C198 in TM3 (Fig. 7 *A* and *B*). The residue occupying the TM8 position influences substrate specificity in GkApcT where mutation of this residue resulted in gain-of-function arginine binding and transport (42). The equivalent position in ApGLNT1 was determined as Q383 (Fig. 7*B*). The side chains of both C198 and Q383 are predicted to be adjacent to the side chain of the glutamine substrate in ApGLNT1, suggesting they could influence substrate binding (Fig. 7*B*). To examine the likely importance of these two binding pocket residues in GLNT1 function, we assessed how conserved they were across GLNT1 orthologs of diverse insects and other aphids by generating homology models of each functionally characterized ortholog using the Alphafold protein database (36) (Fig. 7*C*). The positions of both the Cys residue in TM3 and the Gln residue in TM8 were conserved in all GLNT1 orthologs functionally characterized here (Fig. 7*C*). Furthermore, both residue positions were also conserved in the orthologs from Odonata (*Calopteryx splendens*) and Diptera (*Drosophila melanogaster*) (Fig. 7*D*).

Sequence analysis of TM3 and TM8 in the insect taxa and 9 other aphid species, for which we annotated GLNT1 orthologs, identified both residues to be conserved in all orthologs from within the GLNT1 clade (Fig. 8) and across aphid species spanning the aphid phylogeny (*SI Appendix*, Fig. S7). In contrast, the residue positions were not conserved in the additional, divergent, GLNT1-like sequences from auchenorrhynchan insects (paralogs 2-4) (Fig. 8, *SI Appendix*, Fig. S5, and Dataset S3). The divergence at these two sites, compared with functionally characterized GLNT1 orthologs, suggests that these putative transporters evolved different substrate profiles. In support of GLNT1 being an important gene in insects, in all the auchenorrhynchan species examined, a single conserved GLNT1 sequence (paralog 1 for each species) maintained the Cys and Gln residues in TM3 and TM8, respectively, and was consistently found within the GLNT1 clade (Fig. 8 and *SI Appendix*, Fig. S6).

Sap-Feeding Insects Are Dependent on Endosymbionts for Arginine Supply. *Buchnera* supply of arginine to *A. pisum* is key to the ability of ApGLNT1 to regulate amino acid metabolism in bacteriocytes in response to host demand for essential amino acids, which are metabolically downstream from glutamine. *Buchnera* supply of arginine is supported by loss of arginine biosynthesis through the urea cycle in *A. pisum* and retention of arginine biosynthesis in *Buchnera* (*SI Appendix*, Fig. S8) (8, 21, 25, 47). In fact, this symbiotic complementation in arginine biosynthesis is the general consensus in aphids. In seven additional aphid species representing two subfamilies and three tribes (Dataset S1), the urea cycle is absent (*SI Appendix*, Fig. S8). Complementing this absence, all 52 *Buchnera* strains examined by Manzano-Marín et al. retain functional arginine biosynthesis pathways (47). Host loss of arginine biosynthesis coupled with symbiont retention of arginine biosynthesis in aphids raises the question of whether GLNT1 might also regulate amino acid metabolism in other sap-feeding insects that have lost the urea cycle.

Previous studies in other sap-feeding insects that include *Bemisia tabaci* and the psyllid *Pachypsylla venusta* have also reported functional symbiont-encoded arginine biosynthesis capacity that complements loss of host urea cycle genes (6, 21–25, 48). While in other sap-feeding insects that include *D. citri* and *P. citri*, functional endosymbiont pathways for arginine biosynthesis are also retained (48, 49), host metabolic capacity has not yet been examined (6, 21–25, 48). To investigate the potential of GLNT1 to regulate amino acid metabolism in bacteriomes of *P. citri* and *D. citri*, we annotated urea cycle genes from our *P. citri* and *D. citri* RNA-seq datasets. Like *A. pisum* and *B. tabaci*, *P. citri* and *D. citri* lacked all urea cycle genes of eukaryotic origin except for nitric oxide synthase (NOS) (*SI Appendix*, Fig. S8). In addition, like *P. venusta* (6), *D. citri* encodes a copy of arginosuccinate lyase (ASL)/*argH* of bacterial origin (*SI Appendix*, Table S3 and Fig. S8), indicating that it was secondarily gained through horizontal gene transfer from bacteria. Nevertheless, while *D. citri* may control the final step of arginine biosynthesis, it would

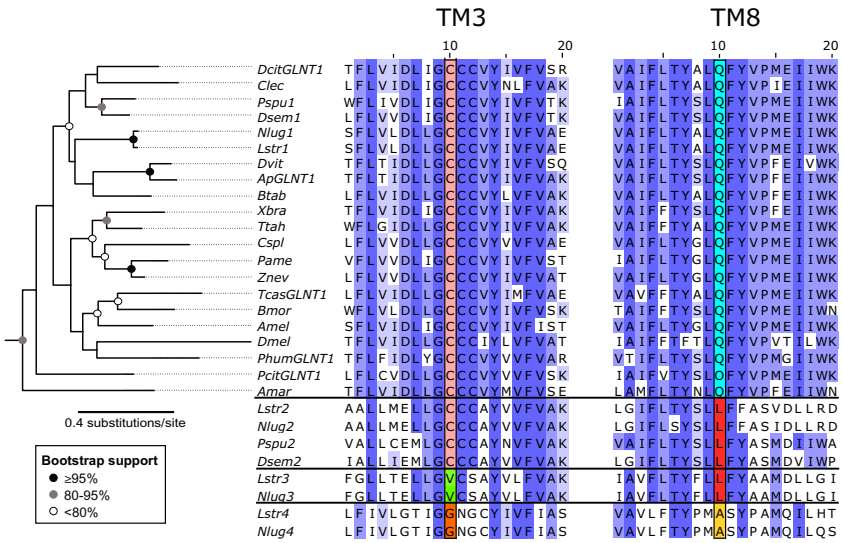

Fig. 8. Key positions within TM3 and TM8 are conserved in GLNT1 orthologs. Alignments of central portions of TM3 and TM8 from GLNT1-related sequences taken from whole protein alignments of various insect taxa. The degree of homology is denoted by shading. Residues in the equivalent position to ApGLNT1 C198 and Q383 are shown in highlighted colors. Phylogeny to the left of the alignment shows relationships between orthologs in the GLNT1 clade. It is part of the full maximum likelihood phylogeny of the insect SLC36 transporter family depicted in *SI Appendix*, Fig. S6. Four-letter taxon IDs for sequences are listed in Dataset S1. Numbers following taxon IDs for auchenorrhynchans (Nlug, Lstr, Pspu, and Dsem) indicate designated paralog number.

still require arginosuccinate from its symbiont, *Carsonella*, in order to synthesize arginine. Loss of host control over complete arginine biosynthesis in *P. citri* and *D. citri*, together with retention of symbiont-supplied arginine, supports the possibility that GLNT1 regulates amino acid metabolism in their bacteriomes.

Discussion

Endosymbiosis has evolved in numerous insect lineages, enabling hosts to feed on nutritionally restrictive diets, such as vertebrate blood or plant sap, through symbiont-mediated synthesis of essential nutrients. Ancient acquisition of these obligate bacterial symbionts followed by tight coevolution has resulted in endosymbiont gene loss and metabolic integration, patterns that have been characterized across many systems (18, 50). Less well understood are the cellular and molecular mechanisms that function to regulate host and endosymbiont metabolic exchange. One such mechanism involves the aphid amino acid transporter, ApGLNT1, which has been proposed to facilitate regulation of amino acid metabolism in aphid bacteriocytes (8). Here, we asked how ApGLNT1 evolved its proposed symbiotic role by examining its evolutionary history through functional characterization of GLNT1 orthologs within a phylogenetic framework.

The Symbiotic Role of ApGLNT1 is Derived from an Ancient, Conserved Role in Insects. By electrophysiological analysis, we investigated the functional characteristics of GLNT1 orthologs in two additional sap-feeding insects (*P. citri* and *D. citri*) and two non-sap-feeding insects (*P. humanus* and *T. castaneum*). We found that, like ApGLNT1, all four GLNT1 orthologs elicit a pH-dependent inward current during exposure to extracellular glutamine (Fig. 3), consistent with proton-coupled glutamine cotransport. Glutamine has affinity constants for GLNT1 orthologs in the mM range, indicating that they are low-affinity transporters that carry glutamine (Fig. 4). Amino acid screens additionally found that all GLNT1 orthologs do not respond to other amino acids at concentrations at which they respond to glutamine (Fig. 5), illustrating that the substrate selectivity of these transporters is narrow. Finally, the response to glutamine in all GLNT1 transporters investigated is inhibited by the presence of arginine (Fig. 6). These characteristics were also found by Price et al. (8, 31) in ApGLNT1, demonstrating both that GLNT1 orthologs are functionally conserved and that GLNT1-like function was present before the evolution of endosymbiosis in sap-feeding insects. Therefore, a symbiotic role for ApGLNT1 must be derived from its ancestral role. The GLNT1 transporter likely has a conserved role across all insects. We found that all representative insects sampled had at least one GLNT1 ortholog (*SI Appendix*, Figs. S5 and S6). Our phylogenetic analysis included 12 insect orders (Dataset S1) that last shared a common ancestor over 400 Mya (28), indicating that GLNT1 is widespread in insects, and ancient. The widespread, ancient distribution of GLNT1 implies that it plays an important, perhaps housekeeping, role in insects.

A single gene was present in almost all tested species. Only four of the insects sampled had multiple GLNT1 orthologs, and they are all members of a monophyletic group (Auchenorrhyncha), suggesting that gene duplication in GLNT1 is specific to that group, possibly occurring in the auchenorrhynchan common ancestor, and are not commonplace among insects. Some GLNT1-like paralogs in the Auchenorrhyncha are more divergent, sharing lower percent identities with ApGLNT1 than other GLNT1 orthologs (Dataset S3), to the point that our maximum likelihood analysis failed to support monophyly of GLNT1s (*SI Appendix*, Fig. S5). Importantly, all auchenorrhynchans have one GLNT1-like paralog that is supported phylogenetically as belonging to a monophyletic group of GLNT1 orthologs, when the divergent paralogs are excluded from the analysis (*SI Appendix*, Fig. S6). Inclusion of these better supported auchenorrhynchan paralogs in the GLNT1 clade is further confirmed by the sequence analysis and modeling of the GLNT1 predicted binding site.

Comparison of the GLNT1 orthologs from multiple insect taxa using a combination of sequence alignment and homology modeling identified conservation of amino acid residues in specific positions (C198 and Q383 in ApGLNT1) within the substrate binding pocket of the transporters. The type of residue residing at the equivalent positions in TM3 and TM8 in other animal and/or bacterial APC Superfamily transporters has previously been shown to influence transporter substrate specificity by dictating the space available within the binding pocket for the amino acid substrate side chain to fit (12, 37, 42). For example, the mammalian AAAP transporter slc36a2 is a H^+-coupled amino acid transporter which transports only dipolar amino acids with small or no side chains such as alanine and glycine (35). Slc36a2 shares approximately 36% protein sequence identity with ApGLNT1 and has a large, aromatic residue (F159) at the key position in TM3 (equivalent to ApGLNT1 C198). When F159 in slc36a2 was substituted by amino acid residues with progressively smaller side chains, the modified slc36a2 transporters were able to accept progressively larger amino acid substrates (37). Notably, the introduction of a Cys residue at this position in TM3 of slc36a2 (to produce slc36a2-F159C) was permissive for both gain-of-function glutamine transport and inhibition by extracellular dibasic amino acid lysine (37). Therefore, there is evidence that a Cys residue in this position in TM3 in other related AAAP transporters is associated with the key GLNT1-like characteristics of both glutamine transport and inhibition by a dibasic amino acid as reported here across all GLNT1 orthologs investigated. Similarly, in the bacterial amino acid transporter GkApcT, mutation of a residue (M321) in the center of TM8 was found to alter substrate specificity (34). In the GLNT1 orthologs, this residue was conserved as a Gln (Q383 in ApGLNT1). In GkApcT, M321 was identified as lining the binding pocket in the resolved protein structure and as being positioned such as to block access of larger amino acids. However, reducing the size of the residue at this position through mutation (M321S) resulted in increased binding pocket space and, when tested functionally, gain-of-function arginine binding (34). The TM3 and TM8 residue positions identified here have been shown to line the binding pocket in other APC Superfamily transporters which also have the core structural 5+5 TM inverted repeat (45, 51).

Overall the conservation of residues in these two key positions in TM3 and TM8 within GLNT1 orthologs from a wide range of aphids and other insect taxa (Fig. 8 and *SI Appendix*, Fig. S7) is consistent with the observed preservation of GLNT1-like substrate specificity. The progressive divergence in these residue positions in the multiple auchenorrhynchan paralogs identified here is consistent with those putative transporters having distinct substrate selectivities compared to GLNT1.

Support for a Novel Symbiotic Regulatory Role in an Ancient, Conserved Transporter. Functional and structural conservation of GLNT1 orthologs from sap-feeding and non-sap-feeding insects coupled with evidence for an ancient origin of GLNT1 in insects implies that the functional characteristics of this amino acid transporter did not evolve to fulfill a specific symbiotic regulatory role in the aphid/*Buchnera* system. We argue that GLNT1 can be ancient and conserved while also playing a novel, symbiotic regulatory role in aphids and perhaps other sap-feeding

insects through recruitment of this transporter to the symbiotic interface. Central to a role for ApGLNT1 in regulating amino acid metabolism in bacteriocytes is that arginine, the substrate inhibiting glutamine uptake, is metabolically downstream from glutamine and synthesized in bacteriocytes by *Buchnera*. Arginine synthesis requires ornithine and carbamoyl phosphate as substrates. While many *Buchnera* strains have lost or pseudogenized the genes required to synthesize ornithine (*argA-argE*) (47), aphids can synthesize ornithine through ornithine aminotransferase (OAT) (*SI Appendix*, Fig. S8). Indeed, despite *Buchnera* retaining *argA-argE* in *A. pisum*, OAT is up-regulated in bacteriocytes (23). Thus, in aphids generally, *Buchnera* most likely uses aphid-synthesized L-ornithine as a substrate for arginine biosynthesis. In contrast, *Buchnera* retains the ability to synthesize carbamoyl phosphate (through *carAB*) (47). Thus, *Buchnera* requires glutamine as an amino donor in arginine biosynthesis. *Buchnera* is not the only source of arginine for aphids, but we posit that it is the most important and dominant source. Unlike many animals, aphids have lost the ability to synthesize arginine through the urea cycle (21) (*SI Appendix*, Fig. S8). The only other source of arginine for aphids, besides *Buchnera*, is their phloem sap diet. Phloem sap indeed carries free arginine at an appreciable concentration, ranging from 4.4 mM to 15.2 mM in eight different host plants of *A. pisum* (52)–comparable to, or even exceeding, arginine IC_{50} values for all GLNT1 orthologs we examined (Fig. 6). Phloem sap levels of arginine, however, are not reflected in *A. pisum* hemolymph, where the arginine concentration is not sufficient to inhibit ApGLNT1 or other GLNT1 orthologs. In contrast, the glutamine concentration in *A. pisum* hemolymph is close to sufficient to saturate the transporter (39). Additional lines of evidence indicate that dietary arginine is an insufficient arginine source for *A. pisum*. Haribal and Jander (53) found evidence that dietary amino acids are generally not incorporated into aphids but rather are catabolized and used for de novo amino acid biosynthesis. In regard to arginine in particular, Dadd & Krieger (54) found, through artificial diet experiments in the aphid *Myzus persicae*, that omitting arginine from the diet did not affect the mean weight of nymphs in the first generation or growth to adulthood in the second generation compared to the complete diet, suggesting that dietary arginine is not necessary. Additionally, the genomes of diverse *Buchnera* strains maintain functional arginine biosynthesis pathways despite the propensity for degenerative evolution in endosymbiont genomes (50). Retention of functional pathways across *Buchnera* strains indicates that arginine biosynthesis by *Buchnera* is a general requirement for the evolutionary maintenance of the aphid/*Buchnera* system.

GLNT1 May Regulate Amino Acid Biosynthesis in Other Sap-Feeding Insect Endosymbioses. Given convergence of features of host/symbiont coevolution in many sap-feeding insects, we hypothesize that GLNT1 also regulates amino acid biosynthesis in the bacteriomes of other sap-feeders. For example, mealybugs, whiteflies, and psyllids rely on their endosymbionts for essential amino acid provisioning, and gene expression studies in all of them have found that amino acid metabolism in their bacteriomes requires some degree of host/symbiont metabolic complementarity (4, 6, 22, 48) (*SI Appendix*, Fig. S8). Conservation of GLNT1 function across insects suggests that, in addition to aphids, other insects with similar biology and needs could have co-opted GLNT1 to regulate amino acid metabolism in their bacteriomes. Importantly, the sap-feeding insects (other than aphids) for which we show arginine biosynthetic pathways in *SI Appendix*, Fig. S8 (*P. citri*, *B. tabaci,* and *D. citri*) minimally experienced ancestral loss of the urea cycle that is complemented by symbiont retention of arginine biosynthesis (4, 22, 48, 49). This pattern is clearest in *P. citri*, which has lost all urea cycle genes except for NOS, the enzyme mediating the conversion between arginine and citrulline, while its symbionts retain between them a functional arginine biosynthesis pathway (48). *B. tabaci* and *D. citri* have both reacquired some urea cycle genes through lateral gene transfer from bacteria to their host insect genomes (22) (*SI Appendix*, Fig. S8), enabling *B. tabaci* to carry out the last two steps and *D. citri* to carry out the last step in arginine biosynthesis. Host control of these last steps (even when functionally analogous genes have been pseudogenized or lost in their symbionts) still requires metabolic precursors of citrulline or arginosuccinate, both of which are solely synthesized by their symbionts. Thus, symbionts are still the source of arginine, meeting an essential requirement for GLNT1 to regulate amino acid biosynthesis in the bacteriomes of these insects. A complete test of our hypothesis that GLNT1 regulates amino acid metabolism in other sap-feeding insect bacteriomes requires localization of GLNT1 orthologs in these insects and data on free arginine concentration in their hemolymph and host plant phloem sap.

Of the insects included in this study, sap-feeding insects are not the only ones that lost urea cycle genes. *P. humanus* has also lost arginosuccinate synthase and ASL, the last two steps in arginine biosynthesis, but retains NOS, Arginase 1, and ornithine transcarbamoylase (OTC). As a human blood feeder, *P. humanus* has a rich dietary supply of amino acids, suggesting that arginine biosynthesis genes in the urea cycle experienced weak or relaxed selective pressure to be maintained, leading to their eventual purge from the *P. humanus* genome. So while *P. humanus* lacks the ability to synthesize arginine, we do not expect that GLNT1 plays a regulatory role in the *P. humanus*/*Riesia* symbiosis because *Riesia* is not bacteriome-associated. Additionally, *Riesia* can make B vitamins but cannot synthesize amino acids (55). *T. castaneum*, which lacks an endosymbiont and has a more complete diet than sap-feeders and blood-feeders, has the most complete urea cycle, retaining all genes except for OTC. The presence of the full complement of urea cycle genes between *T. castaneum* and *P. humanus* indicates that a complete urea cycle was present in the common ancestor of the insects we included in this study. However, the fact that the urea cycle is absent in *P. citri*, *D. citri*, *A. pisum* (21), and other sequenced hemipterans that include *B. tabaci* (22) (*SI Appendix*, Fig. S8) suggests that arginine biosynthesis was lost early in the evolution of sap-feeding insect endosymbiosis.

Concluding Statement

Over the past decade, it has become clear that tight, interdependent metabolic integration is a signature of host/endosymbiont genome coevolution in insect nutritional endosymbiosis (18). Most studies on metabolic integration in sap-feeding insects focus on amino acid and vitamin biosynthetic pathways requiring gene products from hosts and endosymbionts (6, 21–23, 48, 56), while general patterns in the genes regulating nutrient metabolism in these systems remain to be elucidated. This study reveals that the ancient, conserved function of ApGLNT1 does not preclude its co-option to a role in regulating amino acid metabolism in aphid bacteriocytes. Evidence for insufficient dietary supply of arginine (53, 54) and loss of the urea cycle in host aphids (21) (*SI Appendix*, Fig. S8) ensures that downregulation of glutamine supply to aphid bacteriocytes requires *Buchnera*-synthesized arginine. The fact that other phloem-feeders are also dependent on their symbionts for endogenous arginine synthesis (4, 6, 22, 23, 48, 49) (*SI Appendix*, Fig. S8) further suggests the possibility that those insects co-opted

GLNT1 to regulate bacteriome amino acid metabolism. While this work advances our understanding of host/symbiont regulation, it also leaves some questions to be explored. For example, different affinities of sap-feeding insect GLNT1 orthologs for glutamine and arginine may be indicative of nuanced GLNT1 functional dynamics in the push and pull for control of the system that is characteristic of host/symbiont coevolution (57). In addition, our findings pave the way for future work testing whether GLNT1 co-option is a general mechanism facilitating metabolic integration in sap-feeding insect endosymbioses, supporting constraint of ancestral host gene repertoire and function as a signature of genome coevolution in insects with endosymbionts (18).

Materials and Methods

GLNT1 from the pea aphid *A. pisum* and four outgroup insects (Fig. 2) were cloned and functionally expressed in *X. laevis* oocytes. Orthologs from the human body louse *P. humanus* and the flour beetle *T. castaneum* were previously annotated (26), while GLNT1 orthologs from the citrus mealybug *P. citri* and the Asian citrus psyllid *D. citri* were annotated here from RNA-seq data. Two-electrode voltage clamp was used to functionally characterize GLNT1 orthologs based on previously demonstrated ApGLNT1 transport function (8, 31). Structural homology models were constructed using HHPred and Modeller as previously described (12, 37) or AlphaFold (58). Details of taxon sampling, transcriptome sequencing and assembly, GLNT1 annotation, GLNT1 cloning, and functional and structural characterization can be found in *SI Appendix*.

Data, Materials, and Software Availability. Raw RNA-seq reads generated here were deposited in the NCBI Sequence Read Archive under BioProject PRJNA315109. Full-length coding sequences of GLNT1 orthologs used for expression constructs were deposited in GenBank under accession numbers KY448282 (*D. citri*), KY448283 (*P. citri*), KY448284 (*P. humanus*), and KY448285 (*T. castaneum*). Additional raw data (including FASTA files with insect GLNT1 sequences, alignments used for phylogenetic analysis, tree files, homology model files, original microscopy images, and electrophysiology trace files for Figs. 3, 4, and 6) can be accessed in the Zenodo repository (DOI: 10.5281/zenodo.8317238). Previously published data were used for this work (26, 59, 60).

ACKNOWLEDGMENTS. Amy Roda and Scott Weihman helped with collecting *P. citri* from the ARS Subtropical Horticultural Research Station in Miami, FL. Ian Stocks identified *P. citri* based on morphology. Kirsten Pelz-Stelinski provided *D. citri*, Sue Brown provided *T. castaneum*, and John Clark provided *P. humanus* material. Yoland Victor helped with annotating GLNT1 orthologs in insects. We thank Dan Price, Angela Douglas, Alejandro Manzano Marín, and an additional anonymous reviewer for helpful discussion and comments. We additionally thank six anonymous reviewers who gave valuable feedback on previous versions of this manuscript. Devin Kepchia, Ben Sherman, and Suhaila Rahman provided guidance and support with functional characterization in *Xenopus* oocytes. We also thank Ed James and Honglin Feng for the bacteriocyte images in Fig. 1. RNA-seq assemblies were conducted on the Pegasus high-performance computing system at the University of Miami Center for Computational Science. This work was supported by NSF award number IOS-1354154 (A.C.C.W. and C.W.L.), NSF award number DEB-1406631 (R.P.D.), Newcastle University NUAcT fellowship (C.M.H.A.), The Royal Society research grant RGS\R1\221113 (C.M.H.A.), Rank Prize new lecturer grant (C.M.H.A.), and The Physiological Society Momentum Fellowship number 47690-FR (D.T.T.).

Author affiliations: [a]Department of Biology, University of Miami, Coral Gables, FL 33146; [b]School of Natural and Environmental Sciences, Faculty of Science, Agriculture and Engineering, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom; [c]Biosciences Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne NE2 4HH, United Kingdom; and [d]Department of Molecular and Cellular Pharmacology, University of Miami Miller School of Medicine, Miami, FL 33136

Author contributions: R.P.D., C.M.H.A., D.T.T., C.W.L., and A.C.C.W. designed research; R.P.D. and C.M.H.A. performed research; R.P.D., C.M.H.A., D.T.T., C.W.L., and A.C.C.W. analyzed data; and R.P.D., C.M.H.A., D.T.T., C.W.L., and A.C.C.W. wrote the paper.

1. N. A. Moran, Symbiosis as an adaptive process and source of phenotypic complexity. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 8627–8633 (2007).
2. J. Hackett, Migration of the plastid genome to the nucleus in a peridinin dinoflagellate. *Curr. Biol.* **14**, 213–218 (2004).
3. R. Koga, X.-Y. Meng, T. Tsuchida, T. Fukatsu, Cellular mechanism for selective vertical transmission of an obligate insect symbiont at the bacteriocyte-embryo interface. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E1230–E1237 (2012).
4. F. Husnik *et al.*, Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell* **153**, 1567–1578 (2013).
5. A. Nakabachi, K. Ishida, Y. Hongoh, M. Ohkuma, S. Miyagishima, Aphid gene of bacterial origin encodes a protein transported to an obligate endosymbiont. *Curr. Biol.* **24**, R640–R641 (2014).
6. D. B. Sloan *et al.*, Parallel histories of horizontal gene transfer facilitated extreme reduction of endosymbiont genomes in sap-feeding insects. *Mol. Biol. Evol.* **31**, 857–871 (2014).
7. A. K. Hansen, P. H. Degnan, Widespread expression of conserved small RNAs in small symbiont genomes. *ISME J.* **8**, 2490–2502 (2014).
8. D. R. G. Price *et al.*, Aphid amino acid transporter regulates glutamine supply to intracellular bacterial symbionts. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 320–325 (2014).
9. S. Karkar, F. Facchinelli, D. C. Price, A. P. M. Weber, D. Bhattacharya, Metabolic connectivity as a driver of host and endosymbiont integration. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10208–10215 (2015).
10. H. Feng, L. Wang, S. Wuchty, A. C. C. Wilson, microRNA regulation in an ancient obligate endosymbiosis. *Mol. Ecol.* **2**, 1777–1793 (2017).
11. P. Mergaert, Y. Kikuchi, S. Shigenobu, E. C. M. Nowack, Metabolic integration of bacterial endosymbionts through antimicrobial peptides. *Trends Microbiol.* **25**, 703–712 (2017), 10.1016/j.tim.2017.04.007.
12. H. Feng *et al.*, Trading amino acids at the aphid–Buchnera symbiotic interface. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 16003–16011 (2019).
13. M. A. Munson, P. Baumann, M. G. Kinsey, Buchnera gen. nov. and *Buchnera aphidicola* sp. nov., a Taxon Consisting of the Mycetocyte-Associated, Primary Endosymbionts of Aphids. *Int. J. Syst. Evol. Micr.* **41**, 566–568 (1991).
14. A. E. Douglas, The nutritional physiology of aphids. *Adv. Insect. Physiol.* **31**, 73–140 (2003).
15. P. Baumann, Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annu. Rev. Microbiol.* **59**, 155–189 (2005).
16. S. Sudakaran, C. Kost, M. Kaltenpoth, Symbiont acquisition and replacement as a source of ecological innovation. *Trends Microbiol.* **25**, 375–390 (2017).
17. A. E. Douglas, Phloem-sap feeding by animals: Problems and solutions. *J. Exp. Bot.* **57**, 747–754 (2006).
18. A. C. C. Wilson, R. P. Duncan, Signatures of host/symbiont genome coevolution in insect nutritional endosymbioses. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10255–10261 (2015).
19. N. A. Moran, J. P. McCutcheon, A. Nakabachi, Genomics and evolution of heritable bacterial symbionts. *Annu. Rev. Genet.* **42**, 165–190 (2008).
20. A. E. Douglas, Microbial brokers of insect-plant interactions revisited. *J. Chem. Ecol.* **39**, 952–961 (2013).
21. A. C. C. Wilson *et al.*, Genomic insight into the amino acid relations of the pea aphid, Acyrthosiphon pisum, with its symbiotic bacterium *Buchnera aphidicola*. *Insect. Mol. Biol.* **19**, 249–258 (2010).
22. J.-B. Luan *et al.*, Metabolic coevolution in the bacterial symbiosis of whiteflies and related plant sap-feeding insects. *Genome Biol. Evol.* **7**, 2635–2647 (2015).
23. A. K. Hansen, N. A. Moran, Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 2849–2854 (2011).
24. S. J. Macdonald, G. G. Lin, C. W. Russell, G. H. Thomas, A. E. Douglas, The central role of the host cell in symbiotic nitrogen metabolism. *Proc. R Soc. B* **279**, 2965–2973 (2012).
25. S. Shigenobu, A. C. C. Wilson, Genomic revelations of a mutualism: The pea aphid and its obligate bacterial symbiont. *Cell Mol. Life Sci.* **68**, 1297–1309 (2011).
26. D. R. G. Price, R. P. Duncan, S. Shigenobu, A. C. C. Wilson, Genome expansion and differential expression of amino acid transporters at the aphid/Buchnera symbiotic interface. *Mol. Biol. Evol.* **28**, 3113–3126 (2011).
27. H.-L. Lu, C. Chang, A. C. C. Wilson, Amino acid transporters implicated in endocytosis of Buchnera during symbiont transmission in the pea aphid. *Evodevo* **7**, 24 (2016).
28. B. Misof *et al.*, Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**, 763–767 (2014).
29. J. R. Cryan, J. M. Urban, Higher-level phylogeny of the insect order Hemiptera: Is Auchenorrhyncha really paraphyletic? *Syst. Entomol.* **37**, 7–21 (2011).
30. R. A. Dahan, R. P. Duncan, A. C. C. Wilson, L. M. Dávalos, Amino acid transporter expansions associated with the evolution of obligate endosymbiosis in sap-feeding insects (Hemiptera: Sternorrhyncha). *BMC Evol. Biol.* **15**, 52 (2015).
31. D. R. G. Price, A. C. C. Wilson, C. W. Luetje, Proton-dependent glutamine uptake by aphid bacteriocyte amino acid transporter ApGLNT1. *Biochim. Biophys. Acta* **1848**, 2085–2091 (2015).
32. D. T. Thwaites, C. M. H. Anderson, The SLC36 family of proton-coupled amino acid transporters and their potential role in drug transport. *Br J. Pharmacol.* **164**, 1802–1816 (2011).
33. D. C. I. Goberdhan, D. Meredith, C. A. R. Boyd, C. Wilson, PAT-related amino acid transporters regulate growth via a novel mechanism that does not require bulk transport of amino acids. *Development* **132**, 2365–2375 (2005).
34. M. Boll, M. Foltz, I. Rubio-Aliaga, G. Kottra, H. Daniel, Functional characterization of two novel mammalian electrogenic proton-dependent amino acid cotransporters. *J. Biol. Chem.* **277**, 22966–22973 (2002).
35. Z. Chen *et al.*, Structure, tissue expression pattern, and function of the amino acid transporter rat PAT2. *Biochem. Bioph. Res. Commun.* **304**, 747–754 (2003).
36. C. M. H. Anderson *et al.*, H^+/amino acid transporter 1 (PAT1) is the imino acid carrier: An intestinal nutrient/drug transporter in human and rat. *Gastroenterology* **127**, 1410–1422 (2004).
37. N. Edwards *et al.*, Resculpting the binding pocket of APC superfamily LeuT-fold amino acid transporters. *Cell Mol. Life Sci.* **75**, 921–938 (2018).

38. A. M. Evans, K. G. Aimanova, S. S. Gill, Characterization of a blood-meal-responsive proton-dependent amino acid transporter in the disease vector, *Aedes aegypti*. *J. Exp. Biol.* **212**, 3263–3271 (2009).
39. T. Sasaki, H. Ishikawa, Production of essential amino acids from glutamate by mycetocyte symbionts of the pea aphid. Acyrthosiphon pisum. *J. Insect. Physiol.* **41**, 41–46 (1995).
40. H.-L. Lu, D. R. G. Price, A. Wikramanayake, C. Chang, A. C. C. Wilson, Ontogenetic differences in localization of glutamine transporter ApGLNT1 in the pea aphid demonstrate that mechanisms of host/symbiont integration are not similar in the maternal versus embryonic bacteriome. *Evodevo* **7**, 1 (2016).
41. M. H. J. Saier, C. V. Tran, R. D. Barabote, TCDB: The transporter classification database for membrane transport protein analyses and information. *Nucleic Acids Res.* **34**, D181–D186 (2006).
42. K. E. J. Jungnickel, J. L. Parker, S. Newstead, Structural basis for amino acid transport by the CAT family of SLC7 transporters. *Nat. Commun.* **9**, 550 (2018).
43. L. Zimmermann *et al.*, A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* **430**, 2237–2243 (2018).
44. A. Šali, T. L. Blundell, Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
45. A. Yamashita, S. K. Singh, T. Kawate, Y. Jin, E. Gouaux, Crystal structure of a bacterial homologue of Na^+/Cl^-dependent neurotransmitter transporters *Nature* **437**, 215–223 (2005).
46. C. M. H. Anderson, N. Edwards, A. K. Watson, M. Althaus, D. T. Thwaites, Reshaping the binding pocket of the Neurotransmitter: Solute Symporter (NSS) family transporter SLC6A14 ($ATB^{0,+}$) selectively reduces access for cationic amino acids and derivatives. *Biomol.* **12**, 1404 (2022).
47. A. Manzano-Marín *et al.*, Co-obligate symbioses have repeatedly evolved across aphids, but partner identity and nutritional contributions vary across lineages. *Peer. Commun. J.* **3**, e46 (2023).
48. J. P. McCutcheon, C. D. von Dohlen, An interdependent metabolic patchwork in the nested symbiosis of mealybugs. *Curr. Biol.* **21**, 1366–1372 (2011).
49. A. Nakabachi *et al.*, Defensive bacteriome symbiont with a drastically reduced genome. *Curr. Biol.* **23**, 1478–1484 (2013).
50. J. P. McCutcheon, N. A. Moran, Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* **10**, 13–26 (2012).
51. S. K. Singh, C. L. Piscitelli, A. Yamashita, E. Gouaux, A Competitive inhibitor traps LeuT in an open-to-out conformation. *Science* **322**, 1655–1661 (2008).
52. J. Sandström, J. Pettersson, Amino acid composition of phloem sap and the relation to intraspecific variation in pea aphid (Acyrthosiphon pisum) performance. *J. Insect. Physiol.* **40**, 947–955 (1994).
53. M. Haribal, G. Jander, Stable isotope studies reveal pathways for the incorporation of non-essential amino acids in Acyrthosiphon pisum (pea aphids). *J. Exp. Biol.* **218**, 3797–3806 (2015).
54. R. H. Dadd, D. L. Krieger, Dietary amino acid requirements of the aphid, *Myzus persicae*. *J. Insect. Physiol.* **14**, 741–764 (1968).
55. E. F. Kirkness *et al.*, Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 12168–12173 (2010).
56. D. R. G. Price, A. C. C. Wilson, A substrate ambiguous enzyme facilitates genome reduction in an intracellular symbiont. *BMC Biol.* **12**, 110 (2014).
57. A. C. C. Wilson, "Regulation of an insect symbiosis" in *Advances in Insect Physiology: Mechanisms Underlying Microbial Symbiosis, Advances in Insect Physiology*, J. A. Russell, K. M. Oliver, Eds. (Academic Press, Elsevier, 2020), pp. 207–232.
58. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
59. R. P. Duncan *et al.*, Dynamic recruitment of amino acid transporters to the insect/symbiont interface. *Mol. Ecol.* **23**, 1608–1623 (2014).
60. R. P. Duncan, H. Feng, D. M. Nguyen, A. C. C. Wilson, Gene family expansions in aphids maintained by endosymbiotic and nonsymbiotic traits. *Genome Biol. Evol.* **8**, 753–764 (2016).