A Novel Class of Modular Transporters for Vitamins in Prokaryotes

Rodionov, Dmitry A.; Hebbeln, Peter; Eudes, Aymeric; ter Beek, Josy; Rodionova, Irina A.; Erkens, Guus B.; Slotboom, Dirk J.; Gelfand, Mikhail S.; Osterman, Andrei L.; Hanson, Andrew D.

Published in:
Journal of Bacteriology

DOI:
10.1128/JB.01208-08

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2009

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
A Novel Class of Modular Transporters for Vitamins in Prokaryotes

Dmitry A. Rodionov,1,2* Peter Hebbeln,3 Aymeric Eudes,4 Josy ter Beek,5 Irina A. Rodionova,1 Guus B. Erkens,5 Dirk J. Slotboom,5 Mikhail S. Gelfand,2,6 Andrei L. Osterman,1 Andrew D. Hanson,4 and Thomas Eitinger3*

Burnham Institute for Medical Research, La Jolla, California 920371; Institute for Information Transmission Problems, the A. A. Kharkevich Institute, RAS, Bolshoi Karetny pereulok 19, Moscow 127994, Russia2; Institut für Biologie/Mikrobiologie, Humboldt Universität zu Berlin, Chausseestraße 117, 10115 Berlin, Germany3; Horticultural Sciences Department, University of Florida, Gainesville, Florida 326154; Department of Biochemistry, University of Groningen, Groningen Biomolecular Science and Biotechnology Institute, Nijenborgh 4, 9747 AG Groningen, The Netherlands5; and Faculty of Bioengineering and Bioinformatics, Moscow State University, Vorobyovy Gory 1-73, Moscow 119992, Russia6

Received 28 August 2008/Accepted 12 October 2008

The specific and tightly controlled transport of numerous nutrients and metabolites across cellular membranes is crucial to all forms of life. However, many of the transporter proteins involved have yet to be identified, including the vitamin transporters in various human pathogens, whose growth depends strictly on vitamin uptake. Comparative analysis of the ever-growing collection of microbial genomes coupled with experimental validation enables the discovery of such transporters. Here, we used this approach to discover an abundant class of vitamin transporters in prokaryotes with an unprecedented architecture. These transporters have energy-coupling modules comprised of a conserved transmembrane protein and two nucleotide binding proteins similar to those of ATP binding cassette (ABC) transporters, but unlike ABC transporters, they use small integral membrane proteins to capture specific substrates. We identified 21 families of these substrate capture proteins, each with a different specificity predicted by genome context analyses. Roughly half of the substrate capture proteins (335 cases) have a dedicated energizing module, but in 459 cases distributed among almost 100 gram-positive bacteria, including numerous human pathogens, different and unrelated substrate capture proteins share the same energy-coupling module. The shared use of energy-coupling modules was experimentally confirmed for folate, thiamine, and riboflavin transporters. We propose the name energy-coupling factor transporters for the new class of membrane transporters.

Transport proteins residing in the cytoplasmic membrane allow the selective uptake and efflux of solutes and are essential for cellular growth and metabolism (20). Reflecting the importance of transporters, between 3% and 16% of the genes in prokaryote genomes are predicted to encode transporter proteins (26). These transporters form numerous families that are diverse in structure, energy-coupling mechanisms, and substrate specificities (25). As only a small fraction of predicted transporter proteins have known substrates, the functional prediction and annotation of the specificities of transporter proteins in the rapidly growing number of sequenced genomes represent a substantial challenge (25, 36). For example, the uptake of many cofactors and their precursors is essential for the growth of various pathogenic bacteria whose genomes are sequenced, but the transport proteins involved have not yet been identified. The use of computational comparative genomic techniques including gene colocalization, cooccurrence, and coregulation analyses combined with experimental assays is a powerful approach to identify novel transporters and to uncover their cellular role (for a recent review, see reference 11).

The starting point for the present analysis was our recent discovery of multicomponent transport systems for the vitamin biotin (BioYNM) and the transition metals nickel (NikMNQO) and cobalt (CbiMNQO) (14, 30). These transporters all have substrate-specific components (S components), which are integral membrane proteins, and energy-coupling modules. The S components of the biotin transporter (BioY) and the metal transporters (NikMN and CbiMN) are dissimilar in sequence, but the energy-coupling modules contain similar proteins. These modules consist of an ATPase typical of the ATP binding cassette (ABC) superfamily (A component) and a characteristic transmembrane protein (T component), with the stoichiometry being so far undefined. In many prokaryotes, genes encoding energy-coupling AT modules are unlinked to nikMN, chiMN, and bioY. This observation prompted us to hypothesize the existence of AT-module-dependent transporters for additional substrates unrelated to transition metal ions and biotin.

In this study, we combined comparative genomics and experimental techniques to identify new transporters with an AT-module-plus-S-component architecture that are specific for various vitamins and related substrates. The majority of these systems are predicted to share a single AT module, a unique design among membrane transporters. The predicted modular design, substrate specificities of representative members, and shared use of the AT module by various substrate
capture components were confirmed by biochemical analyses of vitamin transport systems of the Firmicutes.

MATERIALS AND METHODS

Bioinformatics analysis and data sources. Prokaryotic genome sequences used for the comparative analysis were obtained from GenBank (1). Metabolic reconstruction, genome context analysis, and functional gene annotation were performed using the SEED comparative genomics resource as described previously (11, 24, 27). The results were captured in the “ECF class transporters” subsystem (http://theseed.ucsd.edu/FIG/subsys.cgi). Candidate DNA regulatory motifs were identified using Genome Explorer (9, 23, 27). Candidate RNA regulatory elements such as riboswitches were identified with the RNA-Pattern program (41) using input profiles from the RNA Families (Rfam) database (13). The Protein Families (Pfam) database was used to identify conserved functional domains (8). Transmembrane domains were predicted using the TMPred server (http://www.ch.embnet.org/software/TMPRED_form.html) (18).

Cloning and expression of L. casei transporters. Lactobacillus casei fodT and thiT were amplified using Taq DNA polymerase (Invitrogen, Carlsbad, CA) with L. casei ATCC 334 genomic DNA as the template and ligated into vector pNZ8048. The resulting constructs were used as templates for the amplification of the thirA promoter, fodT ORF, or thiT, and a terminator using Pfu Ultra DNA polymerase (Invitrogen). Amplicons were ligated into vector pLp252 between the BamHI and XhoI sites. The ecfA/T operon was amplified using Pfu Ultra DNA polymerase and inserted between the NcoI and BglII sites of pARCV (an expression vector derived from pBac181). The resulting constructs were used as templates for the amplification of the E. coli yuaJ gene and ligated into vector pLmEcf between the NcoI and BglII sites of pARCV. The resulting constructs were used as templates for the amplification of the L. casei ATCC 334 genomic DNA as the template and ligated into vector pNZ8048. The recombinant plasmids were cloned into Lactococcus lactis strain NZ9000. Cells were grown at 30°C in supplemented M17 medium (Difco), and for expression, nisin (0.1% [vol/vol] of a culture supernatant of the nisin-producing strain NZ7000) was added.

B. subtilis mutants and roseoflavin inhibition assay. Bacillus subtilis ypdA (ribU), ybaF (ecfA), yuaJ (thiT), and ycfC (niaP) disruption strains were obtained from the Joint Japanese and European Bacillus subtilis consortium (40). For the roseoflavin inhibition assay, B. subtilis cells were grown overnight at 40°C in chemically defined medium containing glucose (4 g/liter), tryptophan (50 mg/liter), K2HPO4 (10 g/liter), KH2PO4 (6 g/liter), sodium citrate (1 g/liter), MgSO4 (0.2 g/liter), K2SO4 (2 g/liter), FeCl3 (4 mg/liter), and MnSO4 (0.2 mg/liter) in the presence of 0.5 mg/liter erythromycin (pMUTIN2 marker). These cultures were diluted to 10- to 20-fold to yield the same cell density (optical density at 600 nm [OD600] of 0.05) in fresh medium and grown in duplicate in the absence or presence of roseoflavin (25 μM).

Vitamin uptake assays. For [3H]5-formyltetrahydrofolate and [3H]thiamine uptake assays, L. lactis cells were washed once with cold phosphate-buffered saline (PBS) and resuspended in PBS at an OD600 of 20. Assays were performed at 30°C with stirring. Cells (500 μl) were precipitated for 5 min with glucose or 2-deoxyglucose (25 mM final concentration). Assays were started by adding 500 μl of PBS containing 2.2 μM [3H]5-formyltetrahydrofolate (specific activity of 4.4 μCi/mmol) or 2.3 μM [3H]thiamine (0.33 μCi/mmol) to the mixture. At intervals, 150-μl aliquots were passed through a Whatman cellulose nitrate membrane filter (0.45 μm). Filters were washed twice with 2 ml of ice-cold PBS and cell-bound radioactivity was quantitated by liquid scintillation counting. For [3H]thiamine uptake assays, Bacillus subtilis cells were cultivated in mineral salts medium at 37°C with vigorous shaking. At an OD600 of 0.5, [3H]thiamine was added (16.6 mM final concentration; 588,000 dpm). Timed aliquots were mixed with 2 ml ice-cold 50 mM potassium phosphate buffer (pH 7.0), and cells were harvested immediately on membrane filters (see above). The filters were washed twice with 2 ml of ice-cold buffer and dried. Radioactivity was determined by liquid scintillation counting.

Cloning and expression of Leuconostoc mesenteroides transporters. Leuconostoc mesenteroides strain ATCC 8293 was grown in supplemented M17 medium (Difco), and for expression, nisin (0.1% [vol/vol] of a culture supernatant of the nisin-producing strain NZ7000) was added.

RESULTS

Comparative genomics of AT-module-dependent transport systems. A comprehensive bioinformatic analysis of 365 prokaryotic genomes using the SEED comparative genomics platform (24) revealed that the abundance and functional diversity of transporters with the novel AT architecture extend far beyond the few cases of metal and biotin transporters noted above. Thus, 432 gene cassettes encoding A and T components (AT modules) were found in 238 genomes, with the A-component genes being very often duplicated (A and A’; see Tables S1 and S2 in the supplemental material). These AT cassette fall into two groups based on their genomic organizations (Fig. 1). Three hundred thirty-five of them (group I, found in diverse microbes) resemble the previously described nickel, cobalt, and biotin transporters (14, 29) in that they occur next to genes encoding small integral membrane proteins. These membrane proteins are candidate S components, but as fewer than half are related to the S components of the nickel, cobalt, or biotin transporters, it is likely that the majority of them have novel substrate specificities. The remaining 97 AT gene cassettes (group II, found mainly in the Firmicutes, the Thermotogales, and some members of the Archaea) do not have adjacent candidate S-component genes. Almost all these cassettes are accompanied by various candidate S-component genes (459 in total) scattered elsewhere in the genome, with some genomes having as many as 12 distinct S-component genes but only one or two copies of the energy-coupling AT module. We predict that in such cases, multiple S components can use the same AT module to form an active transporter complex. The vast majority of these candidate S proteins.
FIG. 1. Distribution and comparative genomic analysis of the new class (ECF class) of prokaryotic transporters. (A) Classification and abundance of group I and group II ECF transporters. Group I transporters have a substrate-specific S component and a dedicated AT module.
**TABLE 1. Functional roles of substrate capture components of ECF transporters**

<table>
<thead>
<tr>
<th>Transporter</th>
<th>Pfam accession no.</th>
<th>Substrate</th>
<th>Evidence for specificity</th>
<th>Evidence for AT module dependence</th>
<th>Reference(s) and/or source</th>
</tr>
</thead>
<tbody>
<tr>
<td>BioY</td>
<td>PF02632</td>
<td>Biotin</td>
<td>E, R, C, O</td>
<td>E, C, O</td>
<td>14, 27, 28, 31</td>
</tr>
<tr>
<td>NikMN</td>
<td>PF01891</td>
<td>Nickel</td>
<td>E, R, C</td>
<td>E, O</td>
<td>29</td>
</tr>
<tr>
<td>CbiMN</td>
<td>PF01891</td>
<td>Cobalt</td>
<td>E, R, C</td>
<td>E, C</td>
<td>29, 33</td>
</tr>
<tr>
<td>YkoE</td>
<td></td>
<td>Thiamine precursors (HMP?)</td>
<td>R, C</td>
<td>C</td>
<td>34</td>
</tr>
<tr>
<td>MtsT</td>
<td>PF07155</td>
<td>Methionine precursors</td>
<td>R, C</td>
<td>C</td>
<td>32</td>
</tr>
<tr>
<td>HtsT</td>
<td>PF09605</td>
<td>?</td>
<td>E, C</td>
<td>E, O</td>
<td>This study</td>
</tr>
<tr>
<td>QrtT</td>
<td></td>
<td>Queuosine precursor</td>
<td>R, C, O</td>
<td>C, O</td>
<td>2, 5, 10, 21, 42, 43; this study</td>
</tr>
<tr>
<td>CbrT</td>
<td></td>
<td>Cobalamin precursor</td>
<td>C, O</td>
<td>C, O</td>
<td>This study</td>
</tr>
<tr>
<td>MtqT</td>
<td></td>
<td>Methylthioadenosine</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>RibU</td>
<td></td>
<td>Riboflavin</td>
<td>E, R, O</td>
<td>E, C, O</td>
<td>2, 5, 10, 21, 42, 43; this study</td>
</tr>
<tr>
<td>FolT</td>
<td></td>
<td>Folate</td>
<td>E, C, O</td>
<td>E, O</td>
<td>This study</td>
</tr>
<tr>
<td>QueT</td>
<td>PF06177</td>
<td>Queuosine precursor</td>
<td>R</td>
<td>C, O</td>
<td></td>
</tr>
<tr>
<td>PanT</td>
<td></td>
<td>Pantothene</td>
<td>C, O</td>
<td>E, C, O</td>
<td></td>
</tr>
<tr>
<td>HmpT</td>
<td></td>
<td>Thiamine precursor (HMP?)</td>
<td>C, O</td>
<td>C, O</td>
<td></td>
</tr>
<tr>
<td>ThiT</td>
<td>PF09515</td>
<td>Thiamine</td>
<td>E, R, O</td>
<td>E, O</td>
<td>7, 34; this study</td>
</tr>
<tr>
<td>ThiW</td>
<td></td>
<td>Thiazole</td>
<td>R, C, O</td>
<td>C, O</td>
<td>34</td>
</tr>
<tr>
<td>NiaX</td>
<td></td>
<td>Niacin</td>
<td>R, O</td>
<td>O</td>
<td>30</td>
</tr>
<tr>
<td>TrpP</td>
<td></td>
<td>Tryptophan</td>
<td>E, R</td>
<td>E, C, O</td>
<td>37, 41</td>
</tr>
<tr>
<td>PdxT</td>
<td></td>
<td>Pyridoxine</td>
<td>C</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>CbiT</td>
<td></td>
<td>Cobalamin precursor (DMB?)</td>
<td>R, C, O</td>
<td>C, O</td>
<td>33</td>
</tr>
<tr>
<td>LipT</td>
<td></td>
<td>Lipolate</td>
<td>C</td>
<td>O</td>
<td></td>
</tr>
</tbody>
</table>

---

*Identifiers of protein families are from the Pfam database (http://pfam.sanger.ac.uk/) (8).

*HMP, hydroxymethylpyrimidine; DMB, dimethylbenzimidazole.

*Evidences for the function of S components include experimental analysis (E), coregulation of their genes in a regulon (R), colocalization on the chromosome (C), or cooccurrence (O) in the genome with metabolic genes from the same pathway.

*Evidences for the AT module dependence of S components include experimental (E), colocalization on the chromosome (C), or cooccurrence (O) in the genome with AT module gene cassettes.

*References describing genomic predictions and experimental validations are in italics and in bold, respectively.

components are again unrelated to S components of the nickel, cobalt, or biotin transporters, and hence, most of them probably act on new substrates.

(i) Functional assignments for S components. The S components were classified into 21 protein families (Table 1). Previously reported experimental data implicated five of these families in transport: the BioY, CbiMN, and NikMN families mentioned above plus one family involved in riboflavin uptake (RibU [YpaA] in *Bacillus subtilis* and *Lactococcus lactis*) and another involved in tryptophan uptake (TrpP in *B. subtilis*) (2, 5, 43, 45). For 15 of the remaining 16 families, we were able to predict substrate specificities by combining genome context analysis and metabolic reconstruction (11, 27). The results show a notable preference for substrates that are vitamins or their precursors, namely, folate, pantothenate, niacin, pyridoxine, lipote, thiamine, and its precursors, and the precursors of cobalamin, methionine, and queuosine (Fig. 1 and Table 1).

The transport systems for folate, thiamine, and biotin in *Lactobacillus casei* were postulated by Henderson et al. in the 1970s to comprise individual substrate-binding membrane proteins plus a common energy-coupling component (named "energy-coupling factor" [ECF]) (15). Below, we provide experimental evidence that these *L. casei* transporters belong to the class of transporters reported here. In recognition of the pioneering work, we accordingly propose the name "ECF" for the new class and *ecfA* and *ecfT* for the genes encoding A and T components of the conserved energy-coupling modules that are shared by multiple S components. A detailed comparative genomic analysis of individual S-component families of the ECF transporters is provided below.

(ii) Nickel and cobalt transporters. The uptake of the transition metals cobalt and nickel, essential components of various enzymes, is mediated by transporters of several different types (6). Previous comparative genomic analyses identified the homologous CbiMNQO and NikMNQO transport systems as being the most widespread transporters for cobalt and nickel ions, respectively (29). The metal specificities were predicted using a combination of the identification of specific regulatory elements, namely, *B. subtilis* riboswitches (*B. subtilis* elements) and binding sites for the nickel-dependent repressor (NikR), and the analysis of colocalization with genes encoding nickel-dependent enzymes or enzymes involved in B12 biosynthesis. The pre-encoded by linked genes. Group II transporters have individual S components and shared AT modules that are unrelated to S components. Composite bar colors indicate the contributions of transporters found in different taxa to the total transporter number. Note that the S components BioY, CbrT, HstT, and QrtT (and, to a lesser extent, RibU, PanT, HmpT, ThiW, QueT, and CbiT) occur in both groups. (B) Comparative genomic analysis of the identified transporter families including their domain compositions, names, predicted substrate specificities, and example gene identifications. Substrate-specific integral membrane components (S) are shown by black rectangles, conserved transmembrane components (T) are shown by blue rectangles, and ATPase domains (A) are shown by red circles. Examples of genome context evidence (e.g., gene coregulation or colocalization) supporting the predicted transporter function are shown on the right.
dicted specificities were experimentally confirmed by metal uptake assays upon the recombinant expression of the cbiMNQO and nikMNQO operons (29). In these transport systems, the transmembrane “O” proteins (T component) and the “Q” proteins (A components) energize their cognate metal-specific “MN” modules (S components). Thus, these systems belong to group I of the ECF transporter class. The “M” components constitute a unique family of integral membrane proteins with two separate subfamilies correlating with specificity for cobalt and nickel, respectively.

(iii) Biotin transporter. A biotin uptake function for the BioY protein family was previously predicted in the comparative genomic reconstruction of the biotin regulon governed by BirA (31). The widespread biotin repressor BirA and two specialized transcriptional factors from different families, BioR in alphaproteobacteria (28) and BioQ in the Actinobacteria (27), control the biotin synthesis genes and/or the biotin transporter gene bioY. In addition to the evidence from multiple regulons, bioY clusters with biotin synthesis genes in some genomes. On the other hand, BirA-regulated bioY genes occur in some organisms without the biotin synthesis pathway. The collective evidence from coregulation, colocalization, and cooccurrence experiments strongly implied that BioY proteins are transporters involved in biotin uptake. Members of the BioY family belong to either the group I transporters that are accompanied by the dedicated AT module BioMN (44 cases, belonging mostly to the Proteobacteria and Actinobacteria) or the group II transporters that depend on the shared EcfAA’T module (81 cases, belonging mostly to the Firmicutes). Some bioY genes (10 cases, mostly in the Archaea but also in Desulfotiboterrum sp.) are located adjacent to ecfAA’T gene cassettes. A detailed phylogenetic analysis of the BioY family was performed recently. In a series of biochemical experiments, BioY from Rhodobacter capsulatus was identified as being a biotin capture and transport protein, which is converted into a high-affinity biotin transporter in the presence of its cognate BioMN module (14).

(iv) Riboflavin transporter. The candidate riboflavin transporter gene ypaA (ribU) was first identified by comparative genomic analysis of the riboflavin biosynthesis regulons governed by RFN (for riboflavin) regulatory elements (10, 21, 42). The RFN-regulated ribU genes were identified both in genomes with the de novo riboflavin synthesis genes and in those of riboflavin auxotrophs that lack the de novo synthesis pathway (e.g., some Streptococcus species). The RFN element serves as the receptor to form a metabolite-dependent riboswitch that directly binds flavin mononucleotide in the absence of proteins (44). The participation of RibU from Bacillus subtilis and Lactococcus lactis in riboflavin uptake was demonstrated by vitamin uptake assays (2, 43). The very high affinity of the RibU protein for riboflavin (Kd [dissociation constant] of 0.6 nM) (5) likely allows the cells to scavenge the vitamin from environments with low concentrations of riboflavin. The group II RibU transporters are widely distributed among the Firmicutes and the Thermotogales (63 cases). Additional RibU transporters attributed to group I were identified in some Archaea and Actinobacteria species, providing genomic evidence for their dependence on an energy-coupling module (Fig. 2). For instance, ribU in Bifidobacterium species forms an operon with a gene encoding a pair of ABC ATPase domains and a transmembrane domain that is homologous to T components of AT modules.

(v) Tryptophan transporter. The candidate tryptophan transporter YhaG (TrpP) was previously identified as being a member of the tryptophan regulon governed by the TRAP attenuation protein in B. subtilis (37, 45). On the other hand, trpP orthologs in the Clostridia are regulated by tryptophan-specific T-box RNA structures (41). The involvement of trpP in tryptophan transport was confirmed by the sensitivity of the trpP mutant to growth-inhibiting levels of tryptophan analogs (37). TrpP transporters belong to either the group I transporters that are accompanied by the dedicated energy-coupling module (in methanogenic Archaea and in some of the Firmicutes) (Fig. 2) or the group II transporters that depend on the shared EcfAA’T module (14 cases, in the Firmicutes).

(vi) Transporters of thiamine and thiamine precursors. Previous comparative genomic analyses of the thiamine regulons governed by THI elements (thiamine-pyrophosphate riboswitches) resulted in the identification of several candidate thiamine-related transporters, including YuaJ/ThiT, YkoEDC, and ThiW (34). Analysis of the distribution of genes involved in thiamine metabolism in the Firmicutes has revealed that thiT is the only thiamine-regulated gene in several pathogens that have no thiamine biosynthesis pathway, suggesting that it is involved in thiamine uptake (34). The YkoEDC system was
predicted to be involved in salvaging the thiamine precursor hydroxymethylpyrimidine, based on the positional clustering of the ykoEDC cassette with the thiaminase II gene tenA, and some other thiamine salvage genes (34). Genetic and microarray studies of B. subtilis mutants in the thiT and ykoD genes demonstrated that both mutations result in a derepression of thiamine-regulated genes, confirming the involvement of thiT and ykoEDC in the uptake of thiamine and its precursors (38). The thiamine-related transporter ThiW was predicted to be involved in the uptake of the thiazole precursor of thiamine based on the colocalization of the thiW with thiM genes, with the latter encoding the thiazole salvage enzyme hydroxyethylthiazole kinase (34). In this work, we used the occurrence profile of ThiT transporters, which are present only in the Firmicutes (34 cases), to attribute them to group II ECF transporters. Thiamine binding activity was verified experimentally for the L. casei ThiT protein (7). ThiW transporters belong to either the group I transporters that are accompanied by the dedicated AT module (in some members of the Archaea and Chloroflexi) (Fig. 2) or the group II transporters that depend on the shared EcfAA’T module (21 cases, mostly in members of the Firmicutes but also in Rubrobacter xylanophilus and Korarchaeum cryptophilum). YkoEDC transporters (group I; 44 cases) always have an integral membrane substrate capture component (YkoE), a dedicated AT module composed of duplicated ABC ATPase domains fused in a single protein (YkoD), and a transmembrane T component (YkoC).

(vii) Folate transporter. The candidate folate transporter FoIT was identified using comparative genomic analyses of the folate biosynthesis subsystem and the previously published amino acid composition and molecular mass data (16) for the L. casei folate binding protein (7). The folIT genes in several Mycoplasma genomes and in Streptococcus suis cluster on the chromosome with the folC gene, encoding the bifunctional enzyme dihydrofolate synthase/folicypolyglutamate synthase, which can add a polyglutamyl tail to folate molecules. The folT gene is widely distributed in the Firmicutes (55 cases) and is also present in a single Thermotogales species; thus, it occurs only in genomes encoding the shared energy-coupling module. Therefore, FoIT was predicted to function as a group II ECF transporter. Folate binding activity was verified experimentally for the L. casei FoIT protein (7).

(viii) Pantothenate transporter. The candidate pantothenate transporter PanT was identified in this study by genome context analysis of the coenzyme A biosynthesis subsystem in the SEED database. Many free-living bacteria are capable of de novo coenzyme A biosynthesis via the precursor pantothenate, whereas numerous pathogens are dependent on exogenous pantothenate. In many cases, the panT gene is colocalized on the chromosome with various pantothenate salvage genes including the phosphopantothenoylcysteine decarboxylase gene coAD, the phosphopantothenoylcysteine synthetase gene coaC, and the pantothenate kinase gene coAX. PanT was found mostly in members of the Firmicutes that require pantothenate for growth (Lactobacillales, Streptococcus, pathogenic members of the Clostridia, and Mycoplasma); thus, it functions presumably as a group II transporter. In contrast, the panT genes in R. xylanophilus and Oenococcus oeni cluster with genes encoding A and T components of the dedicated energy-coupling module (Fig. 2), and consequently, the corresponding PanT transporters were attributed to group I.

(ix) Queuosine precursor transporters. The candidate queuosine precursor transporters QrtT and QueT were first predicted in this work using the observed coregulation of the qrtT and queT genes by PreQ1 riboswitches that sense the modified nucleobase PreQ1 (7-aminomethyl-7-deazaguanine) and control the expression of queuosine synthesis genes (22, 35). The QrtT transporters belong to either the group I transporters that are accompanied by the dedicated AT module QrtUVW (16 cases, in members of the Enterobacteria, Actinobacteria, and Thermotogales) or the group II transporters that depend on the shared EcfAA’T module (8 cases, in the Firmicutes). The QueT transporters were found to belong mostly to group II (41 cases, in the Firmicutes but also in R. xylanophilus, Petrotoga mobilis, and two Archaea species). However, in a single genome of Staphylothermus marinus, the group I QueT transporter is accompanied by the dedicated energy-coupling module (Fig. 2).

(x) Cobalamin precursor transporter. The candidate coenzyme B12-regulated transporter CblT was identified in a previous comparative genomic analysis of the B12 regulon in bacteria (33). Orthologs of the cblT gene in members of the Firmicutes are regulated by B12 riboswitches and colocalized with the cobalamin biosynthesis genes (cbl). Analysis of distributions of genes from the B12 synthesis pathway suggested a possible role of CblT in the uptake of dimethylbenzimidazole (33), a compound attached to the cobinamide precursor to form cobalamin. Here, we attributed most of the CblT transporters in the Firmicutes to group II (11 cases). In a single species of the Firmicutes that lacks the shared energy-coupling cassette (Desulfotibacterium sp.), CblT is accompanied by a dedicated energy-coupling module (Fig. 2), thus constituting a group I transporter.

(xi) Cobalamin transporter. The CbrT transporters belong to either the group I transporters that are accompanied by the dedicated AT module CbrUV (18 cases, in members of the Actinobacteria and some members of the Firmicutes) or the group II transporters that depend on the shared EcfAA’T module (11 cases, in the Firmicutes). The genomic colocalization of cbrT with the cobalamin adenosyltransferase (pduO) and adenosylcobalamin-dependent ribonucleotide reductase (nrdI) genes in four Lactobacillus species suggests the involvement of CbrT in cobalamin uptake. In addition, the cbrT gene of L. reuteri belongs to a large gene cluster encoding the B12 biosynthesis pathway and the B12-dependent pathway of propanediol utilization. With the exception of L. reuteri, other analyzed Lactobacillus genomes lack all known genes involved in adenosylcobalamin biosynthesis and salvage. Genomic analysis of four Lactococcus species that possess the nrdI-cbrT-pduO gene cluster found NrdJ as a single known B12-dependent enzyme, providing an explanation for the observed colocalization of the genes involved in the salvage and utilization of cobalamin.

(xii) Methylthioadenosine transporter. Methylthioadenosine is a degradation product of S-adenosylmethionine and recycled to form methionine. A group I transporter encoded by the mtaTUV gene cassette was identified in only three bacterial genomes and a single archaeal genome. The colocalization of mtaTUV with the minPNKA genes involved in methylthio-
adenosine salvage implicates MtaTUV in the uptake of external methylthioadenosine.

(xiii) Methionine precursor transporter. A group I transporter encoded by the msTUV gene cluster was identified in members of the Proteobacteria and the Firmicutes in some members of the Archaea (33 cases) as an optional member of various methionine regulons, suggesting its involvement in the uptake of a methionine precursor. For example, the msTUV operon is regulated by the S-adenosylmethionine riboswitch (S box) in Bacillus cereus, by the methionine-specific T-box attenuator in Leuconostoc mesenteroides, and by the methionine regulator MtaR in Streptococcus species (32). In proteobacteria, the msTUV cassette was found only in Vibrio species, where it is predicted to be regulated by the methionine repressor MetJ, as it is preceded by candidate MetJ binding sites.

(xiv) Pyridoxine transporter. The group II transporters of the PdxT family were found only in the Firmicutes (11 cases) and are weakly similar (16 to 18% identity) to BioY, ChtT, CbrT, and PanT. The colocalization of pdxT with the pyridoxine kinase gene pdxK and the absence of de novo pyridoxine synthesis in Lactobacillus and Streptococcus suggests a role for PdxT in pyridoxine uptake.

(xv) Thiamine precursor transporter. Most of the HmpT transporters belong to group II, which depends on the shared energy-coupling module (38 cases, in members of the Firmicutes and the Thermotogales and some members of the Archaea). The group I HmpT transporters accompanied by a dedicated AT module (Fig. 2) are present in two members of the Archaea and a single species of the Actinobacteria. Positional analysis of the hmpT genes suggests their likely involvement in the salvage of thiamine or its precursor hydroxymethylpyrimidine. In Streptococcus and Enterococcus species, hmpT is located in a possible operon with a paralog of the phosphomethylpyrimidine kinase gene thiD, whereas in the Archaea, it is colocalized with the thiamine-monophosphate kinase gene thiL.

(xvi) Niacin transporter. The candidate niacin transporter NiaX was previously identified in the comparative genomics study of the niacin regulon controlled by NiaR (30). The NiaR-regulated niaX genes are present in Streptococcus and Enterococcus species that lack genes for de novo NAD biosynthesis, suggesting the involvement of NiaX in niacin uptake. Here, we attributed the NiaX transporters to the group II ECF transporters (18 cases) based on their limited distribution only in the Firmicutes and weak similarity (16 to 19% identity) to OqtU and RibU.

(xvii) Lipoate transporter. The group II transporters of the LipT family were identified only in three Phytoplasma species that possess the shared EcFAA'T module. The LipT proteins are weakly similar (16 to 19% identity) to ThiT, FolT, PdxT, RibU, and HmpT. The colocalization of lipT with the lipoate-protein ligase A gene lplA and the absence of the lipoate synthesis pathway in Phytoplasma suggest a role for LipT in lipoate uptake.

(xviii) Other transporters. Comparative genomic analysis failed to suggest a specific function for group I transporters encoded by the hisTUV genes (named Hts for hypothetical transport system) that were found mostly in animal-associated members of the Firmicutes (e.g., Streptococcus) and the Actinobacteria. The gram-negative oral spirochete Treponema denticola has nine copies of the htsTUV cassette.

The presence of ecfA'T genes encoding candidate energy-coupling modules in Methanosarcina species, which lack genes encoding any of the above-described 21 families of S components, points to the existence of additional ECF transporters with unrecognized specificities.

Experimental validation of AT module sharing by vitamin transporters. To validate our bioinformatic analysis, we tested three of its predictions, namely, that (i) the previously characterized riboflavin transporter RibU depends on the AT module; (ii) the S components FolT and ThiT are specific for folate and thiamine, respectively, and are AT module dependent; and (iii) the multiple S components of group II transporters (Fig. 1) share a common energy-coupling module (EcFAA'T).

(i) RibU interacts with EcFAA'T in B. subtilis. The dependence of riboflavin transport on the S component (RibU) and the energy-coupling module was analyzed in B. subtilis knock-out strains. [3H]Riboflavin transport required functional ribU (ybaF) and ecfT (yceI) genes; the disruption of either gene abolished the uptake of the vitamin (Fig. 3). This result was corroborated by inhibitor studies with the toxic riboflavin analog rosoflavin, which enters cells via the riboflavin transporter (43); disrupting either ribU or ecfT reduced sensitivity to this inhibitor (see Fig. S1 in the supplemental material).

(ii) Substrate specificity and AT module dependency of FolT and ThiT. The predicted substrate and the AT module dependence of FolT were tested first using proteins from Leuconostoc mesenteroides. The folT and ecfA'A'T genes were expressed in a folate-auxotrophic Escherichia coli strain (pabA abgT), which cannot produce the folate precursor 4-aminobenzoate or take up folates. When analyzed for growth on minimal agar medium (19), only transformants containing both folT and the ecfA'A'T cassette were able to utilize 5-formyltetrahydrofolate

![FIG. 3. Riboflavin uptake in Bacillus subtilis. Shown are data for the effect of disrupting ribU (triangles), ecfT (circles), or yceI (squares) on [3H]riboflavin uptake by B. subtilis. (YceI gene served as a control; it encodes a protein unrelated to ECF transporters.) Cells were grown without riboflavin to an OD600 value of 0.5, and [3H]riboflavin was added (17 nM final concentration). At the times indicated, cells were harvested by filtration and washed, and their H contents were determined. Values are means of duplicates; error bars indicate ranges.](image)
energized cells, only those coexpressing drofolate regardless of which proteins they expressed. Among energized cells acquired neither thiamine nor 5-formyltetrahydrofolate (Fig. 4B and C). Deenergized L. casei thiT containing empty vectors (triangles) or carrying (squares). Cells were energized with glucose (black symbols) or deenergized with 2-deoxyglucose (open symbols).

![FIG. 4. Folate and thiamine uptake evidences.](image)

(A) Folate uptake by E. coli cells coexpressing L. mesenteroides folT and ecfAA T. Recombinants containing empty vector or expression plasmids for the production of FolT, EcfAA’ T, or FolT plus EcfAA’ T were spotted (10 μl) after serial 10-fold dilutions onto nonsupplemented minimal medium and onto minimal medium containing 4-aminobenzoate (3.6 μM) or 5-formyltetrahydrofolate (11 μM). Plates were incubated for 48 h at 37°C. (B and C) Uptake of [3H]thiamine (B) and [3H]5-formyltetrahydrofolate (C) by L. lactis containing empty vectors (triangles) or carrying L. casei thiT or folT (circles), ecfAA’ T (diamonds), or thiT or folT coexpressed with ecfAA’ T (squares). Cells were energized with glucose (black symbols) or deenergized with 2-deoxyglucose (open symbols).

(Fig. 4A). This establishes that folate is transported and that transport requires the energy-coupling module.

We then produced the FolT, ThiT, and EcfAA’ T proteins from L. casei in various combinations in Lactococcus lactis and assayed [3H]thiamine and [3H]5-formyltetrahydrofolate uptake in energized or deenergized cells (Fig. 4B and C). Deenergized cells acquired neither thiamine nor 5-formyltetrahydrofolate regardless of which proteins they expressed. Among energized cells, only those coexpressing thiT and ecfAA’ T acquired significant amounts of thiamine. The situation was similar for 5-formyltetrahydrofolate except that cells expressing folT alone acquired some label; this slow uptake may be due to the functional interaction of FolT with the endogenous L. lactis EcfAA’ T module or possibly to the basal activity of the solitary FolT protein. These data confirm activity with the predicted FolT and ThiT substrates and demonstrate the dependence of both thiamine and folate uptake on the same EcfAA’ T module, as inferred in the classical work of Henderson et al. (15).

(iii) Physical interaction of an AT module with various S components. The shared use of the common energy-coupling module was also tested by a series of pullout experiments. His-tagged EcFA, untagged EcFA’, and FLAG-tagged EcFT from L. mesenteroides were coproduced in E. coli with or without FLAG-tagged FolT, PanT, or RibU. Membranes of the recombinant strains were solubilized and subjected to nickel-chelate affinity chromatography, followed by SDS-PAGE and Western blotting. The three EcfAA’ T proteins copurified during chromatography (Fig. 5), as indicated by immunodetection (for His-tagged EcFA and FLAG-tagged EcFT) and electrospray ionization time-of-flight peptide mass fingerprint analysis (for EcFA’) (not shown). As expected, FLAG-tagged FolT, PanT, and RibU did not bind to the affinity resin in the absence of the EcfAA’ T module (Fig. 5, six right-hand lanes). However, each of these S components copurified with the EcfAA’ T complex. The EcFA, EcFA’, and EcFT components thus form a stable tripartite complex and can form quadrupartite complexes with each of three different S components.

**DISCUSSION**

The broad distribution, functional versatility, and modular assembly of the energy-coupling module-dependent transport systems are summarized in Fig. 1 and Fig. S2 in the supplemental material. The ECF transporters form a novel class of membrane transporters that can be classified into two groups: group I includes transporters from diverse microbial lineages (170 species) that have a dedicated AT module encoded in the same gene cluster as an S component, and transporters of group II (a total of 459 transporters in 91 species) employ a universal energy-coupling module (EcfAA’ T) that is encoded by a separate gene cassette and shared by many different unlinked S components. Group II is ubiquitous in the phyla Fircmutes and Thermotogales and also occurs in some members of the Archaea.

The S components identified in this work could be classified into at least 20 protein families that correspond to different substrate specificities (Table 1). Most of them are integral membrane proteins of comparable sizes (155 to 230 residues) that have six predicted transmembrane domains. The NikM and CbiM proteins, which form a single family in the Pfam database, are larger (210 to 250 residues) and are predicted to have seven transmembrane domains. Apart from the CbiM/ NikM family, only five other S-component families (BioY, MtsT, HtsT, QueT, and ThiT) are present in the Pfam data-
base, and all of them are annotated as hypothetical membrane proteins. Sequence comparisons of representative S components from 18 families revealed very little overall pairwise identity between the proteins from different families (see Table S3 in the supplemental material). A detailed phylogenetic comparison will be needed to establish whether different S components are related.

The ECF transporters identified in this study are mechanistically unique. Their substrate specificity is mediated by integral membrane proteins (S components), which form active transporters in the presence of the energy-coupling AT module. The stoichiometry of the transporter components is unknown, but domain fusions in various ECF transport systems give some clues. First, the nik, chi, and bio gene cassettes encode a single A component (ATPase), but as noted above, dual A components are more common. Second, in some cases, the two A components are fused (Fig. 1B and 2). Third, rare fusions of transporter components include two “SAA” fusions, four “TAA” fusions, and one “ST” fusion in the Archaea, the Chloroflexi, and the Actinobacteria (Fig. 2 and see Table S1 in the supplemental material). On this basis, and because shared EcIAA’T components and specific S components formed quadrupartite complexes (Fig. 5), we propose a quadrupartite model in which the S component binds and translocates the substrate across the membrane. The translocation process is coupled to ATP hydrolysis mediated by an AT module that contains two ATPase domains and one transmembrane T component.

How do ECF transporters relate to the ABC transporter superfamily? The latter transporters couple ATP hydrolysis to substrate uptake or efflux (3, 4). ABC importers and exporters share a four-component architecture comprised of two transmembrane and two ATP-hydrolyzing domains. Prokaryotic ABC importers have additional extracytoplasmic soluble proteins that mediate substrate binding and delivery to the respective transmembrane components. Fundamental differences between ECF transporters and classical ABC importers include (i) the absence of extracytoplasmic substrate binding proteins and their replacement by integral membrane proteins and (ii) the shared use of energy-coupling AT modules by many highly diverse S components. Such sharing is occasionally seen in classical ABC transporters, but it always involves very similar substrates and substrate binding proteins (17). A less fundamental but nonetheless marked characteristic of ECF transporters is a predilection for vitamins.

Finally, as noted at the outset, numerous human pathogens such as Mycoplasma, Ureaplasma, and Streptococcus strains rely totally upon transporters to obtain vitamins and other essential metabolites due to the absence of the corresponding de novo biosynthetic pathways. Many of these microorganisms use ECF transporters, and indeed, certain ecf genes have been found to be essential for the growth and survival of Streptococcus pneumoniae and Mycoplasma genitalium (12, 39). All components of ECF transporters, especially the unique S and T proteins, are thus potential targets for antibiotic development. In fact, the centrality of the T component to the uptake of multiple compounds makes it a classic Achilles’ heel.

ACKNOWLEDGMENTS

This work was supported by grants from the Deutsche Forschungsgemeinschaft (EI 374/3-1 to T.E. and P.H.), the National Institute of Allergy and Infectious Diseases (R01-AL066244-01A2 to A.L.O. and D.A.R.), the National Institutes of Health (R01-GM071382 to A.D.H.), The Netherlands Organization for Scientific Research (vidi grant to D.J.S. and Toptalent grant to J.T.B.), and the Molecular and Cellular Biology program of the Russian Academy of Sciences and the Howard Hughes Medical Institute (55005610 to M.S.G.).

We thank A. Pohlmann (Berlin, Germany) for peptide mass analysis and E. Dervyn (INRA, France) for B. subtilis disruptants.

REFERENCES
