JOURNAL OF BACTERIOLOGY, Oct. 1995, p. 6005–6010
0021-9193/95/$04.00+0

Vol. 177, No. 20

Derived Structure of the Putative Sialic Acid Transporter from *Escherichia coli* Predicts a Novel Sugar Permease Domain

JUAN MARTINEZ,[1]† SUSAN STEENBERGEN,[2] AND ERIC VIMR[1,2]*

Departments of Pathobiology[2] and Microbiology,[1] University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Received 20 April 1995/Accepted 4 August 1995

Catabolism of sialic acids by *Escherichia coli* requires the genes *nanA* and *nanT*, which were previously mapped between *argG* and *rpoN* (E. R. Vimr and F. A. Troy, J. Bacteriol. 164:845–853, 1985). This organization is confirmed and extended by physical mapping techniques. An open reading frame beginning 135 bp from the *nanA* translational stop codon could code for a 53,547-Da hydrophobic polypeptide predicted to contain 14 transmembrane segments. Complementation analysis confirmed that *nanT* is required for sialic acid uptake when expressed in *trans*. NanT is homologous to a putative permease encoded by open reading frame 425, which maps between *leuX* and *fecE* in the *E. coli* chromosome. However, unlike this hypothetical permease or previously reported monosaccharide transporters, NanT contains a centrally located domain with two additional potential membrane-spanning segments plus one amphiphilic α-helix that may be important for the structure and function of sialic acid-permease.

The primary distribution of sialic acids (3-deoxy-D-glycero-D-galacto-2-nonulosonates) in Deuterostomia (44) implies that the major environmental sources of these sugars for bacterial nutrition are the many different sialoglycosides found on complex metazoan cell surface glycoconjugates. Microbial use of the terminal sialyl units on these sialoglycoconjugates is thus initiated extracytoplasmically by sialidase (neuraminidase; EC 3.2.1.18)-mediated cleavage, and there are examples of both gram-positive and gram-negative bacteria that synthesize and secrete these hydrolytic enzymes (see reference 38 for a review). *Escherichia coli* and other sialidase-negative bacteria presumably scavenge the free sialic acids released into the environment by sialidase-positive microorganisms (40, 41). The released sialic acids are next actively transported by a specific permease and degraded in the cytoplasm to *N*-acetylmannosamine plus pyruvic acid by *N*-acylneuraminate pyruvate lyase (EC 4.1.3.3), or sialic acid-aldolase; these breakdown products of sialic acid then enter biosynthetic or general catabolic pathways (11, 37, 40, 41). Sialic acid is an inducer of both *nanA* and *nanT*, with induction ratios greater than 1,000 or 90 times higher than the uninduced levels observed for the aldolase and permease activities, respectively (40, 41). The value of sialic acid-aldolase as a reagent for the synthesis of therapeutically useful sialic acid derivatives (5, 6, 16, 18, 43) provided the impetus for the cloning, sequencing, and overexpression of *nanA* by several groups of investigators (14, 15, 26, 27). The three-dimensional structure of native aldolase has also been determined and demonstrates its organization as a homotetramer of 33-kDa polypeptides (12).

Investigation of sialidase-negative *Bacteroides fragilis* mutants in an animal model of extraintestinal disease has indirectly implicated host sialic acids as important metabolites for the growth of this animal pathogen in vivo (7). This conclusion is supported by the primary distribution of sialic acids in higher animals, as previously discussed (11) and the widespread occurrence of inducible sialic acid-aldolase and -permease activities in pathogenic bacteria (23, 24, 34). It has been suggested that sialyl residues on mucins and other sialoglycoconjugates either bathing or attached to the colonic epithelium are in dynamic equilibrium with the intestinal microbiota (4). Hence, sialic acid transport is predicted to be an important physiological process with broad significance to host-microbe interactions. A further indication of the importance of sialic acid transport is highlighted by human Salla disease, an inborn error of metabolism resulting in the loss or inactivation of an acidic sugar transporter in lysosomes (29). The mammalian permease is coupled to the proton gradient across the lysosomal membrane and appears to transport some of the same sialic acid substrates recognized by the bacterial permease (19, 20). In the *E. coli* K1 strain K-235, sialic acid transport was blocked 98% with 1 mM 2,4-dinitrophenol but retained 24 to 34% of its uptake activity even in the presence of 5 mM sodium azide or sodium arsenate, respectively (31), suggesting to us that NanT may be a sugar-cation symporter. The uptake of sialic acid in an unmodified form by an aldolase-deficient mutant showed that the transporter is not a sugar:phosphotransferase (40). Despite these results nothing at the molecular level was known about the sialic acid transporter. In this communication we present the sequence of *nanT* and the derived structure of the putative sialic acid-permease. Homology analysis indicates that *nanT* codes for a membrane protein with weak primary but possibly extensive secondary structural similarity to a range of prokaryotic sugar-cation symporters and eukaryotic facilitators. However, sialic acid-permease contains a novel central domain that is likely to be important for its function and transmembrane topology. This domain is absent in all other known sugar transporters. Homology was detected between the amino- and carboxy-terminal regions of NanT with similar regions of a functionally unidentified *E. coli* open reading frame 425 (ORF425). NanT and the hypothetical permease thus appear to define a new subfamily of transporters.

(The results of this study were submitted by Juan Martinez in partial fulfillment of an undergraduate honor's thesis in the Department of Microbiology, University of Illinois at Urbana.)

Previous cloning of *nanA* from *E. coli* K-12 for the purpose of overproducing *N*-acylneuraminate pyruvate lyase showed

* Corresponding author. Mailing address: 2522 VMBSB, 2001 S. Lincoln Ave., Urbana, IL 61801. Phone: (217) 333-8502. Fax: (217) 244-7421. Electronic mail address: ervimr@ux1.cso.uiuc.edu.

† Present address: Department of Microbiology, Washington University, St. Louis, MO.

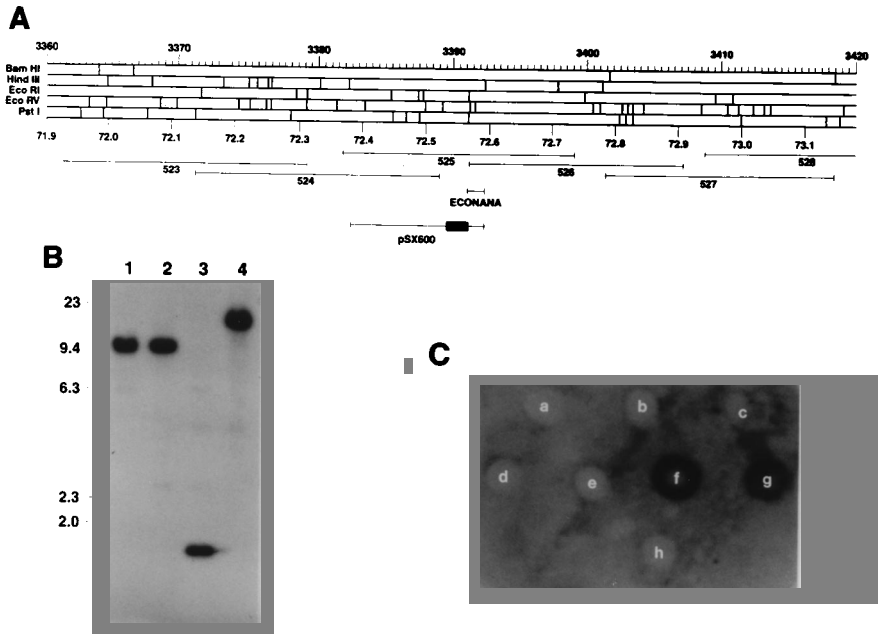

FIG. 1. Physical mapping of *nanA* and characterization of the *Hin*dIII insert in pSX600. (A) The partial restriction endonuclease map of the *E. coli* chromosome was adapted from the realigned genetic and physical maps (32), where the sizes (in kilobases) and locations of the restriction fragments are given in the top line and the corresponding genetic map positions (given in minutes) are given in the bottom line. Numbered segments below the genetic map correspond to λ clones from the Kohara collection (17). ECONANA (32) represents the *nanA*$^+$ subclone described in the text. The region of pSX600 sequenced is shown by the solid bar. (B) Chromosomal DNA isolated from strain EV11 was digested to completion with *Hin*dIII (lane 1), *Eco*RV (lane 2), *Hin*dIII plus *Eco*RV (lane 3), or *Pst*I (lane 4) and probed with the oligonucleotide specific to *nanA* described in the text. The numbers on the left indicate the sizes (in kilobases) of λ DNA *Hin*dIII fragments. (C) The same oligonucleotide described above was used to probe Kohara λ clones 520 to 527, corresponding to phage lysates a through h indicated.

that both sialic acid-aldolase and sialic acid-permease activities are encoded on an approximately 9-kb chromosomal *Hin*dIII fragment (15, 26). Phage P1-mediated transduction experiments indicated that *nanA* and *nanT* are at least 90% linked, mapping between *argG* and *rpoN* in the order *argG-zgj-791*::Tn*10-nanT-nanA-rpoN* (40). The mapping data further indicated that *nanA* and *nanT* are more tightly linked to *rpoN* than to *argG* (40). A 1,243-bp *Hin*dIII-*Eco*RI fragment containing *nanA*$^+$ was sequenced and shown to code for the 32,640-Da aldolase monomer (15, 27). The ATG translational start site located 135 bp 3′ of the *nanA* stop codon defined the start of a potential ORF that could encode the first 45 amino acid residues of sialic acid-permease (15). The occurrence of promoter-like elements 5′ to the *nanA* start site but not in front of the putative *nanT* ORF (27) strongly suggests that *nanA* is the first gene of a sialocatabolic operon. Since the complete sequence of *nanT* was not forthcoming, and because of the limited knowledge about many sialometabolic genes in bacteria and eukaryotes (38), we decided to determine whether the putative sialic acid-permease is similar to the known family of monosaccharide transporters as a first step toward more detailed characterization of sialic acid uptake and the role of this process in microbial physiology and pathogenesis.

To isolate *nanT* for sequence analysis, DNA was prepared from strain EV11 (*thi-1 leuB6 hisG1 argH1 thr-1 lacY1 gal-6 xyl-7 mtl-2 rpsL9 tonA2 supE44 kps-23*) as described previously (39) and digested to completion with *Hin*dIII. Then, fragments with sizes of 6 to 10 kb were ligated into *Hin*dIII-digested pUC18. EV11 is a spontaneous acapsular derivative of our prototypical *E. coli* K-12/K1 hybrid strain EV1 (41). On the basis of phenotypic and genetic (cotransduction) analysis, EV11 is predicted to contain a leaky mutation in *neuS* coding for the polysialyltransferase in region 2 of the *kps* locus for polysialic acid synthesis (39). The *nan* locus in EV11 appears to be identical to the equivalent locus in *E. coli* K-12 (40). Strain HB101 (*supE44 hsdA22 recA13 ara-14 proA2 lacY1 galK2 leuB6 rpsL20 xyl-5 mtl-1*) was transformed with the ligation mixture, and several hundred of the resulting ampicillin-resistant colonies were screened by hybridization with an end-labeled oligonucleotide, 5′-GAAGAACACTGCGATCACTATCGGGCA-3′, complementary to *nanA* and corresponding to residues 116 to 124 of sialic acid-aldolase (26). The generally used conditions for probe labeling and hybridization have been described elsewhere (28). Plasmids from six positive clones were isolated and digested with *Hin*dIII; each appeared to contain the identical insert, and one was saved as pSX600.

Southern hybridization analysis confirmed that the 9.2-kb *Hin*dIII fragment which we cloned had the expected restriction pattern. Figure 1A shows the physical map of a portion of the *E. coli* K-12 chromosome which, on the basis of restriction site analysis, is predicted to include the 1.2-kb *Hin*dIII-*Eco*RI fragment containing *nanA* and was designated ECONANA (32). The size (9.2 kb) of the relevant *Hin*dIII fragment (Fig. 1B, lane 1) is more similar to this map and to the 9.8-kb size reported by Ohta et al. (26) than it is to the size (7.7 kb) of the fragment reported by Kawakami et al. (15). The sizes of the *Eco*RV (9.2 kb), *Eco*RV-*Hin*dIII (1.04 kb), and *Pst*I (11.3 kb) fragments (Fig. 1B, lanes 2 to 4, respectively) were also consistent with those predicted from the physical map shown in Fig. 1A. To confirm these and our previous results (40), a subset of the Kohara collection (17) was probed with the *nanA*-specific oligonucleotide described above, showing that λ clones 525 and 526 contained complementary DNA (Fig. 1C). Together with the genetic and physiologic data already available for the *nan* system, the current results further suggest that *nanA* and *nanT* are organized in a sialocatabolic operon, which is likely to be coordinately regulated by sialic acid availability and positively controlled by the cyclic AMP-catabolite repres-

sor protein complex (31, 37, 38, 40, 41). We are currently attempting to isolate mutants with altered responses to induction by sialic acid as an approach to better define the mechanism of *nan* regulation.

To complete the sequencing of *nanT*, an oligonucleotide, 5′-CGCATCTGTACCCTAC-3′, beginning 44 bp from the predicted *nanA* stop codon (27) was chosen as the initial primer, by using CsCl-purified pSX600 as the template and the T7 polymerase kit from United States Biochemical, as previously described (11). Nineteen additional primers were synthesized from the emerging new sequence data to generate the complete nucleotide order on both strands. Several regions that were difficult to sequence, presumably because of a secondary structure in the template, were determined by using the Ampli*Taq* sequencing kit from Perkin-Elmer, following the manufacturer's instructions. Reactions were carried out in a model 9600 Perkin-Elmer PCR instrument. Sequence data were assembled by using the AssemblyLIGN program in MacVector 5.1 DNA analysis software. As shown in Fig. 2, the second largest potential ORF could code for a 53,547-Da polypeptide predicted to have extensive hydropathic character. Since the ATG translational start codon is preceded by a reasonable ribosome-binding site (35), GGAG, we suggest that this ORF, and not the larger one beginning with GTG 30 bp upstream, codes for NanT.

To provide preliminary evidence that pSX600 codes for an active sialic acid transporter, plasmid transformed into the permease mutants EV55 and EV282 restored growth of these strains on sialic acid as the sole carbon source as well as the uptake of radiolabeled sialic acid (Table 1). As expected, EV50 harboring pSX600 also grew on sialic acid as the sole carbon source (Table 1), since the plasmid is predicted to contain *nanA*$^+$ (Fig. 1). Previous results have identified a putative *nan* promoter located 5′ of *nanA* (27). The relatively high rate of sialic acid uptake by EV50 grown in the absence of exogenously added sialic acid is caused by induction of the operon from within, resulting from endogenous sialic acid synthesis by *neu* biosynthetic gene products of the *kps* locus resident in our K-12/K1 hybrid strains (40). The low levels of sialic acid uptake activity of the *nanA nanT* double mutant EV282 are thus expected to be the same as those in a mutant deficient in only the permease activity (Table 1). In strains containing an active sialic acid-aldolase, uptake assays reflect both sialic acid degradation and transport, and the rate of uptake under this condition was previously shown to be lower than that in an aldolase mutant (40). However, the lack of a dramatic gene dosage effect on permease activity in strain EV50, EV55, or EV282 harboring *nanT*$^+$ in *trans* further implies that there is a relatively low level of expression of *nanT* from the multicopy vector, as previously observed for expression of cloned *nanA* (14, 26) or that there is a limited number of sites for assembly of active permeases. These possibilities are not mutually exclusive. Unexplained vagaries in the expression of cloned transporter genes and inner membrane proteins in general are said to be a problem (9), which may also help to account for the lack of a dramatic gene dosage effect in our uptake assays. In addition, previous results showed that *nanA* expression was increased only about twofold in the presence of an inducer (26), in contrast to increases of greater than 1,000 times in the wild type (40). As predicted, if *nanT* codes for the permease, a pUC18 subclone, pSX601, containing the 3.5-kb *Eco*RI fragment with one restriction enzyme site located 139 bp from the predicted *nanT* ATG start (Fig. 1A and 2) did not complement EV55. This result is consistent with the absence of any independently transcribed ORF required for sialic acid uptake being located immediately downstream of *nanT* but falls short

```
1    AGCGTGTGGGCATCGCCCACCGCGGGAGACTCACAATGAGTACTACAACCCAGAAT
                                    M  S  T  T  T  Q  N    7

57   ATCCCGTGGTATCGCCATCTCAACCGTGCACAATGGCGCGCATTTTCCGCTGCCTGG
      I  P  W  Y  R  H  L  N  R  A  Q  W  R  A  F  S  A  A  W  26

114  TTGGGATATCTGCTTGACGGTTTTGATTTCGTTTTAATCGCCCTGGTACTCACCGAA
      L  G  Y  L  L  D  G  F  D  F  V  L  I  A  L  V  L  T  E  45

171  GTACAAGGTGAATTCGGGCTGACGACGGTGCAGGCGGCAAGTCTGATCTCTGCAGCC
      V  Q  G  E  F  G  L  T  T  V  Q  A  A  S  L  I  S  A  A  64

228  TTTATCTCTCGCTGGTTCGGCGGCCTGATGCTCGGCGCTATGGGTGACCGCTACGGG
      F  I  S  R  W  F  G  G  L  M  L  G  A  M  G  D  R  Y  G  83

285  CGTCGTCTGGCAATGGTCACCAGCATCGTTCTCTTCTCGGCCGGGACGCTGGCCTGC
      R  R  L  A  M  V  T  S  I  V  L  F  S  A  G  T  L  A  C 102

342  GGCTTTGCGCCAGGCTACATCACCATGTTTATCGCTCGTCTGGTCATCGGCATGGGG
      G  F  A  P  G  Y  I  T  M  F  I  A  R  L  V  I  G  M  G 121

399  ATGGCGGGTGAATACGGTTCCAGCGCCACCTATGTCATTGAAAGCTGGCCAAAACAT
      M  A  G  E  Y  G  S  S  A  T  Y  V  I  E  S  W  P  K  H 140

456  CTGCGTAACAAAGCCAGTGGTTTTTTGATTTCAGGCTTCTCTGTGGGGGCCGTCGTT
      L  R  N  K  A  S  G  F  L  I  S  G  F  S  V  G  A  V  V 159

513  GCCGCTCAGGTCTATAGCCTGGTGGTTCCGGTCTGGGGCTGGCGTGCGCTGTTCTTT
      A  A  Q  V  Y  S  L  V  V  P  V  W  G  W  R  A  L  F  F 178

570  ATCGGCATTTTGCCAATCATCTTTGCTCTCTGGCTGCGTAAAAACATCCCGGAAGCG
      I  G  I  L  P  I  I  F  A  L  W  L  R  K  N  I  P  E  A 197

627  GAAGACTGGAAAGAGAAACACGCAGGTAAAGCACCAGTACGCACAATGGTGGATATT
      E  D  W  K  E  K  H  A  G  K  A  P  V  R  T  M  V  D  I 216

684  CTCTACCGTGGTGAACATCGCATTGCCAATATCGTAATGACACTGGCGGCGGCTACT
      L  Y  R  G  E  H  R  I  A  N  I  V  M  T  L  A  A  A  T 235

741  GCGCTGTGGTTCTGCTTCGCCGGTAACCTGCAAAATGCCGCGATCGTCGCTGTTCTT
      A  L  W  F  C  F  A  G  N  L  Q  N  A  A  I  V  A  V  L 254

798  GGGCTGTTATGCGCCGCAATCTTTATCAGCTTTATGGTGCAGAGTGCAGGCAAACGC
      G  L  L  C  A  A  I  F  I  S  F  M  V  Q  S  A  G  K  R 273

855  TGGCCAACGGGCGTAATGCTGATGGTGGTCGTGTTGTTTGCTTTCCTCTACTCATGG
      W  P  T  G  V  M  L  M  V  V  V  L  F  A  F  L  Y  S  W 292

912  CCGATTCAGGCGCTGCTGCCAACGTATCTGAAAACCGATCTGGCTTATAACCCGCAT
      P  I  Q  A  L  L  P  T  Y  L  K  T  D  L  A  Y  N  P  H 311

969  ACTGTAGCCAATGTGCTGTTCTTTAGTGGCTTTGGCGCGGCGGTGGGATGCTGCGTA
      T  V  A  N  V  L  F  F  S  G  F  G  A  A  V  G  C  C  V 330

1026 GGTGGCTTCCTCGGTGACTGGCTGGGAACCCGCAAAGCGTACGTTTGTAGCCTGCTG
      G  G  F  L  G  D  W  L  G  T  R  K  A  Y  V  C  S  L  L 349

1083 GCCTCGCAGCTGCTGATTATTCCGGTATTTGCGATTGGCGGCGCAAACGTCTGGGTG
      A  S  Q  L  L  I  I  P  V  F  A  I  G  G  A  N  V  W  V 368

1140 CTCGGTCTGTTACTGTTCTTCCAGCAAATGCTTGGACAAGGGATCGCCGGGATCTTA
      L  G  L  L  L  F  F  Q  Q  M  L  G  Q  G  I  A  G  I  L 387

1197 CCAAAACTGATTGCCGGTTATTTCGATACCGACCAGCGTGCAGCGGGCCTGGGCTTT
      P  K  L  I  G  G  Y  F  D  T  D  Q  R  A  A  G  L  G  F 406

1254 ACCTACAACGTTGGCGCATTGGGCGGTGCACTGGCCCCAATCATCGGCGCGTTGATC
      T  Y  N  V  G  A  L  G  G  A  L  A  P  I  I  G  A  L  I 425

1311 GCTCAACGTCTGGATCTGGGTACTGCGCTGGCATCGCTCTCGTTCAGTCTGACGTTC
      A  Q  R  L  D  L  G  T  A  L  A  S  L  S  F  S  L  T  F 444

1368 GTGGTGATCCTGCTGATTGGGCTGGATATGCCTTCTCGCGTTCAGCGTTGGTTGCGC
      V  V  I  L  L  I  G  L  D  M  P  S  R  V  Q  R  W  L  R 463

1425 CCGGAAGCGTTGCGTACTCATGACGCTATCGACGGTAAACCATTCAGCGGTGCCGTG
      P  E  A  L  R  T  H  D  A  I  D  G  K  P  F  S  G  A  V 482

1482 CCGTTTGGCAGCGCCAAAAACGATTTAGTCAAAACCAAAAGTTAATCCTGT
      P  F  G  S  A  K  N  D  L  V  K  T  K  S  STOP          496
```

FIG. 2. Nucleotide sequence of *nanT* and derived primary structure of sialic acid-permease. The nucleotide sequence of *nanT* is indicated by the numbers on the left, while the numbers on the right indicate the primary structure of NanT, beginning with the predicted N-terminal methionine residue. One-letter amino acid designations are used. The underlined GGAG sequence indicates the potential ribosome-binding site (35), while the other underlined region indicates the *Eco*RI site used to truncate *nanT* in the pSX601 subclone described in the text.

of formal proof that *nanT* is sufficient for permease activity. To provide direct evidence that *nanT* is required for sialic acid uptake, an alkaline phosphatase-positive Tn*phoA* fusion was isolated in pSX600 by standard procedures (21) and mapped to *nanT* by digestions with *Eco*RI and *Hin*dIII (data not shown). This plasmid did not restore growth of EV282 on sialic acid as

TABLE 1. Complementation of growth defects in mutants deficient in sialic acid-aldolase or -permease activities

Strain (relevant genotype)	Plasmid	Growth on sialic acid[a] (A_{600})	Sialic acid uptake (nmol/min/A_{600})
EV50[b] (*nanA4*)	None	NG[c]	2.04
EV50	pSX600	0.92	5.12
EV55[b] (*nanT1*)	None	NG	0.02
EV55	pSX600	0.40	0.78
EV282[d] (*nanA4 nanT2*)	None	NG	0.05
EV282	pSX600	0.65	4.97

[a] Cells lacking plasmid were grown in 0.4% (wt/vol) glycerol minimal M63 medium (36) prior to performing uptake assays as previously described (40). To test for growth on sialic acid, cells grown in minimal glycerol were diluted 1/50 into M63 containing 1 mg of sialic acid per ml plus appropriate drug supplementation. Cells containing plasmids and grown on sialic acid were used for the relevant uptake assays. The uptake results given are the averages of duplicate samples wherein the difference between replicates was less than 15%.

[b] Chemically induced aldolase (*nanA4*)- and permease (*nanT1*)-deficient mutants (40).

[c] NG, no growth.

[d] Spontaneous permease-deficient mutant selected for growth of EV50 on glycerol in the presence of sialic acid (40).

the sole carbon source. We conclude that pSX600 encodes aldolase and also probably permease in *trans* but that definitive evidence that *nanT* is both necessary and sufficient for sialic acid uptake will require purification or overproduction of the permease, as has been accomplished recently, for example, for FucP (9).

The derived primary structure of NanT (Fig. 2) predicts a hydrophobic polypeptide with 14 transmembrane segments (Fig. 3). By comparing the structure of NanT to the consensus primary structure of the large family of sugar transporters (2), a version (GAMGDRYGRR) of the diagnostic GXXDR/KXGR/K motif (10) was detected between putative spanners 2 and 3; a less conserved version (GFLGDWLGTR) of this motif was detected in the carboxy-terminal half of NanT, consistent with the possible duplication of a primordial transporter gene that coded for six membrane-spanning regions (33). Other features of NanT conserved in the sugar transporter family are an imperfect version of the 42-residue diffused motif through spanners 4 and 5 [R---G---G(A)------P(A)-Y--E(-)-----RG(N)------Q(I)-----G, where deviations from this consensus in NanT are indicated by parentheses], the E------R motif between putative spanners 4 and 5, and the beginning (F----G), in NanT, of a diffused motif through spanners 12 and 13 (10a). A diffused motif in spanner 7 of monosaccharide transporters, L---Q--GIN--FYY, as well as in lactose, melibiose, citrate, and tetracycline transporters (10a), is missing in NanT putative spanner 9, which is predicted to correspond to spanner 7 in sugar transporters with the 6 + 6 topology. The absence of this motif in NanT may be significant because the glutamine and asparagine residues are thought to function in substrate recognition (10a), implying that NanT may use different residues for substrate interactions.

In contrast to the structural similarities between NanT and other sugar transporters noted above, the region between putative spanners 6 and 9 is unlike that of previously described permeases, in which segments located between amino- and carboxy-terminal halves are hydrophilic and exposed in the cytoplasm (2, 3, 10, 10a). Computer-assisted analysis indicates that residues 208 to 228 of NanT have an amphiphilic (average score of 0.56) α-helical character. Helical cylinder projection of this segment indicates that eight hydrophobic residues would be located on one side of an α-helix, similar to the known membrane-interacting helices of apolipoproteins and other amphipathic peptides (13). Positively charged residues would be clustered and separated from negatively charged residues on the side opposite the hydrophobic residues. The putative helix in NanT is of further interest because in one class of sialorecogniton molecules investigated at the tertiary structural level, the sialidases, three active site arginine residues critically interact with the carbon-1 carboxylate of sialic acid (38). Clustering of positively charged residues in the potential amphiphilic region of NanT may likewise indicate the importance of these residues for substrate interactions, especially since the helix is in juxtaposition to putative spanners 7 and 8, which are not found in other sugar transporters. As described above, the anionic sugar permease for sialic acids in the eukaryotic lysosomal membrane transports some of the same substrates as NanT and is inhibited in a substrate-protectable manner by chemical modification with the arginine-modifying reagent phenylglyoxal (20). Given the similar predicted secondary and presumed similar tertiary structures of eukaryotic and prokaryotic sugar transporters (2), it would not be surprising if NanT and the lysosomal sialic acid transporter share common primary structural elements, especially those positively charged residues available for interactions with anionic substrates. Although some progress has been made in reconstituting an active lysosomal transporter (19, 20), the possibility of straightforward genetic techniques in *E. coli* or *Salmonella typhimurium* could provide molecular insight into the mammalian sialic acid transport process as well. Since our results make it likely that an active sialic acid transporter can be produced in *trans*, it should be possible to directly test the

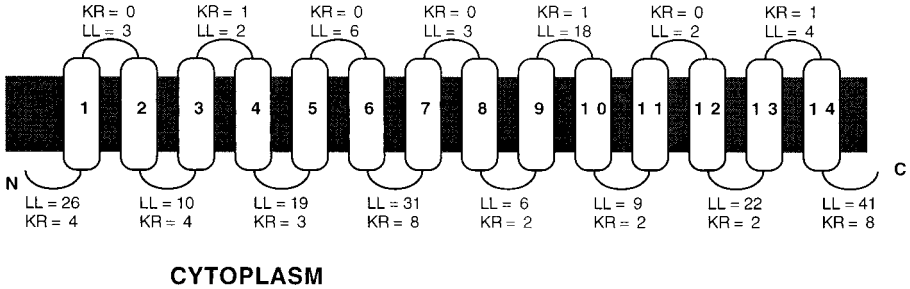

FIG. 3. Topological model of NanT. The algorithm of Kyte and Doolittle (17a) and the positive-inside rule of von Heijne (42) were used to generate the model with the Heidelberg, Germany, suite of programs available from the Internet. KR refers to the number of lysine and arginine residues, and LL refers to the loop length between putative membrane-spanning segments, numbered 1 through 14. Predicted spanners 1, 2, and 4 through 14 were judged as certain, with composite hydropathy scores greater than 1.2, while spanner 3 with a score of 1.1229 was considered putative.

```
NanT MSTTTQNIPWYRHLNRAQWRAFSAAWLGYLLDGFDFVLIALVLTEVQGEFGLTTVQAASLISAAF  65
      ++T     WY+ +N  Q +A  +AWLGY+ DGFDF++I  +L  ++ + G+T +QA  + + AF
O425 TQETNMATAWYKQVNPPQRKALFSAWLGYVFDGFDFMMIFYILHIIKADLGITDIQATLIGTVAF  80

NanT ISRWFGGLMLGAMGDRYGRRLAMVTSIVLFSAGTLACGFAPGYITMFIARLVIGMGMAGEYGSSA 130
     I+R  GG   GAM D+YGR+  M+ +I ++S GT   G A     + + R ++G+GM+GEY  ++
O425 IARPIGGGFFGAMADKYGRKPMMMWAIFIYSVGTGLSGIATNLYMLAVCRFIVGLGMSGEYACAS 145

NanT TYVIESWPKHLRNKASGFLISGFSVGAVVAAQVYSLVVPVWGWRALFPIGULPIIFALWLRKNIP 195
     TY +ESWPK+L++KAS FL+SGFSVG ++AAQ          V+GWR  FPIG+LP++  LW+RK+ P
O425 TYAVESWPKNLQSKASAFLVSGFSVGNIIAAQIIPQFAEVYGWRNSFFIGLLPVLLVLWIRKSAP 210

NanT EAEDWKEKHAGKAPVRTMVDILYRGEHRIANIVMTLAAATALWFCFAGNLQNAAIVAVLGLLCAA 260
     E+++W E++          V+   + +
O425 ESQEWIEQKYKDKSTFLSVERKPHLSIS---------------------------------- 238

NanT IFISFMVQSAGKRWPTGVMLMVVVLFAFLYSWPIQALLPTYLKTDLAYNPHTVANVLFFSGFGAA 314
                              +V    F  +WPI  LLP+YL  D   N   ++ ++  +G G
O425 ----------------MIVFLVCFCLFGANWPINGLLPSYLA-DNGVNTVVISTLMTIAGLGTL 274

NanT VGCCVGGFLGDWLGTRKAYVCSLLASQLLIIPVFAIGGANVWVLGLLLFFQQMLGQGIAGILPKL 379
      G    GF+GD +G +KA+V  L+ S + + P+F I   N  ++GL LF       GIAG++PK
O425 TGTIFFGFVGDKIGVKKAFVVGLITSFIPLCPLFFISVKNSSLIGLCLFGLMFTNLGIAGLVPKF 339

NanT IGGYFDTDQRAAGLGFTYNVGALGGALAPIIGALIAQRLDLGTALASLSFSLTFVVILLIGLDMP 444
     I  YF T  R  G G  YN+GA GG  AP++   I+     LG +L  ++ + + ++ILL+G D+P
O425 IYDYFPTKLRGLGTGLIYNLGATGGMAAPVLATYISGYYGLGVSLFIVTVAFSALLILLVGFDIP 404

NanT SRV 447
      ++
O425 GKI 407
```

FIG. 4. Homology between NanT and the hypothetical transporter encoded by ORF425. Homologous regions between NanT and the putative ORF425 permease were first aligned by using the BLAST program (1), and then gaps (hyphens) were inserted by eye to give the contiguous alignment shown in the figure. Identities are indicated in the center line. +, conservative changes identified by the program.

importance of positively charged residues in the amphiphilic segment by site-directed mutagenesis and that of other residues by random or directed mutagenesis approaches.

Two recent analyses of the natural relationships between eukaryotic and prokaryotic transporters are consistent with a monophyletic origin (8, 22), indicating that these proteins are probably evolutionarily as well as functionally and structurally related. Among the highest pairwise BLAST (1) alignment scores were 216 and 218 between NanT and the putative hydroxy benzoic acid transporters from *Acinetobacter calcoaceticus* (accession no. L05770) and PcaK of *Pseudomonas putida* (accession no. U10895), respectively. Most amino acid identities are in the amino-terminal region containing candidate spanners 1 to 6, which is the most conserved region among known sugar transporters (2, 3, 10a). Less striking though still potentially significant similarity scores were also detected between NanT and *E. coli* proline-betaine transporter (accession no. P30848) and mammalian synaptic vesicle neurotransmitters (accession no. A43344, S27262, and L05435). These results suggest that *nanT* is more closely related to several systems that have nothing to do with sugar transport than to the monosaccharide transporter subfamily. In contrast, the high score of 608, corresponding to 37% identity and 56% similarity when conservative changes were taken into account, was between NanT and the derived primary structure of ORF425 (Fig. 4), a functionally unidentified and unpublished sequence mapping between *leuX* and *fecE* and deposited in the database as part of the complete sequencing of the *E. coli* chromosome (25). The similarity score for the alignment with NanT has a Poisson probability of 1.1×10^{-87}, indicating that *nanT* and ORF425 would code for ancestrally related (homologous) polypeptides. While the function, if any, of ORF425 is unknown, our results indicate that it may code for a permease. However, unlike NanT, the 46.6-kDa polypeptide predicted by ORF425 would lack the central domain, containing instead a hydrophilic structure between putative spanners 6 and 7, reminiscent of the equivalent regions of known hexose transporters (Fig. 4). The unique central domain of NanT is, therefore, likely to function in substrate recognition or some other (unknown) assembly or regulatory property of the permease. Future experiments will test the hypothesis that this region functions in sialic acid recognition.

The sequence identified in this communication has been assigned accession no. U19539. The derived structure of NanT has been assigned accession no. P41036.

This work was supported by NIH research grant AI23039 (E.V.). J.M. was supported in part by a Howard Hughes Undergraduate Research Award, a Colgate-Palmolive Undergraduate Research Award, and an American Society for Microbiology Undergraduate Research Fellowship.

We thank Guy Plunkett III of the University of Wisconsin *E. coli* Genome Sequencing Project (Frederick Blattner, director) for sharing information with us in advance of publication. Special thanks go to Alejandro Martinez for helping us with the protein analysis software package.

REFERENCES

1. **Altschull, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
2. **Baldwin, S. A.** 1993. Mammalian passive glucose transporters: members of a ubiquitous family of active and passive transport proteins. Biochim. Biophys. Acta **1154:**17–49.
3. **Baldwin, S. A., and P. J. F. Henderson.** 1989. Homologies between sugar transporters from eukaryotes to prokaryotes. Annu. Rev. Physiol. **51:**459–471.
4. **Corfield, T.** 1992. Bacterial sialidases—roles in pathogenicity and nutrition. Glycobiology **2:**509–521.
5. **DeNinno, M. P.** 1991. The synthesis and glycosidation of *N*-acetylneuraminic acid. Synthesis **1:**583–593.
6. **Drueckhammer, D. G., W. J. Hennen, R. L. Pederson, C. F. Barbas III, C. M. Gautheron, T. Krach, and C.-H. Wong.** 1991. Enzyme catalysis in synthetic carbohydrate chemistry. Synthesis **1:**499–525.
7. **Godoy, V. G., M. M. Dallas, T. A. Russo, and M. H. Malamy.** 1993. A role for *Bacteroides fragilis* neuraminidase in bacterial growth in two model systems. Infect. Immun. **61:**4415–4426.
8. **Griffith, J. K., M. E. Baker, D. A. Rouch, M. G. P. Page, R. A. Skurray, I. T. Paulsen, K. F. Chater, S. A. Baldwin, and P. J. F. Henderson.** 1992. Membrane transport proteins: implications of sequence comparisons. Curr. Opin. Cell Biol. **4:**684–695.
9. **Gunn, F. J., C. G. Tate, and P. J. F. Henderson.** 1994. Identification of a novel sugar-H^+ symport protein, FucP, for transport of L-fucose into *Escherichia coli*. Mol. Microbiol. **12:**799–809.
10. **Henderson, P. J. F.** 1993. The 12-transmembrane helix transporters. Curr. Opin. Cell Biol. **5:**708–721.

10a. **Henderson, P. J. F., S. A. Baldwin, M. T. Cairns, B. M. Charalambous, H. C. Dent, F. Gunn, W. Liang, V. A. Lucas, G. E. Martin, T. P. McDonald, B. J. McKeown, J. A. R. Muiry, K. R. Petro, P. E. Roberts, K. P. Shatwell, G. Smith, and C. G. Tate.** 1992. Sugar-cation symport systems in bacteria. Int. Rev. Cytol. **137A:**149–208.

11. **Hoyer, L. L., A. C. Hamilton, S. M. Steenbergen, and E. R. Vimr.** 1992. Cloning, sequencing, and distribution of the *Salmonella typhimurium* LT2 sialidase gene, *nanH*, provide evidence for interspecies gene transfer. Mol. Microbiol. **6:**873–884.
12. **Izard, T., M. C. Lawrence, R. L. Malby, G. G. Lilley, and P. M. Colman.** 1994. The three-dimensional structure of N-acetylneuraminate lyase from Escherichia coli. Structure **2:**361–369.
13. **Kaiser, E. T.** 1987. Design of amphiphilic peptides, p. 193–199. *In* D. L. Oxender and C. F. Fox (ed.), Protein engineering. Alan R. Liss, New York.
14. **Kawakami, B., T. Kudo, Y. Narahashi, and K. Horikoshi.** 1986. Genetic and molecular analyses of *Escherichia coli N*-acetylneuraminate lyase gene. J. Bacteriol. **167:**404–406.
15. **Kawakami, B., T. Kudo, Y. Narahashi, and K. Horikoshi.** 1986. Nucleotide sequence of the N-acetylneuraminate lyase gene of *Escherichia coli*. Agric. Biol. Chem. **50:**2155–2158.
16. **Kim, M.-J., W. J. Hennen, H. M. Sweers, and C.-H. Wong.** 1988. Enzymes in carbohydrate synthesis: N-acetylneuraminic acid aldolase catalyzed reactions and preparation of N-acetyl-2-deoxy-D-neuraminic acid derivatives. J. Am. Chem. Soc. **110:**6481–6486.
17. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole E. coli chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50:**495–508.

17a. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.

18. **Lilley, G. G., M. von Itzstein, and N. Ivancic.** 1992. High-level production and purification of *Escherichia coli N*-acetylneuraminic acid aldolase. Protein Expr. Purif. **3:**434–440.
19. **Mancini, G. M. S., C. E. M. T. Beerens, H. Galjaard, and F. W. Verheijen.** 1992. Functional reconstitution of the lysosomal sialic acid carrier into proteolipsomes. Proc. Natl. Acad. Sci. USA **89:**6609–6613.

20. **Mancini, G. M. S., H. R. deJonge, H. Galjaard, and F. W. Verheijen.** 1989. Characterization of a proton-driven carrier for sialic acid in the lysosomal membrane. Evidence for a group-specific transport system for acidic monosaccharides. J. Biol. Chem. **264:**15247–15254.
21. **Manoil, C., and J. Beckwith.** 1985. Tn*phoA*: a transposon probe for protein export signals. Proc. Natl. Acad. Sci. USA **82:**8129–8133.
22. **Marger, M. D., and M. H. Saier, Jr.** 1993. A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. Trends Biochem. Sci. **18:**13–18.
23. **Müller, H. E.** 1974. Neuraminidases of bacteria and protozoa and their pathogenetic role. Behring Inst. Mitt. **55:**34–56.
24. **Nees, S., and R. Schauer.** 1974. Induction of neuraminidase from *Clostridium perfringens* and the correlation of this enzyme with acylneuraminate pyruvate-lyase. Behring Inst. Mitt. **55:**68–78.
25. **Nowak, R.** 1995. Getting the bugs worked out. Science **267:**172–174.
26. **Ohta, Y., M. Shimosaka, K. Murata, Y. Tsukada, and A. Kimura.** 1986. Molecular cloning of the N-acetylneuraminate lyase gene in *Escherichia coli* K-12. Appl. Microbiol. Biotechnol. **24:**386–391.
27. **Ohta, Y., K. Watanabe, and A. Kimura.** 1985. Complete nucleotide sequence of the *E. coli* N-acetylneuraminate lyase. Nucleic Acids Res. **13:**8843–8851.
28. **Petter, J. G., and E. R. Vimr.** 1993. Complete nucleotide sequence of the bacteriophage K1F tail gene encoding endo-*N*-acylneuraminidase (endo-*N*) and comparison to an endo-*N* homolog in bacteriophage PK1E. J. Bacteriol. **175:**4354–4363.
29. **Renlund, M., F. Tietze, and W. A. Gahl.** 1986. Defective sialic acid egress from isolated fibroblast lysosomes of patients with Salla disease. Science **232:**759–762.
30. **Richards, E., M. Reichardt, and S. Rogers.** 1992. Preparation of Genomic DNA from bacteria, p. 2.4.1–2.4.5. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology, vol. 1. Wiley & Sons, New York.
31. **Rodriguez-Aparicio, L. B., A. Reglero, and J. M. Luengo.** 1987. Uptake of N-acetylneuraminic acid by *Escherichia coli* K-235: biochemical characterization of the transport system. Biochem. J. **246:**287–294.
32. **Rudd, K. E.** 1992. Alignment of *E. coli* DNA sequences to a revised, integrated genomic restriction map. p. 2.3–2.43. *In* J. H. Miller (ed.), A short course in bacterial genetics: a laboratory manual and handbook for *Escherichia coli* and related bacteria. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
33. **Saier, M. H.** 1993. Convergence and divergence in the evolution of transport proteins. Bioessays **16:**23–29.
34. **Schauer, R.** 1985. Sialic acids and their role as biological masks. Trends Biochem. Sci. **10:**357–360.
35. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71:**1342–1346.
36. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusion. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
37. **Vimr, E. R.** 1992. Selective synthesis and labelling of the polysialic acid capsule in *Escherichia coli* K1 strains with mutations in *nanA* and *neuB*. J. Bacteriol. **174:**6191–6197.
38. **Vimr, E. R.** 1994. Microbial sialidases: does bigger always mean better? Trends Microbiol. **2:**271–277.
39. **Vimr, E. R., W. Aaronson, and R. P. Silver.** 1989. Genetic analysis of chromosomal mutations in the polysialic acid gene cluster of *Escherichia coli* K1. J. Bacteriol. **171:**1106–1117.
40. **Vimr, E. R., and F. A. Troy.** 1985. Identification of an inducible catabolic system for sialic acids (*nan*) in *Escherichia coli*. J. Bacteriol. **164:**845–853.
41. **Vimr, E. R., and F. A. Troy.** 1985. Regulation of sialic acid metabolism in *Escherichia coli*: role of *N*-acylneuraminate pyruvate-lyase. J. Bacteriol. **164:** 854–860.
42. **von Heijne, G.** 1992. Membrane protein structure prediction: hydrophobicity analysis and the positive-inside rule. J. Mol. Biol. **225:**487–494.
43. **von Itzstein, M., W. Y. Wu, G. B. Kok, M. S. Pegg, J. C. Dyason, B. Jin, T. V. Phan, M. L. Smythe, H. F. White, S. W. Oliver, P. M. Colman, J. N. Varghese, D. M. Ryan, J. M. Woods, R. C. Bethell, V. J. Hotham, J. M. Cameron, and C. R. Penn.** 1993. Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature (London) **363:**418–423.
44. **Warren, L.** 1963. The distribution of sialic acids in nature. Comp. Biochem. Physiol. **10:**153–171.