Protein Science (1992), *1*, 1326–1332. Cambridge University Press. Printed in the USA.

A new subfamily of bacterial ABC-type transport systems catalyzing export of drugs and carbohydrates

JONATHAN REIZER, AIALA REIZER, AND MILTON H. SAIER, JR.
Department of Biology, University of California at San Diego, La Jolla, California 92093-0116
(RECEIVED March 3, 1992; REVISED MANUSCRIPT RECEIVED May 13, 1992)

Abstract

Sequence comparison studies revealed that the drug resistance transporter of *Streptomyces peucetius* (DrrAB) and two nodulation gene products (NodIJ) of *Rhizobium leguminosarum* are homologous to proteins encoded by three sets of genes that comprise capsular polysaccharide export systems in gram-negative bacteria: KpsTM of *Escherichia coli*, BexABC of *Haemophilus influenzae*, and CtrDCB of *Neisseria meningitidis*. These five systems comprise a new subfamily within the family of ATP binding cassette (ABC)-type transporters. We have termed this subfamily the ABC-2 subfamily. For three of the systems comprising this subfamily (Drr, Nod, and Kps) only one integral membrane constituent has been identified, whereas for the other two systems (Bex and Ctr) two dissimilar integral membrane constituents have been found. This observation suggests that the transmembrane channels of ABC-2-type transporters can be formed of homo- or heterooligomers as is true of several other classes of transport systems.

Keywords: ABC-type permeases; ATP hydrolysis; drug resistance; evolution; nodulation; oligomeric channels; phylogenetic tree; polysaccharide export; solute transport

In a recent communication, Guilfoile and Hutchinson (1991) presented the sequences of two genes (*drrA* and *drrB*) that apparently encode a two-component drug resistance transport system from *Streptomyces peucetius*. One of the two proteins (DrrA) was found to be similar to the family of ATP binding proteins that energize solute transport via the ABC (ATP binding cassette) family of permeases (Furlong, 1987; Higgins et al., 1990). The other protein, DrrB, was hydrophobic with six putative transmembrane α-helical segments and was reported to show no significant sequence similarity to other known transport proteins. These two proteins were proposed to function together in the export of daunorubicin and doxorubicin, both of which are produced by *S. peucetius*, and induction of the syntheses of these transport proteins was coordinate with induction of the syntheses of the drug biosynthetic enzymes. A parallel between bacterial drug resistance and mammalian multidrug resistance was proposed (Guilfoile & Hutchinson, 1991). However, no quantitative sequence comparison data were provided.

Reprint requests to: Milton H. Saier, Jr., Department of Biology, 0116, University of California at San Diego, La Jolla, California 92093-0116.

In a recent publication we summarized evidence for the structural and evolutionary relatedness of the integral membrane constituents of several different types of transport systems (Saier & Reizer, 1991). These transporters include (1) a large and diverse class of facilitators that catalyze solute uniport, solute:cation symport, and solute:solute antiport; (2) the ABC-type active transport systems that are energized by ATP hydrolysis; (3) the group translocating sugar transporters of the bacterial phosphotransferase system (PTS), which are energized by phosphoryl transfer from phosphoenolpyruvate to the sugar substrate; and (4) voltage-sensitive ion channels of nerve and muscle cells. All of these transport systems appear to consist of units of six transmembrane α-helical segments although exceptions have been noted (Kerppola & Ames, 1992). In all cases examined to date, at least two such units must be present for construction of a functional permease, but in a few cases only one such unit has been identified suggesting that in these cases, a single transmembrane protomer forms the oligomeric channel. There is precedence for this proposal because the PTS permeases, some voltage-sensitive ion channels, and some facilitators consist of homooligomers, and thus, the transmembrane units in each of these permeases are identical (Saier & Reizer, 1991).

In the present communication we report analyses of the DrrA and DrrB proteins. DrrB was found to exhibit statistically significant sequence similarity to integral membrane proteins of three other transport systems, all of which appear to be energized in processes involving ATP binding proteins resembling those of other ABC-type permeases. These integral membrane proteins are also similar in sequence to the NodJ protein, which in conjunction with the NodI protein most likely functions in oligosaccharide export in *Rhizobium leguminosarum* (S.R. Long, pers. comm., and the reported analyses). Two of the homologous transport systems (BexABC and CtrDCB) may function with the participation of two distinct integral membrane proteins (BexB plus BexC and CtrC plus CtrB, respectively). However, each of the two remaining systems (KpsTM and NodIJ) resembles the DrrAB system in that a gene encoding a second transmembrane constituent of the system has not been discovered. Because these latter transporters apparently have a single six-unit transmembrane transport protein constituent, they resemble homooligomeric facilitators, ion channels, and PTS permeases.

Computer searches revealed that no other proteins in the database—including those comprising facilitators, PTS group translocators, ion channels, or other ABC-type permeases—were demonstrably homologous to these five proteins. Additionally, the DrrA protein is homologous to all four of the ATP binding constituents that presumably energize transport via the four homologous integral membrane transport proteins mentioned above. However, DrrA is also homologous to many other ATP binding proteins including those that energize other ABC-type permeases. These observations suggest that DrrAB and its four homologous systems (BexABC, CtrDCB, KpsTM, and NodIJ) comprise a novel subclass of permeases that are energized by ATP employing a mechanism analogous to that utilized by other well-characterized ABC-type permeases. We have termed this subfamily the ABC-2 family. The evolutionary implications of our quantitative sequence comparisons will be discussed.

Results

Figure 1 presents a multiple alignment of the five homologous integral membrane proteins that are constituents of the ABC-2 family of transport systems. The consensus sequence is provided below the multiple alignment, and residues conserved in all five proteins are indicated by an exclamation mark above the multiple alignment. In addition to the DrrB protein of *S. peucetius* (Sp) (Guilfoile & Hutchinson, 1991), these proteins include the BexB protein of the polyribosylribitol phosphate capsular polysaccharide export system of *Haemophilus influenzae* (Hi) (Kroll et al., 1988, 1990; Kroll & Moxon, 1990), the CtrC protein of the polyneuraminic acid capsular polysaccharide export system of *Neisseria meningitidis* (Nm) (Frosch et al., 1991), the KpsM protein of the polyneuraminic acid capsular polysaccharide export system of *Escherichia coli* (Ec) (Smith et al., 1990; Pavelka et al., 1991), and the NodJ protein of *R. leguminosarum* (Rl) (Evans & Downie, 1986).

Table 1 summarizes the binary comparison scores (in standard deviations) for these five proteins. The KpsM(Ec), BexB(Hi), and CtrC(Nm) proteins are all clearly homologous as reported previously (comparison scores of 18–67 SD). DrrB(Sp) and NodJ(Rl) gave comparison scores of 8 SD and 10 SD, respectively, with KpsM(Ec). The former value is clearly suggestive of homology (a common ancestry), whereas the latter value establishes homology as discussed previously (Doolittle, 1981, 1986). No other protein in the database gave a comparison score with any one of the five homologous proteins of the ABC-2 family, which was suggestive of homology.

The phylogenetic tree for these proteins is shown in Figure 2. CtrC(Nm) and BexB(Hi) are very closely related, as suggested by the comparison score reported in Table 1. KpsM(Ec) is substantially more distant from these two, but DrrB(Sp) and NodJ(Rl) are out on a separate rather distant branch of the tree. The distances shown in this tree are in full agreement with the binary comparison scores reported in Table 1.

Each of the Bex and Ctr capsular polysaccharide export systems appears to possess a second integral inner membrane transport protein. These proteins, designated BexC and CtrB, respectively, were significantly similar to each other but not to any other protein in the data base (data not shown). The C-terminal hydrophobic domain of NodC, which might conceivably function in transport

Table 1. *Binary comparisons of the alignments of the homologous integral membrane constituents of the ABC-2 subfamily of transport systems*

	BexB(Hi) (265)[a]	DrrB(Sp) (283)	CtrC(Nm) (265)	NodJ(Rl) (259)
KpsM(Ec) (258)	27 (247)[b] [22]	23 (248) [8]	25 (246) [18]	25 (44) [10]
BexB(Hi) (265)		18 (256) [4]	70 (261) [67]	40 (40) [9]
Drr(Sp) (283)			22 (37) [3]	29 (45) [7]
CtrC(Nm) (265)				26 (152) [8]

[a] The values in parentheses below the protein designations indicate the numbers of amino acids in the proteins.

[b] Values reported in the table correspond to the percent identities in the segments compared. The number of amino acids in the aligned segment is provided in parentheses. Values in brackets denote the comparison scores in standard deviations higher than those obtained with 100 comparisons of randomized sequences of these protein segments. The FASTA and RDF2 programs (Pearson & Lipman, 1988) were used to assess the similarity scores and their significance. The abbreviations used are defined in the text.

```
                                           !
BexB(Hi)     1   MQYGDQTTFKQSLAIQGRVINALLMREIITRYGRKNIGFLW           LFVEPLLMTFFIVMMWK
CtrC(Nm)     1   MKALHKTSFWESLAIQRRVIGALLMREIITRYGRNNIGFLW           LFVEPLLMTFVIVLMWK
KpsM(Ec)     1      MARSGF    EVQKVTVEALFLREIRTRFGKFRLGYLW           AILEPSAHLLILLGIFG
DrrB(Sp)     1  MTTSPGTVESTTPVSGQLRTVLSAGERPARATAVSATLTHLWRAMMAFKHFPVQLIDIVLMPLIFLLMFT
NodJ(Rl)     1      MGVAT   LPAGGLNWLAVWRRNYLAW KKAALASIL          GNLADPVIYLFGLGAGLG
Consensus       .........F...L..Q.R...AL..REI.TR.G...LG.LW..............EP.LM.F....M..

                                 !   !                                      !
BexB(Hi)    59  FIRADKF  STLNMIAFVMTGYPM  AMMWRNASNRAIGSISANLSLLYHRNVRVLD  TIFTRVLLEVA
CtrC(Nm)    58  FLKADRY  STLNIVAFAITGYPM  LMMWRNASKRAVGSISSNASLLYHRNVRVLD  TILARMILEIA
KpsM(Ec)    52  YIMHRTM  PDISFPVFLLNGLIP  FFIFSSISNRSVGAIEANQGLFNYRPVKPID  TIIARALLETL
DrrB(Sp)    71  YLFGGAFADSTEEYLQFYLPGVTVQAVVMMTVYTGTSLNTDIHKGVFDRFRTLPFWQPATLAGSLLGDVL
NodJ(Rl)    52  VMVGRV   DGVSYTAFLAAGMIA  TSAMTAATFETIYAAFGRMQGQRTWEAMLYTQLTLGDIVVGEMA
Consensus       .........ST....AF...G.......M...AS.R..G.I..N..L...R.V...D..TI..R.LLE.A

BexB(Hi)   123  GASIAQILFMAVLVLIGWIDAPRDVFYMLMAWFLMAMFAFALGLIICAVAQQFDVFGKIWG   TLSFVL
CtrC(Nm)   123  GATIAQIVIMAVLIAIGWIEMPADMFYMLMAWLLMAFFAIGLGLVICSIAFNFEPFGKIWG   TLTFVM
KpsM(Ec)   116  IYVAVYILLMLIVWMAGEYFEITNFLQLVLTWSLLIIILSCGIGLIFMVVGKTFPEMQKVLP  ILLKPL
DrrB(Sp)   141  RYVVALATTVSLGLLLG FRADGGFLGVVGAMLVLIVFGFSVSWIFAALGVVASEPERVSGTSMIVLYPL
NodJ(Rl)   117  WAATKASLAGTGIGIVAAMLGYTHWLALLYALPVIAITGLAFASLGMVVTALAPSYDYFIFYQTLVITPM
Consensus       .A..A.IL.M......G........L..L.AW.L.A.F....GLI...V...F....K..G....L..PL

                   !         !          !  !     !
BexB(Hi)   190  LPISGAFFFVHNLPSQAQSIALWLPMIHGTEMFRHGYFGDTVVT YESIGFLVVSDLALLLMGLVMVKNF
CtrC(Nm)   190  MPLSGAFFFVHNLPPKVQEYALMIPMVHGTEMFRAGYFGSDVIT YENPWYIVLCNLVLLLFGLAMVSKF
KpsM(Ec)   183  YFISCIMFPLHSIPKQYWSYLLWNPLVHVVELSREAVMPGYISE GVSLNYLAMFTLVTLFIGLALYRTR
DrrB(Sp)   210  LFMSNIFVMPETMPGWMQAIVDANPMSHAATASRELMHGTAGFW DVGLVLCVSAGLVAVFAPLTM  RL
NodJ(Rl)   187  LFLSGAVFPVDQLPVAFQQIAAFLPLAHSIDLIRPTMLGQPIANVCLHIGVLCIYIVVPFLVSTALLRR
Consensus       LF.SGAFF.VH.LP...Q.IAL..PM.H..E..R....G............LV...LV.LL.GLAM....

BexB(Hi)   259  SKGIEPQ
CtrC(Nm)   259  SKGVEPQ
KpsM(Ec)   252  EEAMLTS
DrrB(Sp)   277  YRNKNAH
NodJ(Rl)   256   RLMR
Consensus       .......
```

Fig. 1. Multiple alignment of the sequences of the five homologous integral membrane proteins of the ABC-2 family. Residues conserved in all five proteins are indicated by an exclamation mark above the multiple alignment. The consensus sequence (consensus) is provided below the multiple alignment. Numbers to the left of the sequences provide the positions in the proteins of the first residues in each line. The abbreviations of the proteins and species are as defined in the text.

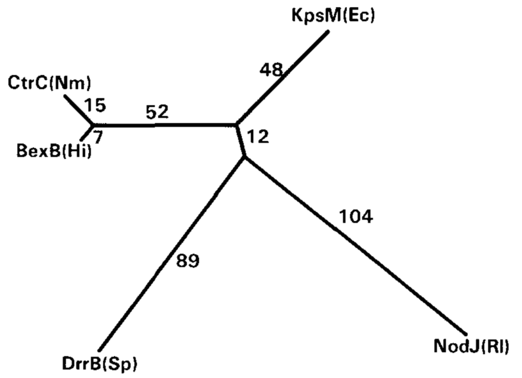

Fig. 2. Phylogenetic tree of the homologous integral membrane proteins of the ABC-2 subfamily of transport systems. The programs used for tree construction, abbreviations, and references are provided in the text.

(Long, 1989), was not homologous to BexC, CtrB, or any other protein in the database.

Multiple alignment of limited portions of the sequences of several homologous ATP binding proteins are presented in Figure 3. The proteins included in this depiction are the BexA(Hi), CtrD(Nm), KpsT(Ec), DrrA(Sp), and NodI(Rl) proteins of the ABC-2 subfamily of transport systems as well as the CysA protein of *E. coli*, the ATP binding constituent of the sulfate–thiosulfate permease of *E. coli* (Sirko et al., 1990), the ORF60-5 protein from the cyanobacterium, *Synechocystis* sp. PCC6803 (Ssp), encoded by a gene adjacent to one encoding a chaperonin protein (Chitnis & Nelson, 1991), and the NosF protein of *Pseudomonas stutzeri* (Ps), a component of the Cu^{2+} processing system (possibly a Cu^{2+} transport system, which, interestingly, may possess only one integral mem-

```
                                                  !             !     !   ! !   !                           !
BexA(Hi)       16   G W K T V L K N I N F E L Q K G E K I G I L G R N G A G K S T L I R L M S G V E P P T S G T I E
CtrD(Nm)       16   G M R T V L D D I N F S L Q K G E K V G I L G R N G A G K S T L V R L I S G V E P P T S G E I K
KpsT(Ec)       16   G R H Y V F K D L N I E I P S G K S V A F I G R N G A G K S T L L R M I G G I D R P D S G K I I
DrrA(Sp)       19   N G T R A V D G L D L N V P A G L V Y G I L G P N G A G K S T T I R M L A T L L R P D G G T A R
NodI(Rl)       23   G G K I V V N D L S F T I A A G E C F G L L G P N G A G K S T I T R M I L G M T S P S V G K I T
CysA(Ec)       13   G R T Q V L N D I S L D I P S G Q M V A L L G P S G S G K T T L L R I I A G L E H Q T S G H I R
ORF60-5(Ssp)   65   Q K V E S L K P L S I A V Q Q G E T F G L L G P N G A G K T T L L K I L L G V I R A S G G R G T
NosF(Ps)       14   G S M T V L H D L N L N L G E G E V L G L F G H N G A G K T T S M K L I L G L L S P S E G Q V K
Consensus           G . . . V L . D L . . . . . . G E . . G . L G . N G A G K S T L . R . I . G . . . P . . G . I .

BexA(Hi)       64                   R S M S I S W P L A F S G       A F Q G S L T G M D N         L R F I C R L Y D
CtrD(Nm)       64                   R T M S I S W P L A F S G       A F Q G S L T G M D N         L R F I C R I Y N
KpsT(Ec)       64                   T N K T I S W P V G L A G       G F Q G S L T G R E N         V K F V A R L Y A
DrrA(Sp)       67   V F G H D V T S E P D T V R R R I S V T G Q Y A S V D E G L T G T E N         L V M M G R L Q G
NodI(Rl)       71   V L G A Q E P G Q V R L A R A K I G I V S Q F D N L D L E F T V R E N         L L V Y G R Y F R
CysA(Ec)       61   F H G T D V   S R L H A R D R K V G F V F Q H Y A L F R H M T V F D N I A F G L T V L P R R E R
ORF60-5(Ssp)   85   L L G     K P I G D R Q T K Q R V G Y L P E N A Y F Y D F L T A W E F         L D Y I G S L F Q
NosF(Ps)       62   V L G   R A P N D P Q   V R R Q L G Y L P E N V T F Y P Q L S G R E T         L R H F A R L K G
Consensus           . . G . . . . . . . . . . . . . . G . . . . . . . F . . . L T G . E N . . . . L . . . . R L . .

                                                        !                           !   !       !
BexA(Hi)       97   V D   P D Y V T R     F T K E F S E L G D     Y L Y E P V K K Y S S G M K A R L A F A L S L S V
CtrD(Nm)       97   V D   I D Y V K A     F T E E F S E L G Q     Y L Y E P V K R Y S S G M K A R L A F A L S L A V
KpsT(Ec)       97   K Q   E E L K E K I E F V E E F A E L G K     Y F D M P I K T Y S S G M R S R L G F G L S M A F
DrrA(Sp)      111   Y S W A R A R E R A A E L I D G F G L G D     A R D R L L K T Y S G G M R R R L D I A A S I V V
NodI(Rl)      115   M S T R E I E T V I P S L L E F A R L E S     K A N T R V A D L S G G M K R R L T L A G A L I N
CysA(Ec)      108   P N A A A I K A K V T K L L E M V Q L A H     L A D R Y P A H V S G G Q K Q R V A L A R A L A V
ORF60-5(Ssp)  155   I G K Q E R Q R R I L A M C D L V G L A Q K T A K K K Q L R Q Y S K G M L Q R V G M A Q A L I N
NosF(Ps)      104   A A L T Q V D         E L L E Q V G L A H     A A D R R V K T Y S K G M R Q R L G L A Q A L L G
Consensus           . . . . . . . . . . . . . . E . . . L . . . . . . . . . . K . Y S . G M . . R L . . A . . L . .

                                  ! !         !
BexA(Hi)      140   E F D C Y L I D E V I A V G D S R F A E K C K Y E L F   E K R K D R   S I I L V S H S P S A M K
CtrD(Nm)      140   E F D C Y L I D E V I A V G D S R F A D K C K Y E L F   E K R K D R   S I I L V S H S H S A M K
KpsT(Ec)      142   K F D Y Y I V D E V T A V G D A R F K E K C A   Q L F K E R H K E S   S F L M V S H S L N S L K
DrrA(Sp)      157   T P D L L F L D E P T T G L D P R S R N Q V W D I V R A L V D A G T   T V L L T T Q Y L D E A D
NodI(Rl)      161   D P Q L L I L D E P T T G L D P H A R H L I W E R L R S L L A R G K   T I L L T T H I M E E A E
CysA(Ec)      154   E P Q I L L L D E P F G A L D A Q V R K E L R R W L R Q L H E E L K F T S V F V T H D Q E E A T
ORF60-5(Ssp)  203   D P E V V F L D E P M S G L D P W G A T K C G K L F F P S R N R A K   P S F S T P T F W R T L E
NosF(Ps)      146   E P R L L L L D E P T V G L D P I A T Q D L Y L L I D R L R Q R G T   S I I L C S H V L P G V E
Consensus           . P . . . . L D E P . . . L D . . . . . . . . . . L . . . . . . . . . . . . L . . H . . . . . .

BexA(Hi)      186   S Y C D N A V V L E N G   I M H
CtrD(Nm)      186   Q Y C D N A M V L E K G   H M Y
KpsT(Ec)      188   E F C D V A I V F K N S Y I I G
DrrA(Sp)      204   Q L A D R I A V I D H G R V I A
NodI(Rl)      208   R L C D R L C V L E A G R K I A
CysA(Ec)      202   E V A D R V V V M S Q G N I E Q
ORF60-5(Ssp)  250   Q I C D R I A I L A R G E L L C
NosF(Ps)      193   A H I N R A A I L A K G C L Q A
Consensus           . . C D R . . V L . . G . . . .
```

Fig. 3. Multiple alignment of major parts of the sequences of the five homologous cytoplasmic ATP binding proteins of the ABC-2 family as well as three closely related proteins: (1) the CysA protein, a constituent of the sulfate-thiosulfate transport system of *Escherichia coli* (Ec), (2) the ORF60-5 protein, a functionally uncharacterized protein of the cyanobacterium *Synechocystis* sp. PCC6803 (Ssp), and (3) the NosF protein, a constituent of the Cu^{2+} processing system of *Pseudomonas stutzeri* (Ps). Residues conserved in all eight proteins are indicated by an exclamation mark above the multiple alignment. The consensus sequence (consensus) is provided below the multiple alignment. Numbers to the left of the sequences provide the positions in the proteins of the first residues in each line. The abbreviations of the proteins and species as well as references for the sequences are provided in the text.

brane protein with six transmembrane helical segments) (Zumft et al., 1990). The identities (indicated by exclamation marks above the alignment) and the consensus sequence (provided below the alignment) illustrate the high degree of sequence similarity exhibited by these proteins.

Table 2 presents binary comparison scores for the proteins included in Figure 3 as well as for an ATP binding domain of a human (Hs) multidrug resistance protein (Mdr3) (Chen et al., 1986). All of these proteins exhibit comparison scores in excess of 10 SD and are therefore homologous, but the Mdr3(Hs) protein is more distant from the ATP binding constituents of the ABC-2 subfamily than they are from each other. It is interesting to note, however, that the Mdr3(Hs) protein domain, like the CysA(Ec) and ORF60-5(Ssp) proteins, is much more similar to DrrA and NodI than it is to KpsT, BexA, or CtrD.

The phylogenetic tree for the proteins, the sequences of which are shown in Figure 3, is presented in Figure 4. In agreement with the tree for the integral membrane constituents of the ABC-2 family, the ATP binding proteins of the polysaccharide transporters, Bex, Ctr, and Kps, are all closely related. The other ATP binding proteins, including DrrA and NodI, are off on a distant branch together with the other members of this family of proteins selected for comparison because of their sequence similarities. The relative distances of the different integral membrane constituents of the ABC-2 subfamily (Fig. 2) are similar to those of the different ATP binding protein constituents of this same subfamily (Fig. 4). This fact suggests that the evolutionary divergence of these two classes of permease proteins may have occurred in parallel, and that an ancestral transport system may therefore have contained both of these proteins. It should be noted, however, that the DrrA, NodI, CysA, and NosF proteins are substantially closer to each other than they are to the capsular polysaccharide transport ATP binding proteins (Fig. 4). Because the integral membrane constituents of the Drr and Nod transporters were not demonstrably homologous to the integral membrane constituents of the Cys or Nos systems, it appears that the constituent proteins of these systems either did not evolve in parallel as parts of complete transport systems derived from a common ancestral system, or that the relative rates of evolutionary divergence of the integral membrane channel forming constituents of these systems were markedly different from those of their ATP binding constituents.

Discussion

In this communication we have provided evidence regarding the structural and evolutionary relatedness of proteins that apparently comprise a subfamily of the ABC family of solute transport systems. Our study began with the finding that the transmembrane constituent of the Drr drug resistance transporter of *S. peucetius*, which pumps daunorubicin and doxorubicin out of the bacterial cell (Guilfoile & Hutchinson, 1991), exhibits striking sequence similarity to corresponding integral membrane protein constituents of polysaccharide export systems of gram-negative bacteria. Of particular interest was the fact that only a single gene, encoding only one integral membrane

Table 2. *Binary comparisons of the alignments of the ATP-binding constituents of the ABC-2 subfamily of transport systems*

	BexA(Hi) (217)[a]	DrrA(Sp) (330)	CtrD(Nm) (216)	NodI(Rl) (311)	CysA(Ec) (365)	ORF60-5 (Ssp) (354)	NosF(Ps) (308)	Mdr3(Hs) (1279)
KpsT(Ec) (224)	45 (217)[b] [61]	27 (195) [23]	44 (217) [60]	26 (211) [15]	46 (61) [13]	24 (230) [17]	27 (210) [20]	33 (61) [6]
BexA(Hi) (217)		30 (151) [11]	82 (213) [95]	26 (195) [13]	40 (47) [12]	23 (218) [15]	29 (217) [26]	33 (61) [6]
DrrA(Sp) (330)			25 (200) [14]	33 (310) [42]	25 (259) [25]	26 (273) [32]	29 (315) [34]	24 (297) [23]
CtrD(Nm) (216)				28 (212) [15]	39 (62) [12]	23 (194) [14]	33 (209) [27]	30 (61) [7]
NodI(Rl) (311)					33 (250) [29]	28 (287) [28]	37 (208) [30]	30 (218) [24]
CysA(Ec) (365)						26 (156) [24]	28 (228) [23]	29 (228) [24]
ORF60-5(Ssp) (354)							32 (278) [33]	23 (222) [11]
NosF(Ps) (308)								26 (261) [15]

[a] The values in parentheses below the protein designations indicate the numbers of amino acids in the proteins.

[b] Values reported in the table correspond to the percent identities in the segments compared. The number of amino acids in the aligned segments is provided in parentheses. Values in brackets denote the comparison scores in standard deviations higher than those obtained with 100 comparisons of randomized sequences of these protein segments. The FASTA and RDF2 programs (Pearson & Lipman, 1988) were used to assess the similarity scores and their significance. The abbreviations used are defined in the text as are references for the sequences compared.

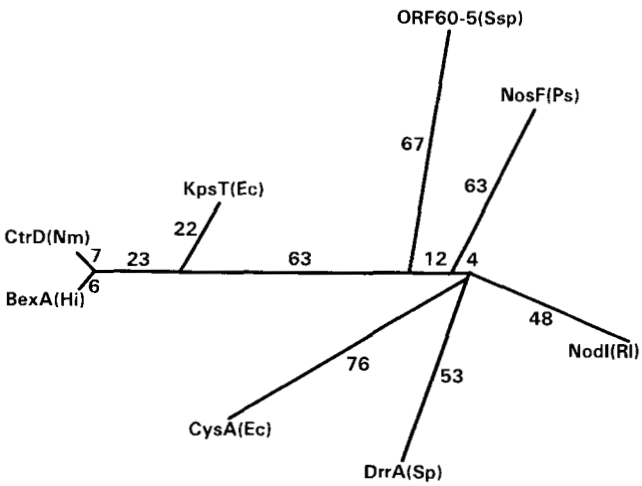

Fig. 4. Phylogenetic tree of eight ATP binding proteins including the ATP binding constituents of the five members of the ABC-2 family of transport systems. The programs used for tree construction, abbreviations, and references are provided in the text.

constituent in each of two of these systems (Drr and Kps), has yet been found. Further, each of these proteins was found to exhibit sequence similarity to the NodJ protein (Fig. 1). Although the biochemical function of NodJ is not known, it appears to act together with NodI to facilitate nodulation of leguminous plants by rhizobial species. Although no direct experimental evidence is available concerning their biochemical function, these proteins might catalyze export of a modified β-1,4-linked *N*-acetylglucosamine oligosaccharide (Long, 1989; S.R. Long, pers. comm.). The sequence analyses reported here clearly support this suggestion. In this connection it is worth noting that no integral membrane protein in addition to NodJ has been found that is likely to function as an additional constituent of this postulated transporter, although most of the *nod* genes of *Rhizobium meliloti* and *R. leguminosarum* (Long, 1989) have been sequenced. Further, only a single candidate for a transmembrane constituent of the Kps system of *E. coli* (KpsM) or of the presumptive copper transport system of *P. stutzeri* (NosY) has been detected (Zumft et al., 1990; Pavelka et al., 1991; R.P. Silver, pers. comm.).

The observations presented in this communication extend and accentuate the suggestion (Kerppola & Ames, 1992) that individual members of the ABC-type transporters, like the members of the various transport families, which include facilitators, ion channels, and PTS permeases, may be capable of functioning either as homooligomers or as a heterooligomers, depending on the system under study. This observation provides further evidence for the postulate (Saier & Reizer, 1991) that these four classes of permeases are evolutionarily, structurally, and functionally related.

Methods

The protein sequence alignments and relative evolutionary distances were determined using the progressive sequence alignment programs of Feng and Doolittle (1990) and Doolittle and Feng (1990), modified for execution in the UCSD VAX/VMS DNA system (Smith, 1988). The Protein Information Resource (PIR) of the National Biomedical Research Foundation (NBRF) database was screened in all homology searches using the FASTA program. The FASTA and RDF2 programs (Pearson & Lipman, 1988) were used to assess the similarity scores and their significance. Comparison scores are recorded in standard deviations higher than obtained with 100 comparisons of randomized sequences of the protein segments analyzed.

Acknowledgments

We thank Drs. C. Richard Hutchinson, Sharon R. Long, and Richard R. Silver for valuable discussions and Melinda Marcus and Mary Beth Hiller for assistance in the preparation of this manuscript. This work was supported by USPHS grants 5R01AI21702 and 2R01AI14176 from the National Institute of Allergy and Infectious Diseases.

References

Chen, C.-J., Chin, J.E., Ueda, K., Clark, D.P., Pastan, I., Gottesman, M.M., & Roninson, I.B. (1986). Internal duplication and homology with bacterial transport proteins in the *mdrl* (P-glycoprotein) gene from multidrug-resistant human cells. *Cell 47*, 381–389.

Chitnis, P.R. & Nelson, N. (1991). Molecular cloning of the genes encoding two chaperone proteins of the cyanobacterium *Synechocystis* sp. PCC #6803. *J. Biol. Chem. 266*, 58–65.

Doolittle, R.F. (1981). Similar amino acid sequences: Chance or common ancestry? *Science 214*, 149–159.

Doolittle, R.F. (1986). *Of Urfs and Orfs.* University Science Books, Mill Valley, California.

Doolittle, R.F. & Feng, D.-F. (1990). A nearest neighbor procedure for relating progressively aligned amino acid sequences. *Methods Enzymol. 183*, 659–669.

Evans, I.J. & Downie, J.A. (1986). The *nodI* gene product of *Rhizobium leguminosarum* is closely related to ATP-binding bacterial transport proteins; nucleotide sequence analysis of the *nodI* and *nodJ* genes. *Gene 43*, 95–101.

Feng, D.-F. & Doolittle, R.F. (1990). Progressive alignment and phylogenetic tree construction of protein sequences. *Methods Enzymol. 183*, 375–387.

Frosch, M., Edwards, U., Bousset, K., Krausse, B., & Weisgerber, C. (1991). Evidence for a common molecular origin of the capsule gene loci in gram-negative bacteria expressing group II capsular polysaccharides. *Mol. Microbiol. 5*, 1251–1263.

Furlong, C.E. (1987). Osmotic-shock-sensitive transport systems. In *Escherichia coli and Salmonella typhimurium* (Neidhardt, F.C., Ingraham, J.L., Magasanik, B., Low, K.B., Schaechter, M., & Umbarger, H.E., Eds., pp. 768–796. American Society for Microbiology, Washington, D.C.

Guilfoile, P.G. & Hutchinson, C.R. (1991). A bacterial analog of the *mdr* gene of mammalian tumor cells is present in *Streptomyces peucetius*, the producer of daunorubicin and doxorubicin. *Proc. Natl. Acad. Sci. USA 88*, 8553–8557.

Higgins, C.F., Hyde, S.C., Mimmack, M.M., Gileadi, U., Gill, D.R., & Gallagher, M.P. (1990). Binding protein-dependent transport systems. *J. Bioenerget. Biomembr. 22*, 571–592.

Kerppola, R.E. & Ames, G.F.-L. (1992). Topology of the hydrophobic membrane-bound components of the histidine periplasmic permease. Comparison with other members of the family. *J. Biol. Chem. 267*, 2329-2336.

Kroll, J.S., Hopkins, I., & Moxon, E.R. (1988). Capsule loss in *H. influenzae* type b occurs by recombination-mediated disruption of a gene essential for polysaccharide export. *Cell 53*, 347-356.

Kroll, J.S., Loynds, B., Brophy, L.N., & Moxon, E.R. (1990). The *bex* locus in encapsulated *Haemophilus influenzae*: A chromosomal region involved in capsule polysaccharide export. *Mol. Microbiol. 4*, 1853-1862.

Kroll, J.S. & Moxon, E.R. (1990). Capsulation in distantly related strains of *Haemophilus influenzae* type b: Genetic drift and gene transfer at the capsulation locus. *J. Bacteriol. 172*, 1374-1379.

Long, S.R. (1989). *Rhizobium*-legume nodulation: Life together in the underground. *Cell 56*, 203-214.

Pavelka, M.S., Jr., Wright, L.F., & Silver, R.P. (1991). Identification of two genes, *kpsM* and *kpsT*, in region 3 of the polysialic acid gene cluster of *Escherichia coli* K1. *J. Bacteriol. 173*, 4603-4610.

Pearson, W.R. & Lipman, D.J. (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA 85*, 2444-2448.

Saier, M.H., Jr. & Reizer, J. (1991). Families and superfamilies of transport proteins common to prokaryotes and eukaryotes. *Curr. Opin. Struct. Biol. 1*, 362-368.

Sirko, A., Hryniewicz, M., Hulanicka, D., & Böck, A. (1990). Sulfate and thiosulfate transport in *Escherichia coli* K-12: Nucleotide sequence and expression of the *cysTWAM* gene cluster. *J. Bacteriol. 172*, 3351-3357.

Smith, A.N., Boulnois, G.J., & Roberts, I.S. (1990). Molecular analysis of the *Escherichia coli* K5 *kps* locus: Identification and characterization of an inner-membrane capsular polysaccharide transport system. *Mol. Microbiol. 4*, 1863-1869.

Smith, D.W. (1988). A complete and integrated system for DNA protein sequence analysis using VAX/VMS computers. *CABIOS 4*, 212.

Zumft, W.G., Viebrock-Sambale, A., & Braun, C. (1990). Nitrous oxide reductase from denitrifying *Pseudomonas stutzeri*. *Eur. J. Biochem. 192*, 591-599.