J. Membrane Biol. 156, 99–103 (1997)

The Journal of
Membrane
Biology

Topical Review

A Novel Family of Ubiquitous Heavy Metal Ion Transport Proteins

I.T. Paulsen, M.H. Saier, Jr.
Department of Biology, University of California at San Diego, La Jolla, CA 92093-0116, USA

Received: 15 July 1996/Revised: 21 October 1996

Abstract. We describe a novel diverse family of metal ion transporter (CDF) proteins (the cation diffusion facilitator (CDF) family) with members occurring in both prokaryotes and eukaryotes. Thirteen sequenced protein members of the CDF family have been identified, several of which have been shown to transport cobalt, cadmium and/or zinc. All members of the CDF family possess six putative transmembrane spanners with strongest conservation in the four N-terminal spanners, and on the basis of the analyses, we present a unified structural model. Members of the family are shown to exhibit an unusual degree of size variation, sequence divergence, and differences in cell localization and polarity. The phylogenetic tree for the CDF family reveals that prokaryotic and eukaryotic proteins cluster separately. It allows functional predictions for some uncharacterized members of this family. A signature sequence specific for the CDF family is derived.

Key words: Transport — Membrane proteins — Cobalt — Cadmium — Zinc — Uptake — Efflux — pmf — Phylogenetic analysis

Introduction

Heavy metal ions such as those of Fe, Co, Ni, Zn and Cd are essential micronutrients, but these ions as well as many others are toxic when present in excess. Consequently, living organisms have evolved transport mechanisms for the active uptake and/or extrusion of these ions, thereby enabling cells to regulate their intracellular concentrations (Silver et al., 1989). Bacterial pathogens in particular have evolved elaborate mechanisms which allow them to effectively compete with their hosts for limiting micronutrients such as iron and copper while protecting themselves against the lethal effects of antimicrobial agents such as mercury, copper and silver (Wooldridge & Williams, 1993; Guerinot, 1994; Brown et al., 1995; Gupta et al., 1995; Odermatt & Solioz, 1995; Solioz & Odermatt, 1995).

Three major families of transport proteins have been identified which include members that catalyze heavy metal ion transport. (i) P-type ATPases are known to specifically catalyze either uptake or extrusion of Cd^{2+} or Cu^{2+} as well as a variety of mono- and divalent ions (H^+, K^+, Na^+, Mg^{2+}, Ca^{2+}) (Odermatt et al., 1993; Bull & Cox, 1994; Fagan & Saier, 1994). (ii) ABC-type transporters are known which transport Ni^{2+}, Mn^{2+}, Fe^{2+} and Mo^{2+}, and some of these proteins are closely related to transporters specific for peptides and sugars (Tam & Saier, 1993; Kuan et al., 1995). (iii) RND transporters include proteins that extrude Ni^{2+}, Co^{2+}, Cd^{2+} and Zn^{2+} (Dinh et al., 1994; Saier et al., 1994). While the ATP-dependent P-type and ABC-type transport systems are ubiquitous, being found in all major classifications of living organisms examined for them, the pmf-dependent RND family proteins have so far been found only in bacteria (Saier et al., 1994). In addition to proteins included within these major families, various heavy metal ion transporters have been identified that have not yet been assigned to established families. These include bacterial transporters specific for chromium, copper, and mercury (Silver & Walderhaug, 1992).

Nies and Silver (1995) recently described a novel family of heavy metal ion transporters with four numbers which they designated the Cation Diffusion Facilitator (CDF) family. In contrast to the three families men-

Correspondence to: M.H. Saier, Jr.

Table 1. Members of the CDF metal ion transporter family

Abbr.	Organism	Size	Function	Acc. No.	Database
CzcD Aeu	*Alcaligenes eutrophus*	316	Regulation of a cobalt/zinc/ cadmium resistance operon	X98451	gp
YglB Bst	*Bacillus stearothermophilus*	108 (partial)	?	P30540	gp
YiiP Eco	*Escherichia coli*	300	?	P32159	sp
P34A Rri	*Rickettsia rickettsii*	301	?	P21559	sp
YdxT Bsu	*Bacillus subtilis*	276	?	P46348	sp
Cot1 Sce	*Saccharomyces cerevisiae*	439	Cobalt uptake	P32798	sp
ZnrP Sce	*Saccharomyces cerevisiae*	442	Zinc/cadmium resistance	P20107	sp
Orf7 Cel	*Caenorhabditis elegans*	549	?	U23529	gp
Orf3 Cel	*Caenorhabditis elegans*	382	?	Z68119	gp
Znt1 Rno	*Rattus norvegicus*	507	Zinc exporter	U17133	gp
Znt1 Mmu	*Mus musculus*	503	Zinc exporter	U17132	gp
Orf2 Sce	*Saccharomyces cerevisiae*	740	?	Z68194	gp
Znt2 Rno	*Rattus norvegicus*	366	vesicular zinc uptake	U50927	gp

tioned above, this family is concerned with the transport of heavy metal ions, and all currently characterized members of the family serve this function. In this study we identify thirteen members of the family have been identified from a variety of bacteria and eukaryotes. Of these, the yeast COT1 and ZRC1 gene products confer cobalt and zinc-cadmium-resistance, respectively, probably due to uptake of the ions into mitochondria (Conklin et al., 1992, 1994). One homologous mammalian zinc-resistance transporter (ZnT-1) has been localized to the plasma membrane and reported to function in Zn^{2+} efflux (Palmiter and Findley, 1995). A second mammalian zinc-resistance transporter (ZnT-2) is localized to the intracellular membrane system and appears to facilitate vesicular sequestration of Zn^{2+} (Palmiter et al., 1996). Finally, the CzcD protein in the bacterium, *Alcaligenes eutrophus*, together with CzcR, allows induction of resistance to cobalt, zinc and cadmium (Nies, 1992). CzcD may catalyze metal ion uptake while CzcR binds the metal ion and promotes transcription of the *czc* operon.

Proteins of the CDF Family

The Table lists the proteins of the CDF family. The protein abbreviations to be used in this study as well as the organismal sources, protein sizes and functions when known are provided. Accession numbers allow easy access to the protein sequences. The functionally characterized transporters from eukaryotes catalyze either uptake or efflux of the divalent heavy metal ions, cobalt, cadmium and/or zinc. A single bacterial protein (CzcD) has been reported to regulate the *czc* (cobalt/zinc/cadmium-resistance) operon of *Alcaligenes eutrophus* (Nies, 1992), probably by catalyzing uptake of the metal ions which interact with the CzcR transcription factor.

The sizes of the CDF family proteins vary dramatically with the bacterial proteins consistently being smaller than the eukaryotic proteins (276–316 residues versus 366–740 residues). In the case of the largest eukaryotic protein (Orf2 Sce), we consider it probable that it is actually somewhat smaller due to incorrect assignment of the N-terminal ATG. All proteins listed in the Table proved to be homologous based on the fact that they exhibited comparison scores with each other that were always in excess of 9 S.D. (Saier, 1994).

Topological and Sequence Analyses of CDF Family Proteins

Figure 1*A* shows an average hydropathy plot of the multiply aligned CDF family proteins while Fig. 1*B* shows a corresponding average similarity plot. In the former plot, the six distinct peaks of hydrophobicity presumably correspond to six transmembrane spanners (I–VI). Peaks III and IV proved to be most hydrophobic while peaks V and VI were least hydrophobic. Examination of the six multiply aligned putative spanners in helical wheel depiction revealed that helices I, II, V and VI were strikingly amphipathic with hydrophobic residues localized to one side of each of these helices and the more conserved semipolar and hydrophilic residues localized to the other side. Helices II, V and VI each contained a single fully conserved aspartyl residue on their hydrophilic sides, and assuming that the hydrophilic sides of these helices face inwards to form a central water-filled channel, the β-carboxyls of these three aspartyl residues might comprise a cation binding site inside the channel. We suggest that the four amphipathic helices comprise an inner core which forms a channel while the two remaining hydrophobic helices (III and IV) are located in the more lipid-exposed outer shell. Site-specific mutagenesis analyses should allow determination of the es-

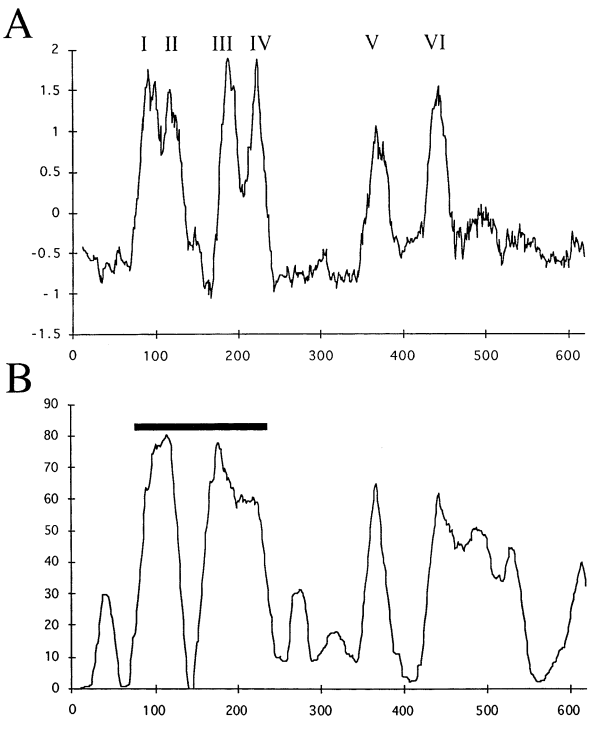

Fig. 1. Average hydropathy (*A*) and average similarity (*B*) plots for the thirteen proteins of the CDF family. A sliding window of 20 residues was used in both plots. The algorithm of Kyte and Doolittle (1982) was used in panel *A.* The six peaks (I–VI) correspond to the six putative transmembrane spanners. In *B*, the bar indicates that portion of the multiple alignment that is presented in Fig. 2.

sentiality of specific residues to cation binding and transport.

A CDF Family-specific Signature Sequence

As shown in Fig. 1*B*, the most conserved region of the CDF family proteins encompasses transmembrane spanners I–IV. The multiple alignment corresponding to this region (represented by the bar in Fig. 1*B*) is reproduced in Fig. 2. The bars above the multiple alignment in Fig. 2 show the positions of the four putative spanners. Two residues are fully conserved in this alignment: the serine in the loop between spanners I and II, and the aspartate located near the center of spanner II.

From this region of the multiple alignment, a CDF family-specific signature sequence (Bairoch, 1992) was derived. It begins with the fully conserved serine shown in Fig. 2 and continues just past the fully conserved aspartate. This signature sequence is:

$$\underline{\underline{S}}\ X\ (ASG)\ (LIVMT)_2\ (SAT)\ (DA)\ (SGAL)\ (LIVFYA)\ (HDN)\ X_3\underline{\underline{D}}X_2\ (AS) \qquad (11)$$

[X = any residue; alternative residues at any one position are in parentheses]

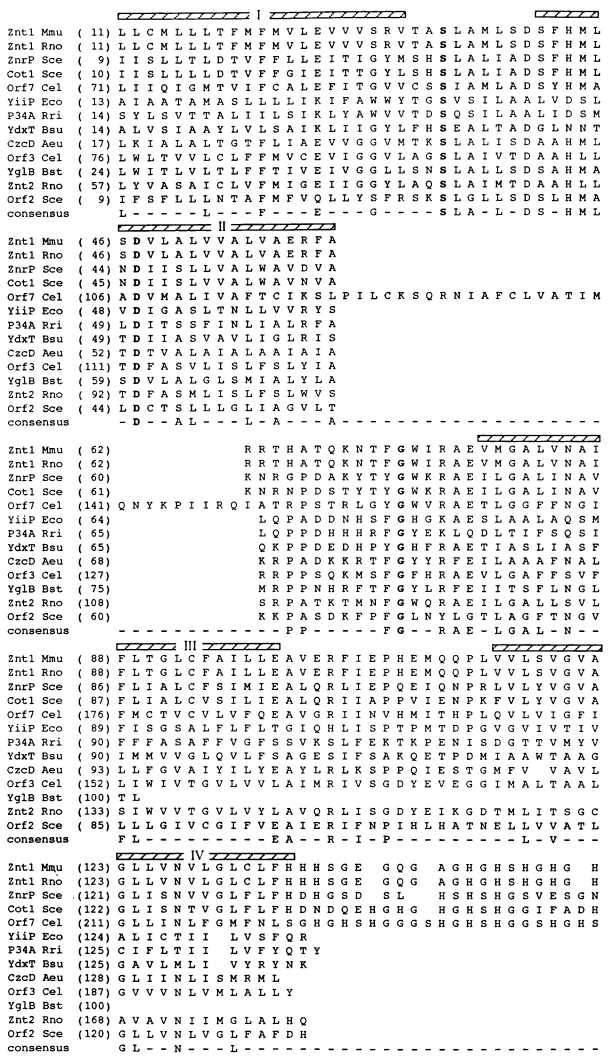

Fig. 2. Multiple alignment of the most conserved portions of the CDF family proteins (corresponding to the region indicated by the bar in Fig. 1*B*). The residue # in each protein is indicated by the number presented in parentheses following the abbreviation of that protein. The consensus sequences (CONSENSUS) is provided below the multiple alignment. The four bars (I–IV) at the top of the alignment indicate the positions of putative transmembrane spanners I–IV. Fully conserved residues are presented in bold print. The signature sequence for the CDF family (*see text*) was derived from the region encompassing the fully conserved S (line 1) and D (line 2) portrayed in bold print.

It was screened against the SwissProt database, and only members of the CDF family were retrieved. It is therefore a bona fide signature sequence which should be useful for the identification of new members of the CDF family as these become sequenced.

Segmental Size Variability Analyses

Proteins of the CDF family exhibit considerable size variability as mentioned above (*see* Table). We local-

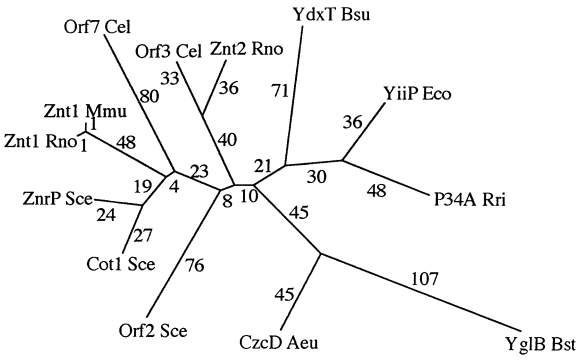

Fig. 3. Phylogenetic tree for the thirteen current members of the CDF family. The eight eukaryotic proteins are shown on the left-hand side of the tree while the five prokaryotic proteins are on the right. The TREE program of Feng and Doolittle (1990) was used for tree construction. Branch length (in arbitrary units) is approximately proportional to phylogenetic distance.

ized this variability to the various hydrophilic regions of the proteins preceding, following and connecting the six putative transmembrane spanners. The N-termini preceding spanners I are usually short (8–23 residues) except for three eukaryotic proteins (Orf7 Cel, Orf3 Cel and Orf2 Sce) in which these hydrophilic regions are 70 residues or more in length. Loops 1 (between spanners I and II) are of invariant length (9 residues) as shown in Fig. 2. Loops 2 are either 16 or 17 residues in length except for Orf7 Cel in which loop 2 is 45 residues long. Loops 3 vary from 13 to 15 residues. However, the variability in loop sizes for loops 4 (10–141 residues) and 5 (6–47 residues) as well as the C-termini (0–171 residues) is very substantial (*not shown*). This variability accounts for the very low degree of similarity observed in Fig. 1*B* for the loop regions in the C-terminal portions of these proteins.

It is noteworthy that the Znt1 Mmu, Znt2 Rno, ZnrP Sce, Cot2 Sce and Orf7 Cel proteins all possess long cytoplasmic loops 4 (108–141 residues) between transmembrane spanners 4 and 5. In the other eukaryotic proteins loops 4 vary in size from 25 to 37 residues, and in the prokaryotic proteins they are uniformly short (10–13 residues). The eukaryotic loops 4 between spanners 4 and 5 consistently contain histidine rich regions with the sequence $(HX)_n$ where X is usually G or C, and $n = 3$–6. The significance of these regions in the eukaryotic proteins is not known, but they could function in metal ion binding, potentially serving either a functional or a regulatory role.

CDF Family Phylogenetic Tree

The phylogenetic tree for the CDF family proteins is shown in Fig. 3. All of the eukaryotic proteins cluster loosely together on the left while all of the prokaryotic proteins cluster loosely together on the right. Orf3 Cel clusters tightly with Znt2 Rno while Orf7 Cel clusters loosely with Znt1 Mmu and Znt1 Rno. It can therefore be proposed that Orf3 Cel, like Znt2, is localized to intracytoplasmic vesicles while Orf7 Cel, like Znt1, is localized to the plasma membrane. Since ZnrP Sce and Cot1 Sce are very closely related, it can be proposed that these two proteins are close isoforms serving similar functions in spite of their differing specificities.

Only one bacterial protein, CzcD Aeu, is at all functionally characterized. Based on its reported functional characteristics (*see* Introduction), this protein presumably transports Co^{2+}, Zn^{2+} and Cd^{2+} (Nies, 1992). It does not cluster with other members of the family. In fact, except for YiiP Eco and P34A Rri, the bacterial proteins are all distantly related. The former two clustered proteins may be orthologues serving the same function while the distant homologues may serve distinct functions.

Conclusions and Perspectives

We have identified thirteen proteins which comprise a novel, ubiquitous family of metal ion transporters (the CDF family). The variability observed for this family of transporters occurs at several levels. (i) They vary in size, ranging in length from 276 to 549 amino acyl residues, a feature that is highly unusual to previously studied transport protein families (Saier, 1994, 1996). (ii) They vary in localization, being present either in the cytoplasmic membrane or in intracellular vesicular membranes of eukaryotes. (iii) They vary in polarity, catalyzing either influx or efflux. (iv) They vary tremendously in sequence, suggesting that some of the functionally uncharacterized proteins may exhibit novel specificities and/or functions. (v) The eukaryotic proteins differ from the prokaryotic proteins in possessing histidine-rich cytoplasmic loops of characteristic sequence between transmembrane spanners 4 and 5 that may serve some unique function, possibly in metal sequestration or in macromolecular recognition.

On the other hand, the CDF family is the first transport protein family so far to be identified that is as yet specific and exclusive for heavy metal ions. The P-type ATPases, the ABC transporters, and the RND transporters all appear to exhibit broader substrate specificities than observed for the currently characterized members of the CDF family, transporting molecular species in addition to heavy metal ions. This observation is particularly surprising in view of the sequence and size variability observed for members of the CDF family, a characteristic that, for example, is not shared by the ABC or RND families (Saier et al., 1994; Kuan et al., 1995).

The fully conserved acidic residues present in am-

phipathic putative transmembrane spanners of the CDF family proteins, may well prove to serve as membranal cation binding sites. Structural and functional analyses of genetically altered members of this family should provide clues as to the molecular bases underlying the transport mechanism characteristic of, and the functional differences exhibited by the protein members of the CDF family. The results presented here should therefore facilitate detailed molecular genetic and biochemical analyses of the structures and functions of these proteins.

We thank Alison M. Beness for assistance with the computer analysis and Mary Beth Hiller for assistance in the preparation of this manuscript. This work was supported by USPHS grants 5RO1AI21702 and 2RO1AI14176 from the National Institute of Allergy and Infectious Diseases.

References

Bairoch, A. 1992. *Nucl. Acids Res.* **20** (Suppl.), S2013–2018

Brown, N.L., Barrett, S.R., Camakaris, J., Lee, B.T.O., Rouch, D.A. 1995. *Mol. Microbiol.* **17:**1153–1166

Bull, P.C., Cox, D.W. 1994. *Trends Genet.* **10:**246–252

Conklin, D.S., Culbertson, M.R., Kung, C. 1994. *Mol. Gen. Genet.* **244:**303–311

Conklin, D.S., McMaster, J.A., Culbertson, M.R., Kung, C. 1992. *Mol. Cell. Biol.* **12:**3678–3688

Dinh, T., Paulsen, I.T., Saier, M.H., Jr. 1994. *J. Bacteriol.* **176:**3825–3831

Guerinot, M.L. 1994. *Annu. Rev. Microbiol.* **48:**743–772

Gupta, S.D., Lee, B.T.O., Camakaris, J., Wu, H.C. 1995. *J. Bacteriol.* **177:**4207–4215

Fagan, M.J., Saier, M.H., Jr. 1994. *J. Mol. Evol.* **38:**57–99

Feng, D.-F, Doolittle, R.F. 1990. *Methods Enzymol.* **183:**375–387

Kuan, G., Dassa, E., Saurin, W., Hofnung, M., Saier, M.H., Jr. 1995. *Res. Microbiol.* **146:**271–278

Kyte, J., Doolittle, R.F. 1982. *J. Mol. Biol.* **157:**105–132

Nies, D.H. 1992. *J. Bacteriol.* **174:**8102–8110

Nies, D.H., Silver, S. 1995. *J. Ind. Microbiol.* **14:**186–199

Odermatt, A., Solioz, M. 1995. *J. Biol. Chem.* **270:**4349–4354

Odermatt, A., Suter, H., Krapf, R., Solioz, M. 1993. *J. Biol. Chem.* **268:**12775–12779

Palmiter, R.D., Cole, T.B., Findley, S.D. 1996. *EMBO J.* **15:**1784–1791

Palmiter, R.D., Findley, S.D. 1995. *EMBO J.* **14:**639–649

Saier, M.H., Jr. 1994. *Microbiol. Rev.* **58:**71–93

Saier, M.H., Jr. 1996. *Microbial Comp Genomics* **1:**129–150

Saier, M.H., Jr., Tam, R., Reizer, A., Reizer, J. 1994. *Mol. Microbiol.* **11:**841–847

Silver, S., Nucifora, G., Chu, L., Misra, T.K. 1989. *Trends Biochem. Sci.* **14:**129–133

Silver, S., Walderhaug, M. 1992. *Microbiol. Rev.* **56:**195–228

Solioz, M., Odermatt, A. 1995. *J. Biol. Chem.* **270:**9217–9221

Tam, R., Saier, M.H., Jr. 1993. *Microbiol. Rev.* **57:**320–346

Wooldridge, K.G., Williams, P.H. 1993. *FEMS Microbiol. Rev.* **12:**325–348