Characterization of an Archaeal Multidrug Transporter with a Unique Amino Acid Composition*

Shira Ninio‡ and Shimon Schuldiner§
From the Alexander Silberman Institute of Life Sciences, Hebrew University of Jerusalem, 91904 Jerusalem, Israel

The Smr family of multidrug transporters consists of small membrane proteins that extrude various drugs in exchange with protons rendering cells resistant to these drugs. Smr proteins identified to date have been found only in Eubacteria. In this work we present the cloning and characterization of an Smr protein from the archaean Halobacterium salinarum, the first Smr in the archaeal kingdom. The protein, named Hsmr, was identified through sequence similarity to the Smr family, and the DNA sequence was cloned into an Escherichia coli expression system. Hsmr is heterologously expressed in a functional form despite the difference in lipid composition of the membrane and the lower salt in the cell and its environment. Cells harboring the Hsmr plasmid transport ethidium bromide in an uncoupler-sensitive process and gain resistance to ethidium bromide and acriflavine. Hsmr binds tetr phenylphosphonium (TPP⁺) with a relatively low affinity (Kd = 200 nM) at low salt concentration that increases (Kd = 40 nM) upon the addition of 2 M of either NaCl or KCl. The Hsmr protein contains many of the signature sequence elements of the Smr family and also a high content of negative residues in the loops, characteristic of extreme halophiles. Strikingly, Hsmr is composed of over 40% valine and alanine residues. These residues are clustered at certain regions of the protein in domains that are not important for activity, as judged from lack of conservation and from previous studies with other Smr proteins. We suggest that this high content of alanine and valine residues is a reflection of a "natural" alanine and valine "scanning" necessitated by the high GC content of the gene. This phenomenon reveals significant sequence elements in small multidrug transporters.

Multidrug transporters actively remove toxic compounds from the cytoplasm of cells. They are widespread from bacteria to man and have been associated with multidrug resistance (1, 2). Many multidrug transporters have been identified to date and classified into several protein families (3, 4). Proteins in the Smr family of small multidrug resistance proteins are ~110-amino acids long and extrude various drugs in exchange with protons, thereby rendering bacteria resistant to these compounds (5, 6). Several Smr proteins have been characterized, purified, and reconstituted in a functional form (7, 8). Most posses a single membrane-embedded charged residue, Glu-14, conserved in more than sixty homologous proteins (8), which was shown to be part of a binding site common to protons and substrates (9, 10). The basic oligomeric structure detected in two-dimensional crystals of EmrE, an E. coli Smr, is a dimer (11). Radiolabeled substrate binding to purified EmrE and negative dominance experiments support the contention that the functional unit of the protein is an oligomer formed by two dimers (12–14).

In the absence of high-resolution structure, it would be useful to study the effect of multiple changes on the function of the protein. This is a complicated task to perform using site-directed mutagenesis, because designing an effective combination of mutations is not straightforward. One way to tackle this problem is by analyzing a set of homologues that provide natural, more complex, and less biased mutants (8). However, also with these mutants it is not always possible to identify suppressor or complimentary substitutions throughout the protein. Here we describe a special case that demonstrates an approach to overcome this problem. We report the first identification and characterization of an Smr protein from Archaea, the third domain of the universal phylogenetic tree (15). This protein, named Hsmr, was cloned from the archaean H. salinarum based on sequence similarity to the Smr family. Hsmr was heterologously expressed in E. coli cells in a functional form despite the difference in lipid composition of the membrane and the lower salt in the cell and its environment. Hsmr renders cells resistant to the effect of multiple drugs and catalyzes the active efflux of ethidium from the cells. Detergent-solubilized Hsmr binds the high affinity substrate tetr phenylphosphonium (TPP⁺)1 in a salt-dependent manner: at low salt with a Kd of 200 nM, and 40 nM at high salt. The Hsmr protein contains many of the signature sequence elements defining the Smr family. However, it also has a highly unusual sequence displaying some unique characteristics. It is composed of over 40% valine and alanine residues, which are distributed throughout the protein, often concentrated at areas where there is no sequence conservation. We suggest that this phenomenon is the outcome of a natural process of alanine and valine “scanning” that is necessitated by the high GC content of the gene.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Plasmids—E. coli BL21(DE3) (Novagen, Madison, WI) were used throughout this work. The plasmids used are pT7–7 (17), pT7–7 EmrE-His (12), and pT7–7 Hsmr-His. The latter was constructed as follows: the hsmr gene (OE2652F in the H. salinarum data base www.halolex.mpg.de) was cloned by polymerase chain reaction using as template genomic DNA

---

* This work was supported by grants from the Deutsche-Israeli Program (Federal Ministry of Education, Science and Research-BMBF, International Bureau at the German Aerospace Center Technology), Grant NS16708 from the National Institutes of Health, and Grant 463/00 from the Israel Science Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This advertisement should be addressed. Tel.: 972-2-6585992; Fax: 972-2-5834625; E-mail: Shimon.Schuldiner@huji.ac.il.

‡ Recipient of a Yeshaya Horowitz Foundation fellowship.
§ This paper is available online at http://www.jbc.org

1 The abbreviations used are: TPP⁺, tetr phenylphosphonium; IPTG, isopropyl-β-d-thiogalactopyranoside; CCCP, carbonyl cyanide m-chlorophenylhydrazone; DDM, n-dodecyl-β-maltoside; Ni-NTA, nickel-nitrilotriacetic acid.
from H. salinarum (provided by Prof. M. Mevarech, Faculty of Life Sciences, Tel Aviv University, Israel). Primers were designed to overlap with each end of the gene and included sites for restriction enzymes NdeI and EcoRI. The gene was cloned into the pT7–7–His vector, which was obtained by removing the emrE gene from vector pT7–7–EmrE–His with restriction enzymes NdeI and EcoRI. The start codon was modified from GTG to ATG.

**Resistance to Toxic Compounds**—Resistance to toxic compounds was tested after overnight growth in 2xYT liquid media consisting of 1.6% (w/v) tryptone, 1% (w/v) yeast extract, and 0.5% (w/v) NaCl.

**Efflux of Ethidium in Whole Cells**—Efflux of ethidium in whole cells was assayed essentially as described (7), except that E. coli BL21(DE3) cells were used. The cells were grown at 37 °C in LB medium, when the culture reached A∞<sub>600</sub> = 0.8, IPTG was added to 0.5 mM. Two hours later the cells were centrifuged and resuspended in minimal medium A supplemented with 0.05% NaCl, 1 mM MgSO<sub>4</sub>, 0.1 mM CaCl<sub>2</sub>, and 0.36% glucose. At this point, ethidium and carbonyl cyanide m-chlorophenylhydrazone (CCCP) were added to a final concentration of 1 μM and 40 μM, respectively, and the cells were incubated for 60 min at 37 °C. The cells were collected by centrifugation and quickly resuspended in medium containing the same ethidium concentration without CCCP, and fluorescence was measured in a PerkinElmer fluorimeter (Luminescence Spectrometer LS-5) with exciting light at 545 nm and emission at 610 nm.

**Overexpression and Purification of Hsmr**—E. coli BL21(DE3) cells bearing plasmid pT7–7–Hsmr–His were grown at 37 °C in 2xYT medium. When the culture reached A∞<sub>600</sub> = 0.8, IPTG was added to 0.5 mM. Two hours later the cells were harvested by centrifugation, washed once with buffer containing 150 mM NaCl, 10 mM Tris, pH 7.5, 250 mM sucrose, 2.5 mM MgSO<sub>4</sub>, and 0.5 mM dithiothreitol, and resuspended in the same buffer containing 15 μg/ml DNaseI (5 μl buffer/g cells). Membranes were prepared by disrupting the cells using a Microfluidics microfluidizer processor (M-110EH). Non-disrupted cells were separated by centrifugation, and the membranes were collected by ultracentrifugation at 300,000 × g for 60 min. The membrane pellet was resuspended in the same buffer, frozen in liquid air, and stored at −70 °C.

Hsmr–His was purified by solubilizing membranes (0.25 mg of total protein belonging to the Smr family (3, 8) and was shown to have a central role in the transport mechanism of EmrE (12, 18). However, although EmrE has only two additional negatively charged residues, Hsmr has seven. They are located in putative hydrophilic loops of the protein, four of them in pairs, a pattern found in other proteins of halophilic organisms (Fig. 1, A and B). The overall number of charged residues in Hsmr is eleven, which is a high number of two histidines and one highly conserved lysine at position 22. A high content of negatively charged amino acids has been reported before in other halophilic proteins, and is suggested to have a stabilizing effect at high salt concentrations (19).

The proportion of hydrophobic residues is similar in the two proteins; however, the composition is very different. Hsmr contains more alanine and valine residues, 19 and 26 compared with 9 and 5 in EmrE, respectively. Thus, Hsmr displays a remarkable amino acid composition of over 40% valine and alanine residues. In parallel, the amount of isoleucine residues also changes from 15 in EmrE to only 3 in Hsmr. A small difference is observed in the number of glycines (11 in Hsmr, 12 in EmrE), leucines (12 and 16, respectively) and phenylalanines (4 and 5, respectively) (not shown). Hsmr contains no cysteine residues and a single tryptophan at position 63, a fully conserved residue. Stevens and Arkin (20) found a strong correlation between the GC content of a genome, and the ratio of (Val + Ala+Ile+Phe) residues of its membrane proteins. The higher the GC content of a genome, the higher the amount of valine and alanine residues. This correlation is explained by a codon bias toward GC-rich codons, together with a need to retain a certain level of hydrophobicity in a membrane-spanning region. In hsmr the codons used for alanine and valine are, exclusively, the ones with the highest GC content: GTG and
Hsmr Catalyzes Ethidium Efflux in Whole Cells—We further explored Hsmr function by directly testing its ability to catalyze the transport of ethidium bromide in whole cells. Entry of ethidium into cells is manifested by an increase in the fluorescence due to binding to nucleic acids (7, 21). Cells overexpressing either Hsmr or mock vector were treated with the uncoupler CCCP to facilitate the entry of ethidium into the cells, as previously described (7). After 60 min of incubation with ethidium (1 μg/ml) and CCCP (40 μM) the cells were collected by centrifugation and rapidly resuspended in medium containing the same concentration of ethidium, but without uncoupler. Upon removal of the uncoupler from Hsmr overexpressing cells, a rapid decrease in fluorescence is observed, even before the tracing begins, and is completed in less than 7 min (Fig. 3, D, left). This decrease in fluorescence, which represents removal of ethidium from the cells against a concentration gradient, is completely reversed upon addition of CCCP (Fig. 3D, left). As expected, in cells harboring a control vector no decrease in fluorescence was observed (Fig. 3D, right). These results demonstrate that Hsmr catalyzes the active removal of ethidium in whole cells.

The Hsmr Dimer Is Resistant to Denaturant Treatment—Expression levels were tested in membranes prepared from IPTG induced *E. coli* BL21 (DE3) cells harboring the plasmid pT7-7 Hsmr-His. The His-tagged Hsmr protein was purified from the membranes using a Ni-NTA column. The purified protein was analyzed on SDS-PAGE, alongside purified samples of EmrE (Fig. 4A). Hsmr reaches high levels of expression in *E. coli* membranes, up to 2 mg per 1 liter culture. As seen in Fig. 4A, the EmrE monomer displays an apparent *M* of 14,000. This value is close to the theoretical value of molecular mass.
Fig. 4. Expression, purification and activity of Hsmr. E. coli BL21 (DE3) cells carrying the plasmid pT7-7 Hsmr-His were grown in rich media, and protein expression was induced using 0.5 mM IPTG for two hours. The cells were harvested, and membranes were obtained by a single pass through a microfluidizer device, followed by ultracentrifugation. A. His-tagged Hsmr was purified from the E. coli membranes by solubilization in SDS-urea buffer and binding to Ni-NTA beads. A sample (equivalent to 25 µg total membrane protein) was analyzed on SDS-PAGE alongside a sample of purified EmrE (0.5 µg). B, increasing amounts or the E. coli membranes expressing Hsmr were solubilized in 1% DDM and assayed for [3H]TPP binding as described under “Experimental Procedures.”

calculated for a protein the size of EmrE or Hsmr (not shown). Surprisingly, Hsmr displays an apparent $M_r$ of 30,000, suggesting that the dimeric complex of Hsmr does not dissociate, even in the presence of SDS. Furthermore, a small amount of protein corresponding to the Hsmr tetramer can also be observed (Fig. 4A). The protein was challenged with an anti-His antibody and also specifically radiolabeled, and in both cases it co-migrated with the Coomassie-stained protein (not shown). Attempts to dissociate the dimer were made by modifying the treatment of the sample before loading on the gel. These treatments included solubilization in various detergents and boiling of the purified protein; however, they yielded no visible monomers on the gel. Monomers were observed only after extraction of diluted protein in a mixture of chloroform and methanol (not shown).

Solubilized Hsmr Binds [3H]TPP and Interacts with Several Other Drugs—A very convenient functional assay of detergent-solubilized EmrE has been previously developed (12). Here we show that DDM-solubilized Hsmr also specifically binds [3H]TPP in a dose-dependent manner (Fig. 4B) in the range up to 50 µg of protein, leveling off at the high protein amounts, because of substrate depletion. To test interaction with other known substrates of EmrE, the compounds were tested for their ability to inhibit the binding reaction. The results are shown in Fig. 5. Methyl viologen had no effect on [3H]TPP binding to Hsmr, as expected from the inability of Hsmr to confer resistance against this compound in living cells (see Fig. 3C). Benzalkonium, ethidium, and acriflavine inhibit [3H]TPP binding to Hsmr in the µM range (Fig. 5), suggesting that they specifically interact with Hsmr. These results are in accordance with the phenotype presented in Fig. 3, and together they demonstrate the function of Hsmr as a multidrug protein, which can interact in a specific manner with different drugs.

Salt Dependence of Hsmr—H. salinarum, from which the Hsmr protein originates, normally resides in extreme environments where the salt concentration reaches very high values of 4–5 M (22). It was therefore interesting to examine the effect of salt on the activity of Hsmr. For that purpose solubilized membranes from cells expressing either Hsmr or EmrE were assayed for [3H]TPP binding at increasing concentrations of NaCl, KCl, or LiCl. Hsmr-binding activity is markedly stimulated in the presence of KCl, NaCl, and LiCl (Fig. 6A). As the concentration of salt rises from zero to 3.2 M an increase in the binding activity is observed, saturating around the high salt concentration. Both KCl and NaCl brought about a 6–7-fold increase in the binding upon addition of about 3.0 M salt. A smaller effect is seen in LiCl that induces only a 4-fold increase at similar concentrations. The binding activity of EmrE is not affected by the salt concentration and remains constant up to 3 M KCl, NaCl, or LiCl (Fig. 6B). To understand the mechanism underlying this salt effect we measured the binding affinity at the different salt concentrations. We found that the binding affinity increases 5-fold, from 200 to 40 nM, upon the addition of 2 M NaCl.

DISCUSSION

The first Smr protein from the archaeon H. salinarum was identified based on sequence similarity. Its unusual alanine and valine content, and the clustering of these residues in
certain domains, suggested possible functional implications. Therefore, we set out to probe the activity of the protein. Hsmr was functionally expressed in E. coli membranes, to high levels. Proteins from extremely halophilic bacteria are adapted to a very high salt environment (above 3 M), and usually denature in the absence of high salt, resulting in a non-functional protein (22). Nonetheless, expression of the hsmr gene in E. coli results in the production of a protein that is functional in vivo as judged from drug resistance phenotype and from a whole cell ethidium efflux assay. Hsmr is active in a wide range of conditions because it does not strictly require salt for its activity. It is also able to function in the E. coli membrane, which has a lipid composition very different from the archaeal membrane it originated from (23, 24). The stability of the protein is also apparent when analyzed on SDS-PAGE. The Hsmr dimer stays intact even at high concentrations of SDS. This unusual phenomenon of an SDS-stable oligomer has been reported before with the dimer of glycoporphine A (25) as well as for the phospholamban pentamer (26). Attempts to dissociate the Hsmr dimer have been made by using different solubilization conditions and by heating the protein sample before loading on the gel. The protein retained its “dimeric” mobility on the gel unless diluted and extracted in an organic solvent. Hsmr is also functional in vitro, in its DDM-solubilized form, where it specifically binds the substrate TPP⁺ in a dose-dependent manner. Although the protein retains its binding activity even in the absence of salt, this activity is greatly enhanced at high salt concentrations as a result of an increase in the affinity of the protein to TPP⁺. This salt effect appears to be a general one because it is observed in both NaCl and KCl as well as in LiCl and MgCl₂ (not shown) to a lesser extent.

The Smr proteins are not as ubiquitous in the Archaea as in the Eubacteria. Only three other homologues can be identified until now, all of them in various Methanosarcinoma species (not shown). Hsmr was identified based on sequence similarity to the Smr family, yet was found to have a unique sequence composition. Compared with other Smr proteins, Hsmr contains many negative charges, located mostly at the hydrophilic loops of the protein. This phenomenon is observed in many halophilic proteins and is said to contribute much to their stability in high salt (27). Hsmr has only one positive charge, Lys-22, which is fully conserved in the Smr family, and two histidine residues facing the opposite side of the membrane. Because there are so few positive charges, which are distributed quite symmetrically on both sides of the membrane, it is difficult to apply the von Heijne rule for topology prediction (28) to this protein. The inapplicability of this rule with Smr proteins has been demonstrated before and was suggested to result from their small size (18).

The most striking feature of Hsmr is the fact that it is built of over 40% valine and alanine residues, clustered at certain regions of the protein. This clustering is in domains that do not seem important for activity. According to Stevens and Arkin (20), and discussed earlier, there exists a strong positive correlation between the GC content of a genome and the amount of valine and alanine residues found in membrane proteins it codes for. This correlation is explained by a codon bias toward amino acids coded predominately by GC nucleotides, together with a need to retain certain physical properties within a trans-membrane domain. The rise in the abundance of valine residues is also thought to be compensating for the lack of AT-coded amino acids, particularly isoleucine, with which it shares similar physical properties. The alanine and valine codons are GCN and GTN, respectively, of which only GGC and GCC (for alanine) and GTG and GTC (for valine) are present in the hsmr gene. Although codon bias could be fulfilled also by glycine and leucine, the frequency of these amino acids does not increase in proteins from organisms with high GC content (20). We speculate that glycine would not be hydrophobic enough and could allow too much flexibility of the helix; leucine may be bulkier than necessary.

The distribution of valine and alanine residues in the first trans-membrane domain of Hsmr correlates well with information deduced from extensive site-directed mutagenesis of the same trans-membrane domain in the homologous protein EmrE. EmrE mutations in residues in the same face of the helix as Glu-14 have effects on the catalytic activity of the protein, whereas mutations on the other face of the helix have no effect (31). Strikingly, the face tolerant to mutations is the one predominated by alanine residues in Hsmr. Without a high-resolution structure of an Smr we can only speculate about the location of an alanine-rich face in a membrane protein. Alanine is a small, relatively hydrophobic amino acid. Because of its size it could provide an interface for tight packing of two helices (30). On the other hand, its hydrophobicity is high enough for interaction with lipids in the membrane.

Hsmr is therefore a case of natural alanine and valine mutation. Multiple sites in the protein, located in regions not important for activity, have become through evolution occupied by valine or alanine residues. The advantage of analyzing sequences such as that of Hsmr is that the outcome of replacements in more than one position can be seen within the same protein. Multiple replacements are more informative than single ones. The latter depend at times on the properties of amino acids in positions other than the ones modified. This could be because of size, charge, or other properties. Multiple replacements are difficult to design, therefore the analysis of valine- and alanine-rich sequences could greatly facilitate their planning or even serve as ready made “mutants,” a quick alternative for experimentation.

We found that the valine and alanine clustering is observed in additional membrane proteins from H. salinarum, and therefore may be part of a general phenomenon. A full genomic and statistical analysis is required to conclude whether this phenomenon can be used to reveal significant sequence elements in membrane proteins in general. If this is indeed a general phenomenon it could be used to devise an algorithm predicting important regions within a Val+Ala-rich membrane protein sequence, without the need for additional information. This prediction, in turn, could be projected on to non-Val+Ala-rich homologues, providing a general tool for a quick preliminary prediction of significant regions within a given membrane protein.
18. Yerushalmi, H., and Schuldiner, S. (2000) J. Biol. Chem. 275, 5264–5269
19. Mevarech, M., Frolov, F., and Gloss, I. M. (2000) Biophys. Chem. 86, 155–164
20. Stevens, T. J., and Arkin, I. T. (2000) Protein. Sci. 9, 505–511
21. Lambert, B., and Le Pecq, J. B. (1984) Biochemistry 23, 166–176
22. Madern, D., Ebel, C., and Zaccai, G. (2000) Extremophiles 4, 91–98
23. Albers, S. V., van de Vossenberg, J. L., Driessen, A. J., and Konings, W. N. (2001) Extremophiles 5, 285–294
24. van de Vossenberg, J. L., Driessen, A. J., and Konings, W. N. (1998) Extremophiles 2, 163–170
25. Furthmayr, H., and Marchesi, V. T. (1976) Biochemistry 15, 1137–1144
26. Watanabe, Y., Kijima, Y., Kadoma, M., Tada, M., and Takagi, T. (1991) J. Biochem. (Tokyo) 110, 40–45
27. Elcock, A. H., and McCammon, J. A. (1998) J. Mol. Biol. 280, 731–748
28. von Heijne, G. (1986) EMBO J. 5, 3021–3027
29. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132
30. Jiang, S., and Vakser, I. A. (2000) Proteins 40, 429–435
31. Gutman, N., Steinr-Morisch, S., and Schuldiner, S. (February 16, 2003) J. Biol. Chem. 10.1074/jbc.M213120200