Characterization of genes for the biosynthesis of the compatible solute ectoine from *Marinococcus halophilus* and osmoregulated expression in *Escherichia coli*

Petra Louis and Erwin A. Galinski

The genes of the biosynthetic pathway of ectoine (1,4,5,6-tetrahydro-2-methyl-4-pyrimidinecarboxylic acid) from the Gram-positive moderate halophile *Marinococcus halophilus* were cloned by functional expression in *Escherichia coli*. These genes were not only expressed, but also osmoregulated in *E. coli*, as demonstrated by increasing cytoplasmic ectoine concentration in response to medium salinity. Sequencing of a 44 kb fragment revealed four major ORFs, which were designated *ectA*, *ectB*, *ectC* and *orfA*. The significance of three of these genes for ectoine synthesis was proved by sequence comparison with known proteins and by physiological experiments. Several deletion derivatives of the sequenced fragment were introduced into *E. coli* and the resulting clones were investigated for their ability to synthesize ectoine or one of the intermediates in its biosynthetic pathway. It was demonstrated that *ectA* codes for \( \gamma\)-2,4-diaminobutyric acid acetyltransferase, *ectB* for \( \gamma\)-2,4-diaminobutyric acid transaminase and *ectC* for L-ectoine synthase. A DNA region upstream of *ectA* was shown to be necessary for the regulated expression of ectoine synthesis in response to the osmolarity of the medium.

**Keywords**: *Marinococcus halophilus*, compatible solutes, ectoine genes, osmoregulation, salt stress

**INTRODUCTION**

Saline environments are characterized by high osmotic strength (low water potential). Most halophilic eubacteria cope with these conditions by accumulating small, highly water-soluble organic compounds, the so-called compatible solutes (Brown, 1976). These osmolytes enable organisms to adapt to a wide range of salt concentrations by adjusting the cytoplasmic solute pool to the osmolarity of the surrounding environment. Ectoines represent the predominant class of osmolytes in aerobic chemoheterotrophic eubacteria (Severin *et al.*, 1992; Frings *et al.*, 1993; Galinski, 1995). The biosynthetic pathway for ectoine has been elucidated at the enzymological level in Gram-negative eubacterial halophiles (Peters *et al.*, 1990; Tao *et al.*, 1992; Galinski & Trüper, 1994). It comprises three steps, the first being the conversion of aspartate semialdehyde, an intermediate in amino acid metabolism, to \( \gamma\)-2,4-diaminobutyric acid. This is followed by acetylation to \( \gamma\)-acetyldiaminobutyric acid. The last step consists of a cyclic condensation reaction to form the tetrahydro-pyrimidine ectoine (Fig. 1).

Expression and regulation of genes involved in osmoadaptation has thus far been investigated almost exclusively in non-halophilic bacteria, especially *Escherichia coli* (for recent reviews see Csonka & Hanson, 1991; Lucht & Bremer, 1994; Galinski, 1995). This organism responds to increased salinity with the rapid accumulation of potassium and concomitant synthesis of glutamate as a counter anion. Subsequently, these charged solutes are partially replaced by endogenous trehalose or compatible solutes accumulated from the medium, if present (Dinnbier *et al.*, 1988). Most studies at the molecular level have so far focused on the various solute uptake systems of *E. coli* (Altenhof & Epstein, 1993; Mellies *et al.*, 1995; Gowrishankar & Manna, 1996). The only investigations concerned with the biosynthesis of compatible solutes covered choline oxidation and trehalose synthesis (Lamark *et al.*, 1996; Strøm & Kaasen, 1993). It has been proposed that relatively non-specific signals, such as intracellular
potential glutamate concentration and DNA supercoiling (Booth & Higgins, 1990), as well as specific regulatory mechanisms of the different genes investigated play a role in osmoregulation.

In this study, we report the identification and sequencing of the three genes responsible for the synthesis of the compatible solute ectoine from the Gram-positive halophile Marinococcus halophilus and their osmoregulated expression in E. coli. Our investigation provides, for the first time, an opportunity to compare the molecular organization of osmoregulatory elements from a true halophile with that of non-halophiles.

**METHODS**

**Bacterial strains, growth conditions and plasmid.** M. halophilus DSM 20408T was grown aerobically at 37°C in LB medium (Miller, 1972) containing 0-3% (w/v) artificial sea salt and 37% (w/v) NaCl. E. coli XL1-Blue (Stratagene) was grown aerobically at 37°C either in LB medium or in medium MM63 (Larsen et al., 1987) with 1.5 ml vitamin solution 1-3 (Imhoff & Trüper, 1977) and 1-5% NaCl. To select for cells bearing plasmid pHSG575 and recombinant derivatives, the medium contained chloramphenicol at a final concentration of 25 μg ml⁻¹. For the selection of cells containing recombinant plasmids on agar plates, isopropyl β-D-thiogalactopyranoside (IPTG, 0.5 mM) and 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-Gal, 40 μg ml⁻¹) were added. For supplementation studies, MM63 with 2-3% NaCl contained 2 mM diaminobutyrate or N-acetyldiaminobutyrate. At the end of the exponential growth phase, cells were harvested (15 000 g, 18°C) and freeze-dried.

Plasmid pHSG575 (Takeshita et al., 1987), a lacZ α-complementing low-copy-number cloning vector, was kindly provided by E. Bremer (Universität Marburg, Germany).

**DNA techniques.** Genomic DNA of M. halophilus was prepared using Qiagen genomic-tips as described by the manufacturer. Small- and large-scale plasmid preparations of E. coli XL1-Blue were performed using Qiaprep spin columns and the Plasmid Midi kit from Qiagen, respectively.

For the construction of a library, genomic DNA from M. halophilus was partially digested with the restriction endonuclease Sau3A and separated on an agarose gel. Fragments ranging from 5 to 15 kb were recovered by electroelution and ligated into the dephosphorylated BamHI site of pHSG575 according to standard techniques (Ausubel et al., 1991). E. coli XL1-Blue was transformed with the recombinant plasmids by the CaCl₂ method (Cohen et al., 1972).

Southern hybridization analysis was performed on nitrocellulose membranes (Schleicher & Schuell) according to standard techniques (Ausubel et al., 1991) using the Nonradioactive Digoxigenin DNA Labeling and Detection kit from Boehringer Mannheim for the preparation of DNA probes.

Plasmids pOSM11, pOSM12, pOSM13, pOSM14 and pOSM18 were constructed by deletion of defined restriction fragments from pOSM1 and religation after the creation of blunt ends with the Klenow fragment of DNA polymerase I. Restriction enzymes and the Klenow fragment were purchased from Boehringer Mannheim with the exception of PpuMI, which was obtained from New England Biolabs.

Nested deletions for sequencing were prepared with the Double Stranded Nested Deletion kit (Pharmacia). Sequencing was performed with the AutoRead 1000 Sequencing kit and the automatic fluorescent sequencer ALF (Pharmacia) in one direction. Sequencing of the opposite strand was performed by Sequisers (München, Germany).

DNA sequences were analysed with the programs GENEPRO, DNASEIS and MACAW. Databank searches were carried out through the National Center for Biotechnology Information (NCBI) with the BLAST program (Altschul et al., 1990) and current versions of the available databases (October 1996).

**Analytical methods.** For the identification of intracellular solutes, freeze-dried cells were extracted with methanol/chloroform/water (10:5:4) by a modification of the technique of Bligh & Dyer (1959), as described by Galinski & Herzog (1990). Extracts were analysed by isocratic and gradient HPLC methods as described previously (Galinski & Herzog, 1990; Kunte et al., 1993).

**RESULTS AND DISCUSSION**

**Cloning and osmotic expression of ectoine genes in E. coli**

E. coli is able to grow at salt concentrations of up to 3% NaCl in the minimal medium MM63 because of its intrinsic ability to adjust its cytoplasmic potassium glutamate and trehalose pools. At 5% NaCl, more efficient compatible solutes, such as betaine or ectoine, are needed for growth. Our aim was to provide E. coli XL1-Blue with genes for the synthesis of the compatible solute ectoine from M. halophilus, and thus increase its osmotolerance. To achieve this, colonies of E. coli XL1-Blue carrying a genomic DNA library of M. halophilus were replica-plated onto medium MM63 containing
pressed almost completely (not shown). Instead, betaine
trehalose levels on the other hand were considerably
with the current view that uptake of compatible solutes
in the ectoine-free host
comparable to that gained through ectoine uptake in
an osmoregulated manner rather than simply excreted
HPLC. Besides trehalose, we detected the compatible
cellularly, we conclude that ectoine was synthezised in
lower in
is preferred over biosynthesis (Dinnbier
NaCl, the observed biosynthesis of ectoine was sup-
E. coli MC4100, as observed by Jebbar
al., 1988). As we could not detect ectoine extra-
the medium. The trehalose pool
Fig. 2. Relationship between the intracellular solute
centration of E. coli(pOSM11) and medium salinity during
growth in minimal medium MM63 as measured by isocratic
HPLC. Mean values and standard deviations from three
independent experiments are shown. Similar results were
obtained with E. coli(pOSM1). Glutamate levels did not exceed
0.12 mmol (g dry wt)$^{-1}$ regardless of the salt concentration
of the medium. The trehalose pool of E. coli XL1-Blue in the
presence of 3% NaCl was 0.35 mmol (g dry wt)$^{-1}$.
this particular sequence in genomic DNA of M. halophilus (not shown). This proves that the cloned DNA
originated from M. halophilus.

**Sequencing of pOSM11**

Restriction analyses of pOSM1 revealed a recombinant
fragment of 5.9 kb. Following the deletion of a 1.5 kb
EcoRI fragment, we were able to demonstrate that the
resulting plasmid pOSM11, containing a 4.4 kb recom-
binant fragment, still displayed osmoregulated ectoine
synthesis (Fig. 2).

The entire recombinant DNA fragment of pOSM11 was
sequenced (Fig. 3). Four major ORFs, oriented in the
same direction, were identified. They are predicted to
encode proteins of 172, 427, 129 and 110 amino acids,
with deduced molecular masses of 19385, 47192, 14796
and 13142 Da, respectively. All reading frames, which
we designated ectA, ectB, ectC and orfA, are preceded
by putative ribosome binding sites (Fig. 3). Downstream
of orfA lies the 5'-end of another ORF (orfB). The lacZ
promoter of the vector is oriented in the opposite
direction, and thus transcription of these reading frames
from the lacZ promoter can be excluded.

The gene for the ectoine synthase from Halomonas sp.
had already been cloned by reverse translation and
sequenced by Min-Yu et al. (1993). We compared their
sequence with our ORFs and found 47% identity to
ectC. An alignment is shown in Fig. 4. Thus, ectC seems
to encode the ectoine synthase from M. halophilus.

Databank searches using the blast algorithm (Altschul
et al., 1990) revealed a high similarity of ORF ectB to
transaminases belonging to subgroup II (Mehta et al.,
1993). All of these transaminases aminate a terminal o xo
group. The same holds true for 1-2,4-diaminobutyric
acid transaminase in the biosynthetic pathway for
ectoine. The highest similarity was observed with $\gamma$
aminobutyrate-, ornithine- and acetylornithine trans-
aminas. An alignment is shown in Fig. 5. Of four residues
found in all transaminases examined previously
(Mehta et al., 1993), three were present in ectB (see $\n$ in Fig. 5). The fourth, an arginine, which binds the $\alpha$
carboxylate of the substrate, is replaced by a lysine in
ectB. Studies with aspartate transaminase have shown
that when arginine is replaced by lysine, which also
carries a positive charge, enzyme activity is maintained,
albeit at a strongly reduced rate (Inoue et al., 1989).
Considering that replacement of arginine with lysine is
likely to have a greater steric effect on $\alpha$-aminating
enzymes like aspartate transaminase, this exchange may
have less relevance for a $\gamma$-aminating transaminase,
especially as this particular region is not very conserved
(Fig. 5). From these sequence comparisons, we conclude
that ectB is the first gene in the biosynthetic pathway of
ectoine, the transaminase converting aspartate-$\beta$-semi-
aldehyde to diaminobutyric acid.

The ectA gene revealed no sequence similarity to known
proteins, whereas orfA and the truncated orfB shared
Fig. 3. For legend see facing page.
Genes for ectoine biosynthesis

ATTAATAATAAAAAATCCCTGATGCCCTTTAGTGCACCCGATATAATGACAGCCAGCCGAAAAAACCAACCCCCAACATCCTGATTATATGATTACTTTAGCTT

Fig. 3. Nucleotide sequence of the recombinant fragment of pOSM11. Deduced amino acid sequences of ORFs designated ectA, ectB, ectC, orfA and orfB (5'-end) are shown below the sequence. Stop codons are marked with asterisks. Putative ribosome-binding sites are boxed. Consensus sequences for two putative σ70-dependent promoters (consensus TTGACA-N_5'-TATAAT) are underlined. Putative consensus sequences for a σB-dependent promoter (consensus GllTAA-N_5'-GGGTAT) are marked with dotted lines. Palindromic sequences are marked by arrows.

similarities with ORFs of the same name believed to encode transposases of IS elements.

Ectoine synthesizing capacity of deletion derivatives of pOSM11

In order to confirm the gene assignments based on sequence similarities and to determine the potential role of ectA and orfA in ectoine synthesis, we constructed several deletion derivatives of pOSM11 and examined them physiologically (Fig. 6). Analysis of pOSM12, which was deficient in the sequences 300 bp downstream of ectC and only retained a part of orfA, revealed the same osmoregulated pattern of ectoine synthesis as pOSM11 in E. coli XL1-Blue. This suggests that orfA is not involved in the ectoine biosynthetic pathway. pOSM13 carried a deletion within ectC but still contained ectA and ectB and the sequences upstream of ectA. No ectoine synthesis was observed; instead the precursor N'-acetyldiaminobutyric acid was accumu-
lated. This confirms that \textit{ectC} encodes ectoine synthase, which converts $N^2$-acylaminobutyric acid to ectoine.

We constructed a subclone containing the complete ORF \textit{ectA} and the upstream sequences (pOSM14). No accumulation of diaminobutyric acid or $N^2$-acylaminobutyric acid was observed, which is a further indication that \textit{ectB} encodes the diaminobutyrate transaminase, the first enzyme in the ectoine pathway. When this subclone was supplemented with diaminobutyric acid, it synthesized $N^2$-acylaminobutyric acid, whereas \textit{E. coli} XL1-Blue, containing only the vector pHGS575, was not able to acetylate diaminobutyric acid but simply accumulated it from the medium. This confirms our assumption that \textit{ectA} is responsible for the acetylation step. In summary, our subcloning studies revealed that \textit{ectA} encodes diaminobutyric acid acetyltransferase, \textit{ectB} diaminobutyric acid transaminase and \textit{ectC} ectoine synthase required for ectoine biosynthesis.

\textbf{Osmoregulation and genetic organization of ectoine genes}

Further deletion derivatives of pOSM11 were constructed to determine the approximate position of osmoregulating sequences in the promoter region. A

\textit{Fig. 5.} Alignment of the sequences of reading frame \textit{ectB} with transaminases. ECGABT, 4-aminobutyric acid transaminase from \textit{E. coli} (Bartsch et al., 1990); BSACOAT, acetylornithine transaminase from \textit{B. subtilis} (O’Reilly & Devine, 1994); BSOAT, ornithine transaminase from \textit{B. subtilis} (Gardan et al., 1995). Identical positions are shown against a dark background and conservative replacements are boxed. Conservative replacements are defined in the legend to Fig. 4. Residues identical in all transaminases examined so far according to Mehta et al. (1993).
deletion of the first 384 bp of the insert upstream of ectA (pOSM15, Fig. 6) had no effect on ectoine synthesis. Deleting a further 190 bp, however, apparently destroyed regulating sequences (pOSM16, Fig. 6), as the latter subclone showed constitutive ectoine synthesis. The cytoplasmic concentration ranged between 0.5 and 0.6 mmol (g dry wt)\(^{-1}\), independent of the salt concentration of the medium. It can therefore be concluded that osmoregulating sequences are located at least 150 bp upstream of the first ORF.

Two possible consensus sequences for \(\sigma^{70}\)-dependent promoters were found upstream of ectB (Fig. 3). A search for consensus sequences of the osmoresponsive promoters for the compatible solute transport systems proU, prop and opuA (Mellies et al., 1994, 1995; Kempf & Bremer, 1995) revealed no matches. As several osmoregulated genes in non-halophiles, including the genes for the biosynthesis of trehalose, are known to be under the control of \(\sigma^{70}\)-dependent promoters (Strøm & Kaasen, 1993; Gordia & Gutierrez, 1996; Manna & Gowrishankar, 1994; Mellies et al., 1995), we also conducted a search for consensus sequences of these (Strøm & Kaasen, 1993), but were unable to find any matches. Instead, we found a sequence similar to the consensus for \(\sigma^{70}\) of Bacillus subtilis (Fig. 3) known to transcribe the so-called 'general stress proteins', which are induced by various stimuli, such as salt stress, ethanol treatment and starvation (Hecker et al., 1996). An imperfect palindrome lies within the probable \(-10\) region.

A perfect palindrome downstream of ectA overlapping with one of the putative \(\sigma^{70}\) promoter sequences (Fig. 3) may represent a Rho-dependent terminator. Therefore, the question arises as to whether the three genes are transcribed separately or as a single operon. The subclone E. coli (pOSM17), with a deletion of both the upstream sequences and part of ectA (Fig. 6), contained the whole reading frame ectB as well as the possible \(\sigma^{70}\) consensus sequences and should therefore produce diaminobutyrate provided that ectB is transcribed from these putative promoters. However, we were not able to detect this compound inside the cells. Furthermore, another subclone containing pOSM18 (Fig. 6), which retained only the complete reading frame ectC, was also unable to produce ectoine when supplemented with \(N^7\)-acetyldiaminobutyrate. These findings indicate that ectB and ectC are not transcribed independently. However, in order to ascertain that the ectoine genes are transcribed as a single operon, further studies at the RNA level are necessary.

In this study we have, for the first time, identified genes for the de novo biosynthesis of a compatible solute (ectoine) in addition to sequences necessary for its osmoregulated expression from a truly halophilic eubacterium. We were able to demonstrate that the osmoregulatory sequences of M. halophilus also function in E. coli. Earlier work on osmoregulated transport systems has also shown the potential of gene transfer. An osmoprotectant transporter from Erwinia chrysanthemi, encoded by the gene ousA and analogous to prop from E. coli, displayed osmoregulated expression in E. coli, although no DNA homology upstream of the coding sequences of prop and ousA was identified (Gouesbet et al., 1996). In addition, the betaine uptake system opuA from B. subtilis (Kempf & Bremer, 1995) was also shown to be under osmotic control in E. coli. In this case, promoter consensus sequences resembling those of prop and proU were identified. These observations support the view that there are features in osmoregulation that are shared between different, even phylogenetically distant, bacteria. Further studies re-
garding the osmoregulation of the ectoine genes from *M. halophilus* will provide a deeper insight into these mechanisms.

**ACKNOWLEDGEMENTS**

We would like to thank Erhard Bremer for providing plasmid pHSG575 and Martin Däumer for his valuable help with the sequencing. We also thank Christiane Dahl for her critical reading of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (Ga 393/3) and by the European Union (Bio2-CT93-0274).

**REFERENCES**


Takeshita, S., Sato, M., Toba, M., Masahashi, W. & Hashimoto-


Received 11 November 1996; accepted 21 November 1996.