An operon encoding a novel ABC-type transport system in *Bacillus subtilis*

Francesco Rodriguez and Guido Grandi

Author for correspondence: Guido Grandi. Tel: +39 2 5205970. Fax: +39 2 52022974.

**Downstream from the surfactin synthetase operon in *Bacillus subtilis* a new operon-type structure has been localized which, on the basis of sequence determination, potentially encodes an ABC-type transport system. The 268 amino acid protein, the product of *orf1*, represents the solute-binding component of the system whereas the *orf2* product, a 234 amino acid protein, is the transmembrane component. Finally *orf3* potentially encodes a typical 241 amino acid ATP-binding protein involved in energy supply. Comparison of the three proteins with the subunits of other ABC-type systems suggests that this new system is involved in amino acid transport.**

**Keywords:** transport systems, amino acid binding, ATP binding, lipoproteins, *Bacillus subtilis*

In Gram-negative bacteria numerous solutes such as sugars, amino acids, peptides, anions and metals are transported across the cytoplasmic membrane by the ABC-type (or ATP-binding cassette type) multi-component transport systems. The common protein components of these systems include two transmembrane proteins that usually span the membrane about six times each, one or two peripheral-membrane ATP-binding protein(s) localized on the cytoplasmic side of the membrane and a high-affinity solute-binding protein which, in Gram-negative bacteria, is periplasmic. The function of the transmembrane component is to channel the solute to the cytoplasm whereas the ATP-binding protein provides energy to the system. Finally, the ligand-binding protein confers specificity and affinity (Higgins, 1992).

The presence of ABC-type transport systems in Gram-positive bacteria has only recently been documented. In particular, among the most extensively characterized are the *amiA* system of *Streptococcus pneumoniae* (Alloing et al., 1990) responsible for the uptake of oligopeptides, the oligopeptide transport systems *spoOK* and *app* of *Bacillus subtilis* (Perego et al., 1991; Rudner et al., 1991; Koide & Hoch, 1994), the *rbs* and *deiA* systems of *B. subtilis* (Woodson & Devine, 1994; Mathiopoulos et al., 1991), the function of which is to transport ribose and dipeptides, respectively, and the glutamine transport system of *Bacillus stearothermophilus* (Wu & Welker, 1991). On the basis of the information obtained from the characterization of these systems it appears that their overall structural organization highly resembles their Gram-negative counterparts. The most relevant difference is found in the solute-binding proteins. In fact in Gram-positive bacteria they are lipoproteins anchored to the external surface of the cell membrane with an N-terminal glycine-cysteine (for a review see von Heijne, 1989).

In the course of our sequencing work on the surfactin synthetase operon (Cosmina et al., 1993; van Sinderen et al., 1993) we identified an operon-type structure potentially encoding three proteins which, on the basis of sequence homology, are likely to form an ABC-type amino acid transport system. This putative operon is located upstream from the *sfp* gene. A partial characterization of this region has been recently reported by Grossman et al. (1993), who found it to be part of a *B. subtilis* chromosomal region able to complement siderophore-deficient mutants of *Escherichia coli*. Interestingly the authors demonstrated that the complementing factor was the product of the *sfp* gene.

To our knowledge, this is the first putative amino acid transport system fully characterized in *B. subtilis*.

A 300 bp *San3A* fragment located approximately 3.3 kbp downstream from the *srfA* operon was inserted into the *BamHI* site of the integrative plasmid pJM103 (Perego et al., 1991). The recombinant plasmid was then used to transform *B. subtilis* strain 168, selecting for chloramphenicol-resistant clones. From one of the transformants which, upon Southern blot and PCR analyses, turned out to have the plasmid correctly inserted (data not shown), the chromosomal DNA was prepared, digested with *PstI* and self-ligated. The ligase mixture was utilized

---

The EMBL accession number for the nucleotide sequence reported in this paper is X77636.
to transform E. coli JM109 competent cells and among the several transformants obtained six were analysed for plasmid content. Plasmid analysis showed that five of the randomly selected clones harboured a plasmid containing the same chromosomal fragment of approximately 6 kb. One of these plasmids, named pSM714, was further characterized.

The sequencing of the 5828 bp fragment was accomplished using the dideoxy chain-termination procedure (Sanger et al., 1977) and the Sequenase version 2.0 enzyme (USB) on pSM714 derivatives obtained upon progressive ExoIII deletions of the cloned region (Henikoff, 1984). The fragment was sequenced on both strands with a mean of three readings per base.

Fig. 1 shows the physical and genetic map of the region within which, approximately 850 bp upstream from sfp, three ORFs (orf1, orf2 and orf3), probably organized in a single transcriptional unit, were found. The nucleotide sequence of orf1, orf2 and orf3 is given in Fig. 2. The 804 bp long orf1 potentially codes for a protein of 268 amino acids whereas orf2 is 702 bp in length (protein of 234 amino acids). The 3' end of orf1 and the 5' end of orf2 overlap by 11 bp. orf3 is 741 bp long and is located 13 bp downstream from the 3' end of orf2. Putative ribosome-binding sites were found upstream from the ATG start codon of both orf1 and orf3 (underlined in Fig. 2) whereas the orf2 start codon appears to be preceded by a less canonical sequence. Two putative rho-independent terminator structures are located downstream from orf3 with a calculated ΔG value of -24.2 kcal mol\(^{-1}\) (\((-101.64)\) kJ mol\(^{-1}\)) and \((-18.2)\) kcal mol\(^{-1}\) (\((-76.44)\) kJ mol\(^{-1}\)), respectively. Finally, a putative \(\sigma^\ast\) promoter (TTGTA\(A\)-18 bp-TAA\(A\)AT) located 23 bp upstream from the orf1 start codon might function as the transcriptional start for the three genes. Although not further discussed here, two additional ORFs, one located upstream from orf1 and the other downstream from orf3, are present in the 6 kb PstI fragment.

When the homologies of Orf1, Orf2 and Orf3 with other proteins were analysed it appeared that the three proteins are likely to form an ABC-type transport system. In particular, Orf1 represents the solute-binding component of the system, sharing homology with the glutamine-binding subunit GlnH of E. coli (31% identity) (Nohno et al., 1986), the arginine- and histidine-binding subunits ArgT and HisJ of Salmonella typhimurium (26% and 25% identity, respectively) (Higgins et al., 1982) and the glutamine-binding subunit of B. stearothermophilus (21% identity) (Wu & Welker, 1991). Relevant homologies were also found with other solute-binding components such as NocT and OccT of Agrobacterium tumefaciens (23% and 21% identity, respectively) (Valdivia et al., 1991; Zanker et al., 1992) and ArtJ and ArtI of E. coli (24% and 27% identity, respectively) (Wissenbach et al., 1993). Interestingly, Orf1 has a typical leader sequence for secretion and the potential cleavage site (LAA-CGA) corresponds to the consensus sequence for the precursors of lipoproteins (von Heijne, 1989).

Orf2 has significant homology with the glnP gene product of E. coli (38% identity) and with the hisQ and hisM gene products of S. typhimurium (27% and 32% identity, respectively). Like these proteins, Orf2 has six putative transmembrane domains, suggesting that it might function as the transmembrane homodimer component of the transport system.

Finally, Orf3 is the typical ATP-binding component of the ABC-type transport systems. It shares 55% identity with GlnQ of E. coli, 54% identity with GlnQ of B. stearothermophilus and 51% identity with HisP of S. typhimurium.

In conclusion, the three gene products here described appear to constitute a new B. stearothermophilus ABC-type transport system. Orf1 is the solute-binding component tethered to the cell membrane with an N-terminal glyceride-cysteine, Orf2 forms the transmembrane complex (probably homodimeric) and Orf3 provides the system with energy.

Tam & Saier (1993) have recently reviewed the structural, functional and evolutionary characteristics of the extracellular solute-binding receptors in bacteria, which were grouped in eight distinct families. According to this classification Orf1 appears to belong to receptor family 3, which recognizes and transports polar amino acids and opines, on the basis of its size, typical for this family (265 ± 12 residues), and of the degree of homology. In particular, one of the two signature sequences which characterize family 3 [GF(DE)(LIV)DLX3(LIVM) (CA)(KE)] is partially conserved in Orf1 between residues 70 and 81.

The homology of Orf1 with the receptor subunits of the polar amino acid and opine transport systems so far characterized together with the striking similarities of Orf2 and Orf3 with the other components of the same systems lead us to hypothesize that this system is involved in amino acid transport.

Unfortunately, we do not have any direct experimental evidence to unequivocally assign a specific function to this
ABC-type transport system in *B. subtilis*

system. When the chromosomal deletion of a 1148 bp HindIII fragment, which destroys both the C-terminus of Orf1 and the N-terminus of Orf3 and removes Orf2 entirely, was introduced into *B. subtilis* 168 and the mutant strain was grown in minimal media containing different amino acids as the sole nitrogen source no difference in the growth curves of the mutant strain with respect to the wild-type strain was observed. However these results do not rule out the involvement of the system in amino acid transport since, as it is the case for other organisms (see for example Townsend & Wilkinson, 1992), each amino acid might have more than one transport system. More detailed physiological and biochemical studies are necessary to identify the function of the system.

**ACKNOWLEDGEMENTS**

We would like to thank Dr A. Sonenshein for helpful discussions. The expert secretarial assistance of E. Iennaco is also acknowledged. This work was financially supported partially by the EEC *Bacillus* Genome Sequencing Project and partially by ENIRICERCHE S.p.A.

**REFERENCES**


Received 5 January 1995; revised 6 March 1995; accepted 22 March 1995.