ELSEVIER

FEMS Microbiology Letters 141 (1996) 129–137

FEMS MICROBIOLOGY LETTERS

A novel amidohydrolase gene from *Bacillus subtilis* cloning: DNA-sequence analysis and map position of *amhX*

Bettina Kempf [a,b], Erhard Bremer [a,b,*]

[a] *Max-Planck-Institut für Terrestrische Mikrobiologie, Karl-von-Frisch Straße, D-35032 Marburg, Germany*
[b] *Philipps-Universität Marburg, Fachbereich Biologie, Laboratorium für Mikrobiologie, Karl-von-Frisch Straße, D-35032 Marburg, Germany*

Received 25 April 1996; accepted 6 May 1996

Abstract

The nucleotide sequence of a new *Bacillus subtilis* gene (*amhX*) was determined that encodes a protein (AmhX) with strong sequence identity to amidohydrolases from both plant and bacterial species and a carboxypeptidase from the archaeon *Sulfolobus sulfataricus*. The *amhX* gene encodes a hydrophilic polypeptide of 383 amino acids with a molecular mass of 41.5 kDa. The *amhX* gene was overexpressed in *E. coli* by using the T7 RNA polymerase/promoter system and the transcription initiation sites for the *amhX* mRNAs in *B. subtilis* were determined by primer extension analysis. Chromosomal *amhX* mutations were constructed by marker replacement and the *amhX* gene was positioned at 25° on the genetic and physical map of the *B. subtilis* chromosome.

Keywords: *Bacillus subtilis*; Amidohydrolase; Gene mapping; Bi-directional terminator; *opuA* operon; Promoter

1. Introduction

Aminoacylases (*N*-acyl-L-amino acid amidohydrolases, EC 3.5.1.14) are found widely in animals, plants, and microorganisms [1–3]. These enzymes remove acyl residues from *N*-acetylated amino acids and catalyze the hydrolysis of carboxylic acid amides to fatty acid anions and L-amino acids [4]. Substrate specificity among the amidohydrolases can vary from very broad to rather restricted enzymatic activity [5]. Amidohydrolases are involved in the degradation of acetylated amino acids [6], and they also participate in the biosynthesis of a number of amino acids [2,5]. They also appear to play an important role in signalling events that control developmental steps in the division and elongation of cells of the dicot *Arabidopsis thaliana* [7]. Amidohydrolases have found important applications in biotechnology. The high chiral specificity of amidohydrolases allows their use in the industrial production of stereoisomers from racemic mixtures [8]. These interesting features of amidohydrolases have been the incentive to search for such enzymes in a number of microorganisms and to clone their structural genes [5,9]. Here we report the characterization of a new gene (*amhX*) from *Bacillus subtilis* which encodes an amidohydrolase exhibiting high sequence identity to amidohydrolases from both procaryotic and eucaryotic organisms.

* Corresponding author. Tel.: +49 (64) 21-28 15 29; Fax: +49 (64) 21-28 89 79; E-mail: bremer@mailer.uni-marburg.de

0378-1097/96/$12.00 Copyright © 1996 Federation of European Microbiological Societies. Published by Elsevier Science B.V.
PII S0378-1097(96)00181-4

2. Materials and methods

2.1. Growth conditions and bacterial strains

Escherichia coli strains were grown aerobically at 37°C either in rich LB medium or in minimal medium A (MMA) with 1% Na succinate as the carbon source. Strain BL21(λDE3) [10] carrying plasmids pPD100, pBKB85, or pBKB86, respectively (Fig. 1), was grown in M9 minimal medium supplemented with 0.2% glucose and 0.2% casaminoacids [11]. *B. subtilis* strains were grown in Spizizen's minimal medium (SMM) with 0.5% glucose, 20 μg/ml L-phenylalanine, 18 μg/ml L-tryptophan and a solution of trace elements [11,12]. The *E. coli arg* mutant strain XS1D2 (F^- Δ101(*argE-ppc*) *nalA rpoB hsdR*) has been described [13]. The arginine requirement of XS1D2 was satisfied by adding L-arginine at a final concentration of 40 μg/ml to the growth medium. The *B. subtilis* strain JH642 (*trpC2 pheA1*; Bacillus Genetic Stock Center, Columbus, OH, USA), is a derivative of the wild-type strain 168 and was constructed by J.H. Hoch. Strains BKB4 [Δ(*opuA-amhX*::*neo*)1] [14] and BKB13 [(*amhX*::*erm*)1] (this

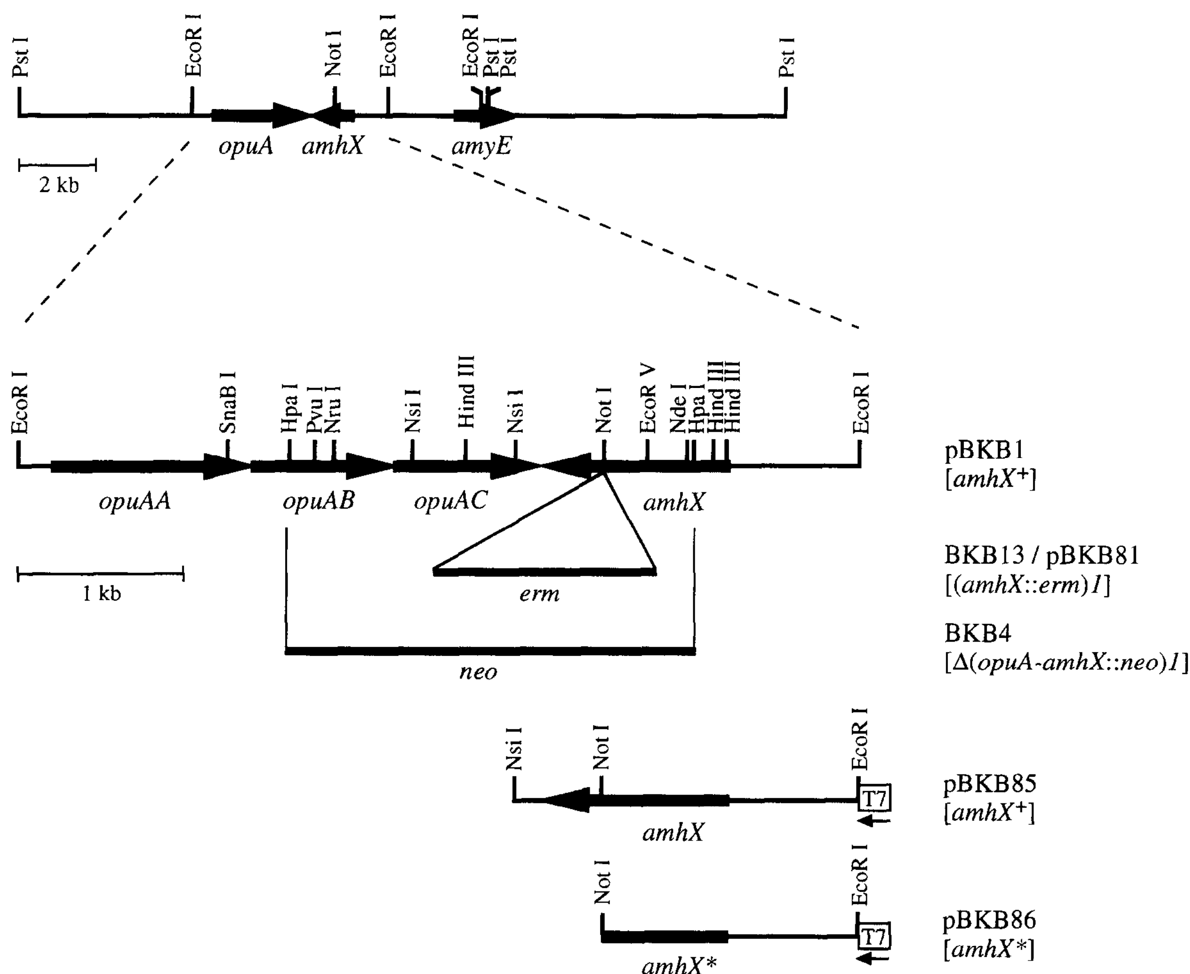

Fig. 1. Genetic and physical organization of the region around 25° on the *B. subtilis* chromosome. The upper part shows the position and direction of transcription of the *opuA* operon and the *amhX* and *amyE* genes together with a few restriction sites. The enlarged region represents the 5.2-kb *Eco*RI restriction fragment from the *B. subtilis* chromosome present in plasmid pBKB1 and the physical structure of the plasmids used in this study.

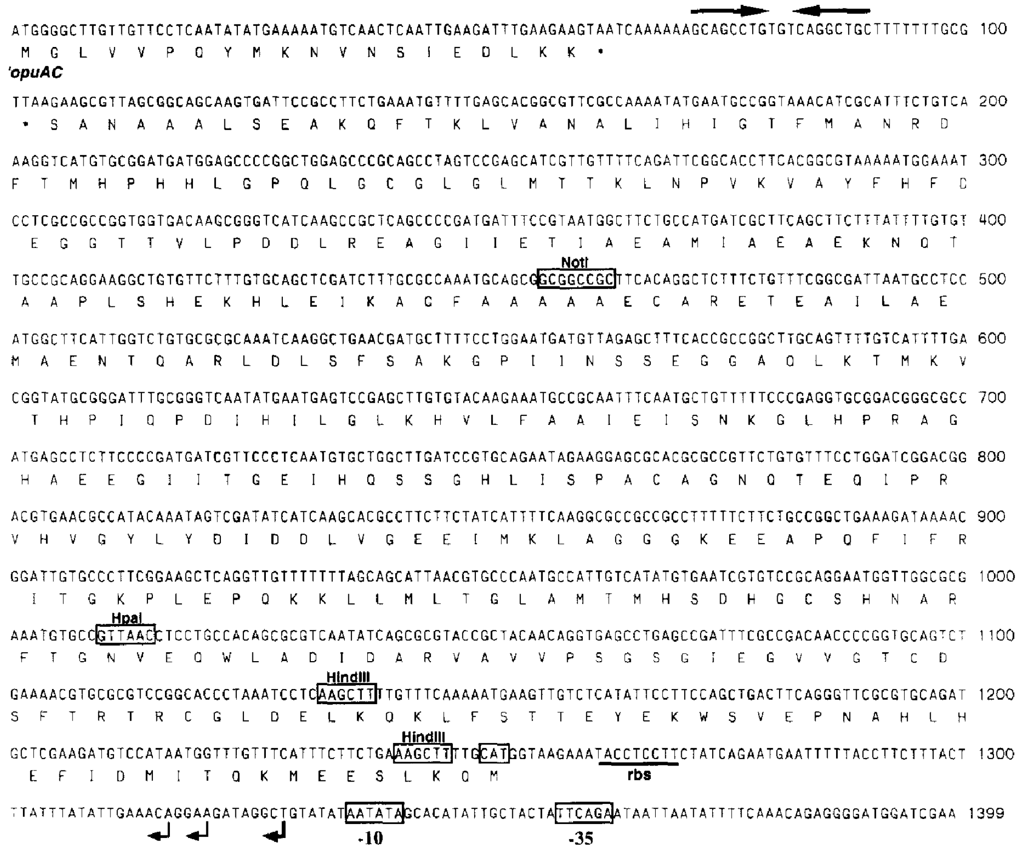

Fig. 2. Nucleotide sequence of the *amhX* region. The DNA sequence of the *amhX* gene and the 3′ end of *opuAC* are presented along with the deduced protein sequences for AmhX and the carboxy-terminal part of OpuAC. The reported DNA sequence was deposited in GenBank® with the accession no. U47860. The shown DNA segment overlaps (bp 1–bp 250) with that reported by us previously [14] for the 3′ end of *opuAC* (accession no. U17292).

study) are derivatives of strain JH642. Strain BKB13 was constructed by transforming strain BKB4 with linearized plasmid DNA of pBKB81 [$opuA^+$ (*amhX*::*erm*)1] and selecting for erythromycin resistance (Ery^r) to identify colonies carrying the (*amhX*::*erm*)1 allele in the chromosome (Fig. 1). Loss of the Δ(*opuA-amhX*::*neo*)1 mutation in these Ery^r transformants was scored by testing their sensitivity to kanamycin and the integrity of the *opuA* operon by transport assays using radiolabeled glycine betaine as the substrate [14]. One of these strains is BKB13 and the physical structure of the (*amhX*::*erm*)1 mutation in this strain was verified by PCR analysis. Antibiotics were used at the following final concentrations in the growth media: ampicillin, 100 μg/ml; chloramphenicol, 30 μg/ml; erythromycin, 0.4 μg/ml and kanamycin, 5 μg/ml.

2.2. Methods used with nucleic acids

All recombinant DNA procedures followed routine methods [15]. The DNA sequence of *amhX* was determined on both DNA strands using the Sequenase version 2.0 system (USB, Braunschweig, Germany). The isolation of the $opuA^+$ $amhX^+$ plasmid pBKB1 (Fig. 1) has been described [14]. A blunt-ended 1.25-kb *Bam*HI-*Cla*I restriction fragment from plasmid pIC56 [16] carrying an erythromycin resistance gene cassette was inserted into the unique *Hin*dII site in the polylinker of pUCBM20 (Boehringer Mannheim, Germany) to yield plasmid pBKB80. The erythromycin resistance gene cassette was cut out from plasmid pBKB80 on a 1.3-kb *Sma*I-*Eco*RV fragment and inserted into the blunt-ended *Not*I restriction site present in plasmid

```
1                                          MQKLSEEMK--QTIMDIFEHLHANPEVS AmhX (Bs)
1   MDFSGRFFVIFVTFFF-LPPLSSAGSYDSGSGLESLAR-GMLHSAKDPEFF--EWMRGIRRKIHENPETG ILR1 (At)
1   MALNNFL--TFQLLLL-LLRVSSESPWIVAGDVSRIPI-NFLELAKSPEVF--DSMVRIRRKIHENPELG ILL1 (At)
1   MALNKLLSLTFQLLLF-LLSVSSESPWI-AEDTSQIQT-KLLEFAKSPEVF--DWMVKIRRKIHENPELG ILL2 (At)
1                                 MDLVEKLKN-DV-----R-EIE-DWIIQIRRKIHEYPELS CpsA (Ss)
1   MNLDLNQLVKWHREFHRFPEIGWSEFWTTSRIADYLEDLDCFEIFLGKQIINPDFVRGRKQAVVDKGLAN HipO (Hi)
1                                     MTKEEIKRLVDEVK--TDVIAWRRHLHAHPELS Ama (Bst)
1                                        MNLIPEILDLQ--GEFEKIRHQIHENPELG HipO (Cj)

27  WKEYETTSFLKQKLEDLGCRTRTFS--DCTGVVGEIGSGSP--VVAVRADIDAL------WQEVNGTFRA AmhX (Bs)
67  FQEFKTSQLVRDELDSLGVKYKYPVA-KTGVVAKIGSCSKPV-FGLRADMDALPLQELVEWESKSKVDGK ILR1 (At)
65  YEEFETSKFIRSELDLIGVKYRYPVA-ITGIIGYIGTGEPPF-VALRADMDALPIQEAVEWEHKSKNPGK ILL1 (At)
66  YEELETSKLIRSELELIGIKYRYPVA-ITGVIGYIGTGEPPF-VALRADMDALPIQEGVEWEHKSKIPGK ILL2 (At)
33  YKEYNTSKLVAETLRKLGVEVEEGVGLPTAVVGKIRGSKPGKTVALRADMDALPVEENTDLEFKSKVKGV CpsA (Ss)
71  AKAYGANEKWLEKMEGY-----------TGCVALFDSGKPGKTIALRFDIDCVNVTETRSPEHIPNKEGF HipO (Hi)
32  FQEEKTAQFVYETLQSFGHLELSRPT-KTSVMARLIGQQPGRVVAIRADMDALPIQEENTFEFASKNPGV Ama (Bst)
29  FDELCTAKLVAQKLKEFGYEVYEEIG-KTGVVGVLKKGNSDKKIGLRADMDALPLQECTNLPYKSKKENV HipO (Cj)

              ▼
87  -------NHSCGHDSHMTMALGTLMLLKKQPELPKGTIRFIFQPAEEK-GGGALKMIEEGVLDDIDYLYG AmhX (Bs)
135 -------MHACGHDTYVAMLLGAAKLLQTTKHLIKGTVKLVFQPGEEGY-AGAYEMLKDEILDDLDGILS ILR1 (At)
133 -------MHACGHDGHVAMLLGAAKILQQHRQHLQGTVVLIFQPAEEGL-SGAKKMREEGALKNVEAIFG ILL1 (At)
134 -------MHACGHDGHVAMLLGAAKILHEHRHHLQGTVVLIFQPAEEGL-SGAKKMREEGALKNVEAIFG ILL2 (At)
103 -------MHACGHDTHVAMLLGGAYLLVKNKDLISGEIRLIFQPAEEDG-GLGGAKPMIEAGVMNGVDYV CpsA (Ss)
130 ASINDGFMHACGHDSHITIGLGVALWIAQNKDKLTGKVKIVFQPAEEGVR-GAAAIAQSGIIDDADYFAS HipO (Hi)
101 -------MHACGHDGHTAMLLGTAKIFSQLRDDIRGEIRFLFQHAEELFPGGAEEMVQAGVMDGVDVVIG Ama (Bst)
98  -------MHACGHDGHTTSLLLAAKYLAS--QNFNGTLNLYFQPAEEGL-GGAKAMIEDGLFEKFDSDY- HipO (Cj)

149 VHVRPIQETQNGAC---APSILHGSSQHIEGTIIGEEAH-GARPHLGKNSIEIAAFLVHKLGLI---HID AmhX (Bs)
197 VHVFPSIPSGGIGS---RPGTVLAGAGLFTVTVHGQGSH-AATPHFSKDPVLAASSAVVALQQIVSRELD ILR1 (At)
195 IHLSPRTPFGKAAS---LAGSFMAGAGAFEAVITGKGGH-AAIPQHTIDPVVAASSIVLSLQHLVSRETD ILL1 (At)
196 IHLSARIPFGKAAS---RAGSFLAGAGVFEAVITGKGGH-AAIPQHTIDPVVAASSIVLSLQQLVSRETD ILL2 (At)
165 FGIHISSSYPSGVFATRK-GPIMATPDAFKIIVHGKGGH-GSAPHETIDPIFISLQIANAIYGITARQID CpsA (Ss)
199 SHISFC---ANTGTVIANPRNFL-STTKIDIRYKGKPAHAGAAPHLGRNALLAAAHTVTQLHGI-ARHGK HipO (Hi)
164 THLWSPLERGKIGIVY---GPMMAAPDRFFIRIIGKGGH-GAMPHQTIDAIAIGAQVVTNLQHIVSRYVD Ama (Bst)
157 VFGWHNMPFGSDKKFYLKKGAMMASSDSYSIEVIGRGGH-GSAPEKAKDPIYAASLLVVALQSIVSRNVD HipO (Cj)

212 PQIPHTVKMTKLQAGGESSNIIPGKASFSLDLRAQTNEAMEALIAETERACEAAAAF----GAKIELHK AmhX (Bs)
263 PLEAGVVTVGYIE-GGHAQNVIPQSAKFGGTFRSLSNDGLLFIQRRIKEISEAQASVYRCKAEVNFEEKK ILR1 (At)
261 PSDSKVVTVTKVN-GGNAFNVIPDSITIGGTLRAFT--GFTQLQERIKEIITKQAAVHRCNASVNLAPNG ILL1 (At)
262 PLDSKVVTVSKVN-GGNAFNVIPDSITIGGTLRAFT--GFTQLQQRVKEVITKQAAVHRCNASVNLTPNG ILL2 (At)
233 PVQPFIISITTIH-SGTKDNIIPDDAEMQGTIRSLDENVRSKAKDYMRRIVSSICGIY----GATCEVKF CpsA (Ss)
264 GMTRINVGVLK---AGEGRNVIPSSAELQLEVRGENKAINEYMTEQVMQIAKGISISFNVAYETEIVGEA HipO (Hi)
230 PLEPLVLSVTQFV-AGTAHNVLPQEVEIQGTVRTFDETLRRTVPQWMERIVKGITEAHGASYEFRFDYGY Ama (Bst)
226 PQNSAVVSIGAFN-AGHAFNIIPDIVTIKMSVRALDNETRKLTEEKIYKICKGLAQAN----DIEIKINK HipO (Cj)

278 EHSLPAATQNKEAEAIMAEAITEILGAE--RLDDPLVTTGGEDFHFYAVKVPNLK----TTMLGLG---C AmhX (Bs)
332 PSLHPVMNNDEGLYEHGKKVAEAMIGKNN--FHDFPVTMGGEDFSFFTQKTKA-A----IFVLGIKNETL ILR1 (At)
328 NQPHFPIVNNMDLYKNFKKVVRDLLGQEA--FVEAVPEMGSEDFSYFAETIPG-H----FSLLGMQDETQ ILL1 (At)
329 REPHPPIVNNKDLYKQFKKVVRDLLGQEA--FVEAAPVMGSEDFSYFAETIPG-H----FSLLGMQDETN ILL2 (At)
298 MEDVYPTTVNNPE----VTDEVMKILSSISTVVETEPVLGAEDFSRFLQKAPG-T----YFFLGTRNEKK CpsA (Ss)
331 VDM----NNDVELIKLIEEI--SLEQPQINNVNSDYAFNASEDATILGRRVQEHGGKAIYFILGADRTA- HipO (Hi)
299 RPVI-------------------------NYDEGDPRHGGNGVRAVRRRGSG-P-------LETEHGRR Ama (Bst)
291 NVVAPVTMNNDEAVDFASEVAKELFGEKNCEF-NHRPLMASEDFGFFCEMKKC-A----YAFLENENDIY HipO (Cj)

339 GLQPGLHHPHMTFDRNAMFTGIHILANAVLKTFQKAESLAAANAS                          AmhX (Bs)
395 GAGKPLHSPYFFVDEEALPVGAALHAAMAVSYLDEHG-HSHEEEVKSEL                      ILR1 (At)
391 GYASS-HSPHYRINEDVLPYGAAIHATMAVQYLKDKASKGSVSGFHDEL                      ILL1 (At)
392 GYASS-HSPLYRINEDVLPYGAAIHASMAVQYLKEKASKGSVSGFHEEL                      ILL2 (At)
359 GCIYPNH-SSKFCVDEDVLKLGALAHALLAVKFSNK                                   CpsA (Ss)
394 ----GHHEAEFDFDENQLLTGVNIYTSLVQKLLS                                     HipO (Hi)
335 RFLRLFAKSARQLFLRRRGQCRKRHRLPAPPPALYD                                   Ama (Bst)
355 -----LHNSSYVFNDKLLARAASYYAKLALKYLK                                     HipO (Cj)
```

Fig. 3. Alignment of AmhX with the amino acid sequences of several amidohydrolases. The amino acid sequence of the AmhX protein from *B. subtilis* (Bs) is aligned with that of the ILR1, ILL1, and ILL2 proteins from *A. thaliana* (At) [7], the CpsA protein from *S. sulfataricus* (Ss), the *H. influenzae* (Hi) HipO protein [22], the Ama protein from *B. stearothermophilus* (Bst) [5] and the HipO enzyme of *C. jejuni* (Cj) [6]. The signature pattern (ARGE_DAPE_CPG2_2, I-x-G-x(2)-[SAG]-H-x-[SAG]-x(1,2)-P-x(2)-[SAGT]) for the aminoacylase family is overlined and the conserved cysteine residue is marked with a triangle.
←

pBKB1, resulting in the [*opuA*$^+$ (*amhX*::*erm*)1] plasmid pBKB81 (Fig. 1). The expression plasmids pBKB85 (*amhX*$^+$) and pBKB86 (*amhX**), carrying the T7ϕ10 promoter were constructed by ligating the 1.9-kb *Eco*RI-*Nsi*I and the 1.4-kb *Eco*RI-*Not*I fragments from plasmid pBKB1 into the low-copy-number expression vector pPD100 [17] cut with *Eco*RI and *Pst*I and *Eco*RI and *Not*I, respectively (Fig. 1). The T7ϕ10-mediated expression of the *amhX*$^+$ and *amhX** genes in strain BL21(λDE3), preparation of total cell extracts and analysis of proteins by SDS-polyacrylamide electrophoresis were performed as described by Dersch et al. [17]. pBKB87 (*amhX*$^+$) is an *E. coli-B. subtilis* shuttle plasmid and was constructed by inserting the 1.9-kb *Eco*RI-*Nsi*I fragment from pBKB1 into plasmid pRB373 [18] cleaved with *Eco*RI and *Pst*I. Total RNA was isolated from log-phase cultures of strain JH 642 (pBKB87) grown in SMM medium with either 0.5% glucose or 1% glycerol as the carbon source as described [14]. For the primer extension reaction, a synthetic primer (5′-CCAGCTGACTTCAGG-GTTCGCG-3′), complementary to the *amhX* mRNA (position 1172–1193 bp; see Fig. 2) was hybridized to 10 μg of RNA and extended with avian myeloblastosis virus reverse transcriptase in the presence of radiolabeled [^{35}S]dATP at a final concentration of 1 μCi/ml. Multiple sequence alignments were carried out at the National Center for Biotechnology Information (NCBI) using the BLAST programs [19] and the current versions of the data bases (January 1996). The DNA sequence of the *amhX* gene (Fig. 2) was assigned the GenBank™ accession number U47860.

3. Results and discussion

3.1. DNA sequence analysis of amhX

Plasmid pBKB1 (Fig. 1) harbours an approx. 5.2-kb chromosomal *Eco*RI restriction fragment from *B. subtilis* carrying the osmoregulated *opuA* operon (*opuAA*, *opuAB*, *opuAC*) encoding the components of the binding-protein-dependent glycine betaine transport system OpuA [14]. We have determined the DNA sequence of a 1336-bp DNA segment downstream from the last gene (*opuAC*) in the *opuA* operon (Fig. 2). Inspection of the DNA sequence revealed a large open reading frame (1149 bp) oriented convergently to *opuA*. We refer in the following to this open reading frame as *amhX* (amidohydrolase) since its deduced protein product (AmhX) exhibits high homologies to a number of amidohydrolases (see below).

The proposed ATG start codon of *amhX* is preceded at an appropriate distance of 9 bp by the sequence 5′-AAGGAGGT-3′, which is identical to the free 3′ end of 16S RNA of *B. subtilis* and thus serves as the ribosome-binding site of *amhX*. The intergenic region between the *amhX* and *opuAC* genes is only 37 bp long (Fig. 2) and contains an inverted repeat which is bordered on both sides by a run of A·T base pairs. This palindrome can potentially form an extended secondary structure with a calculated ΔG(25°C) of −73.8 kJ [20] and exhibits the characteristic features of factor-independent transcriptional terminators [21]. This DNA element might thus serve as a bi-directional terminator for the converging *opuA* and *amhX* transcripts.

3.2. Features of the amhX-encoded protein

The *amhX* reading frame encodes a 383 amino acid residue protein with a calculated molecular mass of 41.5 kDa. It does not exhibit any extended hydrophobic regions and also lacks a signal sequence at its amino terminus. We compared the AmhX sequence to the protein sequences available in the data bases and found striking sequence identities of the AmhX protein to amidohydrolases from plants (*A. thaliana*), Gram-negative (*H. influenzae* and *C. jejuni*)

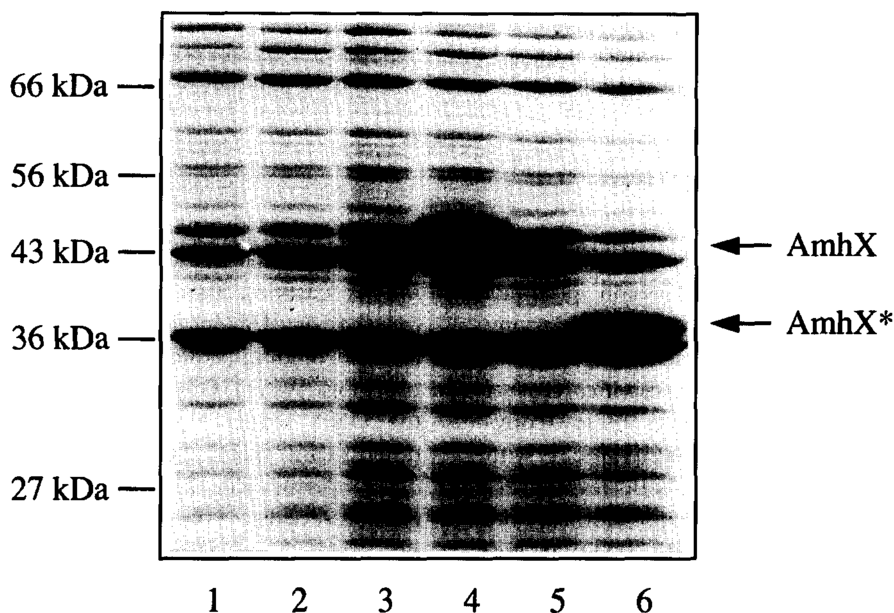

Fig. 4. Expression of the *amhX* and *amhX**-encoded proteins under the control of the T7ϕ10 promoter. T7ϕ10-mediated gene expression was performed in strain BL21(λDE3), carrying the vector, pPD100 (uninduced, lane 1; induced with 1 mM IPTG, lane 2), plasmid pBKB85 (*amhX*$^+$; uninduced, lane 3; induced, lane 4) and plasmid pBKB86 (*amhX**; uninduced, lane 5; induced, lane 6). The proteins were electrophoretically separated on a 12% SDS-polyacrylamide gel and visualized by staining with Coomassie Brilliant Blue.

and Gram-positive (*B. stearothermophilus*) eubacteria and a carboxypeptidase from the archaeon *S. sulfataricus* (Fig. 3). AmhX exhibits strong sequence identity (25%) to the ILR1 protein from the dicot *A. thaliana*, an amidohydrolase suggested to liberate the phytohormone indole-3-acetic acid (IAA) from one of the proposed biologically inactive storage forms of IAA, IAA-Leu [7]. ILR1 is a member of a small gene family present in *A. thaliana*. The two ILR1-related proteins, ILL1 and ILL2, are about 87% identical to one another, 43% identical to ILR1, and 25% identical to the AmhX protein from *B. subtilis* (Fig. 3). In addition to the homology to the plant enzymes, we found that the AmhX protein is related to several amidohydrolases of bacterial origin. A thermostable amidohydrolase (Ama) with a broad substrate specificity from *B. stearothermophilus* [5] and an *N*-benzoylglycine amidohydrolase (hippuricase; HipO) from the human pathogen *C. jejuni* which cleaves *N*-benzoylglycine (hippuric acid) into glycine and benzoic acid [6], show sequence identities to AmhX of 29 and 32%, respectively. The AmhX protein is also related (19% identity) to a putative hippuricase (HipO) from *H. influenzae* [22]. Based upon the extended sequence similarities between AmhX and the amidohydrolases of several species (Fig. 3), we suggest that the *B. subtilis amhX* gene encodes a new amidohydrolase. The detected sequence identities between the enzymes listed in Fig. 3 do not reveal any clues as to the substrate specificity of the AmhX protein.

The ArgE and DapE amidohydrolases from *E. coli* function in the biosynthesis of the amino acids arginine and lysine, respectively, and these enzymes exhibit homology to carboxypeptidase G2 (CPG2) from a *Pseudomonas* sp. The sequence identity detected between the ArgE, DapE, and CPG2 proteins has been interpreted to reflect an evolutionary link between amidohydrolases and carboxypeptidases [2]. Our finding of an extended sequence identity (28%) between AmhX and a carboxypeptidase (CpsA) from the hyperthermophilic archaeon *S. sulfataricus* [23] supports the suggested relationship between these two groups of enzymes (Fig. 3). A signature pattern for the ArgE, DapE, and CPG2 proteins has been proposed by Bairoch [24] (ARGE_DAPE_CPG2_2, I-x-G-x(2)-[SAG]-H-x-[SAG]-x(1,2)-P-x(2)-[SAGT]; PROSITE release 26, version 11.1). This sequence

motif can also be found in the *B. subtilis* AmhX protein and is highly conserved in the sequence of carboxypeptidase CpsA and the various amidohydrolases aligned in Fig. 3. However, AmhX exhibits only very weak homology to the ArgE, DapE, and CPG2 proteins, suggesting that subgroups exist among the 'aminoacylase family' of amidohydrolases [1]. In addition to the sequence homology exhibited by the amindohydrolases, these enzymes have subunit masses between 40 and 50 kDa and cysteine residues appear to be crucial for enzymatic activity [1,24]. AmhX fits into this general pattern with a calculated molecular mass of 41.5 kDa and its six cysteine residues (Fig. 2). However, the position of only a single cysteine residue is conserved in the amino acid sequence of the 7 amidohydrolases and the carboxypeptidase aligned in Fig. 3 and this cysteine is located in a highly conserved block of 9 amino acid residues. It is thus tempting to speculate that this region might be part of the catalytic center of these enzymes.

3.3. Overproduction of the AmhX protein

To visualize the AmhX protein, we constructed the overproduction plasmid pBKB85 (Fig. 1) and expressed the *B. subtilis amhX* gene under the control of the T7ϕ10 promoter in the heterologous host *E. coli*. Induction of transcription in strain BL21(λDE3) carrying pBKB85 resulted in high-level production of a protein with an apparent molecular mass of approx. 43 kDa (Fig. 4; lane 4). The apparent molecular mass of the overproduced protein is in close agreement with that of AmhX deduced from the *amhX* DNA sequence (41.5 kDa). When a mutant *amhX* gene (*amhX**) truncated at its 3′ end was overexpressed in BL21(λDE3), a 37 kDa protein was synthesized (Fig. 4; lane 6). The apparent molecular mass of the AmhX* protein determined from its electrophoretic mobility is in perfect agreement with that calculated from the partially deleted *amhX** gene (37.2 kDa). These findings strongly suggest that the *amhX* open reading frame (Fig. 2) is indeed expressed and represents a functional *B. subtilis* gene. Both the *amhX*$^+$ plasmid pBKB85 and the truncated *amhX** gene carried by pBKB86 (Fig. 1) might prove useful for future biochemical characterization of the enzymatic properties of the *B. subtilis* AmhX amidohydrolase.

3.4. Identification of the transcription initiation sites of amhX

To provide direct experimental proof for the expression of the *amhX* gene in *B. subtilis*, we mapped by primer extension analysis its transcription initiation site(s). Total RNA was isolated from log-phase cultures of the *B. subtilis* strain JH642 carrying the *amhX*$^+$ plasmid pBKB87 grown in minimal medium with either glucose or glycerol as the carbon source. Three *amhX*-specific mRNA species were detected that differed in size at their 5′ ends (Fig. 5). The longest and strongest produced mRNA species initiates at an A residue at position 1329 bp, 77 bp upstream of the *amhX* ATG start codon (Fig. 2). Inspection of the DNA sequence upstream of this initiation site revealed the presence of putative −10 and −35 sequences (Fig. 2) that closely resemble the consensus sequence of promoters recognized by the main vegetative sigma factor (σ^A) of *B. subtilis* [25]. The *amhX* −10 and −35 regions are separated by 17 bp, a typical spacing for σ^A-dependent promoters. In addition to the main *amhX* mRNA, two smaller transcripts were detected whose initiation sites correspond to positions 1321 and 1317 bp (Fig. 2). We are currently unable to say whether these mRNA species are degradation products of the main *amhX* mRNA or represent transcripts initiating from alternative promoters. There was no strong difference in the amount of the three mRNAs produced in glucose- or glycerol-grown cells (Fig. 5).

3.5. Construction of chromosomal amhX mutations

To test whether the *amhX* gene is essential for growth of *B. subtilis*, we constructed two chromosomal *amhX* mutations. In strain BKB4 [Δ(*opuA-amhX*::*neo*)1], most of the *amhX* coding region is deleted and in strain BKB13 [(*amhX*::*erm*)1], an erythromycin resistance cassette is inserted into the unique *Not*I site in *amhX* gene (Fig. 1). We then compared the growth properties of these *amhX* mutant strains with those of the *amhX*$^+$ parent JH642 on rich and minimal agar plates and in liquid cultures at 37°C. We did not observe any growth differ-

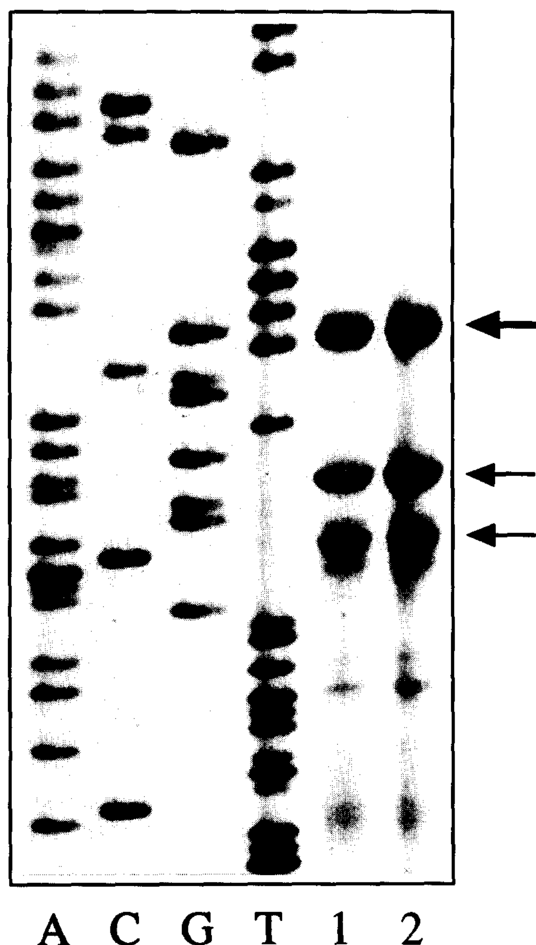

Fig. 5. Mapping of the *amhX* transcription initiation sites. Total RNA was prepared from cells of strain JH642(pBKB87; *amhX*$^+$) grown in minimal medium with either glucose (lane 1) or glycerol (lane 2) as the carbon source. A primer complementary to the *amhX* mRNA was hybridized to 10 μg total RNA and extended with reverse transcriptase. DNA sequencing reactions primed with the same synthetic oligonucleotide used for the primer extension reactions were employed as a standard to size the *amhX* mRNAs.

ences; hence, the *amhX* gene is not essential for *B. subtilis*. Sakanyan et al. [5] have cloned the *ama* structural gene for a thermostable amidohydrolase with broad substrate specificity from *B. stearothermophilus* by heterologous complementation of an *E. coli* mutant (*argE*) defective in acetylornithine deacetylase activity in the arginine biosynthetic pathway. Although the affinity of the Ama protein from *B.* stearothermophilus for *N*-acetylornithine was very low, enough L-ornithine was produced in the recombinant strain with the multicopy *ama*$^+$ plasmid to allow the growth of the *argE* mutant [5]. Since the Ama protein from *B. stearothermophilus* shows significant sequence identity (29%) to the AmhX protein (Fig. 4), we tested whether the *B. subtilis amhX* gene could complement the *argE* defect in the *E. coli* mutant strain XS1D2. We found that the arginine requirement of strain XS1D2 was not abolished by the presence of pBKB1.

3.6. Genetic map location of amhX

We previously performed genetic and physical mapping experiments [14] that positioned the *opuA* operon at 25° on the *B. subtilis* genetic map in the vicinity of the *amyE* gene (Fig. 1). Hence, we can now infer that the *amhX* gene is located between the *opuA* and *amyE* loci and is transcribed in a counter-clockwise fashion on the *B. subtilis* chromosome. The *B. subtilis* genomic DNA is cut infrequently by the restriction enzyme *Not*I, a feature that has been exploited to construct a physical map. One of the *Not*I sites is located upstream of *amyE* [26], and our DNA sequence analysis of the *opuA*-*amhX* region shows that this site is actually located within the *amhX* coding region (Fig. 2). Hence, plasmid pBKB1 can serve as a linking clone for the two adjacently located *Not*I restriction fragments, and the [(*amhX*::*erm*)1] mutation present in strain BKB13 might prove useful for physical and genetic mapping experiments in *B. subtilis*.

Acknowledgments

We thank S. Baumberg, R. Brückner and the late M. Steinmetz for generously providing bacterial strains and plasmids and K.-H. Röhm for helpful discussions. We appreciate the help of V. Koogle and C. Schumacher in preparing the manuscript and thank S. Kneip for photography. Financial support for this study was provided by the Deutsche Forschungsgemeinschaft through the Graduiertenkolleg 'Enzymchemie' and the Fonds der Chemischen Industrie.

References

[1] Palm, G.J. and Röhm, K.-H. (1995) Aminoacylase I from porcine kidney: identification and characterization of two major protein domains. J. Prot. Chem. 14, 233–240.

[2] Boyen, A., Charlier, D., Charlier, J., Sakanyan, V., Mett, I. and Glansdorff, N. (1992) Acetylornithine deacetylase, succinyldiaminopimelate desuccinylase and carboxypeptidase G2 are evolutionarily related. Gene 116, 1–6.

[3] Weiß, H.M., Palm, G.J. and Röhm, K.-H. (1995) Thermostable aminoacylase from *Bacillus stearothermophilus*: significance of the metal center for catalysis and protein stability. Biol. Chem. Hoppe-Seyler 376, 643–648.

[4] Schomburg, D. and Salzmann, M. (1991) Enzyme Handbook, Springer-Verlag, Berlin.

[5] Sakanyan, V., Desmarez, L., Legrain, C., Charlier, D., Mett, I., Kochikyan, A., Savchenko, A., Boyen, A., Falmagne, P., Pierard, A. and Glansdorff, N. (1993) Gene cloning, sequence analysis, purification, and characterization of a thermostable aminoacylase from *Bacillus stearothermophilus*. Appl. Environ. Microbiol. 59, 3878–3888.

[6] Hani, E.K. and Chan, V.L. (1995) Expression and characterization of *Campylobacter jejuni* benzoylglycine amidohydrolase (hippuricase) gene in *Escherichia coli*. J. Bacteriol. 177, 2396–2402.

[7] Bartel, B. and Fink, G.R. (1995) ILR1, an amidohydrolase that releases active indole-3-acetic acid from conjugates. Science 268, 1745–1748.

[8] Chibata, I., Tosa, T., Sato, T. and Mori, T. (1976) Production of L-amino acids by aminoacylase adsorbed on DEAE-sephadex. Methods Enzymol. 44, 746–759.

[9] Mukohara, Y., Ishikawa, T., Watabe, K. and Nakamura, H. (1993) Molecular cloning and sequencing of the gene for a thermostable *N*-carbamyl-L-amino acid amidohydrolase from *Bacillus stearothermophilus* strain NS1122A. Biosci. Biotech. Biochem. 57, 1935–1937.

[10] Studier, F.W. and Moffat, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. 189, 113–130.

[11] Miller, J.H. (1992) A short course in bacterial genetics. A Laboratory Manual and Handbook for *Escherichia coli* and Related Bacteria, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

[12] Harwood, C.R. and Cutting, S.M. (1990) Molecular biological methods for *Bacillus*, Wiley, Chichester.

[13] Mountain, A., Mann, N.H., Munton, R.N. and Baumberg, S. (1984) Cloning of a *Bacillus subtilis* restriction fragment complementing auxotrophic mutants of eight *Escherichia coli* genes of arginine biosynthesis. Mol. Gen. Genet. 197, 82–89.

[14] Kempf, B. and Bremer, E. (1995) OpuA, an osmotically regulated binding protein-dependent transport system for the osmoprotectant glycine betaine in *Bacillus subtilis*. J. Biol. Chem. 270, 16701–16713.

[15] Sambrook, J., Fritsch, E.F. and Maniatis, T.E. (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

[16] Steinmetz, M. and Richter, R. (1994) Plasmids designed to alter the antibiotic resistance expressed by insertion mutations in *Bacillus subtilis*, through in vivo recombination. Gene 142, 79–83.

[17] Dersch, P., Fsihi, H. and Bremer, E. (1994) Low-copy-number T7 vectors for selective gene expression and efficient protein overproduction in *Escherichia coli*. FEMS Microbiol. Lett. 123, 19–26.

[18] Brückner, R. (1992) A series of shuttle vectors for *Bacillus subtilis* and *Escherichia coli*. Gene 122, 187–192.

[19] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

[20] Tinoco, I., Jun., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973) Improved estimation of secondary structure in ribonucleic acid. Nat. New Biol. 246, 40–41.

[21] Platt, T. (1981) Termination of transcription and its regulation in the tryptophan operon of *E. coli*. Cell 24, 10–23.

[22] Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.-F., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J.D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Geoghagen, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O. and Venter, J.C. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269, 496–512.

[23] Tortora, P. (1995) Molecular cloning, nucleotide sequence and expression of a carboxypeptidase-encoding gene from the archaebacterium *Sulfolobus solfataricus* (acc. no. Z48497). unpublished.

[24] Bairoch, A. (1993) The PROSITE dictionary of sites and patterns in proteins, its current status. Nucl. Acids Res. 21, 3097–3103.

[25] Moran, C.P., Jr., Lang, N., LeGrice, S.F.J., Lee, G., Stephens, M., Sonenshein, A.L., Pero, J. and Losick, R. (1982) Nucleotide sequences that signal the initiation of transcription and translation in *Bacillus subtilis*. Mol. Gen. Genet. 186, 339–346.

[26] Itaya, M. and Tanaka, T. (1991) Complete physical map of the *Bacillus subtilis* 168 chromosome constructed by a gene-directed mutagenesis method. J. Mol. Biol. 220, 631–628.