Chromosomal Location of the *Bacillus subtilis* Aspartokinase II Gene and Nucleotide Sequence of the Adjacent Genes Homologous to *uvrC* and *trx* of *Escherichia coli*

By NAI-YONG CHEN,† JING-JUAN ZHANG† AND HENRY PAULUS*

Department of Metabolic Regulation, Boston Biomedical Research Institute, 20 Staniford Street, Boston, Massachusetts 02114, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA

(Received 20 June 1989; revised 20 July 1989; accepted 10 August 1989)

The aspartokinase II (*ask*) operon of *Bacillus subtilis* consists of two in-phase overlapping genes that encode the two subunits of the lysine-sensitive isoenzyme of aspartokinase (ATP:L-aspartate 4-phosphotransferase, EC 2.7.2.4). Transduction mapping of the *ask* operon, inactivated by recombinational insertion of a *cat* marker, indicates a chromosomal location (about 253°) between *leuA* and *aroG*. *ask* is thus remote from *aecB*, eliminating *aecB* as a possible locus for the structural gene of aspartokinase II, but close to *aecA* and *uvrB*. The nucleotide sequence of a 2 kb DNA fragment just upstream of the *ask* operon was determined and found to contain two open reading frames. The deduced amino acid sequence of the distal reading frame exhibits extensive homology with *Escherichia coli* thioredoxin and that of the proximal one, which overlaps with the *ask* promoter, is homologous to the deduced product of the *E. coli uvrC* gene. Insertional mutagenesis of the proximal open reading frame led to a mitomycin-sensitive phenotype, consistent with a role in DNA repair. In conjunction with the data of M. Petricek, L. Rutberg & L. Hederstedt [*FEMS Microbiology Letters* **61**, 85–88] our results define the nucleotide sequence of an 8.8 kb segment of the *B. subtilis* chromosome near 253° and the following order of genes: *trx-uvrB-ask-orfX-sdhC-sdhA-sdhB-orfY-gerE*.

INTRODUCTION

Two classes of mutants of *Bacillus subtilis* resistant to the lysine analogue *S*-(2-aminoethyl)cysteine (AEC) have been characterized (Yeh & Steinberg, 1978; Mattioli *et al.*, 1979). Mutants of the *aecA* class, which map near 250° on the *B. subtilis* chromosome, produce derepressed levels of aspartokinase II and are probably the result of an altered regulatory locus. On the other hand, *aecB* mutants, which map near 290°, have normal levels of aspartokinase II and their resistance to AEC may well be the result of an alteration in one of the threonine biosynthetic enzymes (Mattioli *et al.*, 1979); nevertheless, the *aecB* locus has often been assumed to be the site of the aspartokinase II structural gene (e.g. Piggot & Hoch, 1985).

Recently, we determined the nucleotide sequence of the genes for the subunits of aspartokinase II from *B. subtilis* (Chen *et al.*, 1987). The elucidation of the nucleotide sequence of the aspartokinase II coding region made possible its insertional inactivation and thereby its mapping on the *B. subtilis* chromosome, which we report here. The promoter of the aspartokinase II operon overlaps an unidentified open reading frame (ORF). Sequencing of the

* Please address correspondence to: Dr H. Paulus, Department of Metabolic Regulation, Boston Biomedical Research Institute.

† Permanent address: Institute of Microbiology, Academia Sinica, Beijing, China.

Abbreviations: AEC, *S*-(2-aminoethyl)cysteine; ORF, open reading frame.

region upstream of the aspartokinase II operon revealed the unidentified ORF to have a high degree of homology to the product of the *Escherichia coli* *uvrC* gene and to lie just downstream from the coding region for a protein homologous to *E. coli* thioredoxin.

METHODS

Bacterial strains, phage, and plasmids. *B. subtilis* VS11 (*pheA trpC* Sul^r *aecA ask::cat*) was obtained by the recombinational insertion of a pJH101 derivative into the *ask* locus of *B. subtilis* VB217 as described earlier (Chen & Paulus, 1988). *B. subtilis* GS11 (*pheA trpC ask::cat*) was obtained in an analogous manner from *B. subtilis* GSY225 (Yeh & Steinberg, 1978). A strain with a disruption in the reading frame just upstream of the *ask* locus, *B. subtilis* GPEN, was constructed in a similar manner, using a recombinant pJH101 derivative carrying the *Eco*NI-*Pst*I fragment, which encodes part of that reading frame (Figs 1 and 3). In order to insert the *Eco*NI-*Pst*I segment into the *Eco*RI-*Hind*III sites of pJH101, the former was ligated, after filling the protruding end of the *Eco*NI site using Klenow fragment, in the *Sma*I and *Pst*I sites of the pUC19 polylinker, followed by excision with *Eco*RI and *Hind*III restriction endonucleases. All other strains listed in Table 1 were obtained from the Bacillus Genetic Stock Center, as was bacteriophage PBS1. Plasmid p69-35, a pBR322 derivative with a 3.6 kb *Eco*RI-*Pst*I insert carrying the *B. subtilis* aspartokinase II genes and adjacent regions, has been described (Bondaryk & Paulus, 1985).

Transduction analysis. PBS1 transduction was done by the method of Dubnau *et al.* (1967). The donor phage lysates were prepared as described by Osbourne & Sonenshein (1980).

Materials. The sources of most materials have been described (Chen *et al.*, 1987). DNA sequencing reagents were obtained from US Biochemicals. Deoxyribo-oligonucleotides were synthesized by the phosphoramidite method using a BioSearch Cyclone DNA synthesizer.

Nucleotide sequence determination. The 2 kb *Eco*RI-*Pst*I fragment derived from plasmid p69-35 was cloned into the polylinker of M13mp18, M13mp19 or pUC18 and sequenced by the dideoxynucleotide method (Sanger *et al.*, 1977) as previously described (Chen *et al.*, 1987), except that T7 DNA polymerase (Sequenase) was used. The standard heptadecameric sequencing primer (New England Biolabs) was used for the determination of sequences adjacent to the M13mp18 or M13mp19 cloning sites, but synthetic deoxyribo-oligonucleotides complementary to appropriate sequences already determined were used as primers for the determination of sequences internal to the cloned DNA segment. The sequencing strategy is illustrated in Fig. 1.

Sensitivity to mitomycin C. *B. subtilis* strains were grown to mid-exponential phase in LB medium supplemented with 10 µg chloramphenicol ml⁻¹ (strains GPEN and GS11) or without supplement (strains 1A346 and 1A6). Culture samples (0.1 ml) were then plated at appropriate dilutions (10⁻¹⁰-fold) on LB agar plates without or with mitomycin (0.5 or 0.6 µg ml⁻¹).

RESULTS AND DISCUSSION

Genetic mapping of the aspartokinase II operon

Since mutants of *B. subtilis* that lack aspartokinase II do not have a distinguishable phenotype, on account of the presence of a second aspartokinase enzyme (Paulus, 1984), direct mapping of the aspartokinase gene was not possible. On the other hand, we had earlier inactivated the aspartokinase II genes by recombinational insertion of a plasmid carrying a gene for chloramphenicol transacetylase (Chen & Paulus, 1988) and were thus in the position to map aspartokinase II with chloramphenicol resistance as a selection marker (Haldenwang *et al.*, 1980). Accordingly, the generalized transducing phage PBS1 was used to infect *B. subtilis* strain VS11, in which aspartokinase has been inactivated by insertion of the chloramphenicol-resistance marker (*cat*) from plasmid pJH101 (Chen & Paulus, 1986), and the resulting *ask::cat* lysate was used to transduce various auxotrophic strains. Of special interest was the relationship of the aspartokinase II locus to the *aecA* and *aecB* loci (at about 250° and 290° on the *B. subtilis* genetic map, near *leuA* and *thrA* respectively; Mattioli *et al.*, 1979), both of which have been inferred to influence aspartokinase II (e.g. Piggot & Hoch, 1985). Preliminary transduction experiments showed no significant linkage (<3% cotransduction) between *ask::cat* and *thrA* (290°) but did reveal significant cotransduction with the markers *leuA* (about 85%) and *aroG* (about 25%), which map at 250° and 260°, respectively (data not shown). Three-factor transduction crosses clearly showed that the map position of the *cat* marker inserted in the aspartokinase II coding region lies between *leuA* and *aroG* (Table 2). Comparison of cotransduction frequencies indicated that the *ask::cat* locus is located at approximately 253° on

Fig. 1. Restriction map of the 2 kb *EcoRI*-*PstI* fragment upstream of the aspartokinase II operon and strategy for nucleotide sequence determination. The arrows indicate DNA segments that were sequenced and the direction of sequencing. The designation M13mp18 or M13mp19 at the base of an arrow shows that the sequence was obtained using the standard heptadecameric sequencing primer (New England Biolabs); all other sequences were determined with the synthetic primers indicated by the numbers at the base of the arrows.

Fig. 2. Mapping of the *ask* locus by PBS1 transduction. The transduction donor was *B. subtilis* GS11 (*leuA*⁺ *aroG*⁺ *ask* :: *cat*), which carries a *cat* gene as an insert in the *ask* locus (see text), and the recipient was *B. subtilis* BGSC 1A9 (*leuA* *aroG* Cam^s). The arrows start at the selected marker and the numbers represent cotransduction frequencies between the markers.

Table 1. *B. subtilis* strains

Strain	Genotype	Source or reference
GSY225	<i>pheA1 trpC2</i>	Yeh & Steinberg (1978)
VB217	<i>pheA1 trpC2</i> Sul ^r <i>aecA</i>	Yeh & Steinberg (1978)
GS11	<i>pheA1 trpC2 ask</i> :: <i>cat</i>	This study
VS11	<i>pheA1 trpC2</i> Sul ^r <i>aecA ask</i> :: <i>cat</i>	Chen & Paulus (1988)
GPEN	<i>pheA1 trpC2 uvrB</i> (?) :: <i>cat</i>	This study
BGSC 1A6	<i>ilvA1 pyrD1 trpC2 thyA1 thyB1</i>	Dedonder <i>et al.</i> (1977)
BGSC 1A9	<i>ald-1 aroG932 leuA8 trpC2</i>	Dedonder <i>et al.</i> (1977)
BGSC 1A346	<i>hisH2 thyA1 thyA2 uvrB109</i>	Munakata & Ikeda (1966)
BGSC 1A434	<i>ala-1 leuA8 metB purB thr-5 trpC</i>	Bacillus Genetic Stock Center

the *B. subtilis* genetic map (Fig. 2), close to the tentative map positions of *uvrB* (Munakata, 1977) and *aecA* (Mattioli *et al.*, 1979).

Our mapping results show that the chromosomal location of the aspartokinase II operon is remote from *aecB* but close to *aecA*. *aecB* is thus clearly excluded as the structural gene for aspartokinase II and may well be completely unrelated to aspartokinase II, as originally proposed by Mattioli *et al.* (1979). More information on the relationship between the aspartokinase II operon and the *uvrB* locus is provided by the nucleotide sequence analysis described in the next section.

628	AGG CTT GTG ACG GAA ATC GAG GAT TTT GAA TAT ATT GTG ACG TCC TCC AAT CTT	637	646	655	664	673	1762	1771	1780	1789	1798	1807
	R L V T E I E D F E Y I V T S S N L						AGA AGA CGC TAT ACG AGA GTG CTT CGT GAG AAT CTG CCG CTG CCC GAT CTG ATT					
682	GAA GCG CTT ATT TTA GAA ATG AAT CTG ATC AAA AAG CAT GAT CCG AAA TAC AAC	691	700	709	718	727	1816	1825	1834	1843	1852	1861
	E A L I L E M N L I K K H D P K Y N						ATC ATT GAC GGA GGA AAA GGG CAG ATC AAC GCA GCA AGG GAT GTC ATT GAA AAT					
736	GTC ATG CTC AAA GAC GAC AAA ACC TAT CCT TTC ATA AAA CTC ACC GAT GAA GCG	745	754	763	772	781	1870	1879	1888	1897	1906	1915
	V M L K D D K T Y P F I K L T H E R						I I D G G K G Q I N A A R D V I E N					
790	CAC CCA AGG CTG ATT GTG ACC GCG AAT GTC AAA AAG GAC AAA GGC GCG TAT TTC	799	808	817	826	835	1924	1933	1942	1951	1960	1969
	H P R L I V T R N V K K D K G R Y F						ACC TCA AAT TTA CTG ATC GGT GAT CCG CTG GAG GTG GCG TAT CTG GAA CGA AAC					
844	GCG CCG TAT CCG AAT GTA CAG GCA AGG GAA ACA AAA AAA CTG CTC GAC CGT	853	862	871	880	889	1978	1987	1996	2005	2014	2023
	G P Y P N V Q A A R E T K K L L D R						AGC CAG GAA TTT TAC CTC CTG CAG CGC ATT CAG GAC GAG CTG GAC CGT TTT GCA					
898	CTG TAT CCT CTC AGA AAA TGC TCC AAG CTC CCG GAC AGA GTG TGC CTT TAC TAT	907	916	925	934	943	2032	2041	2050	2059	2068	2077
	L Y P L R K C S K L P D R V C L Y Y						ATC AGT TTT CAC AGG CAA ATC CCG GGA AAA AGC GCG TTT CAA TCC GTT TTG GAC					
952	CAT CTG GCG CAA TGC CTT GCT CCG TGT GTA AAG GAT ATT TCC GAA GAG ACG AAC	961	970	979	988	997	2086	2095	2104	2113	2122	2131
	H L G Q C L A P C V K D I S E E T N						GAC ATC CCG GGT ATC GGA GAG AAA AGA AAG AAA ATG CTG TTA AAG CAT TTC GGT					
1006	AGA GAG CTG GTT GAA AGC ATT ACA CGT TTC TTA AGG GCG GCG TAT AAT GAG GTC	1015	1024	1033	1042	1051	2140	2149	2158	2167	2176	2185
	R E L V E S I T R F L R G G Y N E V						TCC GTT AAA AAA ATG AAG GAA GCA AGC CTT GAG ATC AAA AAA GCC GGT GTT					
1060	AAA AAA GAG CTT GAA GAG AAA ATG CAT GAA GCT GCT GAG AAT CTT GAG TTT GAA	1069	1078	1087	1096	1105	2194	2203	2212	2221	2230	2239
	K K E L E E K M H E A A E N L E F E						CCG CAG CCG CCG CTC AGC TCC TTT ACG ACA AAT TGC AAA AAT AAT GTT GTC CTT					
1114	GGA GCA AAA GAG CTT GGT GAC GAA ATT GCC CAT ATT GAA TCA ACG ATG GAA AAA	1123	1132	1141	1150	1159	2248	2257				
	R A K E L R D Q I A H I E S T M E K						TTA AAT AAG ATC TGA TAA					

Fig. 3. Nucleotide sequence of the 2 kb *EcoRI*-*PstI* fragment upstream of the aspartokinase II operon and deduced amino acid sequences of the major ORFs. Also shown is a 0.28 kb segment downstream of the *PstI* site, which had been sequenced earlier (Chen *et al.*, 1987). The nucleotide sequence shown is that of the DNA strand which corresponds to the sense strand of the *ask* operon and is numbered starting at the *EcoRI* site distal to the *ask* operon. The arrows indicate a region of dyad symmetry which is followed by a run of Ts and resembles a *p*-independent transcription terminator. The asterisks indicate a sequence of residues which are complementary to the 3' end of *B. subtilis* 16S RNA and may constitute a ribosome-binding site.

Table 2. Mapping of the *ask::cat* determinant by *PBS1* transduction crosses with *B. subtilis* GS11 (*pheA trpC ask::cat*) as the donor (D) strain and with *B. subtilis* BGSC 1A9 (*ald-1 trpC2 leuA8 aroG932*) as the recipient (R)

Selected marker	Recombinant type			No. of recombinants	Inferred order
	<i>leu</i>	<i>cat</i>	<i>aro</i>		
Leu ⁺	D	D	D	42	<i>leuA-ask::cat-aroG</i>
	D	R	D	3	
	D	D	R	72	
	D	R	R	33	
Aro ⁺	D	D	D	17	<i>leuA-ask::cat-aroG</i>
	R	D	D	15	
	D	R	D	2	
	R	R	D	131	
Cam ^r	D	D	D	67	<i>leuA-ask::cat-aroG</i>
	R	D	D	7	
	D	D	R	102	
	R	D	R	8	

Nucleotide sequence upstream of the aspartokinase region

The aspartokinase II operon is contained within a 2.9 kb *Pst*I fragment together with part of an ORF which terminates near the aspartokinase II transcription start site (Chen *et al.*, 1987). An adjacent 2 kb *Eco*RI-*Pst*I fragment was obtained from plasmid p69-35, a pBR322 derivative carrying a 3.6 kb *Eco*RI-*Bam*HI fragment with an internal *Pst*I site (Bondaryk & Paulus, 1985). The sequence of the entire 2 kb *Eco*RI-*Pst*I fragment was determined in the hope of identifying the ORF that partially overlaps the promoter of the aspartokinase II operon; it is shown in Fig. 3. The sequence is occupied by two ORFs: nucleotides 11-322 encode a 104-residue polypeptide and nucleotides 461-2254 encode a 598-residue protein. The *ask* promoter partially overlaps the coding sequence of the larger protein, the -35 and -10 consensus sequences corresponding to nucleotides 2235-2240 and 2258-2263, respectively.

Comparison of the deduced amino acid sequences of the ORFs with *E. coli* proteins

The deduced amino acid sequences of the two ORFs in the 2 kb *Eco*RI-*Pst*I fragment were compared with all sequences in the National Biomedical Research Foundation's protein sequence database, using the algorithm of Pearson & Lipman (1985). The putative 104-residue polypeptide showed a high degree of homology with *E. coli* thioredoxin (Holmgren, 1968), with 49 residues (47%) in common (Fig. 4). Especially striking is the near identity of a 20-residue segment (residues 23-42) around the only two cysteines, which correspond to the active-site residues of *E. coli* thioredoxin. The postulated product of the larger ORF was highly homologous to the deduced product of the *E. coli* *uvrC* gene (Sancar *et al.*, 1984), with 225 of 598 amino acid residues (38%) identical (Fig. 5). The close alignment of the products of the two ORFs with those of the *trx* and *uvrC* genes of *E. coli* strongly suggests that the cloned 2 kb *Eco*RI-*Pst*I fragment of the *B. subtilis* chromosome carries the corresponding genes.

Since the putative thioredoxin reading frame begins at residue 11 of the sequenced DNA, possible transcription-control and ribosome-binding sites cannot be defined. Indeed, it is possible that the reading frame extends into the unsequenced region and that the polypeptide product is longer than indicated in Figs 2 and 3. The polypeptide-coding region (Fig. 3, residues 11-325) is followed by a sequence with dyad symmetry and a run of Ts (residues 346-377). The corresponding transcript would form a hairpin loop [$\Delta G = -15.4$ kcal mol⁻¹ (-64.4 kJ mol⁻¹);

Fig. 5. Amino acid alignment of the deduced product of the *E. coli* *uvrC* gene and the deduced polypeptide product encoded by residues 461-2254 of the 2 kb *Eco*RI-*Pst*I fragment shown in Fig. 3. Residues upstream of the postulated translation start site of the *E. coli* *uvrC* product (Sancar *et al.*, 1984) are in italics. Identical amino acid residues are indicated by double dots, highly conserved residues according to the groups (I, L, M, V), (D, E), (K, R), (F, Y), (N, Q) and (S, T) by single dots.

Fig. 4. Amino acid alignment of *E. coli* thioredoxin and the deduced polypeptide product encoded by residues 10–322 of the 2 kb *EcoRI*–*PstI* fragment shown in Fig. 3. Identical amino acid residues are indicated by double dots, highly conserved residues according to the groups (I, L, M, V), (D, E), (K, R), (F, Y), (N, Q) and (S, T) by single dots.

[illegible]

Table 3. *Effect of disruption of the uvrC homologue on sensitivity to mitomycin C*

Strain	Genotype	Percentage of survivors in mitomycin C at:	
		0.05 $\mu\text{g ml}^{-1}$	0.06 $\mu\text{g ml}^{-1}$
GPEN	<i>uvr(?)::cat pheA trpC</i>	0.3	0.004
GS11	<i>ask::cat pheA trpC</i>	29	12.4
BGSC 1A346	<i>uvrB109 thyA thyB hisH</i>	0.038	0
BGSC 1A6	<i>uvrB⁺ thyA thyB ilvA pyrD trpC</i>	18	2.3

Tinoco *et al.*, 1973] followed by a run of Us, characteristic of ρ -independent transcription terminators. The coding region for the *uvrC* homologue that follows probably starts with the AUG codon at residue 461, with a potential ribosome-binding site, AAGGAaGG, at residues 446–453. The calculated free energy of interaction (Tinoco *et al.*, 1973) of 16S ribosomal RNA with this nucleotide sequence is $-14.8 \text{ kcal mol}^{-1}$ ($-61.9 \text{ kJ mol}^{-1}$), a value typical for ribosome-binding sites in *B. subtilis* (Hager & Rabinowitz, 1985). A less likely candidate for a translation initiation site is the AUG codon at residue 521, which corresponds to the postulated start site of the *E. coli uvrC* protein (Sancar *et al.*, 1984); however, this codon is not preceded by a potential ribosome-binding site. Indeed, the site of translation initiation of the *E. coli uvrC* protein merits re-examination in view of the significant similarity between the putative sequence of the homologous *B. subtilis* protein deduced from nucleotides 461–520 and that deduced from the nucleotide sequence preceding the postulated translation start site of the *E. coli uvrC* protein (Fig. 5). The search for potential transcription initiation sites in the region that intervenes between the ρ -independent transcription terminator following the thioredoxin homologue and the translation start of *uvrC* homologue revealed none of the known promoter consensus sequences for the various forms of *B. subtilis* RNA polymerase (e.g. Helmann & Chamberlin, 1988). The mode of transcription termination of the gene homologous to *E. coli uvrC* is also uncertain. Its coding region partially overlaps the *ask* promoter and is not followed by a discernible transcription terminator (Chen *et al.*, 1987). It is thus possible that the transcripts continue either to the *ask* transcription attenuator or to the end of the *ask* operon.

Effect of disruption of the reading frame of the uvrC homologue

To elucidate the function of the gene whose deduced product is homologous to the *uvrC* protein of *E. coli*, we examined the physiological consequences of its disruption. Gene disruption was accomplished by the method of Haldenwang *et al.* (1980), making use of a recombinant pJH101 plasmid (Ferrari *et al.*, 1983) which carried a 0.6 kb fragment of the ORF just upstream of the *ask* operon. The resulting chloramphenicol-resistant recombinant strain, *B. subtilis* GPEN, was indistinguishable from the parent strain in its growth rates in either rich or minimal media but was significantly more sensitive to mitomycin C. Comparison with *B. subtilis* GS11, in which the pJH101 insertion is in the *ask* operon rather than in the adjacent ORF, showed strain GPEN to be several orders of magnitude more sensitive to mitomycin C (Table 3). Comparison of a *uvrB* deletion mutant, *uvrB109* (Munakata, 1977), with a *uvr⁺* strain, although not quite isogenic, revealed a similar difference in mitomycin C sensitivity. It appears, therefore, that the *B. subtilis* gene homologous to *E. coli uvrC* plays a role in DNA repair. The mapping results described in an earlier section place the *ask* operon at 253°, the same general region of the *B. subtilis* chromosome as the *uvrB* gene (Munakata, 1977). Our results are thus consistent with the notion that the gene homologous to *E. coli uvrC* is actually the *uvrB* gene.

CONCLUSIONS

The results described in this paper provide strong evidence that the genes just upstream of the aspartokinase II operon of *B. subtilis* are the *trx* and *uvrB* genes. Recently, Petricek *et al.* (1989) have shown by sequence analysis that the structural gene of aspartokinase II lies just upstream from the *sdhCAB* operon. Together with the data of Petricek *et al.* (1989) and the sequence for *gerE* (Cutting & Mandelstam, 1986), our results define the complete nucleotide sequence of an

Fig. 6. Map of the 8.8 kb region surrounding the aspartokinase II operon near 250° of the *B. subtilis* chromosome. The map is based on sequence data from this paper, Chen *et al.* (1987), Magnusson *et al.* (1986), Phillips *et al.* (1987), and Cutting & Mandelstam (1986). The orientation of the map relative to the *dnaB* and *ilvB* markers is based on the data of Ohne *et al.* (1973) and Hederstedt *et al.* (1982). The direction of transcription is indicated by the solid arrows. The dotted arrows indicate the polarity of unidentified ORFs.

8.8 kb segment of the *B. subtilis* chromosome near 250°, shown in Fig. 6. Except for two small unidentified ORFs, this segment is fully occupied by the following sequence of genes: *trx*–*uvrB*–*ask*–*orfX*–*sdhC*–*sdhA*–*sdhB*–*orfY*–*gerE*.

The exhaustive mapping of this portion of the *B. subtilis* chromosome raises the important question as to the location and the nature of the *aecA* locus, which plays a major role in the control of aspartokinase II levels. Petricek *et al.* (1989) have carried out transformation experiments which show that *aecA* maps near *sdhCAB* and *ask*, probably between the *Bgl*II and *Bam*HI sites in the aspartokinase II operon (Fig. 6). These results are consistent with the notion that *aecA* is a *cis*-acting control element in the aspartokinase II promoter/attenuator region. On the other hand, the possibility that *aecA* is a regulatory gene whose product leads to the derepression of the aspartokinase II operon has not been rigorously eliminated. Experiments are currently in progress to distinguish between these possibilities.

We thank Dr L. Hederstedt for communicating his results prior to publication, Dr A. L. Sonenshein for his advice on transduction techniques and Dr Daniel Zeigler of the Bacillus Genetic Stock Center for strains and helpful suggestions. This research was supported by Grant DMB 88-03644 from the National Science Foundation and by Biomedical Research Support Grant RR 05711 from the National Institutes of Health Research Program. Computer resources used to carry out these studies were provided by the BIONET™ National Computer Resource for Molecular Biology, whose funding is provided by the Biomedical Research Technology Program, Division of Research Resources, National Institutes of Health Grant 1 U41 RR 01685.

REFERENCES

- BONDARYK, R. P. & PAULUS, H. (1985). Cloning and structure of the gene for the subunits of aspartokinase II from *Bacillus subtilis*. *Journal of Biological Chemistry* **260**, 585–591.
- CHEN, N. Y. & PAULUS, H. (1988). Mechanism of expression of the overlapping genes of *Bacillus subtilis* aspartokinase II. *Journal of Biological Chemistry* **263**, 9526–9532.
- CHEN, N. Y., HU, F. M. & PAULUS, H. (1987). Nucleotide sequence of the overlapping genes for the subunits of *Bacillus subtilis* aspartokinase II and their control regions. *Journal of Biological Chemistry* **262**, 8787–8798.
- CUTTING, S., & MANDELSTAM, J. (1986). The nucleotide sequence and transcription during sporulation of the *gerE* gene of *Bacillus subtilis*. *Journal of General Microbiology* **132**, 3013–3024.
- DEDONDER, R. A., LEPESANT, J. A., LEPESANT-KEIZ-LAROVA, J., BILLAULT, R., STEINMETZ, M. & KUNST, F. (1977). Construction of a kit of reference strains for rapid genetic mapping in *Bacillus subtilis* 168. *Applied and Environmental Microbiology* **33**, 989–993.
- DUBNAU, D., GOLDTHWAITE, C., SMITH, I. & MARMUR, J. (1967). Genetic mapping in *Bacillus subtilis*. *Journal of Molecular Biology* **27**, 163–185.
- FERRARI, F. A., NGUYEN, A., LANG, D. & HOCH, J. A. (1983). Construction and properties of an integratable plasmid for *Bacillus subtilis*. *Journal of Bacteriology* **154**, 1513–1515.
- HAGER, P. W. & RABINOWITZ, J. (1985). Transitional

- specificity in *Bacillus subtilis*. In *The Molecular Biology of the Bacilli*, pp. 1–32. Edited by D. A. Dubnau. New York: Academic Press.
- HALDENWANG, W. G., BANNER, C. D. B., OLLINGTON, J. F., LOSICK, R., HOCH, J. A., O'CONNOR, M. B. & SONENSHEIN, A. L. (1980). Mapping a cloned gene under sporulation control by insertion of a drug resistance marker into the *Bacillus subtilis* chromosome. *Journal of Bacteriology* **142**, 90–98.
- HEDERSTEDT, L., MAGNUSSON, K. & RUTBERG, L. (1982). Reconstitution of succinate dehydrogenase in *Bacillus subtilis* by protoplast fusion. *Journal of Bacteriology* **152**, 157–162.
- HELMANN, J. D. & CHAMBERLIN, M. J. (1988). Structure and function of bacterial sigma factors. *Annual Review of Biochemistry* **57**, 839–872.
- HOLMGREN, A. (1968). Thioredoxin. 6. The amino acid sequence of the protein from *Escherichia coli* B. *European Journal of Biochemistry* **6**, 475–484.
- MAGNUSSON, K., PHILLIPS, M. K., GUEST, J. R. & RUTBERG, L. (1986). Nucleotide sequence of the gene for cytochrome *b₅₅₈* of the *Bacillus subtilis* succinate dehydrogenase complex. *Journal of Bacteriology* **166**, 1067–1071.
- MATTIOLI, R., BAZZICALUPO, M., FEDERICI, G., GALLORI, E. & POLSINELLI, M. (1979). Characterization of mutants of *Bacillus subtilis* resistant to S-(2-aminoethyl)cysteine. *Journal of General Microbiology* **114**, 223–225.
- MUNAKATA, N. (1977). Mapping of the genes controlling excision repair of pyrimidine photoproducts in *Bacillus subtilis*. *Molecular and General Genetics* **156**, 49–54.
- MUNAKATA, N. & IKEDA, Y. (1966). Inactivation of transforming DNA by ultraviolet irradiation: a study with ultraviolet-sensitive mutants of *Bacillus subtilis*. *Mutation Research* **7**, 133–139.
- OHNE, M., RUTBERG, B. & HOCH, J. A. (1973). Genetic and biochemical characterization of mutants of *Bacillus subtilis* defective in succinate dehydrogenase. *Journal of General Microbiology* **115**, 738–745.
- OSBOURNE, M. S. & SONENSHEIN, A. L. (1980). Inhibition by lipiarmycin of bacteriophage growth in *Bacillus subtilis*. *Journal of Virology* **33**, 945–953.
- PAULUS, H. (1984). Regulation and structure of aspartokinase in the genus *Bacillus*. *Journal of Biosciences (Bangalore)* **6**, 403–418.
- PEARSON, W. R. & LIPMAN, D. J. (1988). Improved tools for biological sequence comparison. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 2444–2448.
- PETRICEK, M., RUTBERG, L. & HEDERSTEDT, L. (1989). The structural gene for aspartokinase II in *Bacillus subtilis* is closely linked to the *sdh* operon. *FEMS Microbiology Letters* **61**, 85–88.
- PHILLIPS, M. K., HEDERSTEDT, L., HASNAIN, S., RUTBERG, L. & GUEST, J. R. (1987). Nucleotide sequence encoding the flavoprotein and iron-sulfur protein subunits of the *Bacillus subtilis* PY79 succinate dehydrogenase complex. *Journal of Bacteriology* **169**, 864–873.
- PIGGOT, P. J. & HOCH, J. A. (1985). Revised genetic linkage map of *Bacillus subtilis*. *Microbiological Reviews* **49**, 158–179.
- SANCAR, G. B., SANCAR, A. & RUPP, W. D. (1984). Sequences of the *E. coli uvrC* gene and protein. *Nucleic Acids Research* **12**, 4593–4608.
- SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463–5467.
- TINOCO, I., BORER, P. N., DENGLER, B., LEVINE, M. D., UHLENBECK, O. C., CROTHERS, D. M. & GRALLA, J. (1973). Improved estimation of secondary structure in ribonucleic acids. *Nature New Biology* **246**, 40–41.
- YEH, E. C. & STEINBERG, W. (1978). The effect of gene position, gene dosage and a regulatory mutation on the temporal sequence of enzyme synthesis accompanying outgrowth of *Bacillus subtilis* spores. *Molecular and General Genetics* **158**, 287–296.