

Nucleotide Sequence and Characterization of *ptsG* Gene Encoding Glucose-specific Enzyme II of Phosphotransferase System from *Brevibacterium flavum*

Ki-Hong Yoon*

School of Food Science & Biotechnology, Woosong University, Daejeon 300-718, Korea

Received November 4, 2005; Accepted December 8, 2005

Nucleotide sequence of *Brevibacterium flavum ptsG* gene capable of complementing *Escherichia coli* ZSC113 mutations defective to glucose permease activity of phosphotransferase system was completely determined, and the gene product was compared with other glucose-specific enzyme II (EII^{Glc}). A *ptsG* gene of *B. flavum* consisted of open reading frame of 2,025 nucleotides putatively encoding polypeptide of 675 amino acid residues and TAA stop codon. Deduced amino acid sequence of *B. flavum* EII^{Glc} had high homology with EIIs^{Glc} of *Corynebacterium glutamicum*, *C. efficiens*, and *B. lactofermentum*. Arrangement of structural domains, IIBCA, of *B. flavum* EII^{Glc} protein was identical to that of EIIs belonging to glucose-phosphotransferase system.

Key words: *Brevibacterium flavum*, phosphotransferase system, glucose permease gene, nucleotide sequence, comparison

In bacteria, the phosphoenolpyruvate:carbohydrate phosphotransferase system (PTS) is responsible for the uptake and phosphorylation of a number of carbohydrates.¹⁾ The cytoplasmic phosphoproteins, enzyme I (EI) and histidine-containing phosphocarrier protein (HPr), sequentially transfer a phosphoryl group from a phosphoenolpyruvate (PEP) to a membrane-bound EII, which catalyzes the concomitant transport and phosphorylation of the carbohydrates. The EIIs contain three or four structural domains IIA, IIB, IIC, and IID,²⁾ and are divided into four classes³⁾ according to their amino acid sequence similarities. Among them, EII proteins, specific to glucose, sucrose or β -glucoside, and belonging to the glucose-PTS class, contain three domains IIA, IIB, and IIC. The domain IIA, which interacts with a phosphorylated-HPr, exists as a separate protein or as another domains-linked protein. The IIC domain comprises between six and eight putative transmembrane segments showing high hydrophobicity.

The Gram-positive corynebacteria, including *Brevibacterium lactofermentum*, *B. flavum*, *Corynebacterium glutamicum*, *C. efficiens*, and *B. ammoniagenes*, have been used for industrial production of various amino acids and nucleotides. The presence of two PTS systems specific for glucose or fructose was reported in *B. flavum*⁴⁾ and *C. glutamicum*,⁵⁾ from which the genes encoding glucose-specific EII (EII^{Glc}) were recently characterized. The EIIs^{Glc} of *B. ammoniagenes*,⁶⁾ *B. lactofermentum*,⁷⁾ and *C. glutamicum*⁸⁾ contained IIA, IIB, and IIC domains in a single polypeptide with an arrangement of structural domains,

IIBCA, belonging to the sucrose/ β -glucoside subgroup of the glucose-PTS class on the basis of their amino acid sequence alignments. The *ptsG* gene encoding EII^{Glc} was previously cloned from *B. flavum* into *Escherichia coli*.⁹⁾ This work describes the sequencing of a complete *B. flavum ptsG* gene that codes for EII^{Glc}, and compares its deduced amino acid sequence with those of other bacterial strains.

Materials and Methods

Bacterial strains, plasmids and media. *E. coli* XL-1 blue (*supE44 hsdR17 recA1 endA1 gyrA46 relA1 thi lac⁻F[proAB⁺ lac^F lacZ M15 Tn10(tet^r)]*) was used as a host for subcloning experiments, and plasmid pUC19 was used for all cloning and sequencing experiments. MacConkey agar was used to select the *E. coli* transformant carrying recombinant plasmid. *E. coli* was cultured at 37°C in LB broth (10 g tryptone, 5 g yeast extract, 10 g NaCl per liter, pH 7.0). Ampicillin (50 g/ml) was used for the selection of transformants of *E. coli*.

DNA sequencing and computer analysis. The standard procedures of Sambrook *et al.*¹⁰⁾ were used for DNA manipulation. Restriction endonucleases, protease, and RNase obtained from Boehringer Mannheim (Mannheim, Germany), and T4 DNA ligase from Solgent Co. (Daejeon, Korea), were used as recommended by the manufacturers. Restriction endonuclease-generated DNA fragments of the cloned *B. flavum* chromosomal DNA⁹⁾ were subcloned into pUC19. The nucleotide sequences of the fragments were determined with a DNA sequencer (ABI Prism 377, Perkin Elmer Co., Foster, CA, USA). The DNA and protein sequences were analyzed using the DNASIS (Hitachi Software Engineering, Tokyo, Japan) program.

*Corresponding author

Phone: 82-42-630-9742, Fax: 82-42-636-2676

E-mail: ykh@lion.woosong.ac.kr

Results and Discussion

Nucleotide sequence of the *B. flavum ptsG* gene. The *ptsG* gene of *B. flavum* had been previously cloned into pUC19 to form a recombinant plasmid pBFT93 by *in vivo* complementation of an *E. coli* mutant strain lacking EII^{Gk}. For sequencing the *B. flavum ptsG* gene of pBFT93, various restricted fragments were subcloned into pUC19. The 2.9-kb *B. flavum* DNA fragment of pBFT93 was completely sequenced.

The 2,025-nucleotide *ptsG*, beginning at position 228 by an ATG codon and terminating at position 2,255 by the ochre stop codon TAA, was identified in the nucleotide sequence (Fig. 1). The *ptsG* encodes a polypeptide of 675 residues with a calculated M_r of 71,982. The codon usage of *ptsG* exhibited an overall GC content of 52.1%, and a wobble-position GC content of 55.5%. The base composition of a *B. flavum ptsG* gene was similar to those of *C. glutamicum ptsG* (total GC content, 52.3%; wobble-position GC content, 54.6%)⁸⁾ and *B. lactofermentum ptsG* (total GC content, 52.7%; wobble-position GC content, 55.5%).⁷⁾ GC contents of 34 genes from *B. lactofermentum* and *C. glutamicum* ranged from 50 to 62%.¹¹⁾ The *ptsG* gene of *C. efficiens* had a total GC content of 63.9%, with 83.1% GC content at the third base of the codon.¹²⁾ The gene showed high preference for G or C

residues at the third base of the codons. *C. efficiens* can grow at temperature 10°C higher than other strains including *B. flavum*, *B. lactofermentum*, and *C. glutamicum*, and GC content of its total genome is greater than that of *C. glutamicum* genome.¹³⁾ The GAAAGG sequence element, six bases upstream from the first ATG start codon at nucleotide position 216-221, could be the ribosome-binding site for the mRNA. There is a 14-bp palindrome beginning 37 nucleotides downstream of the TAA stop codon. This element could be involved in the rho-independent termination of the *ptsG* gene transcription.

Based on the hydropathy of the deduced amino acid sequence calculated using the method of Kyte and Doolittle¹⁴⁾ (data not shown). EII^{Gk} was predicted to contain a hydrophobic region (IIC domain) corresponding to amino acid residues from 120 to 465 and consisting of six sub-regions with an average hydropathy exceeding 1.0, each comprised of 20 amino acid residues capable of spanning the cytoplasmic membrane. The amino-terminal part (6-84 aa) of the protein was relatively hydrophilic (IIB domain), and the carboxy-terminal part (526-649 aa) was relatively hydrophobic (IIA domain), indicating that the EII^{Gk} of *B. flavum* has an arrangement of IIBCA domains identical to those of EII^{Gk} from *B. lactofermentum*,⁷⁾ *C. glutamicum*,⁸⁾ and *B. ammoniagenes*.⁹⁾

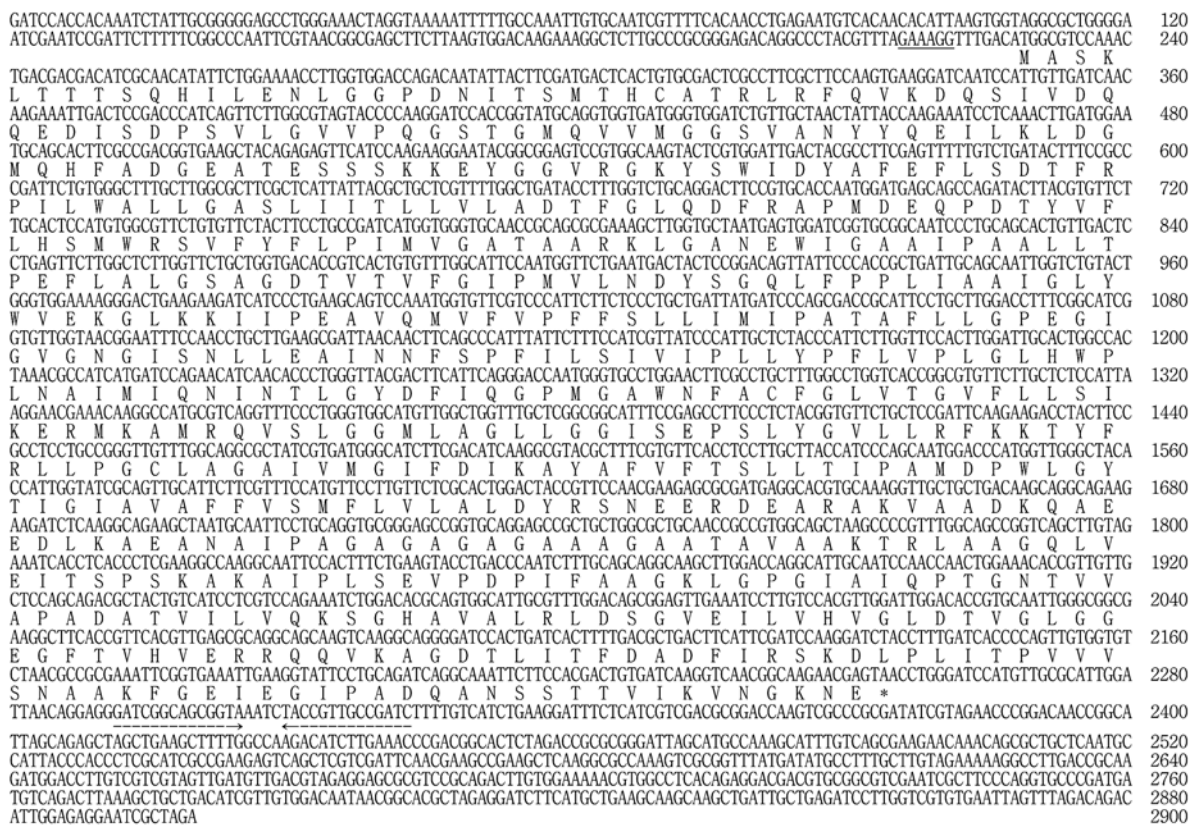


Fig. 1. Nucleotide sequence of 2,900-nucleotide *B. flavum* DNA fragment and deduced amino acid sequence. The underlined sequence preceding the ATG start codon of *ptsG* gene is a putative ribosome-binding site. The TAA stop codon is indicated with asterisks below the sequences. The inverted repeat is underlined by horizontal arrows. Numbers at the end of each line correspond to the nucleotide position. The sequence data have been submitted to the GenBank nucleotide sequence database under accession number DQ267153.

Comparison of the *B. flavum* EII^{Glc} sequence with other EII sequences. Comparison of the deduced amino acid sequence of the *B. flavum* EII^{Glc} with those of other EIIs in the NCBI database using the BLAST search program¹⁵⁾ revealed the present EII^{Glc} was homologous to those belonging to glucose-PTS. The phylogenetic relationships based on amino acid sequences of *B. flavum* EII^{Glc} and other EII of glucose-PTS were investigated using CLUSTAL W software. The *B. flavum* EII^{Glc} exhibited high homologies with *C. glutamicum* EII^{Glc} (97%), *B. lactofermentum* EII^{Glc} (88.3%), and *C. efficiens* EII^{Glc} (81.6%). It was also homologous with the EIIs^{Glc} of *Propionibacterium acnes* (47.4% similarity), *B. ammoniagenes* (44.5% similarity), *C. diphtheriae* (44.3% similarity), and *Bifidobacterium longum* (38.1% similarity). The EII^{Glc} was also similar to the β -glucoside-specific EIIs of *Bacillus halodurans* (32.2% similarity), *B. licheniformis* (32.6% similarity), *Lactobacillus acidophilus* (31.7% similarity), and *L. plantarum* (31.3% similarity). Amino acid sequence of the *B. flavum* EII^{Glc} was aligned with those of three EIIs^{Glc} showing high homologies. Amino acid sequences of the linker region connecting IIA domain and IIB domain exhibited the lowest similarities among the EIIs^{Glc} (Fig. 3). The linker region of *B. flavum* EII^{Glc} consisted of Gly-Ala repeats (495-509), while IIA domain of the other EII^{Glc} was connected to the IIB domain by a PA linker consisting of Pro-Ala repeats.

BFA	1	-----M	ASKLTTTSQHI	LENLGGPDNI	TSMTHCATRL	RFQVKDQSI	VDQQEDI	SDPSVLG	VVPQSG	STMQVVMG	SVANYQEI	ILKLDGM	QHFADGEA
CGL	1	-----M	ASKLTTTSQHI	LENLGGPDNI	TSMTHCATRL	RFQVKDQSI	VDQQEDI	SDPSVLG	VVPQSG	STMQVVMG	SVANYQEI	ILKLDGM	QHFADGEA
BLA	1	-----M	ASKLTTTSQHI	LENLGGPDNI	TSMTHCATRL	RFQVKDQSI	VDQQEDI	SDPSVLG	VVPQSG	STMQVVMG	SVANYQEI	ILKLDGM	QHFADGEA
CEF	1	MTYPVDTP	GPTRSERFD	V*****D	E*****V	A***SALD	A*****K	*****G	R****	*****K	*****K	*****K	*****K
BFA	94	TESSSKKEY	GGVRGKYSW	IDYAFEF	LSDTFRP	ILWALLGAS	LITLLVL	ADTFGL	QDFRAP	MEQPD	TYVFLH	SMWRSV	FYFLP
CGL	94	TESSSKKEY	GGVRGKYSW	IDYAFEF	LSDTFRP	ILWALLGAS	LITLLVL	ADTFGL	QDFRAP	MEQPD	TYVFLH	SMWRSV	FYFLP
BLA	94	TESSSKKEY	GGVRGKYSW	IDYAFEF	LSDTFRP	ILWALLGAS	LITLLVL	ADTFGL	QDFRAP	MEQPD	TYVFLH	SMWRSV	FYFLP
CEF	111	APAAGS	V*****G	*****V	*****A	*****S	*****A	*****K	*****G	*****K	*****K	*****K	*****K
BFA	204	TPEFLAL	GSAGDTVT	VFGIP	MVLNDY	SGQLFP	PLIAA	IGLYV	WEKGLK	KIPEAV	QMVFP	VPFSS	LLIMIP
CGL	204	TPEFLAL	GSAGDTVT	VFGIP	MVLNDY	SGQLFP	PLIAA	IGLYV	WEKGLK	KIPEAV	QMVFP	VPFSS	LLIMIP
BLA	204	TPEFLAL	GSAGDTVT	VFGIP	MVLNDY	SGQLFP	PLIAA	IGLYV	WEKGLK	KIPEAV	QMVFP	VPFSS	LLIMIP
CEF	221	*****L	*****V	*****A	*****S	*****A	*****K	*****G	*****K	*****K	*****K	*****K	*****K
BFA	314	PFLVPL	GLHWPL	NAIMQ	INILGY	DFIQG	PMGAW	NACFGL	VLTGV	FLLSIK	ERMKAM	RQVSL	GGML
CGL	314	PFLVPL	GLHWPL	NAIMQ	INILGY	DFIQG	PMGAW	NACFGL	VLTGV	FLLSIK	ERMKAM	RQVSL	GGML
BLA	314	PFLVPL	GLHWPL	NAIMQ	INILGY	DFIQG	PMGAW	NACFGL	VLTGV	FLLSIK	ERMKAM	RQVSL	GGML
CEF	331	*****L	*****V	*****A	*****S	*****A	*****K	*****G	*****K	*****K	*****K	*****K	*****K
BFA	424	AYAFV	FTSLT	IIPAMP	PWLYG	YITIG	IAVA	FFVSM	FLVLD	YRSNEER	DEARAK	VAAKQ	AEE
CGL	424	AYAFV	FTSLT	IIPAMP	PWLYG	YITIG	IAVA	FFVSM	FLVLD	YRSNEER	DEARAK	VAAKQ	AEE
BLA	424	AYAFV	FTSLT	IIPAMP	PWLYG	YITIG	IAVA	FFVSM	FLVLD	YRSNEER	DEARAK	VAAKQ	AEE
CEF	441	*****A	*****T	*****L	*****FF	*****DA	*****K	*****Q	*****M	*****AE	*****T	*****N	*****TPA
BFA	521	GQLVE	ITSPSK	AKAIPL	SEVDP	IFAAG	KLGP	GAIAI	QPTG	TNTV	VAPAD	ATIL	VQKSG
CGL	529	*EV*D	*V*LEG	*****E	*****E	*****E	*****E	*****E	*****E	*****E	*****E	*****E	*****E
BLA	521	*****L	*****V	*****A	*****S	*****A	*****K	*****G	*****K	*****K	*****K	*****K	*****K
CEF	551	*ETA	*L*A	*LEG	*V*****E	*****V	*****E	*****V	*****E	*****V	*****E	*****V	*****E
BFA	631	IRSKD	LPI	TPVVS	NAAK	FGIE	IGIP	ADQANS	STTV	IKV	NGKNE	(675)	
CGL	639	*****N	*****N	*****V	*****E	*****A	*****DAT	*****AE	(683)				
BLA	631	*****N	*****N	*****V	*****E	*****A	*****DAT	*****AE	(674)				
CEF	661	**A	*G*****T	*Q**S	*V**H**Q	*S	*V	*D	*A**S	*I**T**E	(704)		

Fig. 3. Comparison of *B. flavum* EII^{Glc} (BFL) with *C. glutamicum* EII^{Mn} (CGL), *B. lactofermentum* EII^{Glc} (BLA), and *C. efficiens* EII^{Glc} (CEF). The amino acid sequences of the four polypeptides are shown by a one-letter code and have been aligned by introducing gaps (hyphens) to maximize the similarities. The residues identical to the amino acid sequence of *B. flavum* EII^{Glc} are indicated by asterisks in all other sequences. The histidine and cysteine residues, which may be phosphorylated, are indicated by closed circles above the sequences. The numbers at the beginning of each line correspond to the amino acid positions in the protein.

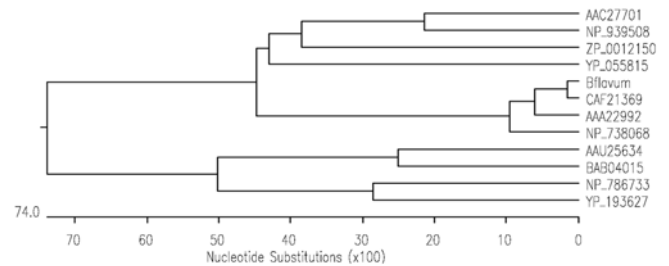


Fig. 2. Phylogenetic analysis of 16S rRNA. Neighbor-joining tree showing the phylogenetic relationships based on amino acid sequences of glucose-specific EIIs and β -glucoside-specific EIIs of several bacteria including *C. glutamicum* (CAF21369), *C. efficiens* (NP_738068), *C. diphtheriae* (NP_939508), *B. lactofermentum* (AAA22992), *B. ammoniagenes* (AAC27701), *Propionibacterium acnes* (YP_055815), *Bifidobacterium longum* (ZP_00121506), *Bacillus halodurans* (BAB04015), *B. licheniformis* (AAU25634), *Lactobacillus acidophilus* (YP_193627), *L. plantarum* (NP_786733), and *B. flavum*. GenBank accession numbers are described in parenthesis. Below the tree scale indicates the number of nucleotide substitutions for protein sequences.

The *B. flavum* EII^{Glc} contained two amino acid residues, which mediate phosphate transfer from phospho-HPr to glucose. One was a histidyl residue at position 594, homologous to the active site, His, of *E. coli* EIIA^{Glc}, which was phosphorylated by the phospho-HPr protein.¹⁶⁾ The other was assumed to be a cystidyl residue (28 aa) in the IIB domain of the protein. The

EII^{Glc} of *E. coli* was phosphorylated at a Cys residue (421 aa) by a soluble phospho-EIIA^{Glc}.¹⁷⁾ The amino acid sequence conservation (HCATRLR) around the Cys residue of *B. flavum* EII^{Glc} was also shared by other EIIs specific for glucose, sucrose, and β -glucoside of glucose-PTS.

References

1. Meadow, N. D., Fox, D. K. and Roseman, S. (1990) The bacterial phosphoenolpyruvate:glycose phosphotransferase system. *Annu. Rev. Biochem.* **59**, 497-542.
2. Saier Jr, M. H. and Reizer, J. (1992) Proposed uniform nomenclature for the proteins and protein domains of the bacterial phosphoenolpyruvate:sugar phosphotransferase system. *J. Bacteriol.* **174**, 1433-1438.
3. Postma, P. W., Lengeler, J. W. and Jacobson, G. R. (1993) Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol. Rev.* **57**, 543-594.
4. Malin, G. M. and Bourd, G. I. (1991) Phosphotransferase-dependent glucose transport in *Corynebacterium glutamicum*. *J. Appl. Bacteriol.* **71**, 517-523.
5. Mori, M. and Shio, I. (1987) Phosphoenolpyruvate:sugar phosphotransferase system and sugar metabolism in *Brevibacterium flavum*. *Agric. Biol. Chem.* **51**, 2671-2678.
6. Yoon, K. -H., Yim, H. and Jung, K. H. (1998) Cloning, expression, and nucleotide sequencing of the gene encoding glucose permease of phosphotransferase system from *Brevibacterium ammoniagenes*. *J. Microbiol. Biotechnol.* **8**, 214-221.
7. Yoon, K. -H., Lee, K. -N., Lee, J. -K. and Park, S. C. (1999) Cloning, nucleotide sequencing, and characterization of *ptsG* gene encoding the glucose-specific Enzyme II of the phosphotransferase system from *Brevibacterium lactofermentum*. *J. Microbiol. Biotechnol.* **9**, 582-588.
8. Lee, J. -K., Sung, M. -H., Yoon, K. -H., Yu, J. -H. and Oh, T. -K. (1994) Nucleotide sequence of the gene encoding the *Corynebacterium glutamicum* mannose enzyme II and analyses of the deduced protein sequence. *FEMS Microbiol. Lett.* **119**, 137-146.
9. Kwon, I., Lee, K. N., Lee, J. -K., Pan, J. -G., Oh, T. -K., Lee, H. H. and Yoon, K. -H. (1995) Cloning and expression of the gene encoding glucose permease of the phosphotransferase system from *Brevibacterium flavum* in *Escherichia coli*. *J. Microbiol. Biotechnol.* **5**, 188-193.
10. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
11. Malumbres, M., Gil, J. A. and Martin, J. F. (1993) Codon preference in corynebacteria. *Gene.* **134**, 15-24.
12. Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E., Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikee, K. and Gojobori, T. (2003) Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of *Corynebacterium efficiens*. *Genome Res.* **13**, 1572-1579.
13. Fudou, R., Jojima, Y., Seto, A., Yamada, K., Kimura, E., Nakamatsu, T., Hiraishi, A. and Yamanaka, S. (2002) *Corynebacterium efficiens* sp. nov., a glutamic-acid-producing species from soil and vegetables. *Int. J. Syst. Evol. Microbiol.* **52**, 1127-1131.
14. Kyte, J. and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132.
15. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402.
16. Presper, K. A., Wong, C. -Y., Liu, L., Meadow, N. D. and Roseman, S. (1989) Site-directed mutagenesis of the phosphocarrier protein, III^{Glc}, a major signal-transducing protein in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **86**, 4052-4055.
17. Meins, M., Jen, P., Mller, P., Richter, W. J., Rosenbusch, J. P. and Erni, B. (1993) Cysteine phosphorylation of the glucose transporter of *Escherichia coli*. *J. Biol. Chem.* **268**, 11604-11609.