The Glucose Transporter of the *Escherichia coli* Phosphotransferase System: Linker Insertion Mutants and Split Variants†

Rudolf Beutler, Markus Kaufmann,‡ Francesco Ruggiero, and Bernhard Erni*

Departement für Chemie und Biochemie, Universität Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

Received November 22, 1999; Revised Manuscript Received January 14, 2000

ABSTRACT: The IICB^{Glc} subunit of the glucose transporter acts by a mechanism which couples vectorial translocation with phosphorylation of the substrate. It contains 8 transmembrane segments connected by 4 periplasmic, 2 short, 1 long (80 residues), cytoplasmic loops and an independently folding cytoplasmic domain at the C-terminus. Random DNase I cleavage, *Eco*RI linker insertion, and screening for transport-active mutants afforded 12 variants with between 46% and 116% of wild-type sugar phosphorylation activity. They carried inserts of up to 29 residues and short deletions in periplasmic loops 1, 2, and 3, in the long cytoplasmic loop 3, and in the linker region between the membrane spanning IIC^{Glc} and the cytoplasmic IIB^{Glc} domains. Disruption of the gene at the sites of linker insertion decreased the expression level and diminished phosphotransferase activity to between 7% and 32%. IICB^{Glc} with a discontinuity in the cytoplasmic loop was purified to homogeneity as a stable complex. It was active only if encoded by a dicistronic operon but not if encoded by two genes on two different replicons, suggesting that spatial proximity of the nascent polypeptide chains is important for folding and membrane assembly.

The glucose transporter of *Escherichia coli* mediates uptake concomitant with phosphorylation of glucose and related hexoses. It consists of the two subunits IICB^{Glc} [1] and IIA^{Glc}. IIA^{Glc} is a cytoplasmic protein; IICB^{Glc} is a polytopic membrane protein of 51 kDa. IIA^{Glc} and IICB^{Glc} are the two terminal components of the phosphorylation cascade (PEP → Enzyme I → HPr → IIA^{Glc} → IICB^{Glc} → Glc) of the PEP: sugar phosphotransferase system (PTS) (for reviews, see refs *1*−*4*). The IICB^{Glc} subunit consists of two domains (*5*). The hydrophobic IIC^{Glc} domain (residues 1 to ∼380) contains the glucose binding site. A topology model based on protein fusions between progressively truncated IIC^{Glc} and alkaline phosphatase or β-galactosidase predicts IIC^{Glc} to span the membrane 8 times (*6*). This model will be used as a framework to present the results described below. The hydrophilic IIB^{Glc} domain (residues ∼390−477) contains the phosphorylation site with the active site residue Cys-421 (*7*). The IIC^{Glc} and IIB^{Glc} domains are linked by the LKTPGRED (residues 381−388) sequence which is highly conserved in the homologous transporters belonging to the glucose family of the PTS (*8*, *9*). IICB^{Glc} could be split at this linker, and IIC^{Glc} and IIB^{Glc} expressed as separate, fully functional polypeptide subunits which retained phosphotransferase activity when combined in vitro (*10*). Subsequently, the structure of the IIB^{Glc} domain could be solved by multidimensional NMR spectroscopy (*11*, *12*). However, the structure of the membrane domain is still unknown. The objective of this work was to investigate whether IIC^{Glc} could be further split into stable subdomains, suitable for the characterization of the molecular mechanism of function and for structure determination. There is accumulating evidence, indeed, that the backbone of polytopic membrane proteins can be cleaved with no or only minor effects on folding and activity. Up to 6 contiguous histidines could be inserted into the interhelix loops of the lactose permease (*13*), and up to 16 residues could be inserted into the surface loops of FhuA, a β-barrel protein of the outer membrane of *E. coli* (*14*). Active lactose permease could be assembled in vivo from discontinuous polypeptide chains (*15*, *16*), and purified complementary membrane fragments of bacteriorhodopsin could be reassembled in vitro (*17*, *18*). The outer membrane porin OmpA could be dissected at its periplasmic turns into fragments which refolded into a structure very similar to the wild-type protein (*19*). To identify sites at which IICB^{Glc} can be split into functional fragments, we first performed random linker insertion mutagenesis to screen for active insertions without preconceived notions as to where they should be. We then inserted oligonucleotide cassettes encoding a translational stop, and a restart sequence to afford four variants of IICB^{Glc} with discontinuous, nonoverlapping polypeptide backbones.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Plasmids, and Growth Media. E. coli K12 ZSC112LΔG (*ptsG::cat manZ glk*) was used as host

† This work was supported by Grant 31-45838.95 from the Swiss National Science Foundation.

* To whom correspondence should be addressed. Tel.: ++41-31-6314346. Fax: ++41-31-6314887. E-mail: erni@ibc.unibe.ch.

‡ Present address: Biochemisches Institut, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland.

[1] Abbreviations: PTS, phosphoenolpyruvate-dependent carbohydrate: phosphotransferase system; IICB^{Glc}, transmembrane subunit of the glucose transporter (EC 2.7.1.69); IIC^{Glc}, transmembrane domain of IICB^{Glc}; IIB^{Glc}, cytoplasmic domain of IICB^{Glc}; CL1−3, cytoplasmic loops 1−3 of IICB^{Glc}; PL 1−4, periplasmic loops 1−4 of IICB^{Glc}; IIA^{Glc}, hydrophilic subunit of the glucose transporter (EC 2.7.1.69); IICBA^{GlcNAc}, *N*-acetylglucosamine transporter (EC 2.7.1.69); HPr, histidine-containing phosphocarrier protein of the PTS; PEP, phosphoenolpyruvate; αMG, α-methyl-D-glucopyranoside.

10.1021/bi992679t CCC: $19.00
Published on Web 03/09/2000

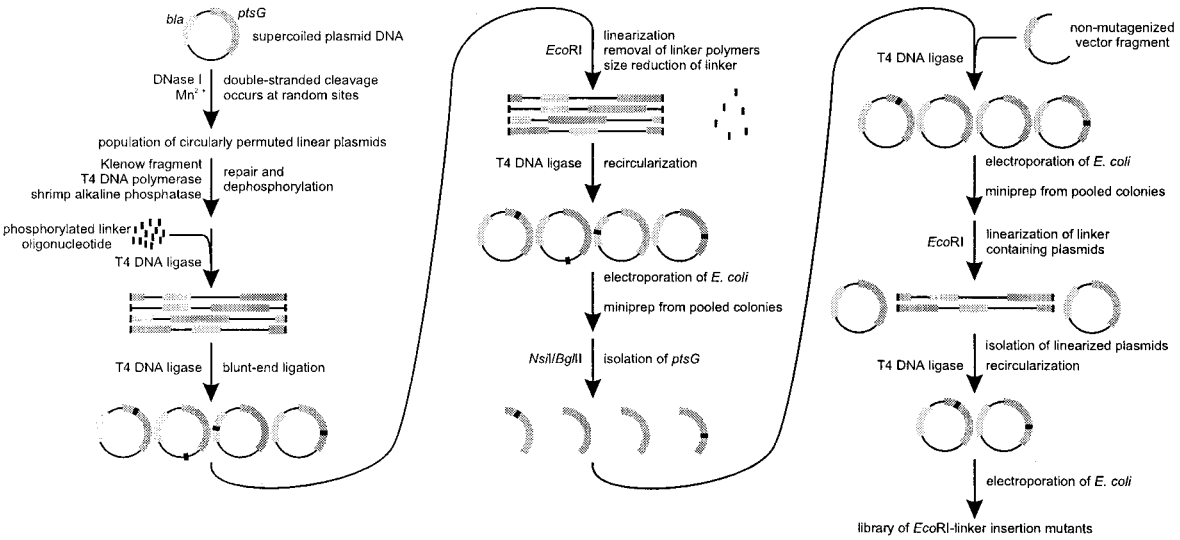

FIGURE 1: Scheme of linker insertion mutagenesis. For details, see Experimental Procedures and Results.

for all experiments (*10*). XL1-Blue (Stratagene) was used for cloning and plasmid amplification. pTSGH11 (*8*) contains *lacI* and encodes under the control of Ptac a $IICB^{Glc}$ with a C-terminal hexahistidine tag. pQEGH12 encodes under the control of its own promoter $IICB^{Glc}$ with a C-terminal hexahistidine tag and carries the genes *bla* and *cat* conferring resistance to ampicillin and chloramphenicol, respectively (*20*). *cat* was removed in order to reduce the size of the target for linker insertion mutagenesis as follows: pQEGH12 was digested with *Hin*dIII and *Nde*I, the ends were made flush with T4 DNA polymerase, and the vector fragment was self-ligated to afford pQEGH12Δcat.

Linker Insertion Mutagenesis. A set of mutants of pQEGH12Δcat containing an *Eco*RI linker anywhere in *ptsG* was constructed as follows (Figure 1): 6 μg of supercoiled pQEGH12Δcat was incubated with 0.25 ng of DNase I in the presence of 1 mM Mn^{2+} at pH 7.5 for 12−20 min at room temperature. The phenolized DNA was treated with T4 DNA polymerase, Klenow fragment, and shrimp alkaline phosphatase. The linearized plasmid was isolated by agarose gel electrophoresis and ligated with a 30-fold molar excess of 5′-phosphorylated heteroduplex linker DNA (GAATTC-CATATGGAATTC). To remove linker polymers, the ligation product was digested with *Eco*RI, phenolized, and religated. *E. coli* XL1-Blue was electroporated with the ligation mixture, and the transformed cells were plated on LB agar containing 100 μg/mL ampicillin. Colonies of an almost confluent lawn were washed from the plate, plasmids were extracted, and the sequence coding for $IICB^{Glc}$ was excised with *Nsi*I and *Bgl*II and ligated with the complementary, nonmutagenized vector fragment of pQEGH12Δcat. XL1-Blue was electroporated with the ligation product and plated as described above, and the plasmid pool digested with *Eco*RI. The linearized plasmids were separated from the undigested circular form by agarose gel electrophoresis, extracted, and self-ligated. ZSC112LΔG was electroporated with the ligation mixture and plated on MacConkey plates supplemented with 0.4% Glc and 100 μg/mL ampicillin to screen for glucose-fermenting (red) colonies which had retained *ptsG* activity.

Construction of Split Variants of ptsG. Plasmids pQEGPL1A, pQEGPL2B, pQEGPL3B, and pQEGCL3 which contained *Eco*RI linkers in *ptsG* were linearized with *Eco*RI, dephosphorylated with shrimp alkaline phosphatase, and ligated with an excess of 5′-phosphorylated heteroduplex DNA (AATTAAGGAGGTATACATATG and AATTCATAT-GTATACCTCCTT) encoding a translation stop, a ribosome binding sequence, and a *Bst*1107I diagnostic restriction site to afford pQEGPL1SS, pQEGPL2SS, pQEGPL3SS, and pQEGCL3SS. pJFGCL3SS was constructed by ligation of the *ptsG* insert fragment of pQEGCL3 and the vector fragment of pTSGH11, both digested with *Nsi*I and *Bam*HI.

Subcloning of the Split ptsG Genes on Different Replicons. pACG′274, encoding residues 1−274 of $IICB^{Glc}$, was constructed by ligation of the insert fragment of pJFGCL3SS with the vector fragment of pABG421 (*10*), both prepared with *Nco*I and *Eco*RI. pJF′G275H, encoding residues 275−476 of $IICB^{Glc}$, was constructed by ligation of the pJFGCL3SS insert fragment with the pMSEH2 vector fragment, both prepared with *Nde*I and *Sca*I. pMSEH2 contains the heteroduplex AATTCTAGAAATAATTTTGTTAACTT-TAAGAAGGAGATATACATATGGAGCT and CCATAT-GTATATCTCCTTCTTAAAGTTAACAAAATTATTTCT-AG inserted between *Eco*RI and *Sac*I in the multiple cloning site of pJF119EH (*21*).

Overproduction and Purification of Proteins. Membranes from ZSC112LΔG(pJFGCL3SS) were prepared, and the proteins were purified as described (*22*) with the following modifications. Membranes from 1 L of cell culture were suspended in 4 mL of buffer A (10 mM Tris-HCl, pH 9.3, 10 mM β-mercaptoethanol) and solubilized by the addition of Cymal-6 (Anatrace) to a final concentration of 15 mM, vortexed, incubated on ice for 15 min, sonicated in a bath-type sonicator for 1 min, and incubated on ice for another 15 min. Insoluble material was removed by ultracentrifugation; the supernatant was mixed with 4 mL of Ni^{2+}−nitrilotriacetic acid agarose (Qiagen), equilibrated with buffer B, pH 8 (50 mM NaP_i, 500 mM NaCl, 10 mM β-mercaptoethanol, 1.2 mM Cymal-6), and incubated for 30 min at room temperature. The slurry was transferred into a chromatography column at 4 °C, and excess solution was drained off. The column was washed at 4 °C with buffer B, pH 8 and pH 6, and with buffer C25 (25 mM imidazole, 50 mM NaP_i, pH 7.5, 500 mM NaCl, 10 mM β-mercaptoethanol,

1.2 mM Cymal-6), and the bound protein was eluted with buffer C80 (80 mM imidazole, 50 mM NaP_i, pH 7.5, 500 mM NaCl, 10 mM β-mercaptoethanol, 1.2 mM Cymal-6).

Assay for PEP:Sugar Phosphotransferase Activity. In vitro phosphorylation of [^{14}C]Glc (1000 dpm/nmol) was assayed by ion-exchange chromatography (*23*, *24*).

Assay for Uptake Activity. Sugar uptake by intact bacteria was determined as described (*25*) with the following modifications. Cells were grown to $A_{550} = 0.5$ in 250 mL of M9 minimal medium supplemented with 100 μg/mL ampicillin, 0.1% casamino acids, 1% glycerol, and 10 μM isopropyl-1-thio-β-D-galactopyranoside to induce a physiologically optimal level of IICB^{Glc}, collected by centrifugation, and resuspended in 3 mL of ice-cold M9 medium and incubated for 30 min on ice. Then 0.33 mL of this cell suspension was diluted with 0.77 mL of M9 medium and aerated for 10 min at room temperature. Uptake was started by the addition of 12 μL of 10 mM [^{14}C]αMG (6000 dpm/nmol). Aliquots of 100 μL were withdrawn after 5, 10, 15, 20, 30, 45, and 60 s, diluted into 8 mL of ice-cold M9 containing 0.4 mM αMG, and filtered through GF/F (Whatman) glass fiber filters under suction. The filters were washed under suction with 20 mL of 0.15 M NaCl, and radioactivity was determined by liquid scintillation counting.

Other Techniques. Protein samples were not boiled in sample buffer before electrophoresis on standard 15−20% polyacrylamide gels (*24*). IICB^{Glc} was visualized on western blots with monoclonal mouse anti-IICB^{Glc} and horseradish peroxidase coupled rabbit anti-mouse IgG (*26*, *27*). Protein concentrations were determined by a modified Lowry assay (*28*) with bovine serum albumin as the standard.

RESULTS

Selection and Amino Acid Sequence of Permissive EcoRI Linker Insertions in ptsG. A library of plasmids with *Eco*RI linkers randomly inserted in the *ptsG* gene encoding IICB^{Glc} was constructed in three steps (Figure 1). Plasmid pQEGH12Δcat was linearized with DNase I in the presence of Mn^{2+}, and then religated in the presence of a 30-fold molar excess of phosphorylated *Eco*RI linker oligonucleotides. *E. coli* XL1-Blue was electroporated and plated on LB agar containing ampicillin. Plasmid DNA from approximately 20 000 pooled colonies was isolated; the *ptsG* gene was excised and ligated with nonmutagenized vector to get rid of unwanted insertions in regions other than *ptsG*. The ligation product was used to transform XL1-Blue by electroporation. Plasmid DNA was extracted from a confluent lawn (containing approximately 175 000 transformants). To enrich for plasmids containing the *Eco*RI linker in *ptsG*, plasmids were linearized with *Eco*RI, and the linear fragments were purified, recircularized, and amplified in XL1-Blue. The plasmid library obtained from 11 500 transformants was used to transform ZSC112LΔG. Transformants were serially diluted and plated on MacConkey plates. Colonies were scored on MacConkey glucose indicator plates. Of the approximately 60 000 colonies, 94% were yellow, 1% deep red, and 4% segmented. Plasmid DNA isolated from 132 deep red colonies was screened for the presence of an *Eco*RI restriction site in *ptsG*. A total of 28 plasmids contained an *Eco*RI linker, 90 did not, and 14 showed a heterogeneous restriction pattern. Of 11 yellow colonies, 10 contained a plasmid with an *Eco*RI site. The 28 plasmids with an *Eco*RI linker were sequenced; 23 had the *Eco*RI linker in the *ptsG* coding region. The relevant amino acid sequences are listed in Figure 2. The *Eco*RI linkers are flanked by insertions (partial duplications) of up to 29 residues or short deletions of less than 4 residues as expected when staggered double-strand breaks produced by DNase I are either filled or resected prior to linker insertion. A total of 9 linker insertions were in the interdomain region between IIC^{Glc} and IIB^{Glc}; 4 of the 9 had distinct sequences, and 5 were exact duplicates and probably of clonal origin. A total of 3 linker insertions with 2 different sequences were in the first periplasmic loop, 8 insertions with 3 different sequences in the second, and 2 insertions in the third periplasmic loop. Only one linker insertion which did not compromise IICB^{Glc} activity was found in a cytoplasmic loop (loop 3) and none in putative membrane-spanning segments.

Construction of Split Variants of IICB^{Glc}. Split variants of *ptsG* were obtained by introducing a cassette encoding a translational stop, ribosome binding, and translational start site into the *Eco*RI sites of the permissive linker insertion mutants. The plasmids pQEGPL1SS, pQEGPL2SS, pQEGPL3SS, and pQEGCL3SS encode variants of IICB^{Glc} split in the first, second, and third periplasmic loops and the third cytoplasmic loop, respectively. In addition, pJFGCL3SS, a plasmid with an inducible promoter, was constructed to facilitate overexpression and purification of the complex split in the cytoplasmic loop.

Functional Characterization of Linker Insertion and Split Variants of IICB^{Glc}. PEP:sugar phosphotransferase activity of the IICB^{Glc} variants was measured in membrane preparations obtained by centrifugation of cell-free lysates. The 12 linker insertion mutants had between 46 and 116% of wild-type IICB^{Glc} activity (Figure 3). The specific activities (per total membrane proteins) of the three split variants obtained from the linker insertion mutants in the periplasmic loops were 10, 18, and 7% of the control. The variant split in the cytoplasmic loop had 32% activity (Figure 3). Taking into account that the split protein is expressed in smaller amounts than the wild type (Figure 5), the two proteins might have similar specific activities (see below). Sugar transport activity of the four split variants of IICB^{Glc} was assayed in intact bacteria with the nonmetabolizable substrate αMG. IICB^{Glc} split in periplasmic loops 1 and 2 had 36 and 24% of wild-type transport activity, the variant split in periplasmic loop 3 had no activity, and the variant split in the cytoplasmic loop retained 57% activity (Figure 4). With the exception of the loop 3 variant, all forms showed a proportionally higher in vivo transport than in vitro sugar phosphotransferase activity. The reduced activity of the split variants can be accounted for by a reduced expression level as shown on the immunblot of membrane extracts (Figure 5). Note, that because a monoclonal antibody directed against an epitope on the IIB^{Glc} domain (*27*) was used, only one fragment is visible of which the electrophoretic mobility increases with decreasing distance between the split site and the C-terminus.

The IICB^{Glc} variant split in the cytoplasmic loop seemed particularly interesting, not only because it had the highest in vivo activity but also because it was split at the end of a 80 residue long segment presumed to be an independently folding cytoplasmic domain. Two experiments were done

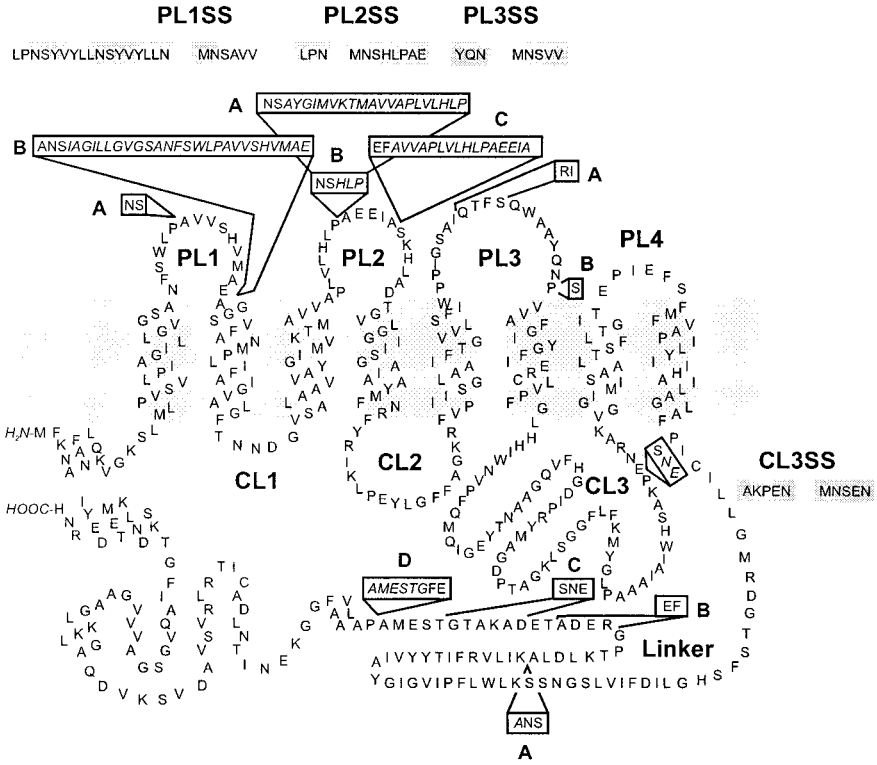

FIGURE 2: Secondary structure model of IICB^{Glc}. The model is based on the analysis of fusion proteins between progressively truncated IICB^{Glc} and alkaline phosphatase/β-galactosidase (*6*). The C-terminal and N-terminal sequences of the split variants (PL1SS−PL3SS and CL3SS) and the sites of discontinuity are indicated by *shaded boxes*. The sequences of the inserts and their location in the periplasmic (PL) and cytoplasmic (CL) loops are indicated by *open boxes*. Segments which are duplications of IICB^{Glc} sequences are indicated in *italics*. Letters A−D refer to Figure 3.

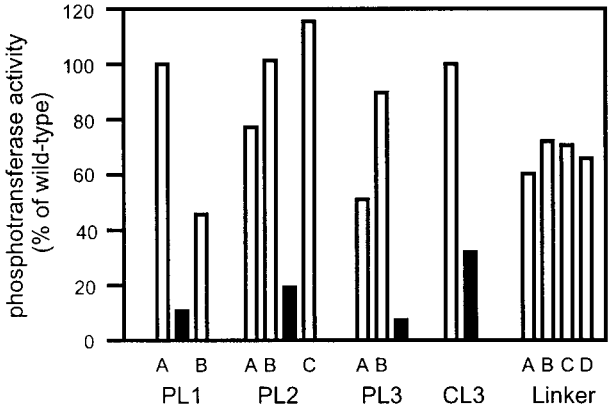

FIGURE 3: In vitro sugar phosphotransferase activities of linker insertion and split variants of IICB^{Glc}. *Open bars*: linker insertion mutants. *Solid bars*: split variants. Letters A−D refer to the linker insertion sequences shown in Figure 2. PL and CL refer to cytoplasmic and periplasmic loops. *Linker* refers to the hinge region between the IIC^{Glc} and the IIB^{Glc} domains. The split variants are derived from the linker insertion mutants located next to their left in this figure. Membranes containing wild-type IICB^{Glc} which were used as reference (100% activity) had between 10 000 and 30 000 nmol of Glc-6-phosphate (mg of membrane protein)$^{-1}$ (30 min)$^{-1}$, depending on the promoter strength.

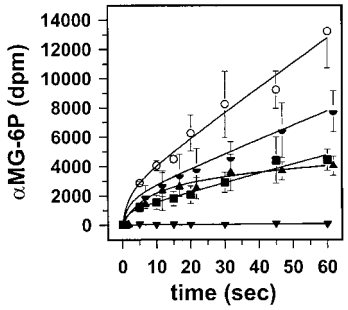

FIGURE 4: Uptake of [^{14}C]αMG by intact cells expressing split variants of IICB^{Glc}. In vivo transport activity of *E. coli* ZSC112LΔG expressing plasmid encoded proteins induced with 10 μM isopropyl-1-thio-β-D-galactopyranoside. The assay was started by the addition of 12 μL of 10 mM [^{14}C]αMG (6000 dpm/nmol) to 100 μL of a cell suspension as described under Experimental Procedures. Wild-type (open circles); split in PL1 (solid squares); split in PL2 (solid upward triangles); split in PL3 (solid inverted triangles); split in CL3 (semicircle, bottom solid).

to further characterize this region. First, chimeric proteins between IICB^{Glc} and the homologous IICBA^{GlcNAc} transporter were constructed by reciprocal exchange of the respective genes at codon 274. Bacteria expressing these proteins did not ferment Glc or GlcNAc on MacConkey indicator plates (results not shown). This suggests that interaction between helices 1−6 on one side and helices 7 and 8 on the other is highly specific and essential either for transport or for phosphorylation, or both activities. Second, the two halves of the split protein were expressed from two independent replicons. The regions coding for the N-terminal and C-terminal halves were cloned in two compatible plasmids and coexpressed, and membranes were prepared as described above. Approximately one-third of protein (Figure 5) and 32% phosphotransferase activity (Figure 3) were detected when the two fragments were encoded by a dicistronic operon. Unexpectedly, PEP:sugar phosphotransferase activity was only 1%, and no protein could be detected on immunblots (not shown) when the fragments were expressed from two replicons. It is likely that in the latter case the proteins are degraded faster than they can be assembled.

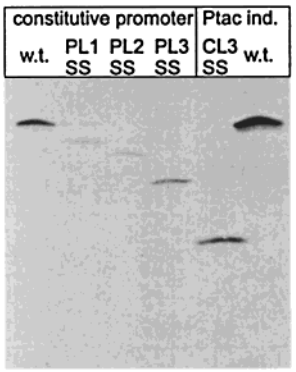

FIGURE 5: Immunblot of membranes containing split variants of $IICB^{Glc}$. Wild-type $IICB^{Glc}$ in lane 1 and split variants PL1SS−PL3SS are expressed under the control of the native *ptsG* promoter. Expression of wild-type $IICB^{Glc}$ in lane 6 and split variant CL3SS was induced with 100 μM isopropyl-1-thio-β-D-galactopyranoside. The blot was probed with a monoclonal antibody against the IIB domain of $IICB^{Glc}$.

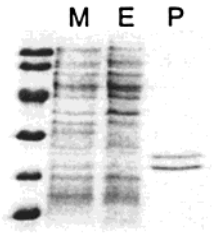

FIGURE 6: Purification of $IICB^{Glc}$ variant split in CL3. *M*, purified membranes; *E*, membrane extract; *P*, purified CL3SS on a Coomassie Blue stained SDS−PAGE. Molecular mass markers: 94, 67, 43, 30, 20.1, and 14.4 kDa.

Purification of the $IICB^{Glc}$ Variant Split in the Cytoplasmic Loop. Because the yield of the CL3 variant was low when the two peptides were expressed from two replicons, and because no protein could be detected, when either of the two fragments was expressed separately (results not shown), we attempted to separate the fragments during purification by chromatography. A hexahistidine tag was added to the C-terminus of the C-terminal fragment; the CL3 variant was overexpressed, solubilized with 15 mM Cymal-6 at pH 9.3, and purified in a single step by Ni^{2+} chelate affinity chromatography (Figure 6). Both, the N- and C-terminal fragments copurified and as a complex had the same specific sugar phosphotransferase activity as the wild-type protein.

DISCUSSION

*Eco*RI linker insertion mutagenesis of *ptsG* and screening for transport-active mutants afforded 23 clones, of which 12 were unique and 11 duplicates of possibly clonal origin. Four out of the 12 insertions were in the hinge region between the IIC^{Glc} and IIB^{Glc} domains which is known to be tolerant to structural modifications. This region is dispensable when the two domains are expressed as independent subunits (*10*), is partially tolerant to alanine scanning mutagenesis (*8*), and can serve as a fusion joint in a chimeric protein between IIC^{Glc} and $IIBA^{GlcNAc}$ (*9*), and as a novel N- and C-terminus of the circularly permuted variant $IIBC^{Glc}$ (*22*). The multiple, independent hits of this region indicate that sampling of the *ptsG* gene was dense enough that most if not all tolerant sites within $IICB^{Glc}$ should have been identified. The remaining eight insertions were clustered, too. Seven insertions were located in three regions assigned to periplasmic loops, and one was in a putative cytoplasmic loop. These results support a topology model of $IICB^{Glc}$ which was derived from the study of fusion proteins between truncated $IICB^{Glc}$ and either alkaline phosphatase or β-galactosidase (*6*). No linker insertions were found in the predicted periplasmic loop 4 and the cytoplasmic loops 1 and 2. The sequence alignment of six transporters homologous to $IICB^{Glc}$ reveals that periplasmic loop 4 and to a lesser extent also cytoplasmic loop 2 are part of conserved regions. In contrast, periplasmic loop 2 and cytoplasmic domain 3 are of very variable length and amino acid sequence in the different proteins. Periplasmic loops 1 and 3 and cytoplasmic loop 1 are inconspicuous. We did not attempt to introduce a linker into these "refractive" loops by site-directed mutagenesis. However, circularly permuted variants of $IICB^{Glc}$ with new N- and C-termini in periplasmic loop 4 and cytoplasmic loops 1 and 2 were found to be transport-negative, strongly indicating that the continuity of these regions indeed is essential for transport activity and/or stable folding of the protein.[2] It is therefore not surprising that linker insertions in these regions went undetected. Similarly, no insertions were found in the IIB^{Glc} domain which is a compact α/β sandwich containing short turns only (*11*).

The linker insertion mutants which were found by screening for a glucose fermentation positive phenotype of bacteria had between 46% and 116% of wild-type sugar phosphorylation activity. Subsequent splitting of the protein at sites of linker insertion reduced the expression level and activity to between 7% and 32%. Except for periplasmic loop 3, splitting of the protein affected nonvectorial phosphorylation more strongly than transport of glucose, although the latter process is mechanistically more demanding than the former.

The variant of $IICB^{Glc}$ split in the third cytoplasmic loop is of particular interest. It is at least 30 times more active if the genes encoding the two polypeptide fragments are cistrons of a single operon than if they are on two independent plasmids. This suggests that assembly and/or membrane insertion is a concerted process which requires spatial proximity of the two partners. It is not known whether the two fragments are inserted independently and then assembled by helix−helix interaction, or whether they are preassembled in the cytoplasm and then inserted into the membrane as a complex. In either case, the time interval for productive assembly might be short, and fragments degrade if they do not meet within this time. There is strong evidence that integral membrane proteins are extruded through a contiguous water-filled channel formed by the ribosome and the translocon (*29*, *30*). If the junction between ribosome and translocon is tight, the C-terminal fragment could sneak past the N-terminal fragment without the need for a restart of signal sequence mediated membrane insertion.

Assembly of functional proteins from two or more fragments has been reported for soluble proteins (e.g., *31*−*33*) and proteins of the inner and outer membrane. Bacteriorhodopsin expressed in *E. coli* and in *S. pombe* tolerates the insertion of a 13 residue peptide in periplasmic loops 2 and 3 and in cytoplasmic loops 1 and 2, but not in periplasmic loop 1 and cytoplasmic loop 3 (*34*). Bacteriorhodopsin could

[2] Beutler, R., Ruggiero, F., and Erni, B. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 1477−1482.

be assembled in *S. pombe* from two contiguous fragments split in the first cytoplasmic loop (*35*), and in vitro from proteolytically produced fragments (*17*, *36*, *37*), and from recombinant, contiguous as well as overlapping fragments split in each of the six interhelical loops (*18*). Peptides of 6 histidine residues were inserted into all 11 interhelical loops of the lactose permease with only 3 insertions (2 cytoplasmic, 1 periplasmic) causing a marked decrease of transport activity (*13*). Functional LacY could be assembled in vivo from contiguous as well as overlapping fragments. LacY with discontinuities in cytoplasmic loop 1 or 3 and in periplasmic loops 1, 2, or 4 were active if the fragments were encoded by two operons or a dicistronic operon on a single plasmid (*15*, *38*−*40*). In contrast, no activity was observed when the two fragments of LacY split in periplasmic loop 1 were expressed from two coexisting plasmids (*16*). This difference is reminiscent of the 30-fold difference described above for $IICB^{Glc}$ with a discontinuity in cytoplasmic loop 3. The correlation between efficiency of subunit assembly and distance of genes in space suggests that complex formation is initiated by short-lived intermediates which only together mature to the stable and functional form. The exact time course of this process will be elucidated by variation of the distance between the genes separated by overlapping stop and start codons, a cassette with a ribosome binding site, a complete cistron in a tricistronic operon, or by a noncoding sequence. In conclusion, the results described here further support a topology model of $IICB^{Glc}$ which was derived from gene fusions; they indicate that some but not all interhelix loops are essential for transport and/or phosphorylation, and they strongly suggest that productive fragment assembly is possible only when the nascent fragments are produced in close proximity.

ACKNOWLEDGMENT

We thank Karin Flükiger for technical assistance.

REFERENCES

1. Meadow, N. D., Fox, D. K., and Roseman, S. (1990) *Annu. Rev. Biochem. 59*, 497−542.
2. Postma, P. W., Lengeler, J. W., and Jacobson, G. R. (1993) *Microbiol. Rev. 57*, 543−594.
3. Postma, P. W., Lengeler, J. W., and Jacobson, G. R. (1996) in *Escherichia coli and Salmonella typhimurium*: *Cellular and Molecular Biology* (Neidhardt, F. C., Curtiss, R., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., Eds.) pp 1149−1174, ASM Press, Washington, DC.
4. Robillard, G. T., and Broos, J. (1999) *Biochim. Biophys. Acta 1422*, 73−104.
5. Erni, B., and Zanolari, B. (1986) *J. Biol. Chem. 261*, 16398−16403.
6. Buhr, A., and Erni, B. (1993) *J. Biol. Chem. 268*, 11599−11603.
7. Meins, M., Jenö, P., Müller, D., Richter, W. J., Rosenbusch, J. P., and Erni, B. (1993) *J. Biol. Chem. 268*, 11604−11609.
8. Lanz, R., and Erni, B. (1998) *J. Biol. Chem. 273*, 12239−12243.
9. Hummel, U., Nuoffer, C., Zanolari, B., and Erni, B. (1992) *Protein Sci. 1*, 356−362.
10. Buhr, A., Flükiger, K., and Erni, B. (1994) *J. Biol. Chem. 269*, 23437−23443.
11. Eberstadt, M., Grdadolnik, S. G., Gemmecker, G., Kessler, H., Buhr, A., and Erni, B. (1996) *Biochemistry 35*, 11286−11292.
12. Gemmecker, G., Eberstadt, M., Buhr, A., Lanz, R., Grdadolnik, S. G., Kessler, H., and Erni, B. (1997) *Biochemistry 36*, 7408−7417.
13. McKenna, E., Hardy, D., and Kaback, H. R. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 11954−11958.
14. Koebnik, R., and Braun, V. (1993) *J. Bacteriol. 175*, 826−839.
15. Wrubel, W., Stochaj, U., Sonnewald, U., Theres, C., and Ehring, R. (1990) *J. Bacteriol. 172*, 5374−5381.
16. Wrubel, W., Stochaj, U., and Ehring, R. (1994) *FEBS Lett. 349*, 433−438.
17. Huang, K. S., Bayley, H., Liao, M. J., London, E., and Khorana, H. G. (1981) *J. Biol. Chem. 256*, 3802−3809.
18. Marti, T. (1998) *J. Biol. Chem. 273*, 9312−9322.
19. Koebnik, R. (1996) *EMBO J. 15*, 3529−3537.
20. Waeber, U., Buhr, A., Schunk, T., and Erni, B. (1993) *FEBS Lett. 324*, 109−112.
21. Fürste, J. P., Pansegrau, W., Frank, R., Blocker, H., Scholz, P., Bagdasarian, M., and Lanka, E. (1986) *Gene 48*, 119−131.
22. Gutknecht, R., Manni, M., Mao, Q., and Erni, B. (1998) *J. Biol. Chem. 273*, 25745−25750.
23. Kundig, W., and Roseman, S. (1971) *J. Biol. Chem. 246*, 1393−1406.
24. Erni, B., Trachsel, H., Postma, P. W., and Rosenbusch, J. P. (1982) *J. Biol. Chem. 257*, 13726−13730.
25. Buhr, A., Daniels, G. A., and Erni, B. (1992) *J. Biol. Chem. 267*, 3847−3851.
26. Towbin, H., Staehelin, T., and Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4350−4354.
27. Meins, M., Zanolari, B., Rosenbusch, J. P., and Erni, B. (1988) *J. Biol. Chem. 263*, 12986−12993.
28. Markwell, M. A., Haas, S. M., Bieber, L. L., and Tolbert, N. E. (1978) *Anal. Biochem. 87*, 206−210.
29. Hamman, B. D., Chen, J. C., Johnson, E. E., and Johnson, A. E. (1997) *Cell 89*, 535−544.
30. Bibi, E. (1998) *Trends. Biochem. Sci. 23*, 51−55.
31. Betton, J. M., and Hofnung, M. (1994) *EMBO J. 13*, 1226−1234.
32. Kippen, A. D., and Fersht, A. R. (1995) *Biochemistry 34*, 1464−1468.
33. Ruiz-Sanz, J., de Prat, G. G., Otzen, D. E., and Fersht, A. R. (1995) *Biochemistry 34*, 1695−1701.
34. Teufel, M., Pompejus, M., Humbel, B., Friedrich, K., and Fritz, H. J. (1994) *EMBO J. 13*, 5794.
35. Hansen, O. K., Pompejus, M., and Fritz, H. J. (1994) *Biol. Chem. Hoppe Seyler 375*, 715−719.
36. Liao, M. J., London, E., and Khorana, H. G. (1983) *J. Biol. Chem. 258*, 9949−9955.
37. London, E., and Khorana, H. G. (1982) *J. Biol. Chem. 257*, 7003−7011.
38. Sahin-Tóth, M., Kaback, H. R., and Friedlander, M. (1996) *Biochemistry 35*, 2016−2021.
39. Bibi, E., and Kaback, H. R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4325−4329.
40. Zen, K. H., McKenna, E., Bibi, E., Hardy, D., and Kaback, H. R. (1994) *Biochemistry 33*, 8198−8206.

BI992679T