Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 3525–3530, March 1999
Biochemistry

Deduction of consensus binding sequences on proteins that bind IIA^{Glc} of the phospho*enol*pyruvate:sugar phosphotransferase system by cysteine scanning mutagenesis of *Escherichia coli* lactose permease

MELISSA SONDEJ*, JIANZHONG SUN†‡, YEONG-JAE SEOK§, H. RONALD KABACK†, AND ALAN PETERKOFSKY*¶

*Laboratory of Biochemical Genetics, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20892; §Department of Microbiology, College of Natural Sciences, Seoul National University, Seoul 151-742, Korea; and †Howard Hughes Medical Institute, Departments of Physiology and Microbiology and Molecular Genetics, Molecular Biology Institute, University of California, Los Angeles, CA 90024-1662

Contributed by H. Ronald Kaback, January 26, 1999

ABSTRACT **Mediated by the protein IIA^{Glc}, the phospho*enol*pyruvate:sugar phosphotransferase system plays a role in the regulation of activity of other sugar transport systems in *Escherichia coli*. By using a direct binding assay, a collection of single-Cys replacement mutants in cytoplasmic loops of lactose permease were evaluated for their capacity to bind IIA^{Glc}. Selected Cys replacements in loops IV/V or VI/VII result in loss of binding activity. Analysis of the mutagenesis results together with multiple sequence alignments of a family of proteins that interacts with IIA^{Glc} provides the basis for developing two regions of consensus sequence in those partner proteins necessary for binding to IIA^{Glc}. The requirement for two interaction regions is interpreted in the regulatory framework of a substrate-dependent conformational change that brings those two regions into an orientation optimal for binding IIA^{Glc}.**

Secondary transport systems use electrochemical gradients to accumulate solutes against a concentration gradient and are found in virtually all biological membranes. The lactose permease (lac permease) from *Escherichia coli* is the most extensively studied secondary transport system (reviewed in ref. 1). This polytopic, cytoplasmic membrane protein catalyzes the coupled, stoichiometric translocation of β-galactosides such as lactose or melibiose and H^+ (2). All evidence suggests that lac permease has 12 transmembrane, α-helical domains that traverse the membrane in a zigzag fashion (Fig. 1), and the tertiary-arrangement of most of the helices has been determined (1, 3, 4). These helical domains are connected by hydrophilic loops with the N and C termini on the cytoplasmic face of the membrane. Substrate binding to lac permease promotes conformational changes in which the helical domains are proposed to undergo sliding motions during substrate turnover (1, 4).

Another type of transport system found in bacteria corresponds to the phospho*enol*pyruvate:carbohydrate phosphotransferase systems (PTSs) (reviewed in ref. 5). The primary function of this multifunctional, multicomponent system is the concomitant phosphorylation and translocation of sugars across the cytoplasmic membrane (6). For glucose transport, the cytoplasmic proteins Enzyme I, HPr, and IIA^{Glc} transfer a phosphoryl group from phosphoenolpyruvate sequentially from one to the other, and then P-IIA^{Glc} passes the phosphate to the membrane protein $IIBC^{Glc}$, which both translocates and phosphorylates glucose. The unphosphorylated form of IIA^{Glc} also plays a role in the inhibition of transport of non-PTS sugars such as lactose, maltose, melibiose, and raffinose (termed inducer exclusion) (7–9). In the case of lactose, IIA^{Glc} binds to lac permease, inhibiting the transport of lactose (10, 11). Other allosteric regulatory functions of IIA^{Glc} include inhibition of the phosphorylation of glycerol by binding to glycerol kinase (GK) (12) and either inhibition or activation of adenylyl cyclase (13). Previous work (14) shows that binding of IIA^{Glc} requires lac permease to undergo a substrate-induced conformational change. Polyhistidine insertion into lac permease cytoplasmic loops IV/V or VI/VII or periplasmic loop VII/VIII does not affect substrate binding, as indicated by transport activity, but almost completely abolishes IIA^{Glc} binding.

Site-directed mutagenesis was used to create a version of lactose permease (15) in which all eight Cys residues were replaced with either Ser or Val (Cys-less permease). Cys-less permease is functional with respect to both lactose transport (15) and its ability to bind IIA^{Glc} (14, 16). This construct was used to create a library of mutants in which each individual amino acid residue in the permease was replaced with Cys (reviewed in ref. 1). Using this library, we examined cytoplasmic loop IV/V and the flanking residues in helix IV and V (Region I) and cytoplasmic loop VI/VII (Region II) from lac permease (see Fig. 1) to determine which residues play an important role in binding IIA^{Glc}.

EXPERIMENTAL PROCEDURES

Materials. Construction of a Cys-less permease and site-directed mutants of *lacY* encoding Cys-less permease have been described (1, 15). [^{3}H]IIA^{Glc} was purified from *E. coli* strain HK759 transformed with pRK248 and pEL110 (14, 17).

Membrane Preparations. Derivatives of *E. coli* K12 strain T184[$lacI^+$ O^+ Z^- Y^- (A) *rpsL met*$^-$ *thr*$^-$ *recA hsdR/F' lacI*^{q} O^+ Z^{D118} (Y^+ A^+)] (18) harboring plasmid pT7-5/cassette LacY encoding Cys-less lac permease (15) with indicated mutations were used for membrane preparations. Mutations in residues 136 and 221 were unavailable for analysis. Luria–Bertani medium (14) supplemented with 50 μg/ml ampicillin and 20 μg/ml streptomycin was inoculated with an overnight culture and grown at 37°C. When the culture reached an A_{600} = 0.5, lac permease expression was induced by addition of a final concentration of 1 mM isopropyl 1-thio-β-D-galactopyranoside, and growth was continued for another 4 hours. Cells were harvested by centrifugation for 10 min at 6,000 × *g*, were

Abbreviations: PTS, phospho*enol*pyruvate:sugar phosphotransferase system; GK, glycerol kinase; NEM, *N*-ethylmaleimide.

‡Current address: Department of Immunology, IMM-18, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92307.

¶To whom reprint requests should be addressed at: National Institutes of Health, Building 36, Room 4C-11, Bethesda, MD 20892. e-mail: alan@codon.nih.gov.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

washed once in 50 mM sodium phosphate buffer (pH 6.3), and were stored as a pellet at −20°C until use. To prepare membranes, cells were resuspended in phosphate buffer and were passed twice through a French press at 12,000 psi (1 psi = 6.89 kPa); cell debris/unbroken cells were removed by centrifugation for 10 min at 10,000 × *g*. Membrane vesicles were collected by centrifugation for 90 min at 100,000 × *g* and were washed once with phosphate buffer, and centrifugation was repeated to pellet the membranes. The final membrane vesicle preparation was resuspended to a protein concentration of ≈20 mg/ml as determined by the bicinchoninic acid protein assay (Pierce).

Binding Assays. [^{3}H]IIA^{Glc} was purified by previously described methods (14). The binding of [^{3}H]IIA^{Glc} to lac permease was performed by a method modified from Seok *et al.* (14). The binding assay contained, in 100 μl: 50 mM sodium phosphate (pH 6.3), 2 mM $MgCl_2$, 2 mM DTT, 750 μg of membrane protein from a given membrane vesicle preparation, and 2.4 μg of [^{3}H]IIA^{Glc} (1,080 cpm/μg). Incubation mixtures were set up in Beckman pollyallomer tubes (5 × 20 mm) with and without the addition of 4 mM melibiose. After 5 min of incubation at room temperature, membranes were collected by centrifugation in a Beckman Airfuge for 30 min at 100,000 × *g*. The supernatant solutions were discarded, and the tubes were rinsed with 100 μl of H_2O. The pellets were resuspended in 100 μl Triton X-100 and were transferred to scintillation vials. The tubes were rinsed with 100 μl of H_2O and were transferred to the vials containing appropriate resuspended pellets. After adding 10 ml of Aquasol (NEN), membrane-bound IIA^{Glc} was quantitated (cpm) by using a Beckman liquid scintillation counter. All results were reproduced at least two times.

Modification of Lac Permease Mutants. Single-Cys lac permease mutants were modified with *N*-ethylmaleimide (NEM). NEM was added to a final concentration of 5 mM to the phosphate buffer, membranes, and H_2O described in the binding assay above and was incubated for 5 min at room temperature. The reaction was terminated by adding to this mixture 10 μl of 0.1 M DTT and incubating for 5 min. The remaining binding assay reagents ($MgCl_2$, [^{3}H]IIA^{Glc}, and melibiose) were added, and the procedure was continued as described above.

RESULTS

Membrane vesicles were prepared from *E. coli* expressing Cys-less lac permease (control) or various site-directed mutants of Cys-less permease. The membranes were analyzed for IIA^{Glc} binding with radiolabeled IIA^{Glc} (14). Because IIA^{Glc} binds to lac permease in a sugar-dependent manner, values obtained in the absence of melibiose were subtracted from values in the presence of melibiose to correct for nonspecific binding. The data for the various single-Cys mutants are expressed as the percent of sugar-dependent binding observed with the Cys-less control (230 cpm).‖

In Region I (residues 126–144; see Fig. 1), two Cys replacement mutants, N137C and A143C, exhibited IIA^{Glc} binding comparable to the Cys-less control (Fig. 2*A*, filled squares). Several mutants in helix IV, A127C, F128C, I129C, and K131C bound a greater amount of IIA^{Glc} than Cys-less permease (321%, 158%, 393%, and 378%, respectively). Mutants E130C, V132C, R135C, E139C, F140C, and R142C exhibited low binding of IIA^{Glc} (13–32% of Cys-less lac permease) whereas mutants S133C, R134C, F138C, and G141C showed moderate binding (36–52% of Cys-less lac permease). Mutants E126C and R144C are unable to bind significant amounts of IIA^{Glc} because they do not bind substrate (19).

‖The substrate-dependent binding obtained with membranes containing Cys-less lac permease is ≈30% of that observed with wild-type lac permease.

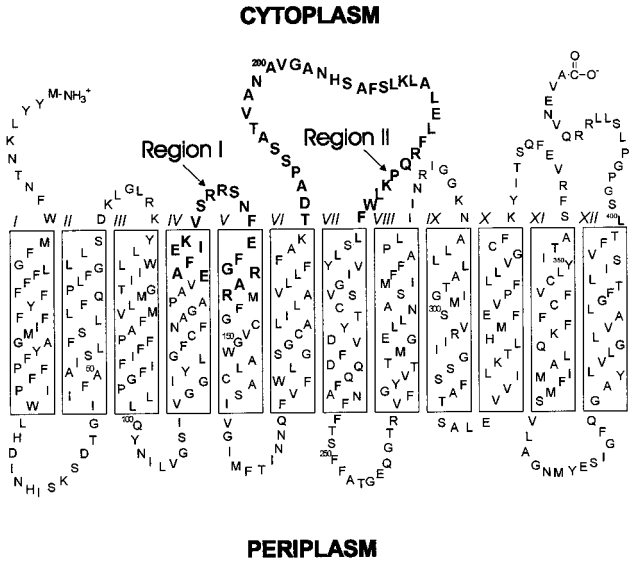

FIG. 1. Secondary-structure model for lactose permease. Two regions of lactose permease (boldface residues) were examined by cysteine-scanning mutagenesis to determine which residues play a role in IIA^{Glc} binding. Region I, cytoplasmic loop IV/V and the flanking residues in helix IV and V. Region II, cytoplasmic loop VI/VII.

In Region II (residues 189–244; Fig. 2*B*), nine mutants (T189C, D190C, A191C, P192C, V201C, A207C, K211C, L212C, and P220C) bound approximately the same amount of IIA^{Glc} as the Cys-less control (79–138% of Cys-less lac permease). Many mutants (F208C, L210C, A213C, L214C, E215C, L216C, F217C, L222C, W223C, and F224C) in the portion of the loop adjacent to helix VII along with A203C (168% of Cys-less lac permease) bound a greater amount of IIA^{Glc} than did the control (165–353% of Cys-less lac permease). No significant IIA^{Glc} binding was observed for mutants N204C and S206C whereas mutants T196C, V197C, N199C, and G202C exhibited very weak binding (6–12% of Cys-less permease). Mutants S193C, S194C, A195C, A198C, A200C, H205C, S209C, R218C, and Q219C exhibited intermediate levels of binding of IIA^{Glc} (35–69% of Cys-less permease).

The possibility was considered that replacement of various bulky residues by Cys was responsible for the enhanced binding of IIA^{Glc} observed in portions of regions I and II. This idea was tested by *N*-ethylmaleimide modification, which should restore a bulky sidechain with a variety of the Cys replacement mutants (shown as open squares in Fig. 2 *A* and *B*). All of the residues that exhibit enhanced binding of IIA^{Glc}, as well as a selected number of control residues throughout Regions I and II, were subjected to NEM modification. This modification had no or only a slight effect (a <50% decrease of the percent of control values for the unmodified membranes) on mutants A127C, F128C, E130C, V132C, N137C, F138C, E139C, F140C, and A143C in Region I and mutants D190C, V201C, A203C, E215C, F217C, and R218C in Region II. For other mutants in Region II (A191C, F208C, L210C, A213C, L214C, L216C, L222C, W223C, and F224C), modification caused a significant decrease in the amount of IIA^{Glc} bound. Surprisingly, two mutants from region I, I129C and K131C, bound an even greater amount of IIA^{Glc} (an additional 150–170% increase in the percent of control) when the Cys side chain was alkylated.

DISCUSSION

A major purpose of this study was to determine which side chains in loops IV/V and VI/VII of lac permease are critical

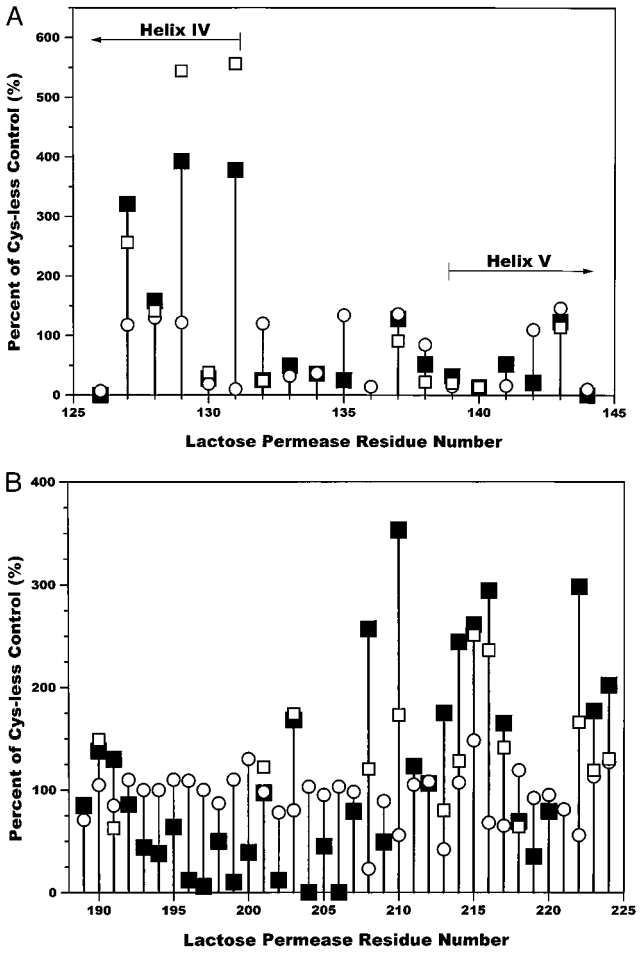

FIG. 2. Effect of cysteine-scanning mutagenesis on IIA^{Glc} binding. Lac permease residues from Region I (*A*) and Region II (*B*) (Fig. 1) were mutated to cysteine. The percent of IIA^{Glc} bound to the various cysteine mutants compared with the Cys-less lac permease control was determined (■); for some of the residues, the single-cysteine residues were modified with NEM before the binding assays (□). ○, the previously determined initial rates of [^{14}C]-lactose transport as percent of the Cys-less control (20–22). The average rate for Cys-less lac permease was 55 nmol·min^{-1}·(mg of protein)$^{-1}$.

for interaction with IIA^{Glc}. The approach used was to compare binding of IIA^{Glc} by lac permease with single-Cys replacements in loops IV/V (as well as some flanking residues) and VI/VII to Cys-less permease. Fig. 2 *A* and *B* shows the results of those studies. Because effective binding of IIA^{Glc} requires a substrate-dependent conformational change, it was essential to eliminate from consideration those mutations that result in loss of transport (i.e., binding) activity. Consequently, transport rates as reported previously (20–22) also are presented as percent of the Cys-less control in Fig. 2 (open circles). In Region I, residues 126, 130, 131, 133, 134, 136, 139, 140, 141, and 144 were eliminated from consideration because of decreased transport activity (Fig. 2*A*).

In Region II, it is interesting that mutation of residues 208, 213, and 222 was reported to result in decreased transport activity (20, 21), but the data in Fig. 2*B* show that the melibiose-dependent binding of IIA^{Glc} was actually enhanced. The decreased transport activity in the mutants is therefore unlikely to result from a defect in substrate binding.

Cys replacements for A127, I129, and K131 in Region I resulted in significantly enhanced binding of IIA^{Glc} (Fig. 2*A*). These residues are all located in Helix IV (19). It seems reasonable to conclude that, in wild-type lac permease, the arrangement of Helix IV places a limitation on IIA^{Glc} binding; this inhibition is relieved by any of these Cys replacements that may change the tilt of Helix IV and influence the orientation of loop IV/V with respect to loop VI/VII.

A similar clustering of residues in which Cys replacement resulted in increased IIA^{Glc} binding was observed in the region of loop VI/VII vicinal to Helix VII (Fig. 2*B*). The Cys replacements showing a substantial stimulation were A203, F208, L210, A213, L214, E215, L216, F217, L222, W223, and F224. Because all of these residues are hydrophobic, it is possible that Helix VII is actually longer, including those residues mentioned above that extend out of the membrane into the cytoplasm (see Fig. 1). In that case, the binding stimulation observed in that region may be explained in the same way as for those seen with the Helix IV changes (see above).

Those cases in which stimulation of binding was observed generally involve replacement of a bulkier side chain by the smaller Cys residue. Experiments were carried out to see whether adding more bulk to the Cys side chain by way of NEM modification results in reversal of stimulation (Fig. 2 *A* and *B*). There was essentially no effect of alkylation on any of the residues tested in Region I (Fig. 2*A*, open squares). However, at some sites in Region II, enhanced binding was reduced by NEM treatment. The positions responding to NEM were A191, F208, L210, A213, L214, L222, W223, and F224.

A combination of inspection of the data in Fig. 2*A* and sequence alignments was used to determine which residues are essential for IIA^{Glc} binding in Region I (Fig. 3*A*). Mutation of either V132, R135, or R142 in lac permease resulted in decreased IIA^{Glc} binding with no effect on transport. Alignment of this region with the corresponding sequences of the permeases for maltose (MalK, the peripheral membrane component), raffinose (RafB), and melibiose (MelB), showed conservation of the Val and two Arg residues (the second Arg residue was replaced with Lys in MalK, RafB, and MelB). In addition, S133 aligned well with conserved Ser, Ala, or Thr residues in the other permeases. Replacement of S133 with Cys caused decreases in both binding of IIA^{Glc} and transport. Because of the apparent conservation of this residue, we suggest that it may be important for IIA^{Glc} binding. In the alignment, a Gly residue always appeared in the position preceding that corresponding to R142 in lac permease. Although replacement of this residue with Cys is associated with decreased transport, as well as decreased IIA^{Glc} binding, it appears likely that the conserved Gly residue is important for IIA^{Glc} binding. This analysis suggests the five-residue consensus for IIA^{Glc} binding associated with Region I shown in Fig. 3*A* [(V, L)(X){0,2}(S, A, T)(X){0,2}(R)(X){0,6}(G)(K, R)].

The crystal structure of the complex of IIA^{Glc} with glycerol kinase has been elucidated. One of the most striking features of the interface is the absence of intermolecular hydrogen bonds (23). The region referred to as contact 1 includes residues in the proposed consensus (Fig. 3*A*). Therefore, the nature of interaction of IIA^{Glc} with the soluble protein glycerol kinase and membrane-bound permeases may be similar.

It previously was shown that a polyhistidine insertion between R134 and R135 (see Fig. 3*A*) of loop IV/V results in decreased IIA^{Glc} binding (14). This is precisely the region in which three of the key residues of the proposed consensus are localized. The other two residues of the consensus (G141 and R142) are located in Helix V proximal to loop IV/V (Figs. 1 and 3*A*). This is consistent with the idea that slight disturbances in the orientation of Helix V with respect to the other helices have substantial effects on IIA^{Glc} binding. In this context, it is worth noting that the interface between Helices IV and V has been identified as an important element in the substrate binding site (19).

A similar approach of multiple sequence alignment and analysis of the data of Fig. 2*B* was used to deduce residues in region II important for binding of IIA^{Glc}. Mutation of either

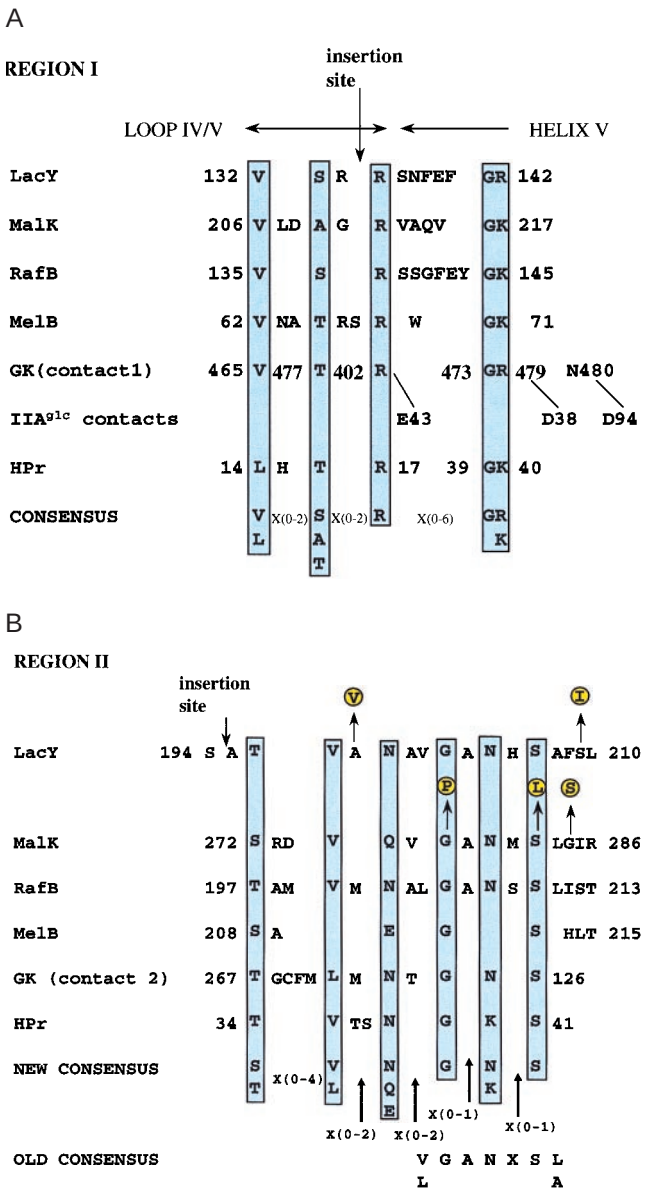

FIG. 3. Multiple sequence alignment of various *E. coli* proteins that interact with IIA^{Glc}. Accession numbers for the various proteins are as follows: LacY, p02920; RafB, p16522; MelB, K01991; GK, M18393; and HPr, J02796. The corrected sequence of MalK is from ref. 28. Residues proposed to be essential for binding to IIA^{Glc} are shown in blue shaded boxes. The IIA^{Glc} contacts with GK are from ref. 12. (*A*) Region I alignment. The site of polyhistidine insertion (labeled insertion site) in loop IV/V of lac permease is from ref. 14. (*B*) Region II alignment. The site of polyhistidine insertion (labeled insertion site) in loop VI/VII is from refs. 14 and 16. Mutants in LacY resulting in loss of inducer exclusion (shown as yellow shaded circles) are from ref. 11. Mutations in MalK resulting in loss of inducer exclusion (yellow shaded circles) are from refs. 8 and 24. The old consensus is from ref. 9.

T196, V197, N199, G202, N204, or S206 of lac permease resulted in essentially complete inhibition of IIA^{Glc} binding with no effect on transport. Alignment of this region with the corresponding sequences of the permeases for maltose (MalK, the peripheral membrane component), raffinose (RafB), and melibiose (MelB) showed absolute conservation of the Gly202 and Ser206 residues. The Thr196 residue was conservatively substituted by Ser in MalK and MelB. The Val197 residue was not found in MelB but was conserved in MalK and RafB. The Asn residue was conservatively substituted by Gln in MalK and by Glu in MelB. The second conserved Asn residue (204) was absent in MelB. Residues associated with contact 2 of glycerol kinase also aligned well with the conserved residues of lac permease. On the basis of this alignment, it is possible to propose a new consensus sequence for the interaction with IIA^{Glc} that consists of six residues [(S, T)(X){0,4}(V, L)(X){0,2}(N, Q, E)(X){0,2}(G)(X){0,1}(N, K)(X){0,1)(S)]. This differs from the previously proposed consensus (9) (labeled old consensus), which was based only on sequence comparisons of LacY, MalK, and RafB.

Further evidence that the region between residues 194 and 210 of lac permease is important derives from the previous finding (14, 16) that a polyhistidine insertion in this region decreases IIA^{Glc} binding. This decrease is probably caused by a steric interference that deters binding. A similar mechanism for binding inhibition may explain the replacement of A198 by Val and S209 (11) by Ile; in both of those cases, smaller residues are replaced by bulkier side chains.

Mutations that result in resistance to inhibition by IIA^{Glc} have been reported for MalK (Fig. 3*B*). Two of these mutations [G278P (8) and S282L (24)] correspond to residues in the suggested consensus. Another site of mutagenesis is residue Gly284 (8), flanking the consensus. In this case, a likely explanation for the inhibitory effect is imposition of a bulky group that interferes with the protein–protein interaction. Some mutations in the carboxyl-terminal tail of MelB result in resistance to PTS-dependent repression (25). Possibly, these mutations lead to long-range conformational changes that affect the orientation of the conserved regions described here.

An important catalytic interaction of IIA^{Glc} in the PTS pathway is with the phosphocarrier HPr. A model for this interaction has been proposed (26). Chen *et al.* (27) used NMR to demonstrate chemical shift changes on formation of the IIA^{Glc}-HPr complex that correspond to hydrophobic surfaces at the active sites. Of interest, scanning of the sequence of HPr revealed regions that correspond very well to the proposed consensus patterns for both Regions I and II (see Fig. 3 *A* and *B*). The active site region of HPr corresponding to Leu14-Arg17 constitutes a loop connecting the first β-strand to the first α-helix and the amino terminus of that helix. That region, together with residues Gly39-Lys40, provides a perfect match to the Region I consensus. Similarly, a β-sheet region (Thr34-Ser41) shows an excellent match to the Region II consensus. The analysis supports the idea that both catalytic and regulatory interactions involving IIA^{Glc} use similar regions.

The proposed consensus sequences for Regions I and II (see Fig. 3) contain various gaps that might result in a lack of specificity such that these sequences would be found in many proteins. To investigate this possibility, the FINDPATTERNS program (see Table 1) was used to search sequences from Fig. 3 for the frequency of occurrence of the Region I and II consensus sequences. Consistent with the layout of Fig. 3, LacY, RafB, and MalK contained only one occurrence each of consensus I and II; MelB contained one instance of the Region I consensus and none of the Region II consensus; and HPr contained one copy of the Region II consensus and none of the Region I pattern. The results of this analysis substantiate the specificity and uniqueness of the two proposed consensus sequences.

A further test of the specificity of the proposed consensus sequences was carried out. The complete *E. coli* genome was surveyed for the occurrence of the Region I and II sequence patterns by using the FINDPATTERNS program. Although there are 1,456 finds of the individual patterns in the 4,413 protein sequences examined, there are only 37 cases in which a sequence contains both patterns (see Table 1). This analysis provides further evidence for the specific and limited occurrence of the proposed recognition sequences.

For the purposes of discussion, the 37 hits listed in Table 1 are divided into five arbitrary groups. There are only four transport proteins identified in the search. The permeases for

Table 1. *E. coli* proteins containing regions I and II consensus patterns

Accession number†	Description of protein
Transport proteins	
p02920	Lactose permease
p02914	malK; maltose permease
p16552	rafB; raffinose permease
p45535	ABC transporter
Kinases	
p14377	Phosphorylates hydrogenase in response to environmental signals
p23837	Phosphorylates acid phosphatase in response to environmental signals
DNA proteins	
p25239	Restriction enzyme
p14565	DNA helicase I
p22706	DNA helicase I
q46802	Transcriptional regulator
Membrane proteins	
p53517	Outer membrane protein involved in assembly and export of afimbrial adhesin subunits
p39180	Controls colony form variation and autoaggregation
p46009	Outer membrane protein involved in assembly of fimbriae subunits
p07110	Member of fimbrial export usher family
p09745	Expressed by a shufflon (biological switch)
p75857	Fimbriae export usher
p25744	Integral membrane protein
p77804	Hypothetical transmembrane protein
p77315	Integral membrane protein; part of transport system
p52143	Hypothetical outer membrane protein
Cytosolic proteins	
p11648	Aminopeptidase
p31459	2-keto-3-deoxy-galactonokinase
p18775	Sulfoxide terminal reductase
p80668	Phenylacetaldehyde dehydrogenase
p52073	Glycolate oxidase subunit
p03827	Required for transposition of insertion element
p19767	Insertion element
p06959	Pyruvate dehydrogenase E2
p03067	Replication initiation protein
p16553	Glycosyl hydrolase
p11585	GTP pyrophosphokinase
p25741	Lipopolysaccharide core biosynthesis protein
p02359	Ribosomal protein S7
p76253	Member of ring-hydroxylating dioxygenase family
p42594	Hypothetical protein
p77689	Hypothetical protein
p32051	Hypothetical lipoprotein

The FINDPATTERNS program (Wisconsin Package Version 9.1, Genetics Computer Group, Madison, WI) was used to search the entire *E. coli* genome in the SwissProt database (sw:*_ecoli) for the consensus sequence patterns denoted Regions I [(V,L)(X){0,2}(S,A,T)(X){0,2}R(X){0,6}G(K,R)] and II [(S,T)(X){0,4}(V,L)(X){0,2}(N,Q,E)(X){0,2}G(X){0,1}(N,K)(X){0,1}S] in Fig. 3. The sequences in the table contain at least one copy of both sequence patterns.
†The accession numbers are from the SwissProt database.

lactose, maltose, and raffinose are cited in Fig. 3. An additional transporter (p45535), a member of the ABC transporter family for which the substrate has not yet been established, may be regulated by IIA^{glc}. There are two related kinases (p14377 and p23837) that are involved in environmental signaling. It is an attractive idea that the PTS, via IIA^{glc}, may be in a communication link with that process. Four proteins that interact with DNA were found to contain the pair of consensus sequences. If IIA^{glc} regulates any of these proteins, this might provide a link between the PTS and transcription. A group of membrane proteins was identified in the search. It is unlikely that the outer membrane proteins would be in the same compartment as IIA^{glc}. However, some of these proteins may be transmembrane and part of transport systems (e.g., p77315). When these proteins are more fully characterized, it would be of interest to search for a regulatory role of IIA^{glc} in their function.

Finally, there is a heterogeneous group of cytosolic proteins identified by the search. There are two noteworthy candidates for regulation by IIA^{glc} in this group. The first, pyruvate dehydrogenase E2 (p06959), is of obvious interest because it is involved in the utilization of pyruvate, one of the products of the PTS pathway. The second, GTP pyrophosphokinase (p11585), is involved in the synthesis of the global regulator ppGpp. It would be of obvious fundamental interest to tie this compound to the PTS pathway via IIA^{glc}.

As indicated above (see Fig. 3), the melibiose permease that is known to be regulated by IIA^{glc} contains only the Region I consensus but does not give a precise match for the Region II pattern. It is therefore possible that some sugar transporters and perhaps other proteins regulated by IIA^{glc} were not detected by the search. In this vein, it is noteworthy that several proteins involved with arabinose transport (p03021, regulatory protein for the arabinose operon; p08531, the inner membrane-associated transporter for L-arabinose; p23910, a protein involved in the transport or processing of arabinose polymers) contain the Region I consensus. Furthermore, two proteins involved in sucrose metabolism (p40715, repressor for Csc operon; p40714, sucrose hydrolase) also contain the

Region I pattern. It is also of interest that the protein involved in the carboxylation of pyruvate to form oxaloacetate (p00864) contains the Region II consensus. Exploration of the possible regulation of these proteins by IIA^{glc} may be a fruitful area for future research.

The proposed consensus sequences in regions I and II are consistent with a model in which interaction of IIA^{Glc} with its partners occurs by a common mechanism, whether the partner is membrane-bound (LacY, RafB, MelB) or soluble (MalK, GK, HPr). Moreover, the idea that there are two regions for interaction with IIA^{Glc} is supported by the previous findings that polyhistidine insertions in either of two cytoplasmic loops of lac permease result in loss of binding activity (14, 16). The model that substrate binding to a regulated protein brings the two binding regions into the appropriate alignment is an attractive interpretation of the data presented here.

We thank Lynn Miller of the Genetics Computer Group (Madison, WI) and Jonathan Reizer (Department of Biology, University of California at San Diego, La Jolla, CA) for helpful discussions dealing with searches for patterns in protein databases.

1. Frillingos, S., Sahin-Tóth, H., Wu, J. & Kaback, H. R. (1998) *FASEB J.* **12,** 1281–1299.
2. Kaback, H. R. (1983) *J. Membr. Biol.* **76,** 95–112.
3. Foster, D. L., Boublik, M. & Kaback, H. R. (1983) *J. Biol. Chem.* **258,** 31–34.
4. Kaback, H. R., Voss, J. & Wu, J. (1997) *Curr. Opin. Struct. Biol.* **7,** 537–542.
5. Postma, P. W., Lengeler, J. W. & Jacobson, G. R. (1996) in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington, DC), pp. 1149–1174.
6. Kaback, H. R. (1968) *J. Biol. Chem.* **243,** 3711–3724.
7. Saier, M. H., Jr., Novotny, M. J., Comeau-Fuhrman, D., Osumi, T. & Desai, J. D. (1983) *J. Bacteriol.* **155,** 1351–1357.
8. Dean, D. A., Reizer, J., Nikaido, H. & Saier, M. H., Jr. (1990) *J. Biol. Chem.* **265,** 21005–21010.
9. Titgemeyer, F., Mason, R. E. & Saier, M. H., Jr. (1994) *J. Bacteriol.* **176,** 543–546.
10. Osumi, T. & Saier, M. H., Jr. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 1457–1461.
11. Wilson, T. H., Yunker, P. L. & Hansen, C. L. (1990) *Biochim. Biophys. Acta* **1029,** 113–116.
12. Hurley, J. H., Faber, H. R., Worthylake, D., Meadow, N. D., Roseman, S., Pettigrew, D. W. & Remington, S. J. (1993) *Science* **259,** 673–677.
13. Peterkofsky, A., Reizer, A., Reizer, J., Gollop, N., Zhu, P. & Amin, N. (1993) *Prog. Nucleic Acid Res. Mol. Biol.* **44,** 31–65.
14. Seok, Y.-J., Sun, J., Kaback, H. R. & Peterkofsky, A. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 13515–13519.
15. van Iwaarden, P. R., Pastore, J. C., Konings, W. N. & Kaback, H. R. (1991) *Biochemistry* **30,** 9595–9600.
16. Hoischen, C., Levin, J., Pitaknarongphorn, S., Reizer, J. & Saier, M. H., Jr. (1996) *J. Bacteriol.* **178,** 6082–6086.
17. Reddy, P., Fredd-Kuldell, N., Liberman, E. & Peterkofsky, A. (1991) *Protein Expression Purif.* **2,** 179–187.
18. Teather, R. M., Bramhall, J., Riede, I., Wright, J. K., Furst, M., Aichele, G., Wilhelm, U. & Overath, P. (1980) *Eur. J. Biochem.* **108,** 223–231.
19. Venkatesan, P. & Kaback, H. R. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 9802–9807.
20. Frillingos, S., Sahin-Tóth, M., Persson, B. & Kaback, H. R. (1994) *Biochemistry* **33,** 8074–8081.
21. Frillingos, S. & Kaback, H. R. (1996) *Biochemistry* **35,** 5333–5338.
22. Frillingos, S., Gonzalez, A. & Kaback, H. R. (1997) *Biochemistry* **36,** 14284–14290.
23. Feese, M. D., Comolli, L., Meadow, N. D., Roseman, S. & Remington, S. J. (1997) *Biochemistry* **36,** 16087–16096.
24. Kuhnau, S., Reyes, M., Sievertsen, A., Shuman, H. A. & Boos, W. (1991) *J. Bacteriol.* **173,** 2180–2186.
25. Kuroda, M., de Waard. S., Mizushima, K., Tsuda, M., Postma, P. & Tsuchiya, T. (1992) *J. Biol. Chem.* **267,** 18336–18341.
26. Herzberg, O. (1992) *J. Biol. Chem.* **267,** 24819–24823.
27. Chen, Y., Reizer, J., Saier, M. H., Jr., Fairbrother, W. J. & Wright, P. E. (1993) *Biochemistry* **32,** 32–37.
28. Dahl, M. K., Francoz, E., Saurin, W., Boos, W., Manson, M. D. & Hofnung, M. (1989) *Mol. Gen. Genet.* **218,** 199–207.