THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 278, No. 27, Issue of July 4, pp. 25191–25206, 2003
Printed in U.S.A.

Solution Structure of the Phosphoryl Transfer Complex between the Signal-transducing Protein IIA^{Glucose} and the Cytoplasmic Domain of the Glucose Transporter IICB^{Glucose} of the *Escherichia coli* Glucose Phosphotransferase System*

Received for publication, March 17, 2003
Published, JBC Papers in Press, April 25, 2003, DOI 10.1074/jbc.M302677200

Mengli Cai‡§, David C. Williams, Jr.‡§¶, Guangshun Wang‡‖, Byeong Ryong Lee, Alan Peterkofsky**, and G. Marius Clore‡ ‡‡**

*From the ‡Laboratory of Chemical Physics, NIDDK, and the **Laboratory of Cell Biology, NHLBI, National Institutes of Health, Bethesda, Maryland 20892*

The solution structure of the final phosphoryl transfer complex in the glucose-specific arm of the *Escherichia coli* phosphotransferase system, between enzyme IIA^{Glucose} (IIA^{Glc}) and the cytoplasmic B domain (IIB^{Glc}) of the glucose transporter IICB^{Glc}, has been solved by NMR. The interface (~1200-Å^2 buried surface) is formed by the interaction of a concave depression on IIA^{Glc} with a convex protrusion on IIB^{Glc}. The phosphoryl donor and acceptor residues, His-90 of IIA^{Glc} and *Cys-35* of IIB^{Glc} (residues of IIB^{Glc} are denoted in italics) are in close proximity and buried at the center of the interface. *Cys-35* is primed for nucleophilic attack on the phosphorus atom by stabilization of the thiolate anion (pK_a ~6.5) through intramolecular hydrogen bonding interactions with several adjacent backbone amide groups. Hydrophobic intermolecular contacts are supplemented by peripheral electrostatic interactions involving an alternating distribution of positively and negatively charged residues on the interaction surfaces of both proteins. Salt bridges between the Asp-38/Asp-94 pair of IIA^{Glc} and the *Arg-38/Arg-40* pair of IIB^{Glc} neutralize the accumulation of negative charge in the vicinity of both the Sγ atom of *Cys-35* and the phosphoryl group in the complex. A pentacoordinate phosphoryl transition state is readily accommodated without any change in backbone conformation, and the structure of the complex accounts for the preferred directionality of phosphoryl transfer between IIA^{Glc} and IIB^{Glc}. The structures of IIA^{Glc}·IIB^{Glc} and the two upstream complexes of the glucose phosphotransferase system (EI·HPr and IIA^{Glc}·HPr) reveal a cascade in which highly overlapping binding sites on HPr and IIA^{Glc} recognize structurally diverse proteins.

In bacteria, carbohydrate transport across the membrane, mediated by the phosphoenolpyruvate:sugar phosphotransferase system (PTS),[1] involves the tight coupling of translocation and phosphorylation. The PTS is a classical example of a signal transduction pathway involving phosphoryl transfer (1), whereby a phosphoryl group originating on phosphoenolpyruvate is transferred to the translocated carbohydrate via a series of three bimolecular protein-protein complexes. The first two steps of the PTS are common to all sugars: enzyme I (EI) is autophosphorylated by phosphoenolpyruvate and subsequently donates the phosphoryl group to the histidine phosphocarrier protein HPr. The proteins downstream from HPr are sugar-specific, comprising four distinct families of IIA permeases (2–4). In the case of the glucose branch of the PTS, the phosphoryl group is transferred from HPr to IIA^{Glc} and thence from IIA^{Glc} to the C-terminal cytoplasmic domain (IIB^{Glc}) of the glucose transporter IICB^{Glc}. In addition to their function within the PTS cascade, the PTS proteins also serve to regulate other pathways (2). Thus, dephosphorylated enzyme I inhibits bacterial chemotaxis (5); dephosphorylated HPr functions as a positive regulatory subunit of glycogen phosphorylase (6); dephosphorylated IIA^{Glc} is a negative regulator of glycerol kinase (7), as well as a variety of non-PTS permeases (2), whereas phosphorylated IIA^{Glc} is a positive regulator of adenylyl cyclase (8); finally, the dephosphorylated form of IICB^{Glc} sequesters the global repressor Mlc, thereby initiating PTS gene transcription in response to the uptake of glucose from the extracellular environment (9–12). This multiplicity of interactions in which individual proteins specifically recognize a wide variety of structurally diverse targets serves as a paradigm for understanding protein-protein interactions and the factors determining their specificity.

The glucose-specific transporter, IICB^{Glc}, comprises an N-terminal transmembrane domain thought to consist of eight membrane-spanning helices (residues 17–323) connected to a C-terminal cytoplasmic domain (IIB^{Glc}, residues 401–476) via a long flexible linker (13–15). Recently, we have solved the solution NMR structures of the initial protein-protein complex of the *Escherichia coli* PTS between the N-terminal domain of enzyme I (EIN) and HPr (16) and the subsequent complexes of

* This work was supported in part by the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (to G. M. C.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and experimental NMR restraints (code 1O2F) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers Uiversity, New Brunswick, NJ (http://www.rcsb.org/).

§ These two authors contributed equally to this work.

¶ Recipient of a Pharmacology Research Associate Training postdoctoral fellowship from NIGMS, National Institutes of Health.

‖ Present address: Eppley Institute, University of Nebraska Medical Center, Omaha, NE 68198-6805.

‡‡ To whom correspondence should be addressed: Laboratory of Chemical Physics, Bldg. 5, Rm. B1-30I, NIDDK, National Institutes of Health, Bethesda, MD 20892-0510. Tel.: 301-496-0782; Fax: 301-496-0825; E-mail: mariusc@intra.niddk.nih.gov.

[1] The abbreviations used are: PTS, phosphoenolpyruvate:sugar phosphotransferase system; EI, enzyme I; EIN, N-terminal domain of enzyme I; HPr, histidine-containing phosphocarrier protein; IIA^{Glc}, glucose-specific enzyme IIA; IICB^{Glc}, the glucose-specific transporter; IIB^{Glc}, cytoplasmic B domain of IICB^{Glc}; IIB^{Suc}, the cytoplasmic B domain of the sucrose-specific transporter; NOE, nuclear Overhauser effect; r.m.s., root mean square; PTP, protein-tyrosine phosphatase.

HPr with both IIA^{Glc} (17) and IIA^{Mtl} (18). In this paper, we present the structure of the final cytoplasmic protein-protein complex of the glucose arm of the *E. coli* PTS, that between IIA^{Glc} and the C-terminal cytoplasmic domain (IIB^{Glc}) of the IICB^{Glc} transporter.

EXPERIMENTAL PROCEDURES

Expression and Purification of Proteins—The following constructs of *E. coli* IIA^{Glc} and IIB^{Glc} were employed in the present study: IIA^{Glc} (residues 1–168 (19) and 11–168 with an additional N-terminal glycine) and IIB^{Glc} (residues 1–90 and 12–90 with an additional N-terminal glycine and a P17A mutation). IIB^{Glc}-(1–90) is identical to the construct described by Buhr *et al.* (14), with the exception of the absence of a C-terminal His tag, and comprises residues 1–4 and 391–476 of full-length IICB^{Glc}. The numbering scheme for IIB^{Glc} is that employed in previous structural studies (15, 20), with the active site cysteine at position 35 corresponding to residue 421 of full-length IICB^{Glc}.

IIA^{Glc}-(1–168) and IIB^{Glc}-(1–90) were expressed, purified and uniformly labeled with ^{15}N (>95%) and/or ^{13}C (>95%) as described previously (19, 21).

IIA^{Glc}-(11–168) and IIB^{Glc}-(12–90/P17A) were subcloned into the PET32a expression vector as fusions with thioredoxin and a 6-His tag. The protein-coding sequences were amplified with PCR primers that included an N-terminal thrombin cleavage site designed to add only a single N-terminal glycine to the native sequence after thrombin digestion. The original thrombin cleavage site encoded by the PET32a vector was removed by site-directed mutagenesis (Arg → Gln). All constructs were verified by DNA sequencing. *E. coli* strain BL21-DE3 (Novagen) was transformed with the appropriate expression vector, grown in Luria Bertani or minimal media (with $^{15}NH_4Cl$ and/or $^{13}C_6$-glucose as sole nitrogen and carbon sources, respectively), and induced with 1 mM isopropyl-β-D-thiogalactopyranoside at an A_{600} ~1 (22). The bacteria were harvested by centrifugation after 2–4 h of induction. The cell paste was resuspended in 20 mM Tris (pH 7.5), 150 mM NaCl (100 ml/liter of cell culture), and the bacteria were lysed by several passages through a microfluidizer (Microfluidics Corp., Newton, MA). The lysate from a 2-liter culture was clarified by centrifugation and passed over a nickel-Sepharose column (~10 ml; Amersham Biosciences), and the fusion protein was subsequently eluted with a 100-ml gradient of imidazole (0.25 mM to 0.5 M). Both fusion proteins were then dialyzed against 20 mM Tris (pH 7.5) and 200 mM NaCl and subsequently digested with thrombin (10 NIH units/mg of protein) for at least 18 h at room temperature. Thrombin was removed by passage over a benzamidine-Sepharose column (1 ml; Amersham Biosciences) followed by the addition of 1 mM phenylmethylsulfonyl fluoride. The cleaved 6-His-thioredoxin was removed by passing the digested proteins back over a nickel-Sepharose column.

IIA^{Glc}-(11–168) was further purified by ion exchange chromatography using MonoQ resin (Amersham Biosciences) and eluted with a 0–1 M NaCl gradient. Pooled protein was then passed over a Sephadex-75 gel filtration column (Amersham Biosciences) equilibrated with 10 mM potassium phosphate (pH 7.2).

Similarly, IIB^{Glc}-(12–90/P17A) was passed over a Sephadex-75 gel filtration column equilibrated with 10 mM potassium phosphate (pH 7.2) and 1 mM dithiothreitol. The pH of the pooled protein fractions was lowered rapidly (pH <3) with trifluoroacetic acid, and the protein was incubated at room temperature for several hours to allow for spontaneous dephosphorylation of the active site cysteine. IIB^{Glc}-(12–90/P17A) was further purified by reverse phase chromatography using a C8 column and eluted with a 25–75% acetonitrile gradient in 0.1% trifluoroacetic acid. The protein was then dialyzed against 10 mM potassium phosphate (pH 7.2) and 5 mM dithiothreitol.

All purified proteins were >98% pure as judged by SDS-PAGE and ^{1}H-^{15}N correlation spectra. The composition of all purified proteins was confirmed by mass spectrometry.

NMR studies were carried out on free IIB^{Glc}-(1–90) and IIB^{Glc}-(12–90/P17A) and on 1:1 complexes of IIA^{Glc}-(1–168)·IIB^{Glc}-(1–90) and IIA^{Glc}-(11–168)·IIB^{Glc}-(12–90/P17A). NMR samples contained ~1 mM protein or protein-protein complex in 10 mM phosphate buffer, pH 7.2, and 20 mM deuterated dithiothreitol. The following complexes were employed (only the presence of ^{15}N and ^{13}C isotopes is indicated; if no C or N isotope is mentioned, then the sample contained ^{12}C or ^{14}N at natural isotopic abundance): IIA^{Glc}(^{15}N)·IIB^{Glc}, IIA^{Glc}·IIB^{Glc}(^{15}N), IIA^{Glc}(^{15}N/^{13}C)·IIB^{Glc}, and IIA^{Glc}·IIB^{Glc}(^{15}N/^{13}C). It should be noted that residues 1–18 of IIA^{Glc} and 1–14 of IIB^{Glc} are disordered in solution in both the free proteins and the complex. The chemical shifts of residues 19–168 of IIA^{Glc} are identical for the shorter and longer constructs.

Although all initial experiments were carried out with IIB^{Glc}-(1–90), we noticed that a number of cross-peaks in the ^{1}H-^{15}N correlation spectra of both free and complexed IIB^{Glc} were doubled and that the relative intensity of the two components was temperature-dependent. Examination of the calculated structure of IIB^{Glc} (see below) suggested that this was, in all likelihood, due to a cis-trans isomerization of Pro-17, located at the N terminus of helix 1, coupled with the interaction of Pro-17 with the aromatic ring of Tyr-87, located in helix 3. This was confirmed by examination of the Pro-17 → Ala mutation, which completely removed all cross-peak heterogeneity. The only significant backbone $^{1}H_N$/^{15}N chemical shift differences (~0.1 ppm for $^{1}H_N$ or ~1 ppm for ^{15}N) between IIB^{Glc}-(12–90/P17A) and IIB^{Glc}-(1–90) involve residues 16 and 18–20 in the immediate vicinity of the P17A mutation; the cross-peaks for residues 15 and 21–90 in IIB^{Glc}-(12–90/P17A) correspond to the major (trans-Pro-17) form of IIB^{Glc}-(1–90). Thus, the P17A mutation results in no structural change other than the removal of a cis-trans isomerization at the N terminus of helix I. The binding characteristics of IIB^{Glc}-(1–90) and IIB^{Glc}-(12–90/P17A) to IIA^{Glc} are also identical, as judged by the same chemical shift perturbation and the same pattern of intermolecular NOEs. The quality of the spectra, both free and complexed, obtained with IIB^{Glc}-(12–90/P17A), however, are significantly superior, permitting us to observe many more intermolecular NOEs in the IIA^{Glc}-(11–168)·IIB^{Glc}-(12–90/P17A) complex than in the IIA^{Glc}-(1–168)·IIB^{Glc}-(1–90) complex. Thus, the intermolecular NOEs observed in the IIA^{Glc}-(11–168)·IIB^{Glc}-(12–90/P17A) complex comprise all of those observed in the IIA^{Glc}-(1–168)·IIB^{Glc}-(1–90) complex plus additional intermolecular NOEs.

NMR Spectroscopy—All spectra were recorded at 35 °C on Bruker DMX500, DMX600, DRX600, DMX750, and DRX800 spectrometers equipped with either *x*,*y*,*z*-shielded gradient triple resonance probes or a *z*-shielded gradient triple resonance cryoprobe. Spectra were processed with the NMRPipe package (23) and analyzed using the programs PIPP, CAPP, and STAPP (24). ^{1}H, ^{15}N, and ^{13}C sequential assignments were obtained using three-dimensional double and triple resonance through-bond correlation (25–27). ^{3}J N-Cγ, C′-Cγ, and Cα-Cδ couplings were measured using quantitative *J* correlation spectroscopy (28). Interproton distance restraints were derived from multidimensional NOE spectra with mixing times ranging from 75–120 ms. Three-dimensional experiments used for sequential assignments included HNCO, HNCACB, CBCA(CO)NH, HBHA(CBCACO)NH, C(CCO)NH, H(CCO)NH, HCCH-COSY, and HCCH-TOCSY experiments. NOE experiments included three-dimensional ^{15}N-separated NOE, ^{13}C-separated NOE, ^{13}C-separated/^{13}C-separated NOE, and ^{13}C-separated/^{12}C-filtered NOE spectra, and four-dimensional ^{13}C/^{13}C-separated NOE spectra.

Long range ^{1}H-^{15}N correlation spectra to correlate the Nδ1 and Nε2 ^{15}N shifts with the Hδ2 and Hε1 ^{1}H shifts of the imidazole ring (29) were used to confirm the tautomeric states of the histidine residues in the complex.

Residual Dipolar Couplings—Residual dipolar couplings were measured by taking the difference in the corresponding *J* splittings in oriented and isotropic media. $^{1}D_{NH}$ dipolar couplings were obtained using either a two-dimensional in-phase/anti-phase ^{15}N-^{1}H HSQC experiment (30) or a two-dimensional sensitivity-enhanced E.COSY-type ^{15}N-^{1}H HSQC experiment, which simultaneously provides $^{1}D_{NC'}$ and $^{1}D_{HNC'}$ couplings (31).

$^{1}D_{NH}$, $^{1}D_{NC'}$, and $^{1}D_{HNC'}$ couplings were measured for free IIB^{Glc}-(12–90/P17A) in a liquid crystalline medium of filamentous phage (32, 33), and the magnitude of the alignment tensor was determined from the distribution of normalized dipolar couplings (34). At low salt, IIB^{Glc} interacts with phage pf1, resulting in significant line broadening. This interaction was abolished by increasing the ionic strength to 0.5 M NaCl, permitting a high quality set of backbone dipolar couplings to be obtained. Unfortunately, we were unable to find a suitable alignment medium that did not interact with one or the other protein in the IIA^{Glc}·IIB^{Glc} complex. Thus, dihexanoylphosphatidylcholine/dimyristoylphosphatidylcholine bicelles (35) interacted with IIA^{Glc}; and phage pf1 (33), polyethylene glycol/hexanol (36), and polyacrylamide gel (37) interacted with IIB^{Glc}. This is manifested by selective line broadening of only one of the two components, even in the presence of a significant excess (up to 4-fold) of the unlabeled over ^{15}N/^{13}C-labeled component whose dipolar couplings were being measured. Thus, the dipolar couplings measured on the line-broadened component within the complex contain a small contribution from the free state interacting with the alignment medium. As a consequence, the magnitude of the alignment tensors, for all alignment media employed, were found to be sufficiently noncoincident for the two proteins within the complex to preclude their utility in accurately determining relative protein-protein orientation.

Further, under high salt conditions, where IIB^{Glc} and pf1 do not interact, complex formation is abolished.

Nevertheless, the $^1D_{NH}$ dipolar couplings measured for the IIA^{Glc}·IIB^{Glc} complex can be used as a cross-validation tool to demonstrate the absence of any significant backbone conformational changes upon complex formation. This involves singular value decomposition best fitting of observed and calculated dipolar couplings using an appropriate reference structure, namely the 2.1-Å resolution crystal structure coordinates of free IIA^{Glc} (38) and the NMR coordinates of IIB^{Glc} determined using dipolar couplings measured in pf1 for free IIB^{Glc}, and interproton distance and torsion angle restraints derived from combined measurements on both free and complexed IIB^{Glc} (see below). The values of the dipolar coupling R-factors, R_{dip} (39), obtained in this manner for complexed IIA^{Glc}, are 15.7, 16.1, and 20.1% in polyethylene glycol/hexanol, polyacrylamide gel, and pf1 media, respectively. (The slightly higher value in pf1 is due to the slightly lower quality of the pf1 data). Likewise, the R_{dip} values obtained for complexed IIB^{Glc} are 17.6 and 18.5% in polyacrylamide gel and pf1, respectively. In addition, the maximum combined $^1H_N/^{15}N$ shift perturbation upon complexation, $\Delta_{H/N}$ (in Hz) given by $[(\Delta\delta^1H_N)^2 + (\Delta\delta^{15}N)^2]^{1/2}$, is only 170 Hz for IIA^{Glc} and ~130 Hz for IIB^{Glc} at 600 MHz, with only 5 residues of IIA^{Glc} having a $\Delta_{H/N}$ between 100 and 170 Hz and 4 between 60 and 100 Hz and 2 residues of IIB^{Glc} having a $\Delta_{H/N}$ between 100 and 140 Hz and 9 between 60 and 100 Hz. From the combined dipolar coupling and chemical shift data, one can safely conclude that any possible structural differences in the backbone between free and complexed IIB^{Glc} are too small to be determined by NMR.

Structure Calculations—NOE-derived interproton distance restraints were classified into four ranges: 1.8–2.8, 1.8–3.5, 1.8–5.0, and 1.8–6.0 Å (40). An additional 0.5 Å was added to the upper bound for NOEs involving methyl groups, and distances involving nonstereospecifically assigned protons were represented by a $(\Sigma r^{-6})^{-1/6}$ sum. χ_1 and χ_2 torsion angle restraints were derived from analysis of heteronuclear ^3J couplings and NOE/ROE experiments (26). Structures were calculated using the Xplor-NIH NMR molecular structure determination package (41).

An ensemble of 60 simulated annealing structures for IIB^{Glc} was first calculated using torsion angle dynamics (42) based on NMR data recorded on both free and complexed IIB^{Glc}: specifically, $^1D_{NH}$, $^1D_{NC'}$, and $^1D_{HNC'}$ dipolar couplings measured on free IIB^{Glc}; NOE-derived interproton distance restraints obtained from spectra recorded on both free and complexed IIB^{Glc} (note that the pattern of intramolecular NOEs is identical for free and complexed IIB^{Glc}, so that higher quality NOE spectra from the free protein were used to confirm those observed in the complexed protein); backbone ϕ,ψ torsion angle restraints derived from backbone (N, C′, Cα, Cβ, Hα) chemical shifts using a data base search procedure with the program TALOS (43) (identical for free and complexed IIB^{Glc}, since the chemical shift perturbation upon complexation is minimal); and side chain torsion angle restraints (no difference found between free and complexed IIB^{Glc} with the possible exception of *Arg-40*, which appears to be rotamer-averaged in the free state but has a g^- χ_1 rotamer in the complex). The target function comprises square well potential terms for the NOE-derived interproton distance restraints and torsion angle restraints (44); harmonic potentials for the $^{13}C\alpha/\beta$ chemical shift restraints (45), dipolar coupling restraints, (46) and covalent geometry; and a quartic van der Waals repulsion potential (44), a torsion angle data base potential of mean force (47), and a radius of gyration term (48) to represent the nonbonded contacts. The radius of gyration term is a weak overall packing potential, and the target value is given by $2.2N^{0.38}$, where N is the number of ordered residues (48).

The structure of the IIA^{Glc}·IIB^{Glc} complex was then determined by conjoined rigid body/torsion angle dynamics (18, 42) on the basis of intermolecular NOE data and NMR data (NOEs, torsion angle restraints, and dipolar couplings) relating to the backbone region of IIB^{Glc} encompassing the IIA^{Glc} binding site and to the interfacial side chains of IIA^{Glc} and IIB^{Glc}. The initial coordinates of IIA^{Glc} comprise the x-ray structure (with protons added) of *E. coli* IIA^{Glc} (RCSB accession code 2F3G, molecule 2, 2.13-Å resolution) (38). The initial coordinates for IIB^{Glc} comprise the complete ensemble of 60 simulated annealing structures calculated as described above for IIB^{Glc} in the absence of any intermolecular restraints (*i.e.* each structure calculation for the complex makes use of a different initial structure for IIB^{Glc}). The coordinates for the backbone and noninterfacial side chains of IIA^{Glc} are held fixed; the overall IIB^{Glc} molecule is allowed rotational and translational degrees of freedom, with the coordinates of the noninterfacial backbone and side chain atoms of IIB^{Glc} treated as a rigid body; the backbone region of IIB^{Glc} encompassing the binding site for IIA^{Glc} (residues *34–40*, *59–63*, and *72–80*) and all interfacial side chains of IIA^{Glc} and IIB^{Glc} are given their full torsional degrees of freedom. This procedure, rather than one in which IIB^{Glc} is given complete torsional degrees of freedom, is employed to enhance convergence and to ensure that structural distortions of IIB^{Glc} are not introduced as a result of any potential underestimation in the upper bounds of some NOE-derived intermolecular distance restraints.

Structures were visualized and analyzed with the program VMD-XPLOR (49). Figures were generated using VMD-XPLOR (49), RIBBONS (50), and GRASP (51). The ensemble distributions of side chains were depicted by isosurfaces of the reweighted atomic density maps (52) drawn at a value of 20% of the maximum with a uniform radius of 1 Å. Optimized Cα alignments and superpositions for different proteins were carried out with the program O (53). Sequence searches and alignments were carried out with the program BLAST (54).

RESULTS AND DISCUSSION

Structure Determination—The IIA^{Glc}·IIB^{Glc} complex is in fast exchange on the chemical shift time scale, and the lower limit for the dissociation rate constant (as judged by the maximum observed ^1H_N chemical shift difference between the free and bound states) is ~800 s^{-1}.

The solution structure of the IIA^{Glc}·IIB^{Glc} complex was solved by multidimensional heteronuclear NMR spectroscopy. A combination of isotopically (^{15}N and/or ^{13}C) labeled proteins was used to simplify the spectra for assignment purposes and to specifically observe intermolecular NOE contacts between IIA^{Glc} and IIB^{Glc} (27). An example of the quality of the data is shown in Fig. 1, which displays selected strips from a three-dimensional ^{13}C-separated/^{12}C-filtered NOE experiment recorded on a IIA^{Glc}($^{12}C/^{14}N$)·IIB^{Glc}($^{13}C/^{15}N$) sample in which NOEs are specifically observed from protons attached to ^{13}C on IIB^{Glc} to protons attached to ^{12}C on IIA^{Glc}.

The structure of the complex was solved using conjoined rigid body/torsion angle dynamics simulated annealing (18, 42). The $^1H_N/^{15}N$ chemical shift and dipolar coupling data indicate that no significant perturbation in the backbone of either IIA^{Glc} or IIB^{Glc} occurs upon complex formation (*i.e.* any structural differences in the backbone between the bound and free proteins are too small to be determined by NMR; see "Experimental Procedures"). In the case of IIA^{Glc}, a 2.1-Å resolution crystal structure of IIA^{Glc} is available (Protein Data Bank accession code 2F3G) (38), and hence the coordinates for the backbone and noninterfacial side chains of IIA^{Glc} could be readily employed without any need to redetermine their structure. Although an NMR structure for IIB^{Glc} has been published (15), the agreement between measured and calculated dipolar couplings measured on free IIB^{Glc} is very poor, with a backbone ^1H-^{15}N dipolar coupling R-factor (39), R_{dip}, of 77 ± 3% for the ensemble of 11 deposited structures (Protein Data Bank accession code 1IBA). (Note that R_{dip} scales between 0 and 100%, with 0% for a perfect fit and 100% representing a random orientation of N-H vectors (39)). Consequently, the solution structure of IIB^{Glc} was redetermined.

The solution structure of the IIA^{Glc}·IIB^{Glc} complex was determined on the basis of 1769 experimental NMR restraints, including 113 intermolecular NOE-derived interproton distance restraints. A summary of the structural statistics is provided in Table I, a stereoview of a best fit superposition of the backbone for the final ensemble of 60 simulated annealing structures is shown in Fig. 2*A*, and a reweighted atomic density map (52), derived from the complete ensemble, is shown in Fig. 2*B* to illustrate the precision with which the interfacial side chains are determined.

Overall Description of the IIB^{Glc} Structure—A ribbon diagram of IIB^{Glc} is shown in Fig. 3*A*. (To distinguish between residues of IIA^{Glc} and IIB^{Glc}, residues of IIB^{Glc} are denoted by *italic type* throughout). IIB^{Glc} is an α/β protein with a trigonal pyramid-like appearance. The base is formed by a four-stranded antiparallel β-sheet (β_1, residues *29–35*; β_2, *38–43*;

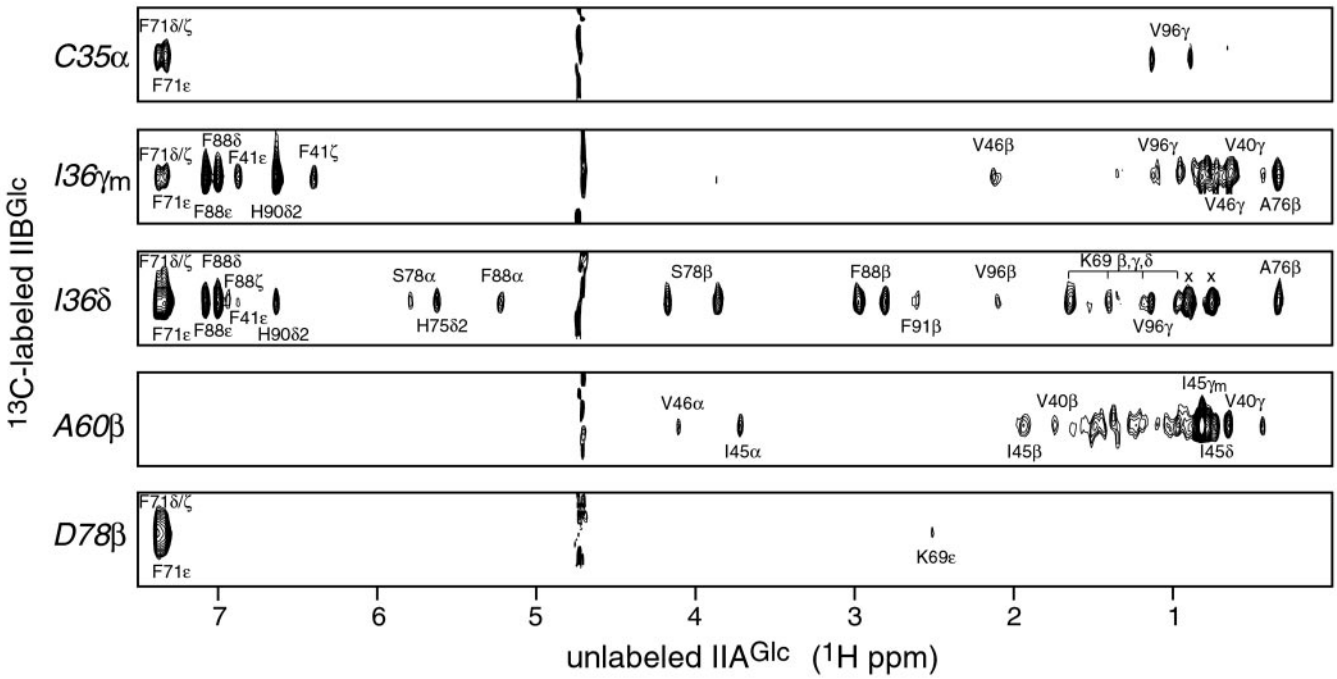

FIG. 1. **Intermolecular NOEs in the IIA^{Glc}·IIB^{Glc} complex.** Shown are strips from a three-dimensional ^{13}C-separated/^{12}C-filtered NOE spectrum recorded at 800 MHz on a 1:1 IIA^{Glc}(^{12}C/^{14}N)·IIB^{Glc}(^{13}C/^{15}N) complex illustrating NOEs from protons attached to ^{13}C on IIB^{Glc}-(12–90/P17A) to protons attached to ^{12}C on IIA^{Glc}-(11–168).

β_3, *61–65*; and β_4, *68–72*) arranged in a 1, 2x, −1 topology on top of which sit three α-helices (α_1, residues *15–23*; α_2, *51–57*; α_3, *76–90*). Helix-helix contacts occur between helices 1 and 2 and helices 1 and 3 with corresponding interhelical angles of −160° and −81°, respectively. There are two classical antiparallel β-bulges, the first comprising residues *30* and *31* of β_1 and residue *42* of β_2 and the second comprising residues *60* and *61* of β_3 and *72* of β_4. The β-turn between strands β_1 and β_2 (residues *35–38*) includes the active site cysteine at position *35* and is a distorted type I turn (with ϕ/ψ angles of −95°/−40° and −130°/10° for *Ile-36* and *Thr-37*, respectively). The β-turn connecting strands 3 and 4 is a type II′ turn. There are two 3^{10} helices (residues *25–28* and *45–48*), which interact with each other through a network of hydrogen bonds involving the side chain of Asn-28 (Fig. 3*B*). In this regard, the Hδ21 resonance of *Asn-28* is downfield shifted by ~1.3 ppm relative to that of the other Asn residues of IIB^{Glc}. Residues *1–13* are disordered in solution.

Whereas the topology of the present IIB^{Glc} structure is the same as that determined by Eberstadt *et al.* (15), the overall backbone r.m.s. difference between the two structures is ~5 Å. A large contribution to this difference arises from ~45° differences in the orientations of helices α_1 and α_3, as well as some large displacements in loops and turns that were poorly defined in the older structure. The Cα atoms of only 40 of a total of 76 residues can be superimposed with an atomic r.m.s. difference of 2.1 Å, namely residues *26–33*, *38–46*, *50–57*, and *59–73*, comprising the first 3^{10} helix, helix α2, strands β_1–β_4, and the β_3/β_4 turn. It should be noted that the difference between the present and previous structures of IIB^{Glc} is not attributable to any change in conformation upon complexation with IIA^{Glc}: the $^1D_{NH}$ dipolar couplings measured on free IIB^{Glc} are in very poor agreement with the structure of Eberstadt *et al.* (15), as already discussed above; the backbone chemical NH and ^{15}N chemical shift changes upon complexation are small (see "Experimental Procedures" and Ref. 21); and the pattern of intramolecular NOEs within IIB^{Glc} is unchanged upon complexation. Rather, it is a reflection of improved NMR technology to resolve spectral overlap, improved spectral quality due to the use of the P17A mutant to remove heterogeneity arising from cis-trans proline isomerization (see "Experimental Procedures"), and the use of residual dipolar couplings to provide long range orientational information. In this regard, it is worth noting that the first and second generation WHATIF packing *Z*-scores (55, 56), which provides a good independent measure of structure quality, are −0.4 and +0.98, respectively, for the present structure, compared with −3.1 and −3.0, respectively, for the Eberstadt *et al.* (15) structure. (For reference, a first generation packing score of larger than −0.5 is considered perfect, whereas a value below −3.0 represents either a wrong structure or a very bad model; similarly, a second generation packing score greater than 0 is considered to be a good structure, whereas a value less than −3 represents a doubtful structure (56).)

The structure of IIB^{Glc} is very similar to that of the C-terminal oligomerization domain of the arginine repressor from *Bacillus stearothermophilus* (57), although there is no significant sequence identity between the two domains. A structural best fit superposition with respect to the folded IIB^{Glc} core (residues *15–90*) yields a Cα backbone r.m.s. difference of 2.0 Å for 59 of a total of 76 residues, with a sequence identity of ~13% for the structurally aligned residues. The structure of IIB^{Glc} is also very similar to the oligomerization domain of the arginine repressor of *E. coli* (58), except that the fragment crystallized was truncated after helix α_1 (Cα atomic r.m.s. difference of 2.1 Å for 49 residues, with a sequence identity of ~10%).

The Environment of the Active Site Cys-35 of IIB^{Glc}—The pK_a of *Cys-35* in free IIB^{Glc} was determined by monitoring the ^{1}H-^{15}N cross-peak of *Thr-37* as a function of pH. The pH dependence of the ^1H_N and ^{15}N chemical shifts of *Thr-37* follows simple Henderson-Hasselbalch behavior with a total excursion of ~1.5 and ~4 ppm, respectively. The ^{1}H-^{15}N cross-peaks of two other residues, *Ile-36* and *Arg-38*, also undergo large pH-dependent changes but are more difficult to follow over the complete titration curve. The pK_a of *Cys-35* is ~6.5, several pH units lower than the usual range observed in both small pep-

TABLE I
Structural statistics

	⟨SA⟩[a]	$(\overline{SA})r$ [a]
Number of experimental restraints		
Intramolecular interproton distance restraints		
IIB^{Glc}	987	
IIA^{Glc} interfacial side chains	30	
Intermolecular interproton distance restraints	113	
Backbone hydrogen bond restraints for IIB^{Glc} (two per hydrogen bond)	72	
Torsion angle restraints		
IIB^{Glc}	221	
IIA^{Glc} interfacial side chains	34	
Residual dipolar couplings for IIB^{Glc}	174	
$^{13}C\alpha/\beta$ chemical shift restraints for IIB^{Glc}	138	
r.m.s. deviation from distance restraints (Å)[b]	0.016 ± 0.003	0.011
r.m.s. deviation from torsion angle restraints (degrees)[b]	0.46 ± 0.04	0.36
r.m.s. deviation from $^{13}C\alpha/\beta$ shifts for IIB^{Glc} (ppm)	1.02 ± 0.11	1.01
R-factors for residual dipolar couplings for IIB^{Glc} (%)[c]		
$^1D_{NH}$	5.3 ± 0.3	4.9
$^1D_{NC'}$	15.4 ± 0.5	16.0
$^2D_{HNC'}$	12.8 ± 0.5	13.4
Deviations from idealized covalent geometry for IIB^{Glc}		
Bonds (Å)	0.002 ± 0	0.003
Angles (degrees)	0.248 ± 0.011	0.410
Impropers (degrees)	0.462 ± 0.076	0.471
Percentage of residues for IIB^{Glc} in most favorable region of Ramachandran map[d]	91.7 ± 1.1	95.5
Coordinate precision (Å)		
IIB^{Glc} alone[e]		
Backbone	0.23	
All heavy atoms	0.71	
Overall complex[f]		
IIA^{Glc} + IIB^{Glc} backbone	0.31	
IIA^{Glc} + IIB^{Glc} interfacial side chains	0.67	

[a] The notation of the NMR structures is as follows: ⟨SA⟩, the final 60 simulated annealing structures; $(\overline{SA})r$, the restrained regularized mean structure derived from the ensemble of simulated annealing structures using the procedure described in Ref. 18.

[b] None of the structures exhibited interproton distance violations of >0.5 Å or torsion angle violations of >5°. The intramolecular interproton distance restraints for IIB^{Glc} comprise 189 intraresidue and 273 sequential $|i - j| = 1$, 218 medium range $i < |i - j| \leq 5$ and 307 long range $|i - j| > 5$ restraints; the intramolecular restraints within IIA^{Glc} involving interfacial side chains only comprise 4 intraresidue, 3 sequential, 7 medium range, and 16 long range restraints. The torsion angle restraints for IIB^{Glc} comprise 152 ϕ, ψ backbone torsion angle restraints and 69 side chain torsion angle restraints.

[c] The dipolar coupling R-factor (R_{dip}) is defined as the ratio of the r.m.s. deviation between observed and calculated values to the expected r.m.s. deviation if the vectors were randomly oriented. The latter is given by $(2D_a^2[4 + 3\eta^2]/5)^{1/2}$, where D_a is the magnitude of the axial component of the alignment tensor, and η is the rhombicity (39). The values of D_a^{NH} and η for free IIB^{Glc} dissolved in 15 mg/ml phage pfl and 500 mM NaCl, derived from the distribution of normalized dipolar couplings (34), are 10.5 Hz and 0.2, respectively.

[d] The percentage of residues in the most favorable region of the Ramachandran plot (78) for the X-ray structure of IIA^{Glc} (molecule 2 of 2F3G) (38) is 89%. The dihedral angle G-factors (78) for ϕ/ψ, χ_1/χ_2, χ_1, and χ_3/χ_4 for IIB^{Glc} are 0.23 ± 0.03, 0.63 ± 0.11, 0.25 ± 0.12, and 0.29 ± 0.22, respectively. The WHATIF (55) first and second generation overall packing Z scores (56) for IIB^{Glc} are −0.4 and +0.98, respectively.

[e] Defined as the average r.m.s. difference between the final 60 simulated annealing structures and the mean coordinates for IIB^{Glc} only (residues 14–90). Residues 1–13 of IIB^{Glc} are disordered in solution. The backbone atoms comprise the N, Cα, C′, and O atoms.

[f] Defined as the average r.m.s. difference between the final 60 simulated annealing structures and the mean coordinates for the whole complex (residues 19–168 of IIA^{Glc} + residues 14–90 of IIB^{Glc}). Note that since IIA^{Glc} is treated as a rigid body, this does not take into account the errors in the X-ray coordinates of IIA^{Glc}. Note that residues 1–18 of IIA^{Glc} are disordered in the present complex, as well as in the crystal (38) and solution (79) structures of free IIA^{Glc} and the solution structure of the HPr-IIA^{Glc} complex (17).

tides (pK_a ~9) and proteins (pK_a ~8–11) (59). Thus, at physiological pH, *Cys-35* is predominantly in the thiolate state, primed for nucleophilic attack on the phosphorus atom of phosphorylated IIA^{Glc}. The environment around *Cys-35* is depicted in Fig. 3*C*. The χ_1 side chain torsion angle of *Cys-35* is in the g^+ rotamer, and the negative charge on the thiolate is stabilized by four NH—Sγ hydrogen bonds involving the backbone amides of *Ile-36*, *Thr-37*, and *Arg-38* and the side chain NϵH amide of *Arg-38*. Further stabilization of the thiolate species may be afforded by close proximity to the positive pole of the N terminus of helix α_3 (Fig. 3*A*). The distance (4.6 Å) between the Oγ1 atom of *Thr-37* and the Sγ atom of *Cys-35* is too long to permit the formation of a direct OH—Sγ hydrogen bond, but a water-mediated interaction is possible. Indeed, it may be the case that the increased nucleophilicity of such a water molecule, afforded by hydrogen bonding to the hydroxyl group of *Thr-37* could be responsible for spontaneous hydrolysis of the Sγ—P bond. At neutral pH, dephosphorylation of IIB^{Glc} is slow, with a half-life of ~40 h, consistent with the stabilization of the thiolate anion (20)[2]; at pH values below 3.5, where *Cys-35* is entirely in the thiol, protonated state, dephosphorylation occurs very rapidly.[2]

There are features of the active site region of IIB^{Glc}, including the stabilization of the thiolate anion, that are reminiscent of the active site loop of protein-tyrosine phosphatases (PTPs) (60–62). A superposition of the two active sites is shown in Fig. 3*D*. The active site cysteine of PTP1B (62) is located at position 215 and has a pK_a of ~4.7 (63). The Cα atomic r.m.s. difference between residues 213 and 218 of PTP1B and *33* and *38* of IIB^{Glc} is ~0.8 Å. The side chain conformation of the active site cysteine and the backbone conformation of the active site β-turn are the same for the two proteins. The backbone amides of Ser-216, Ala-217, Gly-218, and Gly-220 of PTP1B (62) are located in exactly the same positions as the backbone amides of *Ile-36*, *Thr-37*, and *Arg-38* and the NϵH group of *Arg-38*, respectively, thereby permitting identical hydrogen bonding in-

[2] M. Cai, D. G. Williams, Jr., G. Wang, B. R. Lee, A. Peterkofsky, and G. M. Clore, unpublished data.

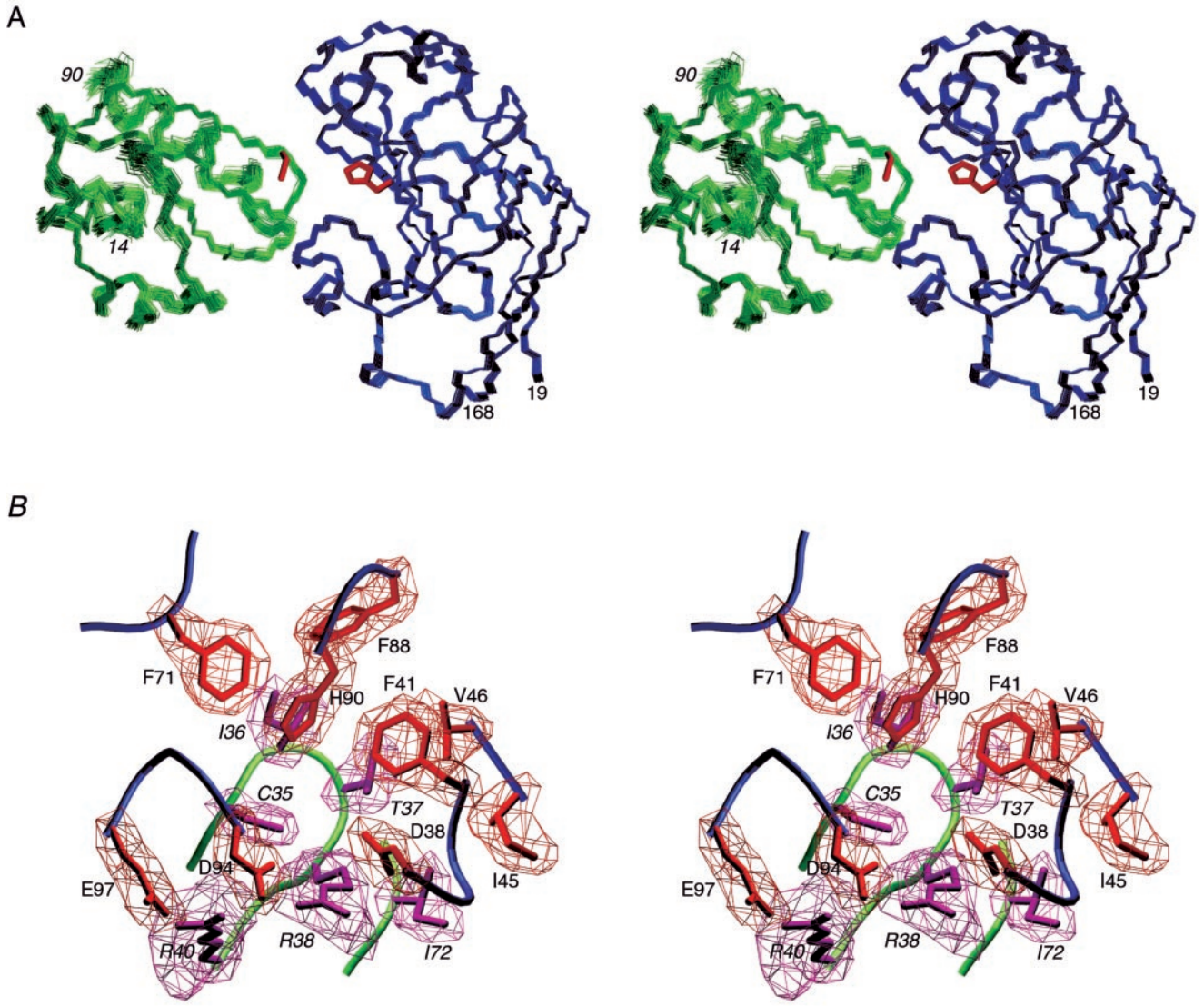

FIG. 2. **The structure of the *E. coli* IIA^{Glc}·IIB^{Glc} complex.** *A*, stereoview of a superposition of the final 60 simulated annealing structures best fitted to the backbone of residues 19–168 of IIA^{Glc} (in *blue*) and 14–90 of IIB^{Glc} (in *green*); the side chains of the active site histidine of IIA^{Glc} (His-90) and the active site cysteine (*Cys-35*) of IIB^{Glc} in the restrained regularized mean coordinates are shown in *red*. Residues 1–18 of IIA^{Glc} and 1–13 of IIB^{Glc} are disordered in solution. *B*, isosurface of the reweighted atomic density map drawn at a value of 20% maximum, calculated from the final 50 simulated annealing structures, for selected side chains of IIA^{Glc} (*red*) and IIB^{Glc} (*purple*); the backbones of IIA^{Glc} (*blue*) and IIB^{Glc} (*green*) are displayed as *tubes*. The side chain coordinates displayed within the atomic density map are those of the restrained regularized mean structure. Residues from IIB^{Glc} are denoted in *italic type*.

teractions with the Sγ atom of the active site cysteine. The two polypeptides diverge at Gly-218 of PTP1B and *Arg-38* of IIB^{Glc}, with the former going into a loop and the latter continuing in a β-strand. Consequently, there is no functional group in IIB^{Glc} that is similarly located to the invariant Arg-221 of PTPs whose guanidino group interacts with the phosphoryl group bonded to Cys-215. In the case of PTB1B, there are two additional hydrogen bonding interactions from the hydroxyl and backbone amide groups of Ser-222 to the Sγ atom of Cys-215. Whereas the NεH group of *Arg-40* of IIB^{Glc} is located in a similar region of space to the backbone amide and hydroxyl group of Ser-222, it is too far from the Sγ atom of *Cys-35* to permit a direct hydrogen bonding interaction, but an indirect interaction via a bridging water molecule is possible.

Overall Description of the IIA^{Glc}·IIB^{Glc} Structure—Ribbon diagrams of the overall IIA^{Glc}·IIB^{Glc} complex and of the protein-protein interface are shown in Figs. 4 and 5*A*, respectively. Surface representations of the binding sites are shown in Fig. 6*A*. *E. coli* IIA^{Glc} (shown in *blue* in Figs. 4 and 5) is a predominantly β-sheet sandwich protein comprising two six-stranded antiparallel β-sheets, one on each side of the molecule (39, 63). The surface on IIA^{Glc} that interacts with IIB^{Glc} is approximately circular (about 23 × 20 Å) and concave in shape, with the active site His-90 located at the bottom of the surface depression (Fig. 6*A*, *left*), and comprises three strands (β_5, β_6, and β_7) of a six-stranded antiparallel β-sheet (β_5, β_6, β_7, β_{10}, β_2, β_3) bounded by two short helical elements, helix α_1 and a 3^{10} helix (Fig. 5*A*). The interaction surface on IIB^{Glc} is complementary to that on IIA^{Glc}; it is also circular (about 19 × 17 Å) but convex, with the active site *Cys-35* located close to the apex of the protrusion (Fig. 6*A*, *right*). The IIB^{Glc} binding site comprises an almost contiguous stretch of residues from *Asp-33* to *Arg-40*, consisting of the C-terminal end of strand β_1, the β_1/β_2 turn (which contains the active site *Cys-35*), and strand β_2, bounded by strand β_3, the turn connecting strand β_4 to helix α_3, and the N-terminal end of helix α_3 (Figs. 4 and 5*A*). There are a total of 33 residues at the protein-protein interface, 19 from IIA^{Glc}, and 14 from IIB^{Glc}. The total accessible surface area

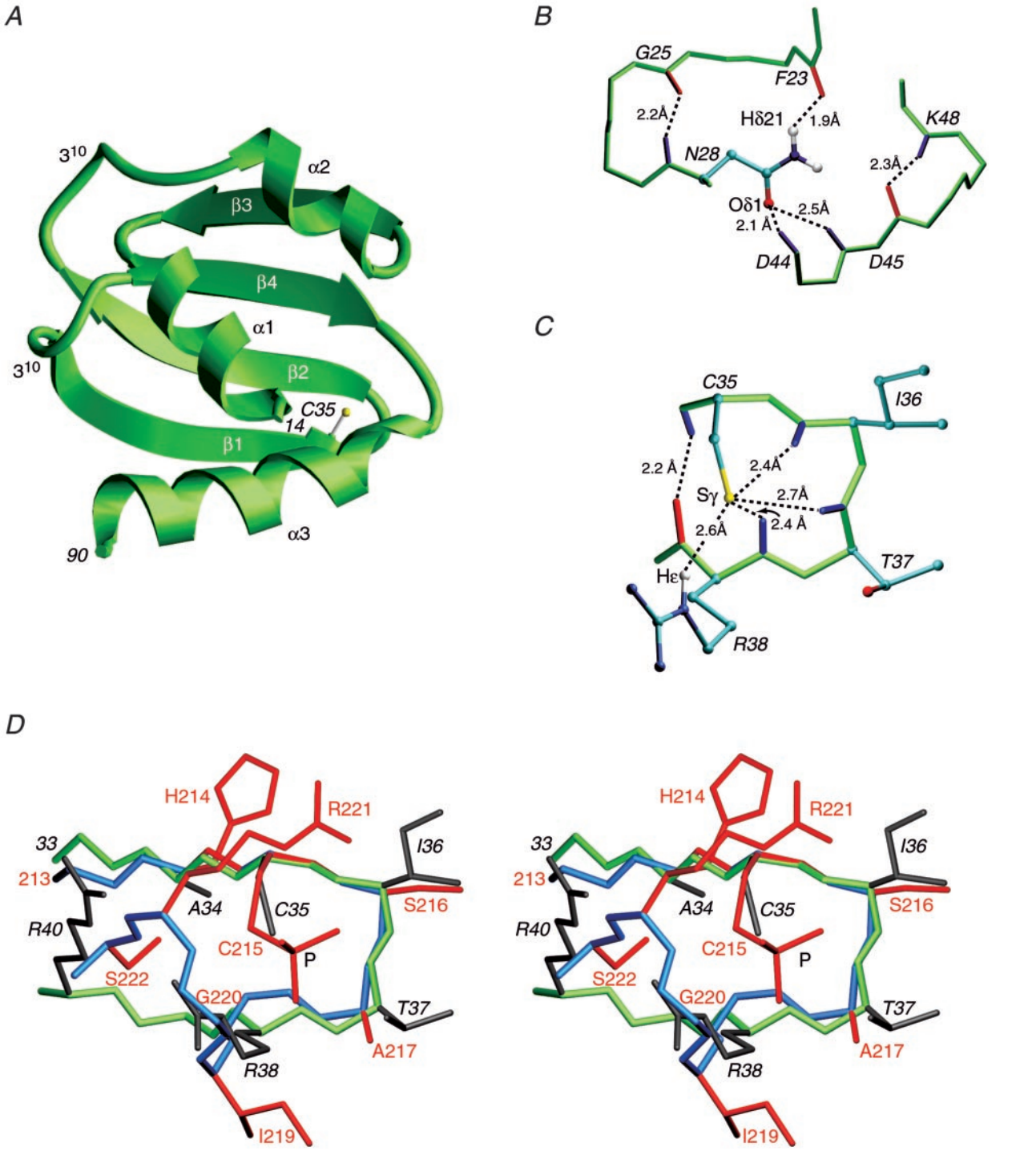

FIG. 3. **The structure of IIB^{Glc}.** *A*, ribbon diagram including the side chain of the active site *Cys-35*. *B*, detailed view showing hydrogen bonding interactions involving the side chain of *Asn-28* that bridge the two 3^{10} helices. *C*, detailed view of the hairpin turn containing the active site *Cys-35*. Note that the thiolate anion of *Cys-35* (pK_a ~6.5) is stabilized by hydrogen bonding interactions involving the backbone amide groups of *Ile-36*, *Thr-37*, and *Arg-38* as well as the NεH of the *Arg-38* side chain. *D*, stereoview showing a best fit superposition of the active sites of IIB^{Glc} and the protein-tyrosine phosphatase PTP1B in its phosphorylated state. The backbone and side chain atoms are shown in *green* and *gray*, respectively, for IIB^{Glc} and in *blue* and *red*, respectively, for PTP1B. The coordinates of phosphorylated PTP1B (RCSB accession code 1A5Y) are from Ref. 62. (Residue labels are in *black italic type* for IIB^{Glc} and in *red* for PTP1B).

buried upon complexation is 1208 Å^2, of which 570 Å^2 originates from IIA^{Glc} and 638 Å^2 from IIB^{Glc}.

A detailed view of the protein-protein interface and a summary of the intermolecular contacts are provided in Fig. 7, *A* and *B*, respectively. Both active site residues, His-90 of IIA^{Glc} and *Cys-35* of IIB^{Glc}, are surrounded by two partially overlapping semicircles. The first comprises hydrophobic residues surrounding His-90 and *Cys-35* from above (in the view shown in Fig. 6*A*), and the second consists of charged residues, both positive and negative, surrounding His-90 and *Cys-35* from below (in the view shown in Fig. 6*A*). Thus, in the complex, both His-90 and *Cys-35* are completely buried in the middle of the interface.

The hydrophobic cluster of residues on the interaction surfaces of both proteins (Val-39, Val-40, Phe-41, Ile-45, Val-46, Phe-71, Phe-88, and Val-96 of IIA^{Glc} and *Ala-34*, *Ile-36*, *Thr-37*,

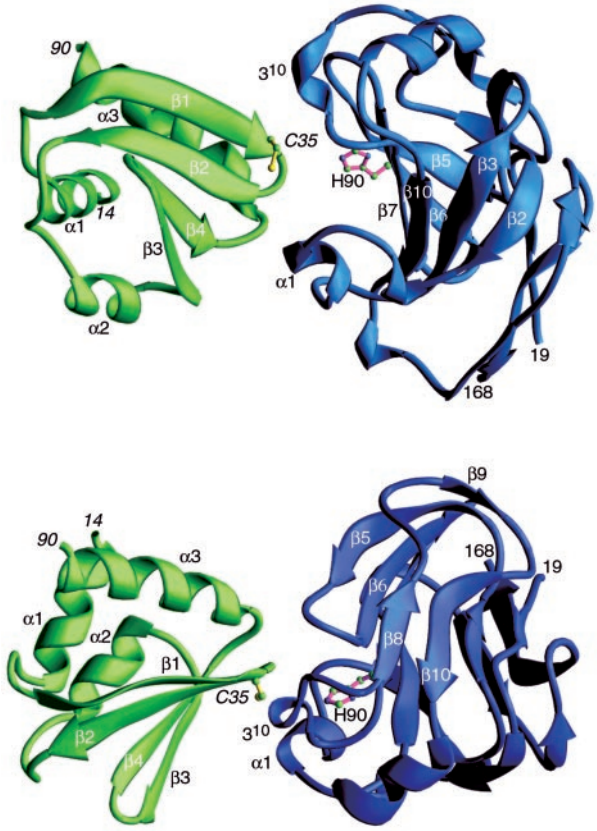

FIG. 4. **Two views of a ribbon diagram of the restrained regularized mean coordinates of the IIA^{Glc}·IIB^{Glc} complex.** IIA^{Glc} is shown in *blue*, and IIB^{Glc} is *green*. The bonds of the side chains of His-90 of IIA^{Glc} and *Cys-35* of IIB^{Glc} are displayed in *purple* and *yellow*, respectively. Residues from IIB^{Glc} are denoted in *italic type*.

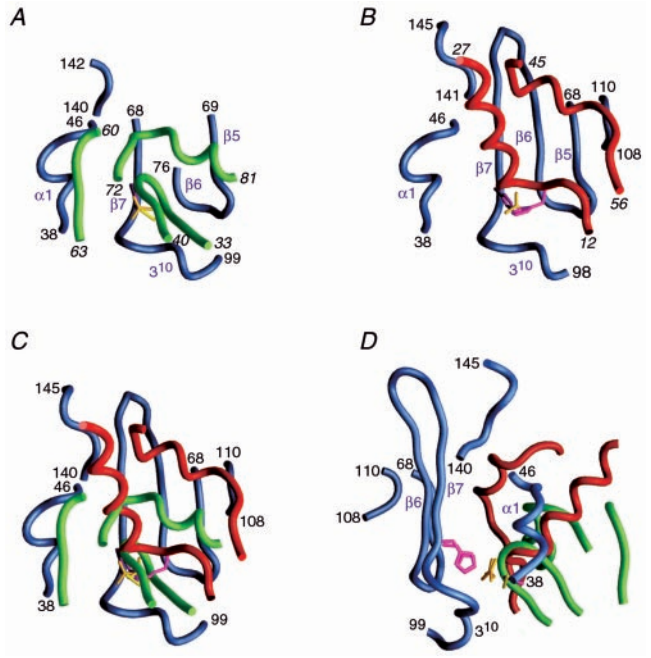

FIG. 5. **Comparison of the backbone of IIB^{Glc} and HPr involved in binding to IIA^{Glc}.** The backbones are shown as *tubes* with IIA^{Glc} in *blue*, IIB^{Glc} in *green*, and HPr in *red*. The side chains of His-90 of IIA^{Glc}, *Cys-35* of IIB^{Glc}, and His-15 of HPr are shown in *purple*, *yellow*, and *purple*, respectively; the phosphoryl group in the putative transition state is also displayed in *yellow*. *A*, IIA^{Glc}·IIB^{Glc}, *B*, IIA^{Glc}·HPr; *C* and *D*, two views illustrating a best fit superposition to IIA^{Glc} of IIB^{Glc} and HPr in the two complexes. The coordinates for the HPr·IIA^{Glc} complex are taken from Ref. 17 (RCSB accession code 1GGR). Residues from IIB^{Glc} and HPr are denoted in *italic type*.

Ile-72, and *Thr-75* of IIB^{Glc}) are perfectly apposed in the complex. Particularly striking are the extensive interactions between *Ile-36*, located at the apex of the convex binding site of IIB^{Glc}, with the dense cluster of interfacial aromatic side chains of IIA^{Glc} (Phe-41, Phe-88, Phe-41, His-75, and His-90) (*cf.* the intermolecular NOE data in Fig. 1) (Fig. 7*A*).

The semicircle of charged residues on both protein interaction surfaces display a complementary pattern of alternating negative and positive charges (Fig. 6*A*). Asp-38 and Asp-94, which are located immediately adjacent to His-90, form salt bridges with *Arg-38* (Figs. 7*A* and 8, *B* and *C*). Asp-94, together with Glu-97, is also salt-bridged to *Arg-40* (Figs. 7*A* and 8*B*). At the edges of the binding sites, there are three additional salt bridges between Lys-99 and *Asp-33*, Glu-72 and *Lys-81*, and Lys-69 and *Asp-78* (Fig. 7*A*).

It is interesting to note that IIB^{Glc} and the structurally homologous oligomerization domain of the arginine repressor (57, 58) make use of different interaction surfaces to fulfill their respective functions. Thus, the interaction surface employed by IIB^{Glc} is centered around active site *Cys-35* located in the β_1/β_2 turn and surrounded by spatially adjacent elements of the structure (Figs. 4, 5*A*, and 6*A* (*right*)). The oligomerization domain of the Arg repressor forms a trimer of trimer (57, 58); the trimer interface comprises all four strands of the antiparallel β-sheet (*i.e.* strands β_1, β_2, β_3, and β_4); the interface between the trimers is made up by helices α_1 and α_2 and the turn between strands β_3 and β_4. The oligomerization domain of the Arg repressor also contains the arginine binding site formed by helix α_2 and strand β_3 (58). From the perspective of structural genomics, this suggests that it is not necessarily straightforward to deduce function from structure, that the same protein fold can possess different functionalities, and that different, completely nonoverlapping surfaces on the same structural scaffold can be adapted for different types of interactions.

The Phosphoryl Transition State Intermediate—Odd and even numbers of phosphoryl transfer steps in the PTS proceed with inversion and retention of configuration of the phosphorus, respectively, indicating that the transition state involves a pentacoordinate phosphoryl group in a trigonal bipyramidal geometry with the donor and acceptor atoms in apical positions and the oxygen atoms lying in the equatorial plane (64).

Phosphoryl transfer in the IIA^{Glc}·IIB^{Glc} complex involves the transfer of phosphorus from the Nϵ2 atom of His-90 (29, 65, 66) to the Sγ atom of *Cys-35* (67). Although the phosphorylated state of IIB^{Glc} is somewhat more stable than that of the other components of the PTS, it is not sufficiently long lived to permit a complete NMR structure determination of the phosphorylated complex. Moreover, in the presence of equimolar IIA^{Glc}, rapid dephosphorylation of IIB^{Glc}-P occurs.[2] Nevertheless, the phosphoryl transition state intermediate can readily be modeled by introducing a phosphoryl group into the coordinates for the restrained regularized mean IIA^{Glc}·IIB^{Glc} structure and submitting these to further regularization subject to both covalent geometry restraints relating to the trigonal bipyramidal geometry at the phosphorus and all of the experimental NMR restraints, essentially as described for the HPr·IIA^{Glc} (17) and HPr·IIA^{Mtl} (18) complexes.

The Sγ–Nϵ2 distance in the transition state is mechanism-dependent and could vary from ~3.8 Å (the sum of N–P and N–S bond lengths) in the case of a fully associative transition state to ~7.2 Å (the sum of the N, 2 × P, and S van der Waals radii) in the case of a fully dissociative transition state.

In the unphosphorylated IIA^{Glc}·IIB^{Glc} complex, the Nϵ2–Sγ distance between *Cys-35* and His-90 is 5.75 Å, the χ_1 torsion angle of Cys-35 is g^+ (~63°), and the χ_1/χ_2 torsion angles of

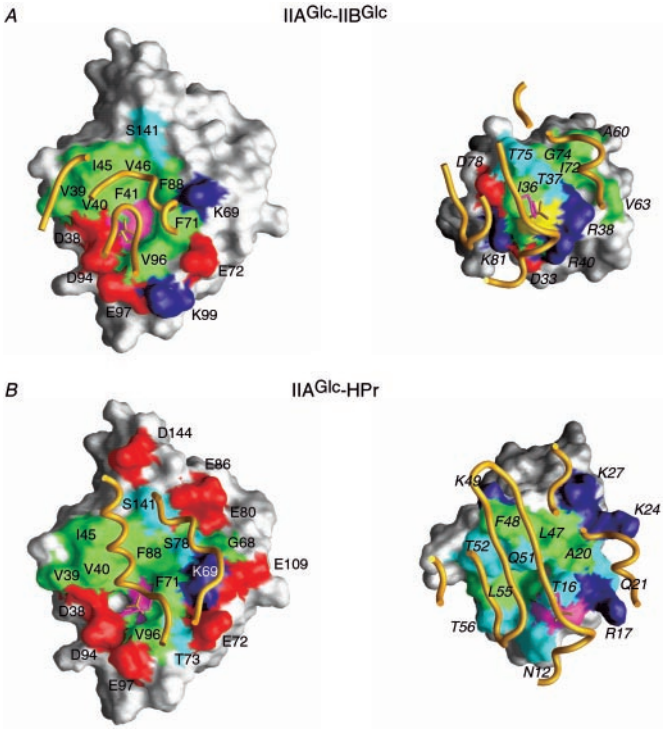

FIG. 6. **Surface representations illustrating the binding surfaces involved in the IIA^{Glc}·IIB^{Glc} (*A*) and IIA^{Glc}·HPr (*B*) complexes.** The binding surfaces on IIA^{Glc} for IIB^{Glc} and HPr are shown in the *left* of *A* and *B*, respectively; the binding surfaces on IIB^{Glc} and HPr for IIA^{Glc} are shown in the *right* of *A* and *B*, respectively. The binding surfaces are *color-coded*, with hydrophobic residues in *green*, polar residues in *light blue*, the active site histidines (His-90 for IIA^{Glc} and *His-15* of HPr) in *purple*, the active site cysteine (*Cys-35*) of IIB^{Glc} in *yellow*, positively charged residues in *dark blue*, and negatively charged residues in *red*. The relevant portion of the backbone of the partner protein is shown as a *gold ribbon*. Residues from IIB^{Glc} and HPr are denoted in *italic type*. The coordinates for the HPr·IIA^{Glc} complex are taken from Ref. 17 (RCSB accession code 1GGR).

His-90 are t/g^-. Upon regularization with no restriction on the Nε2–P and Sγ–P bond lengths in the transition state but with a planarity restraint to ensure that the imidazole ring of His-90, the P atom, and the Sγ atom of *Cys-35* lie in one plane and angle restraints to maintain planar trigonal geometry for the Nε2 atom of His-90, the Nε2–Sγ distance is reduced to 5.4 Å (*i.e.* substantial dissociative character), the rotameric states of *Cys-35* and His-90 are unaltered (χ_1 of *Cys-35* is increased to 77°, and χ_1/χ_2 of His-90 change by ≤5°), and there is no change in the backbone (Fig. 8*A*). If an additional distance restraint of 4 Å between the Sγ and Nε2 atoms is imposed, corresponding to an S_N2 mechanism (*i.e.* ~50% associative), the Nε2-Sγ distance is reduced to 4.1 Å, the χ_1 angle of *Cys-35* remains g^+ but is further increased to 87°, the χ_1/χ_2 angles of His-90 remain unperturbed, and there are only minimal changes in the backbone coordinates of Val-89–Phe-91 of IIA^{Glc} and *Ala-34–Ile-36* of IIB^{Glc} with an atomic displacement of only ~0.1 Å (Fig. 8*A*). Thus, the present structure of the unphosphorylated IIA^{Glc}·IIB^{Glc} can readily accommodate both types of mechanism.

Detailed views of the environment surrounding the phosphoryl group in the transition state complex are shown in Fig. 8, *B* and *C*. The phosphoryl group is completely buried within a cavity at the center of the protein-protein interface. In the view shown in Fig. 8*B*, the floor and walls of the cavity are predominantly hydrophobic (Val-40, Phe-41, Val-46, Phe-88, Phe-71, and Val-96 of IIA^{Glc} and *Ile-36* and the methyl group of *Thr-37* of IIB^{Glc}), whereas the roof consists of a cluster of interacting charged residues (Asp-38 and Asp-94 of IIA^{Glc} and *Arg-38* and *Arg-40* of IIB^{Glc}). There is a dense network of hydrogen bonds to the phosphoryl group; on the IIA^{Glc} side, these originate from the backbone amide of Asp-94 and possibly the Nε2H atom of His-75 (indirectly via a potential bridging water molecule), and on the IIB^{Glc} side, they originate from the backbone amides of *Ile-36* and *Thr-37*, the hydroxyl group of *Thr-37*, and the guanidino group of *Arg-38*. The multifunctional nature of the side chain of Arg-38 plays a crucial role; its guanidino group also partially neutralizes the negatively charged carboxylates of Asp-38 and Asp-94, which are in very close proximity to the phosphoryl group (~4.3 and ~3.6 Å, respectively), and its NεH group is part of the cluster of NH groups that is hydrogen-bonded to the Sγ atom of *Cys-35* (Figs. 3*C* and 8*C*).

The interactions between IIB^{Glc} and the phosphoryl group in the transition state are fully consistent with the observed backbone amide ^{15}N chemical shift differences, $|\Delta(\delta^{15}N)|$, between phosphorylated and unphosphorylated IIB^{Glc}: specifically, very large perturbations with $|\Delta(\delta^{15}N)|$ ~5–6 ppm for Ile-36 (downfield shifted upon phosphorylation), *Thr-37*, and *Arg-38* (upfield shifted upon phosphorylation); smaller perturbations with $|\Delta(\delta^{15}N)|$ ~ 1.5–2 ppm for *Cys-35*, *Leu-39*, and *Arg-40*; and insignificant perturbations with $|\Delta(\delta^{15}N)| < 0.5$ ppm for the remaining residues (20).[2] In addition, in the phosphorylated state of IIB^{Glc}, the Nε-Hε cross peak for the side chain of Arg-38 is detectable in the ^1H-^{15}N HSQC spectrum (at 85.4/7.4 ppm).[2]

A mechanistic scheme summarizing the structural results is shown in Fig. 9. Phosphoryl transfer from IIA^{Glc} to IIB^{Glc} is initiated by nucleophilic attack by the thiolate anion of Cys-35 (pK ~6.5). In phosphorylated IIA^{Glc}, the Nδ1H atom of His-90 is hydrogen-bonded to the backbone carbonyl of Gly-92, and the Nε2H of His-75 can readily come into direct hydrogen bonding distance of the phosphoryl group by small changes in the χ_2 angle of His-90. In the transition state, the distance from the Nε2 of His-75 to the phosphoryl group is lengthened, possibly destabilizing the Nε2(His-90)–P bond. We suggest that the transfer of the phosphoryl group from IIA^{Glc} to IIB^{Glc} is favored predominantly as a consequence of the greater chemical stability of the Sγ(Cys-35)–P bond relative to the Nε2(His-90)–P bond arising from three factors: (*a*) the very close proximity of the negatively charged carboxylates of Asp-38 and Asp-94 of IIA^{Glc} to the phosphoryl group (Fig. 8, *B* and *C*), (*b*) the extensive network of intramolecular hydrogen bonds from amide groups to the sulfur atom of *Cys-35* (Fig. 3*C*), which are readily maintained in the transition state (Fig. 8*C*), and (*c*) the larger number of stabilizing hydrogen bonding interactions to the phosphoryl group originating from IIB^{Glc} than IIA^{Glc} both in the transition state and in the uncomplexed phosphorylated forms of the two proteins (Fig. 8*C*). This interpretation of the structural data is consistent with the observed overall equilibrium constant ([IIB^{Glc}-P][IIA^{Glc}]/[IIB^{Glc}][IIA^{Glc}-P]) of ~3 for the transfer of phosphorus from IIA^{Glc} to IIB^{Glc} (68).

Correlation with Biochemical Data—Sequence comparisons of IIB^{Glc} and IIA^{Glc} from Gram-negative (*E. coli*) and Gram-positive (*Bacillus subtilis*) bacteria, as well as from mycoplasma (*Mycoplasma capricolum*), are shown in Fig. 7*C*. The 14 interfacial residues of IIB^{Glc} are highly conserved, significantly more so than the 19 interfacial residues of IIA^{Glc}. In addition, the substitutions are generally conservative in nature. Moreover, within each bacterial class, the percentage identity of interfacial residues is extremely high: 100% for both IIA^{Glc} and IIB^{Glc} from Gram-negative bacteria (*E. coli*, *Vibrio cholera*, *Yersinia pestis*, and *Salmonella typhimurium*) and 90–95 and 100% for IIA^{Glc} and IIB^{Glc}, respectively, from Gram-positive bacteria (*B. subtilis*, *Staphylococcus carnosus*, and *Staphylococcus aureus*).

From the sequence comparisons, it is clear that whereas the absolute identity of individual hydrophobic interfacial residues

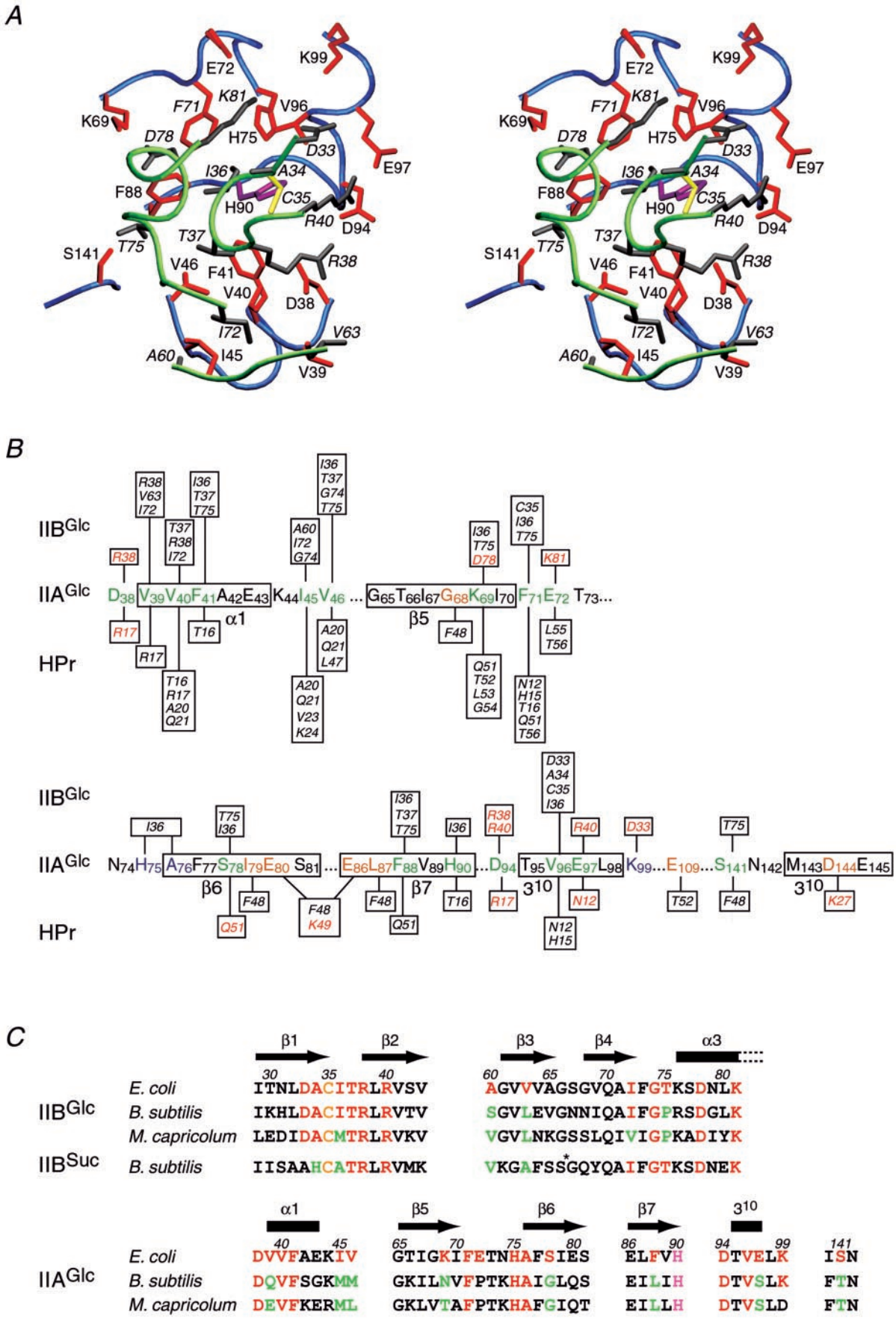

FIG. 7. **Intermolecular interactions in the unphosphorylated IIA^{Glc}·IIB^{Glc} complex.** *A*, stereoview of the unphosphorylated IIA^{Glc}·IIB^{Glc} interface. The backbones of IIA^{Glc} and IIB^{Glc}, depicted as *ribbon diagrams*, are shown in *blue* and *green*, respectively; the side chains of IIA^{Glc} and IIB^{Glc} are shown in *red* and *gray*, respectively; and the active site His-90 of IIA^{Glc} and *Cys-35* of IIB^{Glc} are depicted in *purple* and *yellow*, respectively. *B*, diagrammatic summary of interfacial contacts observed in the IIA^{Glc}·IIB^{Glc} and IIA^{Glc}·HPr complexes. Residues of IIB^{Glc} and HPr

may vary, the network of intermolecular hydrophobic interactions themselves is preserved. For example, in the *E. coli* system, Val-46 of IIA^{Glc} is in contact with *Ile-36*, *Thr-37*, *Gly-74*, and *Thr-75* of IIB^{Glc}. Val-46 is substituted by Met in *B. subtilis* and Leu in *M. capricolum* IIA^{Glc}; *Ile-36* is substituted by Met in *M. capricolum* IIB^{Glc}; and Thr-75 is substituted by Pro in both *B. subtilis* and *M. capricolum* IIB^{Glc}. Thus, the substitutions are in general compensatory, and the packing density at the interface remains largely unperturbed.

Although the five charged residues (*Asp-33*, *Arg-38*, *Arg-40*, *Asp-78*, and *Lys-81*) at the IIB^{Glc} interface are conserved in the three bacterial classes, only two (Asp-38 and Asp-94) of the 6 charged residues at the IIA^{Glc} interface are conserved. Thus, Lys-69 is substituted by Asn or Thr, Glu-72 is substituted by Pro, Glu-97 is substituted by Ser, and Lys-99 is substituted by Asp (in *B. subtilis*). As a result, only three intermolecular salt bridges (Asp-38–*Arg-38*, Asp-94–*Arg-38*, and Asp-94–*Arg-40*) are invariant. The other four salt bridges (Glu-97–*Arg-40*, Lys-69–*Asp-78*, Glu-72–*Lys-81*, and Lys-99–*Asp-33*) are all located at the periphery of the protein-protein interface. With the exception of the E72P substitution, a hydrophilic functional group is retained, and consequently, the direct salt bridging interactions in the *E. coli* complex may be replaced by weaker, indirect water-bridged interactions. Comparison of the apparent K_m values for homologous *versus* heterologous phosphoryl transfer from IIA^{Glc} to IIB^{Glc} can therefore be used to assess the impact of these peripheral electrostatic interactions. The K_m of *E. coli* IIA^{Glc} for *E. coli* IIB^{Glc} (~1.7 μM) is 15-fold lower than that of *B. subtilis* IIA^{Glc} for *E. coli* IIB^{Glc} (~25 μM) and 10-fold lower than that of *E. coli* IIA^{Glc} for *B. subtilis* IIB^{Suc} (K_m ~17 μM) and *B. subtilis* IIA^{Glc} for *B. subtilis* IIB^{Suc} (K_m ~18 μM) (69). (Note that no data are available for *B. subtilis* IIB^{Glc}, but *B. subtilis* IIB^{Suc} (70) is highly homologous to IIB^{Glc}, particularly with respect to the interfacial residues; see Fig. 7*C*). This suggests that the four peripheral electrostatic interactions present in the *E. coli* IIA^{Glc}·IIB^{Glc} complex but absent in the other three complexes contribute ~1.5 kcal·mol^{-1} to the binding energy.

The impact of the four peripheral electrostatic interactions on binding affinity is cumulative. Thus, mutations in IIA^{Glc} of Lys-69 to Leu, Lys-69 to Glu, and Lys-99 to Glu only reduce phosphoryl transfer activity from IIA^{Glc} to IIB^{Glc} by 25, 70 and 35%, respectively (71). The K69E mutation removes the intermolecular salt bridge with *Asp-78* but preserves intermolecular hydrophobic interactions with *Ile-36* and *Thr-75*. Electrostatic repulsion between K69E and *Asp-78* and between K99E and *Asp-33* is mitigated to some extent by the reduction in side chain length of glutamic acid relative to lysine. Moreover, the increased separation between the negatively charged carboxylate pairs affords the possibility of weak attractive intermolecular interactions mediated by one or more bridging water molecules.

Interestingly, even mutation of the conserved Asp-94 of IIA^{Glc} to Gly only reduces phosphoryl transfer activity from IIA^{Glc} to IIB^{Glc} by ~30% relative to wild type IIA^{Glc} (71). This is perhaps not surprising, since the D94G mutation does not introduce any electrostatic repulsion between the two proteins, and the remaining intermolecular electrostatic interactions, coupled with shape complementarity, still provide the necessary orientational specificity to guide the correct docking of the two proteins.

A key feature of the IIA^{Glc}·IIB^{Glc} complex is that any given interaction only makes a small contribution to the whole. Consider, for example, Phe-71 of IIA^{Glc}. Whereas Phe-71 makes extensive intermolecular van der Waals contacts with residues of IIB^{Glc} (*Cys-35*, *Ile-36*, and *Thr-75*), substitution of Phe-71 by either Lys or Ser only reduces phosphoryl transfer activity from IIA^{Glc} to IIB^{Glc} by ~30% (71). Although Phe-71 is spatially adjacent to the active site His-90, it is located close to the periphery of the interface (Fig. 6*A*). In the case of the F71K mutation, the long aliphatic portion of the lysine side chain still permits intermolecular hydrophobic interactions, and the NζH$_3^+$ group can make a potential intermolecular salt bridge with *Asp-78* (Figs. 6*A* and 7*A*). In the case of the F71S mutation, only the van der Waals interaction between the β-methylene group of residue 71 and the δ methyl of *Ile-36* is preserved, but the peripheral location of residue 71 ensures that an energetically unfavorable cavity is not introduced at the protein-protein interface (Fig. 7*A*). Thus, the reduction in IIA^{Glc} phosphoryl transfer activity is relatively small and can readily be attributed to somewhat suboptimal interfacial packing in the vicinity of residue 71.

Mutational data for IIB^{Glc} are only available for three residues: *Cys-35*, *Arg-38*, and *Arg-40* (72, 73). The *C421S* mutant of IICB^{Glc} (equivalent to *C35S* in the isolated IIB^{Glc} domain) does not support glucose transport *in vivo* and does not catalyze phosphoryl exchange between glucose and glucose-6-phosphate at equilibrium (72). In addition, the *C35S* mutant of IIB^{Glc} is only poorly phosphorylated by IIA^{Glc}.[2] This is hardly surprising, since the hydroxyl group of serine is a poor nucleophile. *Arg-38* and *Arg-40* of IIB^{Glc} are invariant and involved in intermolecular salt bridges (Figs. 7*A* and 8, *B* and *C*). In addition, the guanidino group of *Arg-38* participates in stabilization of phosphorylated *Cys-35* through interactions with both the sulfur atom and the phosphoryl group (Fig. 8*C*). Whereas mutation of either *Arg-38* to *Lys* or *Arg-40* to *Lys* abolishes phosphoryl transfer from IIB^{Glc} to glucose, both mutations still support phosphoryl transfer from IIA^{Glc} to IIB^{Glc} (73). These observations can be rationalized as follows. From the perspective of affinity of IIB^{Glc} for IIA^{Glc}, the R38K and R40K mutations still permit salt bridges to be formed with Asp-38 and Asp-94 and, hence, are unlikely to have any major effect on K_m. In the context of the phosphoryl transfer reaction from IIA^{Glc} to IIB^{Glc}, the *R40K* mutation does not impact the stabilization of the phosphorylated cysteine, and the *R38K* mutation only removes one of four hydrogen bonding interactions with the sulfur atom of *Cys-35* and one of five interactions with the phosphoryl group (Figs. 8*C* and 9). Thus, one can conclude that the multifunctional groups of the Arg-38 side chain are not a prerequisite for stabilization of the thiolate form of *Cys-35* required for nucleophilic attack on phosphorylated His-90 of IIA^{Glc}. Since the structure of the membrane-spanning IIC^{Glc}

involved in intermolecular hydrogen bonding or electrostatic interactions are denoted in *red*. Residues of IIA^{Glc} shown in *green* participate in intermolecular interactions in both complexes; residues of IIA^{Glc} shown in *blue* are only involved in the IIA^{Glc}·IIB^{Glc} complex; residues of IIA^{Glc} shown in *orange* are only involved in the IIA^{Glc}·HPr complex. *C*, sequence comparison of interfacial residues of IIB^{Glc} (*top*) and IIA^{Glc} (*bottom*) from Gram-negative (*E. coli*) and Gram-positive (*B. subtilis*) bacteria and from mycoplasma (*M. capricolum*). Interfacial residues are color-coded as follows: *red*, identical to residue in *E. coli* protein; *green*, conservative substitution relative to *E. coli* protein (note that the substitutions of Val-39 in *E. coli* IIA^{Glc} to Gln or Glu in *B. subtilis* and *M. capricolum* IIA^{Glc} are indicated as conservative, since the hydrophobic interactions involving Val-39 are preserved by the methylene groups of the Gln and Glu side chains). The active site cysteine of IIB^{Glc} and histidine of IIA^{Glc} are shown in *yellow* and *purple*, respectively. The percentage sequence identities, relative to *E. coli*, for the interfacial residues of IIB^{Glc} (14 residues) and IIA^{Glc} (19 residues) are 79 and 53%, respectively, for *B. subtilis* and 64 and 47%, respectively, for *M. capricolum*. The corresponding percentages for the *B. subtilis versus M. capricolum* comparison are 78 and 74%, respectively. Also shown for comparison is the sequence for the interface of IIB^{Suc} from *B. subtilis*. The percentage identity between the interfacial residues of *E. coli* IIB^{Glc} and *B. subtilis* IIB^{Suc} is 64%.

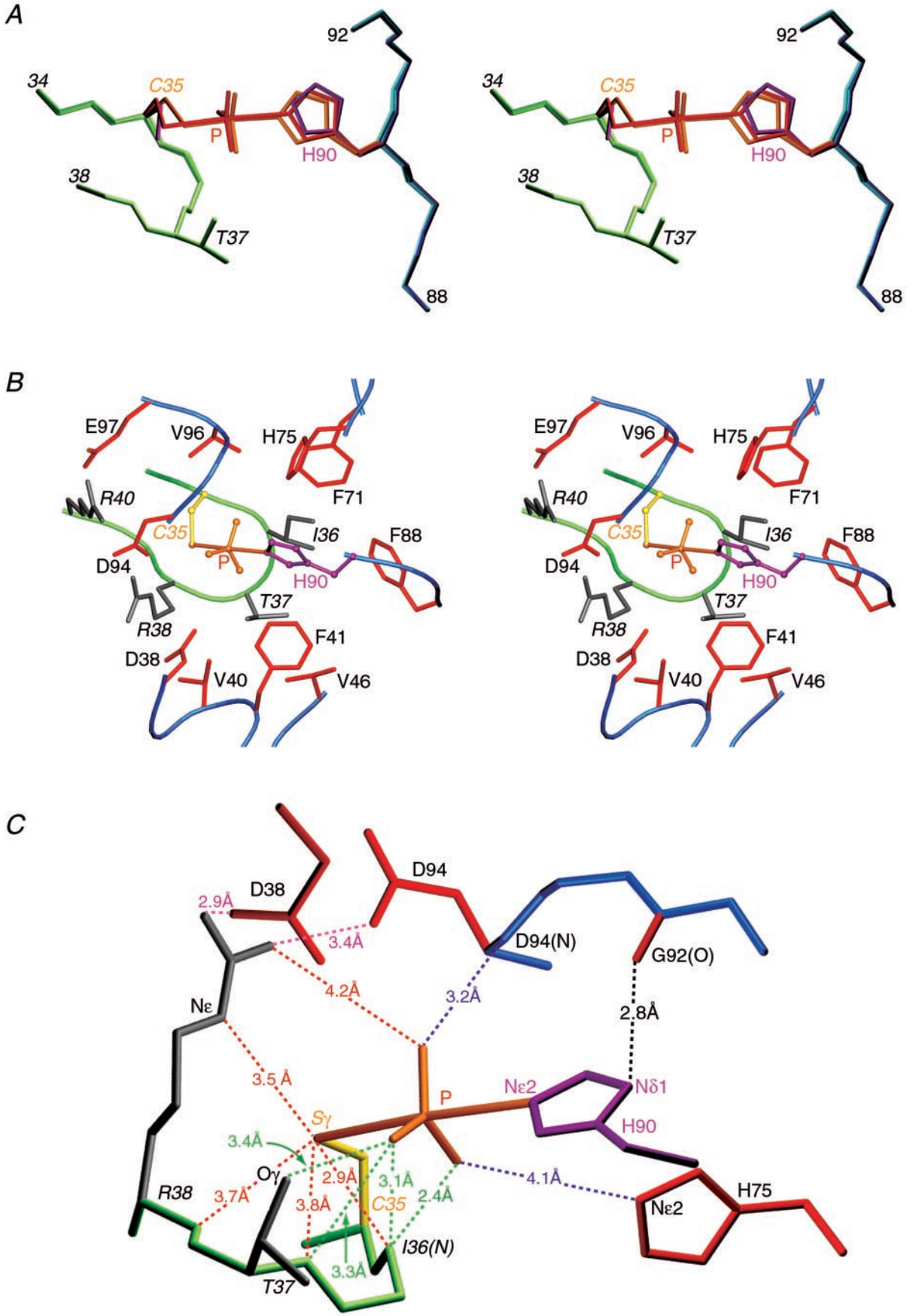

FIG. 8. **The phosphoryl transition state of the IIA^{Glc}·IIB^{Glc} complex.** *A*, detailed view (stereo) around the active site His-90 and *Cys-35*, illustrating the backbone and side chain positions in the unphosphorylated complex, the dissociative transition state, and the associative transition state. The backbones of IIA^{Glc} and IIB^{Glc} are shown in *dark blue* and *dark green*, respectively, for the unphosphorylated complex and in *light blue* and *light green*, respectively, for the two transition state complexes. The side chains of His-90 and *Cys-35* and the pentacoordinate phosphoryl

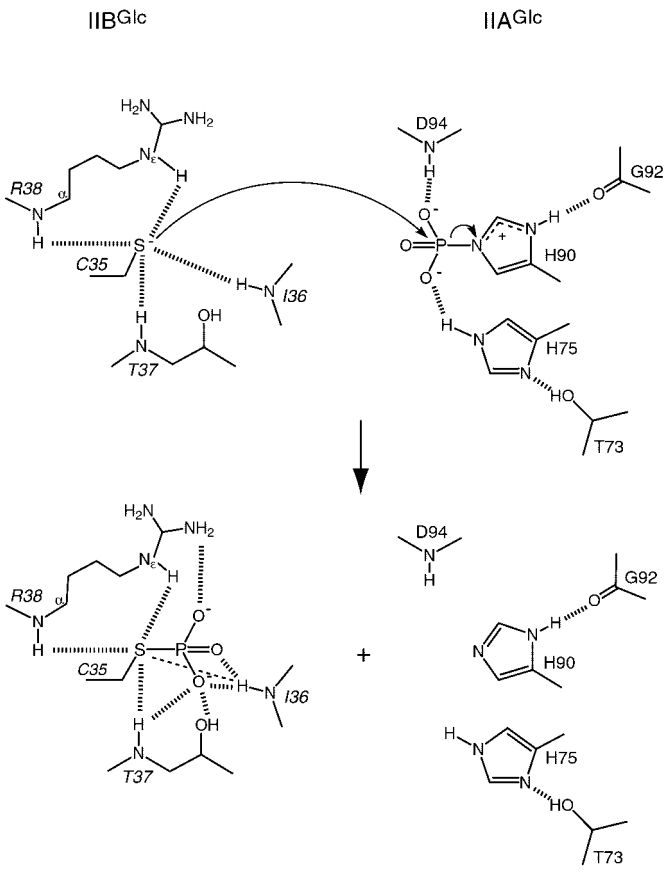

FIG. 9. **Schematic of the mechanism of phosphoryl transfer from IIA^{Glc} to IIB^{Glc}.**

domain of the glucose permease is unknown, one cannot ascertain the exact role of the *Arg-38* and *Arg-40* side chains in phosphoryl transfer to glucose other than to state that the multifunctional character of these arginine side chains must be critical for the reaction.

IIB^{Glc} Uses Similar Interaction Surfaces for IIA^{Glc} and the Mlc Repressor—In addition to participating in phosphoryl transfer within the glucose arm of the PTS, the IIB domain of IIBC^{Glc}, in its dephosphorylated form, binds tightly to the global transcriptional repressor Mlc, thereby inducing transcription of a number of genes involved in sugar metabolism and transport (9–11). Recent mutational analysis of surface residues in the vicinity of the active site *Cys-35* indicates that IIA^{Glc} and Mlc bind to overlapping sites on IIB^{Glc} (12). In the view of the interaction surface for IIA^{Glc} depicted in Fig. 6*A* (*right*), the mutational data suggests that the binding surface for Mlc comprises the right half of the IIA^{Glc} binding surface: specifically, *Cys-35*, *Thr-37*, *Arg-38*, *Gln-70* (which is located between *Arg-38* and *Val-63*, depicted in Fig. 6*A*), and *Ile-72*. Thus, *R38A*, *R38H*, *R38K*, and *Q70A*/*I72A* mutations completely abolish derepression of Mlc-regulated genes; *C35S*, *C35D*, *T37A*, and *Q70A* partially abolish derepression (activity ranging from 40 to 80% of wild type unphosphorylated IIB^{Glc}); and *D33A*, *I36A*, *R40A*, and *I72A* mutations behave identically to wild type unphosphorylated IIB^{Glc} (12). (Note that although the single point mutation *I72A* has no effect on activity, *Ile-72* must be part of the Mlc binding site, since *I72A* acts synergistically with *Q70A* in the context of the double *Q70A*/*I72A* mutations.)

Comparison of the IIA^{Glc}·IIB^{Glc} and HPr·IIA^{Glc} Interfaces—A comparison of the interaction of IIA^{Glc} with IIB^{Glc} and HPr (17) is provided in Figs. 5 and 6. The two protein complexes are characterized both by common features and features that make them distinct from one another. (Residues of both IIB^{Glc} and HPr are indicated in italic type.)

Since His-90 of IIA^{Glc} accepts a phosphoryl group from *His-15* of HPr and donates a phosphoryl group to *Cys-35* of IIB^{Glc} (29, 65, 66), it is evident that the binding surfaces for IIB^{Glc} and HPr on IIA^{Glc} must overlap. Moreover, the concave interaction surface on IIA^{Glc} is complemented by convex interaction surfaces on both IIB^{Glc} and HPr. However, the three-dimensional structure of IIB^{Glc} and HPr, as well as the structural elements of their respective binding surfaces for IIA^{Glc}, are entirely different (*cf.* Fig. 5). Thus, HPr employs two helices to interact with IIA^{Glc} (17) (Fig. 5*B*), whereas IIB^{Glc} makes use of β-strands and a β-hairpin, supplemented by a small stretch of helix (Fig. 5*A*). In the context of a phosphoryl transition state intermediate, it is noteworthy that His-90 of IIA^{Glc}, the pentacoordinate phosphoryl group, and the Sγ atom of *Cys-35* of IIB^{Glc} and the Nδ1 atom of His-15 of HPr, occupy the same spatial positions in the two complexes (Fig. 5, *C* and *D*). It is also worth noting that histidine cannot substitute for *Cys-35* in IIB^{Glc}, since the longer histidine side chain could not be accommodated in the context of a phosphoryl transition state IIA^{Glc}·IIB^{Glc} complex.

The binding surfaces for IIB^{Glc} and HPr on IIA^{Glc} overlap extensively (Fig. 6). 23 residues of IIA^{Glc} participate in the interaction with HPr, and 19 participate in the interaction with IIB^{Glc}; of these, 16 residues of IIA^{Glc} are involved in both protein-protein complexes (Fig. 7*B*). The hydrophobic component of the IIB^{Glc} and HPr binding surfaces on IIA^{Glc} is virtually identical and principally made up of Val-39, Val-40, Phe-41, Ile-45, Val-46, Phe-71, Phe-88, and Val-96 (Fig. 6, *A* and *B*, *left panels*). The complementary patch of hydrophobic residues on the two partner proteins comprises *Ala-20*, *Leu-47*, *Phe-48*, the aliphatic portion of the side chain of *Gln-51* and *Leu-55* for HPr (Fig. 6*B*, *right*), and *Ile-72*, the methyl groups of *Thr-37* and *Thr-75*, and *Ile-36* for IIB^{Glc} (Fig. 6*A*, *right*). The triad of three negatively charged residues consisting of Asp-38, Asp-94, and Glu-97 on IIA^{Glc} is preserved in both binding sites (Figs. 6, *A* and *B*, *left*), and their charges are neutralized by interaction with *Arg-17* and *Asn-12* of HPr (Fig. 6*B*, *right*) and *Arg-38* and *Arg-40* of IIB^{Glc} (Fig. 6*A*, *right*). Indeed, *Arg-17* of HPr and *Arg-38* of IIB^{Glc} are involved in functionally equivalent intermolecular interactions with Asp-38 and Asp-94 in the two complexes.

Despite the overlap and similarities, the differences between the binding surfaces in the two complexes are striking and reside in the distribution and nature of charged residues at the outer edges of the binding surfaces. The binding surface on HPr

group (in the case of the transition state complexes) are shown in *purple* for the unphosphorylated complex, in *red* for the dissociative transition state (Nϵ2–Sγ distance of ~5.4 Å between His-90 and *Cys-35*), and in *orange* for the associative transition state (Nϵ2–Sγ distance of ~4.1 Å between His-90 and *Cys-35*). *B*, detailed view (stereo) of the active site in the putative associative transition state. The backbone (depicted as a *tube*) of IIA^{Glc} and IIB^{Glc} is shown in *blue* and *green*, respectively; the side chains of IIA^{Glc} and IIB^{Glc} are shown in *red* and *gray*, respectively; His-90 is in *purple*, *Cys-35* in *yellow*, and the pentacoordinate phosphoryl group in *orange*. *C*, the dense network of hydrogen bonding or electrostatic interactions involving the phosphoryl group and the Sγ atom of *Cys-35* in the transition state (distances given relate to the putative associative transition state). There is a network of intermolecular interactions that stabilize the phosphoryl group (*dashed blue lines* originate from IIA^{Glc} and *dashed green lines* from IIB^{Glc}); a set of intramolecular interactions that stabilize the thiolate (pK_a ~6.5) of *Cys-35* (*dashed red lines*); intermolecular salt bridges between *Arg-38* of IIB^{Glc} and Asp-38 and Asp-94 of IIA^{Glc} (*dashed purple lines*); and a single intramolecular hydrogen bond stabilizing the Nδ1 tautomeric state of His-90 (*dashed black line*). The *color coding* is the same as in *B*. Residues from IIB^{Glc} are denoted in *italic type*.

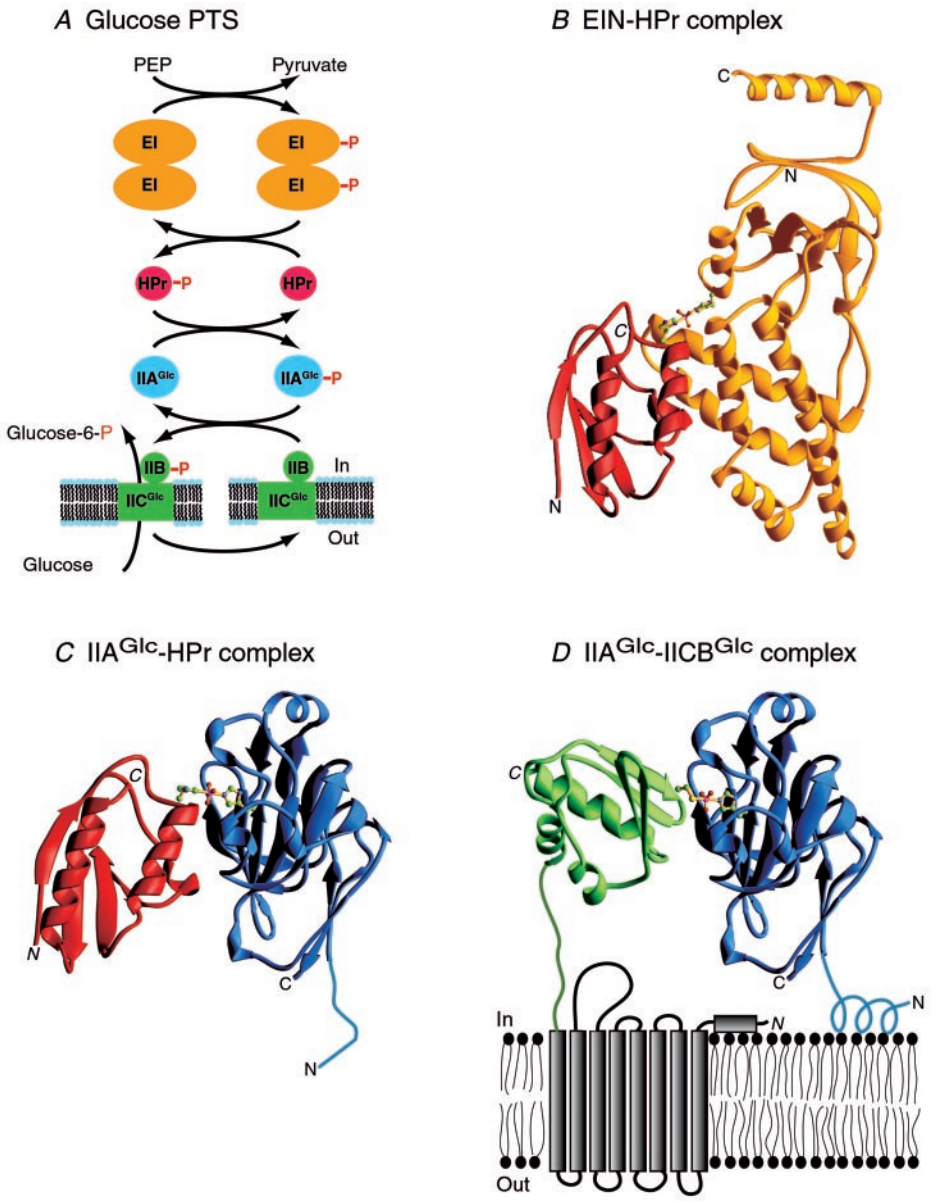

FIG. 10. **Summary of the glucose arm of the *E. coli* PTS.** *A*, diagrammatic illustration of the PTS cascade illustrating the transfer of phosphorus originating from phosphoenolpyruvate and ending up on glucose through a series of bimolecular protein-protein complexes between phosphoryl donor and acceptor molecules. Shown are *ribbon diagrams* of the first (EIN·HPr) (*B*), second (HPr·IIA^{Glc}) (*C*), and third (IIA^{Glc}·IICB^{Glc}) (*D*) complexes of the glucose PTS. EIN is shown in *gold*, HPr in *red*, IIA^{Glc} in *blue*, and the IIB^{Glc} domain of IICB^{Glc} in *green*. Also shown in *yellow* are the active site histidine residues of EIN (His-189), HPr (*His-15*), and IIA^{Glc} (His-90) and the active site cysteine (*Cys-35*) of IIB^{Glc}, together with the pentacoordinate phosphoryl group (*red atoms*) in the putative transition states of the complexes. IIB^{Glc} constitutes the C-terminal cytoplasmic domain of IICB^{Glc}. The transmembrane IIC^{Glc} domain of IICB^{Glc} is thought to comprise eight transmembrane helices (shown diagrammatically in *black*). Note that the N-terminal end of IIA^{Glc} (residues 1–18) is disordered in free solution (*C*), but upon interaction with a lipid bilayer, residues 2–10 adopt a helical conformation (*D*), thereby further stabilizing the IIA^{Glc}·IIB^{Glc} complex, by partially anchoring IIA^{Glc} to the lipid membrane (76). Coordinates for the EIN·HPr and IIA^{Glc}·HPR complexes are taken from Refs. 16 (RCSB accession code 3EZA) and 17 (RCSB accession code 1GGR), respectively.

has four positively charged residues (*Arg-17*, *Lys-24*, *Lys-27*, and *Lys-49*) and is completely devoid of any negatively charged ones (Fig. 6*B*, *right*). The binding surface on IIB^{Glc}, on the other hand, has both positively (*Arg-38*, *Arg-40*, and *Lys-81*) and negatively (*Asp-33* and *Asp-78*) charged residues (Fig. 6*A*, *right*). On the IIA^{Glc} side, the cluster of four negatively charged residues (Glu-109, Glu-80, Glu-86, and Asp-144) that constitutes part of the binding site for HPr (Fig. 6*B*, *left*) is completely absent from the interaction with IIB^{Glc} (Fig. 6*A*, *left panel*). Likewise, Lys-99 of IIA^{Glc} is only utilized in the IIA^{Glc}·IIB^{Glc} complex (Fig. 6*A*, *left panel*). The differences between the two binding surfaces on IIA^{Glc} are mirrored in the IIA^{Glc} binding sites on HPr and IIB^{Glc}; specifically, *Lys-27* and *Lys-49* of HPr form salt bridges with Glu-80/Glu-86 and Asp-144, respectively, of IIA^{Glc}, whereas *Asp-33* of IIB^{Glc} is salt-bridged to Lys-99. In addition, although Lys-69 of IIA^{Glc} is used in both binding surfaces, it serves different roles. In the IIA^{Glc}·IIB^{Glc} complex, Lys-69 is salt-bridged to *Asp-78*; in the IIA^{Glc}·HPr complex, however, the aliphatic portion of the Lys-69 side chain is in van der Waals contact with *Gln-51* and *Thr-52*, and its NζH$_3^+$ group is hydrogen-bonded to the backbone carbonyl of *Leu-53*.

A Cascade of Proteins for Phosphoryl Transfer to Glucose—Great effort has been expended in the structural elucidation of proteins involved in the PTS (summarized in the reviews cited in Refs. 3, 4, and 74). What was lacking was structural data on protein-protein complexes for this system. Fig. 10*A* shows the cascade of protein-protein complexes involved in phosphoryl transfer in the glucose-specific arm of the PTS. Since we published structures for both EIN·HPr (16) and IIA^{Glc}·HPr (17) complexes, the present report of the structure for the IIA^{Glc}·IIB^{Glc} complex signals the completion of our objective to understand the phosphotransfer protein-protein interfaces of the entire cascade for glucose phosphorylation and transport in *E. coli* (Fig. 10). These complexes shed light on understanding fundamental aspects of protein-protein recognition, mechanisms for phosphoryl transfer between proteins, and the diversity of structural elements recognized by a single protein. Specificity of the protein-protein interaction surfaces is characterized by geometric and chemical complementarity, coupled with extensive redundancy to permit the effective recognition of multiple partners. There is little or no conformational change in the protein backbone before and after association. Some interfacial side chains, however, adopt different conformations (side chain

conformational plasticity) depending on the interacting partner so as to achieve optimal intermolecular interactions (17, 18). A consequence of these properties is increased velocity in signal transduction by eliminating any unnecessary time delay required for significant conformational change.

The interaction surfaces for HPr on EI and IIA^{Glc} are very similar despite the fact that their underlying structures are completely different in terms of linear sequence, secondary structure (helices for EI, β-strands for IIA^{Glc}), and topological arrangement of structural elements. Thus, both interaction surfaces are concave and circular in appearance, comprising a central hydrophobic core surrounded by a ring of negatively charged residues (*cf*. Fig. 6 of Ref. 18). The single positively charged residue on the interaction surface of EI, Arg-126, is slightly displaced relative to Lys-69 on IIA^{Glc} and interacts with the side chain of *Gln-51* of HPr instead of the backbone of *Leu-53* (16, 17). HPr makes use of essentially the same surface to interact with both its upstream partner EI and its downstream partner IIA^{Glc} (17). Concomitantly, the binding sites for IIB^{Glc} and HPr on IIA^{Glc} overlap extensively (~85% of the binding site for IIB^{Glc} constitutes part of the binding site for HPr). One might therefore anticipate that IIB^{Glc} could also interact with EI. However, we find absolutely no evidence of any interaction between the N-terminal domain of EI and IIB^{Glc} by NMR.[2] From a functional perspective, this is clearly very important, since it ensures that the PTS cascade is not bypassed. In addition, prevention of the potential shortcut between EI and IICB^{Glc} for glucose phosphorylation is also necessary, since these proteins also regulate the functions of proteins in other pathways (2). We postulate that the structural basis for specificity and discrimination lies precisely in the different charge distributions on the interaction surfaces of HPr and IIB^{Glc} noted above. Model building suggests that the inability of IIB^{Glc} to bind to EI is due to electrostatic repulsion. Thus, if the relative orientation of HPr and IIB^{Glc} observed in the complexes with IIA^{Glc} was preserved in the complexes with EI, there would be highly unfavorable, repulsive electrostatic interactions between *Asp-33*, *Asp-78*, and *Lys-81* of IIB^{Glc} and Asp-120, Asp-82/Glu-83/Glu-84, and Arg-126, respectively, of EI (*cf*. Fig. 6 of Ref. 18). The likely importance of electrostatic interactions as a discriminator and determinant of specificity is suggested by sequence comparisons. *Asp-33*, *Asp-78*, and *Lys-81* are fully conserved in IIB^{Glc} from Gram-negative, Gram-positive bacteria and mycoplasma, despite the fact that their complementary, oppositely charged partners on IIA^{Glc} are not (Fig. 7*C*); similarly, Asp-82 and Arg-126 of EI are fully conserved in the three classes of microorganisms, and Glu-84 and Asp-120 are partially conserved (*cf*. Fig. 5 of Ref. 16).

Earlier, Roseman and colleagues (65, 75) observed that the N-terminal tail of IIA^{Glc} is unimportant for phosphoryl transfer from HPr to IIA^{Glc} but crucial for phosphotransfer from IIA^{Glc} to IIB^{Glc}. Structurally, the N-terminal tail of IIA^{Glc} is disordered in both the IIA^{Glc}·HPr (Fig. 10*C*) and IIA^{Glc}·IIB^{Glc} complexes, suggesting that it is not involved in the PTS protein-protein interactions. We have previously shown that the N-terminal tail of IIA^{Glc} is capable of associating with the *E. coli* membrane by forming an amphipathic helix (76). We suggest that the interaction of the N-terminal tail of IIA^{Glc} with the membrane provides additional stabilization of the IIA^{Glc}·IICB^{Glc} complex, thereby enhancing the ability of IIA^{Glc} to efficiently donate its phosphoryl group to IICB^{Glc}. In the native state, IIB^{Glc} is covalently linked to IIC^{Glc} via a ~75-residue linker (13). In the structure of the IIA^{Glc}·IIB^{Glc} complex, the first ordered residues of IIA^{Glc} (Thr-19) and IIB^{Glc} (*Met-14*) are located on the same face, which readily allows IIA^{Glc} to bind the membrane as depicted in Fig. 10*D*.

Upon completion of phosphoryl transfer from IIA^{Glc} to IIB^{Glc}, IIA^{Glc} must dissociate from IIB^{Glc} so that glucose can move to the active site Cys-421 of IICB^{Glc} (corresponding to *Cys-35* of IIB^{Glc}) to be phosphorylated. What triggers the departure of IIA^{Glc} from the IICB^{Glc} surface? The present NMR data indicate that interaction between IIA^{Glc} and IIB^{Glc} is clearly transient in nature ($k_{off} \geq 800\ s^{-1}$). Although the data presented in the current paper do not touch upon the subject of side chain dynamics, we propose that the intermolecular salt bridges between IIA^{Glc} and IIB^{Glc} are transient in nature. We also suggest that rapid formation and breaking of the salt bridges between the *Arg-38/Arg-40* pair of IIB^{Glc} and the Asp-39/Asp-94 pair of IIA^{Glc} may result in transient repulsions between the Asp-38/Asp-94 pair of IIA^{Glc} and the bianionic phosphate group. This may serve two purposes. First, it may promote phosphoryl transfer from IIA^{Glc} to IIB^{Glc} by increasing the lability of the Nϵ2(His-90)–P bond. Second, once IIB^{Glc} is phosphorylated, the same repulsions may promote departure of IIA^{Glc} from the IICB^{Glc} surface for the next cycle of phosphoryl transfer or for the regulation of other proteins.

Finally, the structural knowledge accumulated here will be useful for antibacterial drug design (77), since the PTS has not yet been found in humans (or any other eukaryotic organism).

Acknowledgments—We thank Carole Bewley for very helpful discussions; Charles Schwieters, Dan Garrett, Frank Delaglio, and John Kuszewski for software support; and Dusty Baber for support of NMR spectrometer hardware.

REFERENCES

1. Kundig, W., Ghosh, S., and Roseman, S. (1964) *Proc. Natl. Acad. Sci. U. S. A.* **52,** 1067–1074
2. Postma, P. W., Lengeler, J. W., and Jacobson, G. R. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Neidhardt, F. C., ed) pp. 1149–1174, American Society for Microbiology Press, Washington, D. C.
3. Robillard, G. T., and Broos, J. (1999) *Biochim. Biophys. Acta* **1422,** 73–104
4. Siebold, C., Flükiger, K., Beutler, R., and Erni, B. (2001) *FEBS Lett.* **504,** 104–111
5. Lux, R., Jahreis, K., Bettenbrock, K., Parkinson, J. S., and Lengeler, J. W. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 11583–11587
6. Seok, Y.-J., Sondej, M., Badawi, P., Lewis, M. S., Briggs, M. C., Jaffe, H., and Peterkofsky, A. (1997) *J. Biol. Chem.* **272,** 26511–26522
7. Novotny, M. J., Frederickson, W. L., Waygood, E. B., and Saier, M. H., Jr., (1985) *J. Bacteriol.* **162,** 810–816
8. Peterkofsky, A., Reizer, A., Reizer, J., Gollop, N., Zhu, P.-P., and Amin, N. (1993) *Prog. Nucleic Acids Res. Mol. Biol.* **44,** 31–65
9. Tanaka, Y., Kimata, K., and Aiba, H. (2000) *EMBO J.* **19,** 5344–5352
10. Lee, S.-J., Boos, W., Bouché, J.-P., and Plumridge, J. (2000) *EMBO J.* **19,** 5353–5361
11. Nam, T.-W., Cho, S.-H., Shin, D., Kim, J.-H., Jeong, J.-Y., Lee, J.-H., Roe, J.-H., Peterkofsky, A., Kang, S.-O., Ryu, S., Seok, Y.-J. (2001) *EMBO J.* **20,** 491–498
12. Seitz, S., Lee, S.-J., Pennetier, C., Boos, W., and Plumbridge, J. (2003) *J. Biol. Chem.* **278,** 10744–10751
13. Buhr, A., and Erni, B. (1993) *J. Biol. Chem.* **268,** 11599–11603
14. Buhr, A., Flükiger, K., and Erni, B. (1994) *J. Biol. Chem.* **269,** 23437–23443
15. Eberstadt, M., Grdadolnik, S. G., Gemmecker, G., Kessler, H., Buhr, A., and Erni, B. (1996) *Biochemistry* **35,** 11286–11292
16. Garrett, D. S., Seok, Y.-J., Peterkofsky, A., Gronenborn, A. M., and Clore, G. M. (1999) *Nat. Struct. Biol.* **6,** 166–173
17. Wang, G., Louis, J. M., Sondej, M., Seok, Y.-J., Peterkofsky, A., and Clore, G. M. (2000) *EMBO J.* **19,** 5635–5649
18. Cornilescu, G., Lee, B. R., Cornilescu, C. C., Wang, G., Peterkofsky, A., and Clore, G. M. (2002) *J. Biol. Chem.* **277,** 42289–42298
19. Reddy, P., Fredd-Kuldell, N., Liberman, E., and Peterkofsky, A. (1991) *Protein Expression Purif.* **2,** 179–187
20. Gemmecker, G., Eberstadt, M., Buhr, A., Lanz, R., Grdadolnik, S. G., Kessler, H., and Erni, B. (1997) *Biochemistry* **36,** 7408–7417
21. Garrett, D. S., Seok, Y.-J., Liao, D.-I., Peterkofsky, A., Gronenborn, A. M., and Clore, G. M. (1997) *Biochemistry* **36,** 2517–2530
22. Cai, M, Huang, Y, Sakaguchi, K, Clore, G. M., Gronenborn, A. M., and Craigie, R. (1998) *J. Biomol. NMR* **11,** 97–102
23. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6,** 277–293
24. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) *J. Magn. Reson.* **95,** 214–220
25. Clore, G. M., and Gronenborn, A. M. (1991) *Science* **252,** 1390–1399
26. Clore, G. M., and Gronenborn, A. M. (1998) *Trends Biotechnol.* **16,** 22–34
27. Bax, A., and Grzesiek, S. (1993) *Acc. Chem. Res.* **26,** 131–138
28. Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R., and Zhu, G. (1994) *Methods Enzymol.* **239,** 79–106
29. Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1993) *Protein Sci.* **2,** 543–558

30. Ottiger, M., Delaglio, F., and Bax, A. (1998) *J. Magn. Reson.* **131,** 373–378
31. Ding, K., and Gronenborn, A. M. (2002) *J. Magn. Reson.* **158,** 173–177
32. Clore, G. M., Starich, M. R., and Gronenborn, A. M. (1998) *J. Am. Chem. Soc.* **120,** 10571–10572
33. Hansen, M. R., Rance, M., and Pardi, A. (1998) *J. Am. Chem. Soc.* **120,** 11210–11211
34. Clore, G. M., Gronenborn, A. M., and Bax, A. (1998) *J. Magn. Reson.* **133,** 216–222
35. Tjandra, N., and Bax, A. (1997) *Science* **278,** 1111–1114
36. Rückert, M., and Otting, G. (2000) *J. Am. Chem. Soc.* **122,** 7793–7797
37. Chou, J. J., Gaemers, S., Howder, B., Louis, J. M., and Bax, A. (2001) *J. Biomol. NMR* **21,** 377–382
38. Feese, M. D., Comolli, L., Meadow, N. D., Roseman, S., and Remington, S. J. (1997) *Biochemistry* **36,** 16087–16096
39. Clore, G. M., and Garrett, D. S. (1999) *J. Am. Chem. Soc.* **121,** 9008–9012
40. Clore, G. M., and Gronenborn, A. M. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 5891–5898
41. Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) *J. Magn. Reson.* **160,** 65–73
42. Schwieters, C. D., and Clore, G. M. (2001) *J. Magn. Reson.* **152,** 288–302
43. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* **13,** 289–302
44. Nilges, M., Gronenborn, A. M., Brünger, A. T., and Clore, G. M. (1988) *Protein Eng.* **2,** 27–38
45. Kuszewski, J., Qin J., Gronenborn, A. M., and Clore, G. M. (1997) *J. Magn. Reson. Ser. B* **105,** 92–96
46. Clore, G. M., Gronenborn, A. M., and Tjandra, N. (1998) *J. Magn. Reson.* **131,** 159–162
47. Clore, G. M., and Kuszewski, J. (2002) *J. Am. Chem. Soc.* **124,** 2866–2867
48. Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1999) *J. Am. Chem. Soc.* **121,** 2337–2338
49. Schwieters, C. D., and Clore, G. M. (2001) *J. Magn. Reson.* **139,** 239–244
50. Carson, M. (1991) *J. Appl. Crystallogr.* **24,** 958–961
51. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins* **11,** 281–296
52. Schwieters, C. D., and Clore, G. M. (2002) *J. Biomol. NMR* **23,** 221–225
53. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr.* **47,** 110–119
54. Altschul, S. F., Gish, W., Miller, W, Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410
55. Vriend, G. (1990) *J. Mol. Graphics* **8,** 52–56
56. Vriend, G., and Sander, C. (1993) *J. Appl. Crystallogr.* **26,** 47–60
57. Ni, J., Sakanyan, V., Charlier, D., Glansdorff, N., and van Duyne, G. D. (1999) *Nat. Struct. Biol.* **6,** 427–432
58. van Duyne, G. D., Gosh, G., Maas, W. K., Sigler, P. B. (1996) *J. Mol. Biol.* **256,** 377–391
59. Fersht, A. R. (1985) *Enzyme Structure and Mechanism*, 2nd Ed., p. 156, W.H. Freeman & Co., New York
60. Barford, D., Flint, A. J., and Tonks, N. K. (1994) *Science* **263,** 1397–1404
61. Stuckey, J. A., Schubert, H. L., Fauman, E. B., Zhang, Z.-Y., Dixon, J. E., and Saper, M. A. (1994) *Nature* **370,** 571–575
62. Pannifer, A. D. B., Flint, A. J., Tonks, N. K., and Barford, D. (1998) *J. Biol. Chem.* **273,** 10454–10462
63. Zhang, Z.-Y., and Dixon, J. E. (1993) *Biochemistry* **32,** 9340–9345
64. Begley, G. S., Hanse, D. E., Jacobson, G. R., and Knowles, J. R. (1982) *Biochemistry* **21,** 5552–5556
65. Meadow, N. D., and Roseman, S. (1982) *J. Biol. Chem.* **257,** 14526–14537
66. Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1992) *Biochemistry* **31,** 5215–5224
67. Meinz, M., Jenö, Müller, D., Richter, W. J., Rosenbusch, J. P., and Erni, B. (1993) *J. Biol. Chem.* **268,** 11604–11609
68. Rohwer, J. M., Meadow, M. D., Roseman, S., Westerhoff, H. V., and Postma, P. W. (2000) *J. Biol. Chem.* **275,** 34909–34921
69. Reizer, J., Sutrina, S. L., Wu, L. -F., Deutscher, J., Reddy, P., and Saier, M. H., Jr. (1992) *J. Biol. Chem.* **267,** 9158–9169
70. Fouet, A., Arnaud, M., Klier, A., and Rapoport, G. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 8773–8777
71. Sondej, M., Seok, Y.-J., Badawi, P., Koo, B.-M., Nam, T.-W., and Peterkofsky, A. (2000) *Biochemistry* **39,** 2931–2939
72. Nuoeffer, C., Zanolari, B., and Erni, B. (1988) *J. Biol. Chem.* **263,** 6647–6655
73. Lanz, R., and Erni, B. (1998) *J. Biol. Chem.* **273,** 12239–12243
74. Herzberg, O., and Klevit, R. (1994) *Curr. Opin. Struct. Biol.* **4,** 814–822
75. Meadow, N. D., Coyle, P., Komoryia, A., Anfinsen, C. B., and Roseman, S. (1986) *J. Biol. Chem.* **261,** 13505–13509
76. Wang, G., Peterkofsky, A., and Clore, G. M. (2000) *J. Biol. Chem.* **275,** 16401–16403
77. Rognan, D., Mukhija, S., Folkers, G., and Zerbe, O. (2001) *J. Comput. Aided Mol. Des.* **15,** 103–115
78. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
79. Pelton, J. G., Torchia, D. A., Meadow, N. D., Wong, C. Y., and Roseman, S. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 3479–3483