Structure, dynamics and biophysics of the cytoplasmic protein–protein complexes of the bacterial phosphoenolpyruvate: sugar phosphotransferase system

G. Marius Clore and Vincenzo Venditti

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, USA

The bacterial phosphotransferase system (PTS) couples phosphoryl transfer, via a series of bimolecular protein–protein interactions, to sugar transport across the membrane. The multitude of complexes in the PTS provides a paradigm for studying protein interactions, and for understanding how the same binding surface can specifically recognize a diverse array of targets. Fifteen years of work aimed at solving the solution structures of all soluble protein–protein complexes of the PTS has served as a test bed for developing NMR and integrated hybrid approaches to study larger complexes in solution and to probe transient, spectroscopically invisible states, including encounter complexes. We review these approaches, highlighting the problems that can be tackled with these methods, and summarize the current findings on protein interactions.

The PTS as a paradigm for understanding complex protein interactions

The bacterial phosphoenolpyruvate: sugar phosphotransferase system (PTS) is the key signal transduction pathway involved in the regulation of central carbon metabolism in bacteria [1–7]. The PTS comprises a sequential cascade of bimolecular protein–protein complexes whereby a phosphoryl group originating on phosphoenolpyruvate (PEP) is transferred onto incoming sugars, thereby coupling phosphoryl transfer to active sugar transport across the membrane. The first two steps of the PTS are common to all sugars: Enzyme I (EI) is autophosphorylated by PEP and subsequently transfers the phosphoryl group to the histidine phosphocarrier protein HPr. HPr donates the phosphoryl group to the A component of the sugar-specific Enzymes II. There are four classes of Enzymes II: glucose (the first branch of the PTS to have been discovered in 1964 [1]), mannitol, mannos, and lactose/chitobiose (Figure 1A). The organization of the sugar-specific Enzymes II is similar: there are two cytoplasmic domains, IIA and IIB, and a transmembrane domain IIC, which may be supplemented by a fourth transmembrane domain IID. These domains may occur either as isolated proteins or be covalently joined by long linkers in a single contiguous protein (Figure 1A). IIA accepts the phosphoryl group from HPr and donates it to IIB; finally the transmembrane domain IIC catalyzes the coupled translocation and phosphoryl transfer from IIB to the incoming sugar. Despite the similar organization of Enzymes II, the IIA domains from the four sugar classes bear no similarity to one another in terms of sequence, or secondary, tertiary, or quaternary structure [8–14]. The IIB domains from the mannitol [15,16] and chitobiose [17,18] branches are similar in terms of secondary and tertiary structure but have no significant sequence similarity outside of the active site loop. The active site loop of the IIB domain of the glucose branch [19] bears some similarities to that of the IIB domains from the mannitol and chitobiose branches, but displays no overall similarity in sequence, or secondary or tertiary structure. Lastly the IIB domain of the mannose pathway [20–22] bears no similarity at all to that of the other three branches.

In addition to their function within the PTS cascade, components of the PTS are also involved in the regulation of many other proteins, including glycogen phosphorylase, adenylate kinase, glycerol kinase, various non-PTS permeases, and the global repressor Mlc [3,7,23–27]. In addition, under conditions of nitrogen limitation, competitive inhibition of EI by α-ketoglutarate [28], the carbon substrate for ammonia assimilation, blocks sugar transfer across the cell membrane, thereby providing a direct

Corresponding author: Clore, G.M. (marius@mail.nih.gov).
Keywords: signal transduction; protein–protein recognition; bacterial phosphotransferase system; NMR spectroscopy; hybrid methods in structure determination; residual dipolar couplings; solution X-ray scattering; sparsely populated states; encounter complexes.
0968-0004/3 – see front matter.
Published by Elsevier Ltd. http://dx.doi.org/10.1016/j.tibs.2013.08.003
biocatalytic link between central nitrogen and carbon metabolism in bacteria [29].

Many reviews have been written on the biology of the PTS and on the structures of individual components of the PTS [2–7]. Although crystal structures [8–13,17,21,22,30–35] have been solved for many of the cytoplasmic isolated proteins or domains of the PTS (the remainder being solved by NMR [14–16,18–20,36–41]), crystallization of PTS complexes has proved refractory. This is largely due to the fact that the complexes are transient and rationally weak with equilibrium dissociation constants (K_d) ranging from micro- to millimolar. Fortunately, this is not an impediment to the structure determination of such complexes in solution by NMR.

The multiplicity of protein–protein interactions in the PTS provides a paradigm for studying the factors governing specific recognition of multiple diverse targets. The current review focuses specifically on the NMR work in our laboratory, initiated in 1997 with the structure determination of the N-terminal domain of EI [39], aimed at understanding the structural basis of specific protein–protein recognition within the PTS; the fundamental biophysical mechanisms underlying protein–protein interactions; and the nature of large conformational rearrangements in multidomain proteins. In terms of structural biology and biophysics, these studies of protein–protein complexes of the PTS have provided a test bed for developing hybrid methods for solving the structures of large complexes and proteins in solution combining NMR and, in some instances, solution X-ray scattering, with crystal or NMR structures of individual domains, and for developing NMR methods for detecting and characterizing sparsely populated, transient encounter complexes formed by random collisions between partner proteins that precede the formation of the specific complex.

Solving the 3D solution structures of PTS complexes

The first structure of a PTS protein–protein complex to be determined was that between the N-terminal domain of EI (EIN) and HPr [42], for which the traditional approach involving a complete NMR structure determination of the entire complex was used (Box 1). It was rapidly realized, however, that this time-consuming approach could be both speeded up and rendered more accurate by making use of the available crystal or NMR structures of the free proteins [43]. This assumes that the backbone conformation of the proteins undergoes minimal changes upon complex formation, which can be readily ascertained both from the measurement of residual dipolar couplings (RDCs; see Glossary) as well as the observation of small backbone chemical shift changes upon complexation. The remainder of the PTS complexes were therefore solved by a procedure known as conjoined rigid body/torsion angle simulated annealing (Box 2), in which the proteins are treated as rigid bodies that are free to translate and rotate relative to one another, and interfacial side chains are given torsional degrees of freedom [43,44]. The driving force for these calculations resides in intermolecular nuclear Overhauser enhancement (NOE)-derived interproton distance restraints combined, where possible, with RDCs to provide orientational restraints and heteronuclear scalar couplings to derive torsion angle restraints for the interfacial side chains. This approach does not require that the complete backbone be treated as a rigid body: if there is evidence that a region of the backbone undergoes a conformational change upon complexation, then it is a simple matter to give that region torsional degrees of freedom whose conformational space is dictated by experimental NMR restraints (NOEs and RDCs measured on the complex) [45].

Whether one chooses to do a full NMR structure determination or to make use of known structures of the free protein, it is essential to distinguish unambiguously intermolecular NOEs from intramolecular NOEs within the individual proteins. This is readily accomplished by making use of appropriate isotopic labeling of the individual proteins [46]. For example, by 13C labeling one protein and having the other protein at natural isotopic abundance one
Traditionally, the geometric information used to solve 3D NMR structures of macromolecules relies principally on short (<6 Å) interproton distances derived from NOE measurements [91,92]. Many protons that are close in space are far apart in sequence, and hence, short interproton distance restraints are conformationally restrictive and sufficient, in their own right, to determine a 3D structure. Loose distance restraints, typically ranging from 1.8 to 2.7 Å, 1.8 to 3.5 Å, 1.8 to 5 Å, and 1.8 to 6 Å, corresponding to strong, medium, weak, and very weak NOEs, are adequate. Indeed, attempts to make use of more precise interproton distance restraints can often lead to a reduction in accuracy, owing to a phenomenon known as spin diffusion, whereby an NOE from say proton i to proton j is subsequently transferred to other protons in close proximity to proton j but far (>5 Å) from proton i. Supplementary structural restraints on backbone torsion angles can be obtained from three-bond scalar couplings [93] and backbone chemical shifts [94]. However, NOE and torsion angle restraints are short range in nature as they require close spatial proximity of the relevant atoms, and are therefore not good at specifying long-range order.

Long-range information can be obtained from RDCs measured in a dilute liquid crystalline medium [95,96] such as bichelles [97] or phage pf1 [98]. These media induce a small degree (~10^{-5}) of alignment with respect to the magnetic field such that linewidths remain narrow but internuclear dipolar couplings are no longer averaged to zero by Brownian rotational diffusion. The value of an RDC (D) for an interatomic vector (such as a backbone N-H bond) is given by a simple geometric relationship that is dependent on the magnitude of the principal component of the alignment tensor D_{\text{anis}}, the rhombicity of the tensor \eta, and two angles \theta and \phi describing the orientation of the vector relative to the x, y, and z axes of the alignment tensor (Figure I). RDCs are especially powerful for determining the relative orientation of two proteins within a complex as the two components share the same alignment tensor, and if the proteins can be treated as rigid bodies, only a small number of RDCs are required [43]. As a result it is possible to obtain an accurate structure of a protein–protein complex on the basis of only a few intermolecular NOEs to provide translational information, supplemented by RDCs for orientation.

![Figure I](image)

**Box 1. Brief overview of NMR structure determination**

<table>
<thead>
<tr>
<th>Nature of RDCs</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>D_{\text{anis}}</td>
<td>\frac{1}{5} (3 \cos^2 \theta - 1) + \frac{3}{2} \eta \sin^2 \theta \cos^2 \theta</td>
</tr>
</tbody>
</table>

**Solution structures of the PTS complexes – specific recognition of structurally diverse partners**

The structures of all nine soluble complexes of the PTS are presented in Figure 1. HPr interacts with five different proteins (EIN, IIA^{Glc}, IIA^{Mtl}, IIA^{Man}, and IIA^{Chb}) that display no similarities with regard to primary, secondary, tertiary, or quaternary structure. The only apparent commonality is that the active site residue is a histidine, and in each case a pentacoordinate phosphorl transition state can be modeled without necessitating any significant change in backbone coordinates. For the four IIA–IIA-HPr complexes, three involve phosphoryl transfer from a histidine on IIA to a cysteine on IIB. In the IIA^{Man}_–IIA^{Man} complex, phosphoryl transfer occurs from a histidine to a histidine. As in the case of the complexes with HPr, the phosphoryl transition state can be modeled with no significant changes in backbone coordinates. Thus, phosphoryl transfer for all cytoplasmic complexes of the PTS can proceed with maximum energetic efficiency without incurring an energetic penalty from any significant conformational changes required for either partner protein.

Despite the apparent lack of similarity in the complexes at the level of the ribbon diagrams shown in Figure 1B, the interaction surfaces are actually similar despite the different underlying secondary and tertiary structural elements. This is apparent from the molecular surface representations of the interfaces shown in Figure 2. All the protein–protein interfaces are characterized by large buried accessible surface areas ranging from 1200 to 1800 Å^2, with approximately equal contributions from both partners. The interaction surfaces for HPr and Enzymes IIB are convex, whereas those of EIN and Enzymes IIA are concave, providing shape complementarity. In general, each interaction surface comprises a predominantly hydrophobic central region interspersed with scattered polar residues, and asymmetrically distributed charged residues at judicious locations along the outer edges of the interaction.
surface. The active site residue is offset from the center of the binding site. Surface complementarity is largely provided by the hydrophobic residues, whereas electrostatic interactions serve to modulate affinity and direct the exact relative orientation of the partner proteins within a given complex. Sequences comparisons over several species indicate that although the absolute identity of the hydrophobic residues at the interface may vary, the network of intermolecular hydrophobic interactions is preserved and substitutions are in general compensatory, such that the packing density at the interface across different species remains largely unperturbed [19,42].

HPr uses essentially the same convex surface to interact with all its partners, and all charged residues on the interaction surface of HPr are positive (Figure 2A, top row). A key feature of the interactions of HPr with its partner proteins is redundancy of charged residues, such that the disposition of the complementary negatively charged residues on the partner proteins need not be identical. Thus, not all charged residues present at the protein–protein interface are involved in salt bridge interactions. In the same vein, each Enzyme IIA uses largely the same binding surface to interact with HPr and its IIB partner (compare the middle row of Figure 2A with the upper row of Figure 2B).

One might therefore ask why it is that the Enzymes IIB cannot bypass the corresponding Enzymes IIA and interact directly with EIN? The is because the binding surfaces on all the Enzymes IIB differ from that of HPr in one very significant way: although the majority of charged residues in the IIB binding site are positive, there are one or two negatively charged residues (Figure 2B, middle row) that complement positively charged residues on the corresponding IIA (Figure 2B, upper row), but would be repulsed by negatively charged residues on the surface of EIN (Figure 2A, upper row).

Another key feature of the binding surfaces that permits specific recognition of a wide variety of partners is side chain conformational plasticity, illustrated in Figure 3A for several side chain interactions between HPr and EIN or IIA$^{Glc}$ [47]. For example, the side chain of Phe48 undergoes a conformational switch from one rotamer in the EIN–HPr complex to another in the IIA$^{Glc}$–HPr complexes, thereby permitting Phe48 to interact with hydrophobic components on the partner proteins, specifically the methyl groups of Leu79, Leu85 and Ile108 for EIN, and the backbone residues of a β-sheet for IIA$^{Glc}$ (Figure 3A, left panel). Likewise, Arg17 of HPr interacts in one side conformation with Glu67 and Glu68 of EIN and in another conformation with Asp38 and Asp94 of IIA$^{Glc}$ (Figure 3A, right panel).

The phosphoryl transition state intermediate

Phosphoryl transfer in the PTS complex occurs via in-line phosphoryl transfer in which the donor atom, the phosphoryl, and the acceptor atom lie along a straight line [53]. Isotope labeling experiments have shown that odd and even numbers of phosphoryl transfer steps result in inversion and retention, respectively, of the configuration of the phosphorus [54,55], indicating that the transition state involves a pentacoordinate phosphoryl transition state in a trigonal bipyramidal geometry, with the donor and acceptor atoms in apical positions, and the oxygen atoms lying in the equatorial plane. The phosphoryl transition states can be modeled by introducing a phosphoryl group subject to constraints related to trigonal bipyramidal geometry at the phosphorus group. Phosphorylation occurs at the His(Ne2) atom for EIN and the four Enzymes IIA, at the His(Nε1) atom for HPr and IIB$^{Max}$, and at the Sy atom of cysteine for the remaining Enzymes IIB. The phosphoryl transition state can be formed without altering the position of the two partner proteins and with minimal changes in backbone coordinates immediately adjacent to the active site histidine or cysteine residues (Figure 2A,B, bottom row). The distances between the acceptor and donor atoms in the transition state can potentially range from the sum of the donor atom–phosphorus atom and phosphorus atom–acceptor atom bond lengths to the sum of the van der Waals radii of the donor, phosphorus (×2) and acceptor atoms, corresponding to fully associative and fully dissociative mechanisms, respectively. The N–P and S–P distances in the modeled transition states are consistent with substantial dissociative character, as predicted from a comparison of linear free energy relations for nonenzymatic and enzymatic phosphoryl transfer reactions [56]. Although the resolution of the structures of the complexes is limited, a fully associative mechanism would require substantial distortions and strain in the backbone adjacent to the donor and acceptor residues.

The phosphoryl transition states are shown in the lower rows of Figure 2A,B. The phosphoryl group and the active site histidines and/or cysteines lie on a bed of hydrophobic residues, and the phosphoryl group itself is stabilized by hydrogen bonding interactions involving polar (Thr, Ser, and His) or charged (Arg) residues. The number of hydrogen bonds to the phosphoryl group from each partner (both in the transition state and in models of the individual
phosphorylated proteins), at least in the case of the glucose pathway where measurements of the various equilibria are available, correlates with the directionality of phosphoryl transfer. Thus, the phosphoryl transfer from EIN to HPr and from IIA\textsuperscript{Glc} to IIB\textsuperscript{Glc} are favored by factors of ~10 and ~3, respectively, over the reverse transfers, whereas the phosphoryl transfers from HPr to IIA\textsuperscript{Glc} and from IIA\textsuperscript{Glc} to HPr are equally favorable [57], consistent with the larger number of hydrogen bonds stabilizing the phosphoryl group originating from HPr than EIN and from IIB\textsuperscript{Glc} than IIA\textsuperscript{Glc}, whereas the same number of hydrogen bonds stabilize the phosphoryl group in HPr and IIA\textsuperscript{Glc}.

Impact of linkers connecting PTS domains on the efficiency of phosphoryl transfer

Many of the domains of Enzymes II are connected by long flexible linkers. For example, the A, B, and transmembrane C domains of Enzyme IIA\textsubscript{ABC\textsuperscript{Mtl}} and the A and B domains of Enzyme IIAB\textsuperscript{Man} are expressed as single proteins, as are the IIB\textsuperscript{Glc} and transmembrane IIIB\textsuperscript{Glc} domains of IIABC\textsuperscript{Glc} [3]. In addition, many of the PTS complexes have \( K_d \) values in the high micromolar to millimolar range [20,45,48–51]. The \( K_d \) for the interaction of the isolated IIA\textsuperscript{Mtl} and IIB\textsuperscript{Man} domains is 3–4 mM [50], whereas that of the isolated IIA\textsuperscript{Man} and IIB\textsuperscript{Man} domains is ~0.5 mM [20].
Tethering two domains increases their effective local concentration, thereby increasing the probability of complex formation. The A and B domains of II^Mtl are connected by a 21-residue flexible linker, from which one can calculate, using well-established polymer chain theory, that the effective local concentration of the A and B domains is ~4 mM [50]. This value is consistent with ^15N relaxation dispersion measurements on IIAB^Mtl that yield a population of ~50% for the associated state with a unimolecular association rate constant of ~20 000 s⁻¹ for the interaction.
of the tethered A and B domains [58]. The latter corresponds to an apparent bimolecular association rate constant of $5 \times 10^6 \text{M}^{-1} \text{s}^{-1}$ [58] which is within the range typically observed for specific protein–protein interactions ($0.5 \times 10^9$–$5 \times 10^9 \text{M}^{-1} \text{s}^{-1}$) and in excellent agreement with the predicted value of $2 \times 10^6 \text{M}^{-1} \text{s}^{-1}$ for a purely diffusive process obtained from Brownian dynamics simulations [59]. Thus, the linker serves to tune the system optimally to achieve reasonable high occupancy coupled with rapid association and dissociation ($\sim20 \text{,000 s}^{-1}$ in both directions) to carry out efficiently three sequential phosphoryl transfer steps from HPr to IIAGlc, from IIAGlc to IIABMtl, and finally from IIABMtl onto the incoming sugar bound to the cytosolic side of IIICMtl. The phosphoryl transfer rate between the A and B domains of IIAMtl determined from NMR lineshape analysis is $\sim500 \text{ s}^{-1}$ in both directions [58]. This value is $\sim40$-fold lower than the rate constants for intramolecular domain–domain association and dissociation, therefore, one can calculate that $\sim80$ association/dissociation events take place for every phosphoryl transfer reaction. Thus, the rate-limiting step for phosphoryl transfer between the A and B domains of IIABMtl is governed by the chemistry of the phosphoryl transfer reaction itself rather than the rate of association to form the specific complex or the rate of dissociation to permit the next phosphoryl transfer reaction in the pathway to take place.

The glucose Enzymes II exhibit a variant of this phenomenon. Although the B and C domains are tethered by a $\sim75$-residue linker, the A domain is expressed as a separate protein. The first 18 residues of IIAGlc are disordered in solution and although the presence or absence of the N-terminal tail has no effect on phosphoryl transfer between HPr and IIAGlc, the presence of the N-terminal tail is critical for efficient phosphoryl transfer to IIIBGlc in vivo.
[60,61]. It turns out that residues 2–10 of the N-terminal tail of IIA\textsuperscript{Gle} associate with \textit{Escherichia coli} membranes to form an amphiphatic helix [62], thereby bringing IIA\textsuperscript{Gle} in close proximity to IIB\textsuperscript{Gle} and stabilizing the IIA\textsuperscript{Gle–IIB\textsuperscript{Gle}} interaction by effectively increasing the local concentration of IIA\textsuperscript{Gle} and IIBC\textsuperscript{Gle} (Figure 3B).

**Solution structure of intact EI and its complex with HPr – approaches to solving structures of larger (>100 kDa) complexes in solution**

The N-terminal domain (EIN) of EI can transfer a phosphoryl group to and accept a phosphoryl group from HPr but cannot be autophosphorylated by PEP [63–66]. Autophosphorylation of EI requires the presence of the C-terminal dimerization domain EIC. Intact EI is a 128-kDa dimer and therefore large by NMR standards. The probability of successfully solving a structure of this size using conventional NMR methodology is small, and even if it could be done, the coordinate accuracy would be low. To solve the solution structure efficiently and accurately of such a system therefore requires the development of hybrid methodology that makes use of prior available structural information combined with limited solution RDC and small (SAXS) and wide (WAXS) angle X-ray scattering [40].

The crystal structure of a trapped phosphorylated intermediate of intact \textit{E. coli} (Figure 4A, middle panel, and Figure 4B, left panel) has been solved by crystallizing EI from a solution containing PEP and Mg\textsuperscript{2+} and then quenching the autophosphorylation reaction using the inhibitor oxalate [33]. A comparison of the structure of the EIN domain in the EI phosphorylated intermediate with that of the isolated EIN domain, both free (X-ray [30] and NMR [39]) and complexed to HPr [42], reveals major conformational changes in the disposition of the two subdomains, \(\alpha\) and \(\alpha/\beta\), of EIN (Figure 4A). The EIN\textsuperscript{a} subdomain provides the interaction surface for HPr, whereas the EIN\textsuperscript{ab} subdomain contains the active site histidine, His189 [42]. The position of the active site histidines of HPr (His15) and EIN (His189) in the EIN–HPr complex allows for in-line phosphoryl transfer between EIN and HPr without requiring any substantial conformational changes (Figures 2A and 4A, left panel) [42]. In the phosphorylated EI intermediate [33], the EIN\textsuperscript{a} subdomain undergoes a reorientation of about \(\sim 70^\circ\) relative to the EIN\textsuperscript{ab} subdomain, such that the Ca–Ca distance between the active site histidines of HPr and EIN would be increased from \(\sim 12\ \text{Å}\) in the EIN–HPr complex to \(\sim 30\ \text{Å}\) (Figure 4A, middle panel); a distance too large to permit phosphoryl transfer from EIN to HPr. However, the active site His189 in the EIN\textsuperscript{ab} subdomain of the phosphorylated EI intermediate is optimally positioned for phosphoryl transfer from PEP bound to the EIC domain to His189 (Figure 4A, left panel) [33]. If the orientation of the EIN\textsuperscript{ab} subdomain relative to the EIC dimer in the crystal structure of the phosphorylated intermediate were preserved, the EIN\textsuperscript{a} subdomain would interpenetrate the EIC domain in the configuration found in the isolated EIN domain (Figure 4A). Thus, the transition from free EI to phosphorylated EI must be accompanied by two major rigid body conformational rearrangements involving reorientation of EIN\textsuperscript{ab} relative to EIC and of EIN\textsuperscript{a} relative to EIN\textsuperscript{ab}.

RDCs measured on the EIN domain of intact EI (both free and complexed to HPr) are in excellent agreement with the orientation of the EIN\textsuperscript{a} and EIN\textsuperscript{ab} subdomains in the isolated EIN domain, but are inconsistent with that in the crystal structure of the EI phosphorylated intermediate. Thus, one can conclude unambiguously that the relative orientation of the EIN\textsuperscript{a} and EIN\textsuperscript{ab} subdomains in free and HPr-bound full-length EI remains unchanged relative to that in the isolated EIN domain [40].

The hybrid approach used to solve the structure of free EI and the EI–HPr complex made combined use of RDCs and SAXS/WAXS. The RDCs restrain the relative orientations of the EIN\textsuperscript{a} and EIN\textsuperscript{ab} subdomains within each subunit and of the two symmetry-related EIN domains in the dimer. The structure of the EIC domain dimer is the same in several EI and isolated EIC structures [33–35,67,68]; therefore the orientation of the symmetry-related EIN domains relative to the EIC dimer can be determined from RDCs located only in the EIN domain, as one of the principal axes of the alignment tensor must coincide with the \(C_2\) symmetry axis of the dimer (since the RDCs for the two identical subunits are the same). RDCs alone are not sufficient to determine a unique structure, owing to the existence of several equivalent solutions arising from the intrinsic equivalence of 180° rotations about the axes of the RDC alignment tensor. However, when the RDCs are combined with stereochemical and covalent geometry restraints within the linker connecting the EIN and EIC domains together with shape and translational information contained within the SAXS/WAXS profiles, only a single solution emerges from RDC and SAXS/WAXS-driven simulated annealing calculations, in which the domains are treated as rigid bodies and only the linker (residues 255–261) is allowed to vary in conformation [40,41].

The RDCs and SAXS/WAXS profiles [40] do not agree with the crystal structures of phosphorylated \textit{E. coli} EI (Figure 4B, left panel) [33] or free \textit{Staphylococcus aureus} [35] and \textit{Staphylococcus carnosus} [34] EI. The \(\chi^2\) for the fits to the SAXS/WAXS data (from scattering vector values of \(q = 0.014–0.44\ \text{Å}^{-1}\)) are 128, 56, and 30 for the three crystal structures, respectively. The same is true of small angle neutron scattering (SANS) with \(\chi^2\) values of 62, 54, and 30, respectively. Moreover, the calculated Svedberg sedimentation coefficient for phosphorylated EI (\(S = 6.74\)) and free \textit{S. aureus} EI (\(S = 6.45\)) are significantly larger than the experimental value (\(S = 5.68\)) determined by sedimentation velocity, indicating that the structures are too compact, whereas that for \textit{S. carnosus} EI (\(S = 5.55\)) is too small [40], reflecting an overly expanded structure. In addition, the structure of the EIN domain in the \textit{S. carnosus} crystal structure is partially disordered.

The solution structures of \textit{E. coli} free EI and EI complexed to HPr determined by conjoined rigid body/torsion angle/Cartesian simulated annealing driven by RDCs and SAXS/WAXS are shown in Figure 4B [40]. Both structures are independently validated by agreement to both WAXS at high \(q (0.44–0.8\ \text{Å}^{-1})\) and SANS. The latter provides an independent validation tool for the EI–HPr complex as contrast-matching was used to render the HPr component invisible (by using a complex of deuterated EI and protonated HPr in 40.4% D\textsubscript{2}O).
The relative orientation of EIN relative to EIC in free EI and the EI–HPr complex are not identical (Figure 4B, right panel). For free EI, the EIN domain makes extensive contacts (~300 Å² of buried accessible surface area) with the EIC domain, with several complementary charge–charge interactions. In this configuration, HPr bound to the EINα subdomain would partially overlap with the EIC domain. Binding of HPr to the EINα subdomain is accompanied by additional movement of the EIN domain away from the EIC domain to make room for HPr, such that HPr is sandwiched between the EIN and EIC domains. There are a small number of contacts between HPr and EIC involving several complementary electrostatic interactions with a buried accessible surface area of ~100 Å². The conformational transition of the EIN domain from the free EI configuration to that of the EI–HPr configuration is achieved by only minor changes in backbone torsion angles within the linker connecting the EINαβ subdomain to the EIC domain.

Figure 5 depicts the postulated catalytic cycle for EI [40]. The change in configuration of the EIN domain from that of the phosphorylated EI intermediate to that of free EI and the EI–HPr complex must involve two sequential or concerted conformational transitions. First, a ~70° reorientation of EINαβ relative to EIC, accompanied by a ~34 Å rms displacement of the EINαβ subdomain. Second, a ~70° reorientation of EINα relative to EINαβ through backbone torsion angle changes within the two linkers joining the EINα to EINαβ subdomains. The latter cannot occur first because EINα would clash with EIC in the absence of any accompanying reorientation of EINαβ.

Although the SAXS/WAXS, SANS and RDC data for EI and the EI–HPr complex can be satisfied by a single structure, it is likely that interdomain motions and sparsely
populated states are present. However, because the RDC and X-ray scattering data are sensitive to different types of motion, the former being dependent on orientation and the latter on molecular size and shape, one can be confident that the calculated structures are representative of the predominant average structure in solution.

**Interplay between conformational dynamics and dimerization of the EIC domain**

Binding of PEP and Mg$^{2+}$ to EI results in a ~30-fold decrease in the equilibrium dissociation constant ($K_{dimer}$) for the dimer [69,70]. The decrease in $K_{dimer}$ upon ligand binding is thought to play an important role in regulating the PTS because only intact dimeric EI can be auto-photophosphorylated by PEP [70]. Binding of PEP to the dimeric isolated EIC domain results in large $^1$H/$^15$N chemical shift perturbations around the PEP binding site and in the adjacent $\beta3\alpha3$ turn located at the dimer interface [71].

Examination of EIC in various crystal structures [33–35,67,68] suggests that the $\beta3\alpha3$ turn exists in open and closed conformations, with the latter corresponding to the conformation in the phosphorylated EI intermediate (Figure 6). $^{15}$N relaxation dispersion measurements reveal the existence of a dynamic equilibrium between major (97%) and minor (3%) species with an overall exchange rate of $\sim$1550 s$^{-1}$ [71]. The backbone $^{15}$N chemical shift differences between the major and minor species in free EIC determined from the relaxation dispersion data correspond with the chemical shift differences observed upon addition of PEP [71]. Thus, it is likely that binding of PEP occurs via conformational selection of a low-population state corresponding to the closed state of the $\beta3\alpha3$ turn.

**Encounter complexes in the PTS**

Specific protein–protein recognition generally proceeds via a two-step process involving the initial formation of an
ensemble of short-lived encounter complexes via diffusion-controlled intermolecular collisions, followed by translations and rotations of the two partner proteins down a 2D funnel-like energy landscape, resulting in the formation of a well-defined specific complex stabilized by a complementary set of electrostatic and van der Waals interactions [72,73]. Encounter complexes are thought to play an important functional role in fine-tuning reaction fluxes inside the cell [74] by enhancing association on-rates through an increase in the interaction cross-section and a reduction in the conformational search space on the path to the specific complex [75–78]. Encounter complexes are generally extremely difficult to study experimentally because they are short-lived, highly transient, and sparsely populated, and therefore invisible to conventional structural and biophysical methods.

Transient, sparsely populated states, however, can be studied by NMR using paramagnetic relaxation enhancement (PRE) (Box 3). A comparison of the experimental PRE profiles for the EIN–HPr complex with those calculated from the structure of the specific EIN–HPr complex is shown in Figure 7A [79]. The paramagnetic tag (EDTA-Mn⁵⁺) that gives rise to the PRE is covalently attached to three sites (E5C, E25C, and E32C) on HPr (at natural isotopic abundance) using surface engineered cysteines, whereas EIN is uniformly labeled with ¹⁵N. This isotopic labeling scheme permits one to detect specifically intermolecular PRE effects originating from the paramagnetic label attached to HPr on the backbone amide resonances of EIN. It can be seen that although there are features of the experimental intermolecular PRE profiles (black circles) that match the calculated profiles for the specific complex, there are regions where the observed intermolecular PREs (purple circles) are larger than the calculated values. Thus, there are regions on EIN that spend a small proportion of time closer to the paramagnetic tags on HPr than in the specific complex. Similar observations have been made for complexes of HPr with IIA²⁺ and IIA³⁺ [80].

The discrepancies between observed and calculated intermolecular PREs can be fully accounted for by the presence of a small population (5–10%) of transient encounter complexes whose distribution can be calculated using PRE-driven rigid body ensemble simulated annealing [80] (Box 2). The calculated distribution of HPr on the surface of EIN is largely correlated with surface electrostatics, which has been confirmed experimentally by the observation that the intermolecular PREs arising from the encounter complexes are significantly more sensitive to ionic strength than those from the specific complex [81].

Further analysis of the intermolecular PRE intensities as a function of added paramagnetically labeled HPr has yielded further insights into the nature of the encounter complexes [79]. The PREs attributable to the specific
**Box 3. PRE and the detection of sparsely populated states**

PRE is an NMR technique that involves covalently attaching a paramagnetic tag (such as a nitroxide or EDTA-Mn\(^{2+}\)) to an engineered surface exposed cysteine residue \([80,99,100]\). The PRE effect (measured by the difference in \(^1\)H transverse relaxation rates between samples with and without the paramagnetic tag) is proportional to the \(-r^3\) separation between the paramagnetic tag and the protons of interest. \(^1\)H transverse relaxation rates reflect the linewidths of the \(^1\)H resonances in the NMR spectrum, the larger the \(^1\)H transverse relaxation rate, the broader the resonances. The magnetic moment of an unpaired electron is large, and therefore, the PRE at short distances is correspondingly large. For example, for a 20–30-kDa molecule, the transverse PRE rate \(T_2\) for a 30-Å distance would be \(-2\) s\(^{-1}\), whereas that for an 8-Å distance would be \(-6000\) s\(^{-1}\). In a system comprising two species, one populated at 99% with a distance of 30 Å, the other at 1% with a distance of 8 Å, that exchange fast on the PRE time scale (defined as an exchange rate larger than the difference in PRE rates for the two species), the PREs observed on the spectrum of the major species will be a population weighted average of the PRE rates for the major and minor species; in this particular example, the PRE rate measured on the spectrum of the major species would therefore be around 50–80 s\(^{-1}\). As a result, the imprint of the invisible minor species will be reflected in the PRE profiles measured on the spectrum of the major species, providing there are paramagnetic tag-proton distances that are shorter in the minor species than the major one.

Complex titrate as a simple one-site binding isotherm with a \(K_d\) of \(~7\) µM, in agreement with the \(K_d\) determined by isothermal titration calorimetry. The PREs arising from the encounter complexes, however, titrate as three classes (Figure 7B). Class I PREs display the same titration behavior as the intermolecular PREs arising from the specific complex. Thus, these intermolecular PREs arise from encounter complexes that are exclusive with the specific complex (i.e., the binding sites must overlap such that the specific complex and the class I encounter complexes cannot occur together). Class II PREs follow the concentration of free HPr, and therefore arise from encounter complexes that are nonexclusive with the specific complex (i.e., class II encounter complexes coexist with the specific complex). Finally, class III PREs exhibit a mixture of class I and II behavior. When the three classes of encounter complex PREs are mapped onto the surface of EIN, it is clear that class I PREs report on encounter complexes interactions near the active site of EIN that are occluded by the specific complex (Figure 7C). The class II PREs report on ternary HPr\(_{\text{non-specific}}$/EIN/HPr complexes that predominate when the active site is occupied by the specific complex (Figure 7C). These finding can be summarized by the equilibrium binding model shown in Figure 7D.

---

**Figure 7.** Characterization of transient sparsely populated encounter complexes for the interaction of EIN and HPr. (A) Comparison of experimental backbone amide intermolecular PREs \(^{15}\)N-labeled EIN and arising from covalently attached paramagnetic tags (EDTA-Mn\(^{2+}\)) located at two positions on HPr (ESC and E32C) with the PRE profiles calculated from the structure of the specific complex (black line). Black and purple circles indicate PREs attributable to the specific complex and to an ensemble of encounter complexes, respectively. (B) Intermolecular PREs as a function of added paramagnetically-labeled HPr(ESC) illustrating three types of titration behavior. (C) Mapping of intermolecular PREs attributable to the specific complex (black) and to the encounter complexes (class I, blue; class II, red; mixture of classes I and II, purple) and encounter complex PREs that are too large to measure accurately (pink). (D) Equilibrium binding model for the EIN/HPr association pathway. Adapted from \([79]\). Abbreviations: EIN, N-terminal domain of Enzyme I; HPr, histidine phosphocarrier protein; PRE, paramagnetic relaxation enhancement.
What is the possible role of the transient HPr\textsubscript{non-specific}/EIN/HPr ternary complexes? Is it possible that these ternary complexes may help EI compete for the cellular pool of HPr, even while phosphotransfer is occurring at the EIN active site, thereby facilitating higher rates of sugar uptake when substrate is transiently abundant [79]? In other words, the ternary complexes may be important for efficiently reloading the EIN active site with HPr when demand for sugar transport is high. Given the intracellular concentrations of EI and HPr (5 and 20–100 μM, respectively), one can estimate that the population of ternary complex ensembles in vivo is ~1%. This estimate may be revised higher if intracellular crowding and compartmentalization further favors the formation of transient ternary complexes.

The II\textsubscript{AMan}−II\textsubscript{BMan} interaction provides an extreme example of a relatively long-lived, highly-populated (~50% occupancy) and well-defined encounter complex [20], as opposed to the invisible sparsely populated ensemble of transient states seen for the complexes of HPr with EI, II\textsubscript{AMan}, and II\textsubscript{AMan} [79,80]. In the case of II\textsubscript{AMan}−II\textsubscript{BMan}, the intermolecular NOE data report on a mixture of two species comprising a productive, phosphotransfer competent complex (Figure 8, left panel) and a nonproductive complex in which the two active site histidines (His10 of II\textsubscript{AMan} and His175 of II\textsubscript{BMan}) are separated by 25 Å (Figure 8, right panel). The productive and nonproductive complexes are in fast exchange with one another so that only a single set of resonances are observed, but the observed intermolecular NOEs are inconsistent with the existence of a single complex. The structural transition between productive and nonproductive complexes involves a 90° rotation coupled to a ~37 Å translation of II\textsubscript{BMan} relative to II\textsubscript{AMan}. The interaction surface, however, on II\textsubscript{AMan} for the nonproductive complex comprises a subset of residues located in the central region of the interface in the productive complex. Likewise, the interaction surfaces for the productive and nonproductive complexes on II\textsubscript{BMan} also partially overlap. Thus, the productive and nonproductive complexes are mutually exclusive, as are both of these complexes with the upstream II\textsubscript{AMan}−HPr complex.

The nonproductive II\textsubscript{AMan}−II\textsubscript{BMan} complex (or more accurately intermolecular NOEs attributable to the nonproductive complex) can be largely eliminated by introducing a phosphomimetic H10E mutation in II\textsubscript{AMan}. The equivalent phosphomimetic His175E in II\textsubscript{BMan}, however, has no effect. The selection of the productive complex by II\textsubscript{AMan}(H10E) can be attributed to charge neutralization through interaction of the negatively charged carboxylate of H10E (equivalent to phosphorylated His10) with the positively charged guanidino group of Arg172 located at the center of the II\textsubscript{BMan} binding surface [20]. This is supported by mutation of Arg172 to Gin which results in a substantial increase in the population of the productive complex; the nonproductive complex, however, is not completely eliminated as the mutation still leaves an unfavorable polar residue in the middle of the interface. These observations are consistent with the finding that II\textsubscript{BMan}(R172Q) is less efficiently phosphorylated than the wild type by II\textsubscript{AMan} [82].

What is the role of the nonproductive II\textsubscript{AMan}−II\textsubscript{BMan} complex? It could represent an extreme example of an encounter complex, where the nonproductive complex could facilitate formation of the specific complex in a similar manner as highly transient, diffuse encounter complex ensembles.

**Concluding remarks**

The complexes of the PTS provide a paradigm for studying protein–protein interactions and understanding the determinants of specificity in a multifaceted and complex signal transduction system that allows for interactions between many partner proteins. The complexes of the PTS, which range from ~30 to ~150 kDa, have also served as a framework for extending NMR methodology to larger and more complex systems, for establishing integrative hybrid
approaches combining RDCs and solution X-ray scattering to tackle systems in excess of 100 kDa, and for developing novel biophysical techniques based on NMR paramagnetic relaxation enhancement to uncover the existence of transient, spectroscopically invisible, sparsely populated encounter complexes that constitute the first step towards specific complex formation.

There remain two open questions/challenges that still need to be addressed with respect to the structural biology and biophysics of the PTS. The first relates to EI and the mechanism and dynamics of large interdomain structural rearrangements that must occur during the course of the catalytic cycle. How, for example, are small local structural changes in the EIC domain upon binding to PEP transmitted to the EIN domain? In the absence of PEP does the open free state coexist with a small population of spectroscopically invisible closed state (Figure 4), and if so, what is the population of the latter; what are the interconversion rates between the species; and how are these populations and rates modulated by different ligands (e.g., the substrate PEP versus the inhibitor α-ketoglutarate, both of which bind to the same site on EIC [28]). These questions can largely be addressed by multidimensional hetronuclear NMR spectroscopy, including the application of relaxation dispersion and paramagnetic relaxation enhancement measurements, but are rendered especially challenging owing to the large size (by NMR standards) of EI. The second area relates to high-resolution structures of the transmembrane sugar transporters (IIC and IID; Figure 1) of the PTS, their interaction with Enzymes IIB, and the mechanism of selective sugar transport across the membrane. Although there have been some low-resolution electron microscopy (EM) and cysteine crosslinking studies on the IIC transporters from the glucose [83,84] and mannitol [85,86] branches, this field remains largely an open book. Encouragingly crystals of IIC\textsuperscript{Glc} diffracting to 4.5 Å resolution have been recently obtained [87]. Solving structures of the transmembrane transporters will require extensive biochemical work to obtain either crystals diffracting to high resolution or suitable preparations for single molecule analysis by EM.

Acknowledgments

G.M.C. thanks members of his laboratory, past and present, and colleagues who have made major contributions to the work on complexes of the PTS, in particular, D. Garrett, M. Cai, G. Wang, G. Cornilesuc, D. Williams, J. Hu, K. Hu, J-Y. Suh, Y-S. Jung, C. Tang, J. Iwahara, Y. Takayama, A. Grishaev, and C. Schwieters. This work was supported by funds from the Intramural Program of the NIH, NIDDK, and the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the NIH (to G.M.C.).

References

13. Sliz, P. et al. (1997) The structure of enzyme IIA\textsuperscript{lactose} from Lactococcus lactis reveals a new fold and points to possible interactions of a multicomponent system. Structure 5, 775–788
30 Liao, D.I. et al. (1996) The first step in sugar transport: crystal structure of the amino terminal domain of enzyme I of the E. coli PEP phosphotransferase system: a model of the phosphotransfer complex with HPr. Structure 4, 861–872
34 Marquez, J. et al. (2006) Structure of the full-length enzyme I of the phosphoenolpyruvate-dependent sugar phosphotransferase system. J. Biol. Chem. 281, 32508–32515
39 Garrett, D.S. et al. (1997) Solution structure of the 30 kDa N-terminal domain of enzyme I of the Escherichia coli phosphoenolpyruvate:sugar phosphotransferase system by multidimensional NMR. Biochemistry 36, 2517–2530
45 Cornilescu, G. et al. (2002) Solution structure of the phosphoryl transfer complex between the cytoplasmic A domain of the mannitol transporter P₈M₉M₉ and HPr of the Escherichia coli phosphotransferase system. J. Biol. Chem. 277, 42289–42298
47 Wang, G. et al. (2000) Solution structure of the phosphotransfer complex between the signal transducing proteins HPr and IIAChlorella of the Escherichia coli phosphoenolpyruvate:sugar phosphotransferase system. EMBO J. 19, 5635–5649
48 Williams, D.C., Jr et al. (2005) Solution NMR structure of the 48-kDa IIAChlorella-HPr complex of the Escherichia coli mannose phosphotransferase system. J. Biol. Chem. 280, 20775–20784
50 Suh, J.Y. et al. (2006) Solution structure of a post-transition state analog of the phosphotransfer reaction between the A and B cytoplasmic domains of the mannitol transporter P₈M₉M₉ of the Escherichia coli phosphotransferase system. J. Biol. Chem. 281, 27510–27517
75 Northrup, S.H. et al. (1988) Brownian dynamics of cytochrome c and cytochrome c peroxidase association. Science 241, 67–70
89 Grishaev, A. et al. (2008) Refined solution structure of the 82-kDa enzyme malate synthase G from joint NMR and synchrotron SAXS restraints. J. Biomol. NMR 40, 95–106