THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 279, No. 37, Issue of September 10, pp. 39115–39121, 2004
Printed in U.S.A.

Three-dimensional Solution Structure of the Cytoplasmic B Domain of the Mannitol Transporter $II^{Mannitol}$ of the *Escherichia coli* Phosphotransferase System*

Received for publication, June 17, 2004
Published, JBC Papers in Press, July 15, 2004, DOI 10.1074/jbc.M406764200

Patricia M. Legler‡, Mengli Cai‡, Alan Peterkofsky§, and G. Marius Clore‡¶

From the ‡Laboratory of Chemical Physics, NIDDK, National Institutes of Health and the §Laboratory of Cell Biology, NHLBI, National Institutes of Health, Bethesda, Maryland 20892-0520

The solution structure of the cytoplasmic B domain of the mannitol (Mtl) transporter (II^{Mtl}) from the mannitol branch of the *Escherichia coli* phosphotransferase system has been solved by multidimensional NMR spectroscopy with extensive use of residual dipolar couplings. The ordered IIB^{Mtl} domain (residues 375–471 of II^{Mtl}) consists of a four-stranded parallel β-sheet flanked by two helices (α_1 and α_3) on one face and helix α_2 on the opposite face with a characteristic Rossmann fold comprising two right-handed $\beta_1\alpha_1\beta_2$ and $\beta_3\alpha_2\beta_4$ motifs. The active site loop is structurally very similar to that of the eukaryotic protein tyrosine phosphatases, with the active site cysteine (Cys-384) primed in the thiolate state ($pK_a \leq 5.6$) for nucleophilic attack at the phosphorylated histidine (His-554) of the IIA^{Mtl} domain through stabilization by hydrogen bonding interactions with neighboring backbone amide groups at positions i + 2/3/4 from Cys-384 and with the hydroxyl group of Ser-391 at position i + 7. Modeling of the phosphorylated state of IIB^{Mtl} suggests that the phosphoryl group can be readily stabilized by hydrogen bonding interactions with backbone amides in the i + 2/4/5/6/7 positions as well as with the hydroxyl group of Ser390 at position i + 6. Despite the absence of any significant sequence identity, the structure of IIB^{Mtl} is remarkably similar to the structures of bovine protein tyrosine phosphatase (which contains two long insertions relative to IIB^{Mtl}) and the cytoplasmic B component of enzyme II^{Chb}, which fulfills an analogous role to IIB^{Mtl} in the *N,N′*-diacetylchitobiose branch of the phosphotransferase system. All three proteins utilize a cysteine residue in the nucleophilic attack of a phosphoryl group covalently bound to another protein.

In bacteria, the translocation of sugars across the cytoplasmic membrane is coupled to concomitant phosphorylation, a process that is mediated by the phosphoenolpyruvate:sugar phosphotransferase system (PTS)[1] signal transduction pathway (1). The initial steps of the PTS are common to all sugars in that enzyme I transfers a phosphoryl group from phosphoenolpyruvate to the histidine phosphocarrier protein (HPr). Thereafter, the phosphoryl group is transferred from HPr to the sugar-specific enzymes II. There are four general classes of enzymes II (2–4), namely glucose, mannitol (Mtl), mannose, and lactose/chitobiose (Chb). The domain organization of the sugar-specific enzymes II are similar, and the domains may or may not be covalently linked to one another. There are two cytoplasmic domains, IIA and IIB, and a transmembrane IIC domain (and in some cases IID as well). IIA accepts the phosphoryl group from HPr and subsequently donates it to IIB; IIC catalyzes the translocation of the sugar across the cytoplasmic membrane and its phosphorylation by IIB. The IIA and IIB cytoplasmic domains for the different sugar classes bear no sequence similarity to one another and, in the case of those whose structures have been solved, no structural resemblance either (2–4). Crystal and NMR structures of the N-terminal domain of enzyme I (EIN) (5, 6) and HPr (7–11) have been determined. Crystal and/or NMR structures from a variety of species have been solved for representatives of the IIA domains from all four classes (12–16) and the IIB domains from three of the four classes (17–22), the exception being IIB^{Mtl}. In addition, four cytoplasmic protein-protein complexes of the PTS have been solved by NMR: EIN-HPr (23), IIA^{Glc}-HPr (24), IIA^{Glc}-IIB^{Glc} (22), and IIA^{Mtl}-HPr (25). As part of our continuing structural work on the PTS, we present the solution structure of IIB^{Mtl} using multidimensional NMR spectroscopy.

EXPERIMENTAL PROCEDURES

Expression and Purification of IIB^{mtl}—The mannitol transporter (II^{Mtl}) is composed of three domains expressed as a single protein containing, from the N- to the C-terminal end, the domains C-B-A (26–30). Although C is embedded in the membrane, B and A are cytoplasmic but attached to the inner membrane. The B domain contains the second phosphorylation site (Cys-384) of the II^{Mtl} complex. *Escherichia coli* chromosomal DNA was used as the template to amplify, by PCR, the region corresponding to the B domain together with the C-B linker (26–30). The forward PCR primer 5′-AAAGAAGAAGATGATCATATGGCAGCAACTCGTCGT-3′, spanning the region from bases 1155 to 1190 of the NCBI file V01503, contained an engineered NdeI site (underlined), and the reverse primer 5′-CGCGCCTAGGTCGACCAGGTT*TCATCA*GGAATCGTCAAAGCTGTC-3′, spanning the

* 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 1VKR) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

¶ To whom correspondence should be addressed: Laboratory of Chemical Physics, Bldg. 5, Rm. B1-30I, NIDDK, National Institutes of Health, Bethesda, MD 20892-0520. Tel.: 301-496-0782; Fax: 301-496-0825; E-mail: mariusc@intra.niddk.nih.gov.

[1] The abbreviations used are: PTS, phosphotransferase system; Mtl, mannitol; HPr, histidine-containing phosphocarrier protein, IIA^{Mtl}, IIB^{Mtl}, and IIC^{Mtl}, A, B, and C domains, respectively, of the mannitol-specific transporter II^{Mtl}; Chb, chitobiose; IIB^{Chb}, cytoplasmic B component of the *N,N′*-diacetylchitobiose-specific transporter II^{Chb} (also referred to in the literature as $II^{Cellobiose}$); BPTP, bovine protein tyrosine phosphatase; β-ME, β-mercaptoethanol; NOE, nuclear Overhauser effect; r.m.s., root mean square.

TABLE I
Structural statistics

Experimental restraints	⟨SA⟩[a]	$(\overline{SA})_r$[a]
R.m.s. deviations from experimental restraints[a]		
Distances (Å) (1554)	0.007 ± 0	0.009
Torsion angles (°) (214)	0.40 ± 0.05	0.47
$^{13}C\alpha/\beta$ chemical shifts (ppm) (189)	1.09 ± 0.02	1.09
$^{3}J_{HN\alpha}$ couplings (Hz) (70)	0.71 ± 0.03	0.70
Residual dipolar coupling *R*-factors[b] (%)		
$^{1}D_{NH}$ (74) (%)	2.6 ± 0.1	2.8
$^{1}D_{NC'}$ (62) (%)	14.9 ± 0.3	14.9
$^{2}D_{HNC'}$ (64) (%)	13.1 ± 0.5	13.5
Residues in most favorable region of Ramachandran map[c] (%)	90.4 ± 1.5	91
Number of bad contacts per 100 residues	5.7 ± 1.4	4.0
Deviations from idealized covalent geometry[c]		
Bonds (Å)	0.002 ± 0	0.004
Angles (°)	0.53 ± 0.01	0.69
Improper torsion (°)	0.40 ± 0.05	1.16
Coordinate precision[d] (Å)		
Backbone (residues 375–471)		0.29
All atoms (residues 375–471)		0.89

[a] ⟨SA⟩ are the final 80 simulated annealing structures. $(\overline{SA})_r$ is the restrained regularized mean structure derived from the mean coordinates obtained by averaging the coordinates of the 80 stimulated annealing structures best fitted to each other (with respect to residues 375–471). The number of terms for the various restraints is given in parentheses. None of the structures exhibits interproton distance violations >0.3 Å or torsion angle violations >5°. There are 1444 structurally useful interproton distance restraints comprising 419 sequential ($|i - j| = 1$), 347 medium range ($1 < |i - j| \leq 5$), and 517 long range ($|i - j| > 5$) restraints, 50 sequential/medium range and 21 long range ambiguous restrains (represented by a $(\Sigma r^{-6})^{-1/6}$ sum), and 90 intraresidue restraints. 110 distance restraints for 55 backbone hydrogen bonds located in helices and sheets were added during the final stages of refinement. The torsion angle restraints comprise 97 ϕ, 94 ψ, and 68 χ restraints.

[b] The dipolar coupling *R*-factor, which scales between 0 and 100%, 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 distributed as given by $[2D_a^2(4 + 3\eta^2)/5]^{1/2}$, where D_a is the magnitude of the principal component of the alignment tensor and η the rhombicity (75). The values of D_a^{NH} and η, derived from the distribution of normalized dipolar couplings (76), are −15.8 Hz and 0.02, respectively.

[c] Calculated with the program PROCHECK (77). The dihedral angle G factors for ϕ/ψ, χ_1/χ_2, χ_1, and χ_3/χ_4 are 0.17 ± 0.03, 0.63 ± 0.09, −0.17 ± 0.12, and 0.12 ± 0.17, respectively. The WHATIF first and second generation packing scores (78) are −0.38 and 0.14, respectively; values larger than −0.5 and 0, respectively, are considered to represent high quality structures.

[d] The precision of the coordinates is defined as the average atomic r.m.s. difference between the individual 80 simulated annealing structures and the corresponding mean coordinates obtained by best fitting residues 375–471. Residues 367–374 and 472–489 are disordered in solution.

region from bases 1631 to 1587, contained an engineered SalI site (underlined) and a tandem pair of in-frame termination codons (italics). This PCR product was designed to express a peptide encoding amino acid residues Ala-349 to Ser-489 of II^{Mtl} preceded by a Met residue. The NdeI- and SalI-cut PCR product was purified and cloned into the corresponding sites of the vector pRE1 (25). The selected clone was verified by DNA sequencing. A fragment of the original clone in pRE1 encoding residues 366 to 489, which corresponds to the IIB^{Mtl} domain, was then subcloned into a modified pET-32-a vector (Novagen, Inc., Madison, Wisconsin) to form a thioredoxin fusion protein with a His_6 tag. The original thrombin cleavage site was mutated, and a new thrombin cut site was inserted at the N-terminal end of the protein to remove the thioredoxin and the His_6 tag. The new thrombin cleavage site added a glycine residue before Ser-366. The selected clone was verified by DNA sequencing. The IIB^{mtl}-fusion protein was expressed at 37 °C in *E. coli* BL-21(DE3) cells in minimal medium using $^{15}NH_4Cl$ and [$^{13}C_6$]glucose as the sole nitrogen and carbon sources, respectively.

The cells were microfluidized in 50 mM Tris-HCl, pH 7.5, 200 mM NaCl, and 2 mM β-mercaptoethanol (β-ME), and cell debris was removed by centrifugation at 10,000 × *g* for 30 min. The crude extract was loaded onto a 15-ml column of chelating Sepharose charged with nickel sulfate. The column was washed with 50 mM Tris HCl, pH 7.5, 200 mM NaCl, 5 mM imidazole, and 2 mM β-ME. The protein was eluted with the same buffer containing 0.7 M imidazole. Thrombin (25 units) (Novagen Inc.) was added to the IIB^{mtl}-thioredoxin fusion protein, and the mixture was dialyzed overnight in 50 mM Tris-HCl, pH 7.5, 200 mM NaCl, and 2 mM β-ME at room temperature to remove the imidazole and allow thrombin cleavage of the thioredoxin and His_6 tags. A suspension (1 ml) of Chelex-100 in 40 mM sodium phosphate buffer, pH 7, was added to the dialysis buffer to remove excess nickel. After overnight digestion, the protein was passed over a 3-ml column of benzamidine-Sepharose (Amersham Biosciences) to remove the thrombin. A second nickel column removed the His_6-thioredoxin cleavage product. The combined column fractions were concentrated and loaded onto a G-75 Sephadex column (50 mM Tris-HCl, pH 7.5, 200 mM NaCl, 2 mM EDTA, and 5 mM β-ME). One peak eluted with an apparent molecular mass of ~14 kDa. The fractions were collected and concentrated using Centricon YM-3 (3000 molecular weight cutoff) ultrafiltration units to a final volume of ~2 ml. To dephosphorylate IIB^{Mtl}, the pH was lowered to 2 in the presence of 6 M guanidine hydrochloride, 20 mM sodium phosphate buffer, and 2 mM β-ME. The denatured protein was incubated in a water bath for 2 h at 37 °C. The pH was raised to 7.5, and the protein was refolded by dialysis in 200 mM sodium phosphate buffer, pH 7.5, overnight at 4 °C. The buffer was exchanged to 40 mM sodium phosphate, pH 7.1, 25 mM dithiothreitol, and either 95% H_2O/5% D_2O or 99.996% D_2O by ultrafiltration for NMR experiments. The protein concentration in the NMR samples varied from 0.7 to 1.5 mM. The composition of the purified IIB^{Mtl} domain was verified by mass spectrometry.

NMR Spectroscopy—All spectra were recorded at 35 °C on Bruker DMX500, DMX600, DRX600, DMX750, and DRX800 spectrometers equipped either with *x*, *y*, *z*-shielded gradient triple resonance probes or *z*-shielded gradient triple resonance cryoprobes. Spectra were processed with the NMRPipe package (31) and analyzed using the programs PIPP, CAPP, and STAPP (32). Sequential resonance assignments were derived from analysis of the following double and triple resonance three-dimensional NMR experiments (33–35): three-dimensional HNCO, HNCACB, CBCA(CO)NH, HBHA(CBCACO)NH, H(CCO)NH, C(CCO)NH, and HCCH-total correlation spectra. NOE interproton distance restraints were derived from three-dimensional ^{15}N-, ^{13}C-, and $^{13}C/^{13}C$-separated NOE and four-dimensional $^{13}C/^{13}C$-separated NOE experiments. Three-bond *J* couplings ($^{3}J_{HNH\alpha}$, $^{3}J_{NC\gamma}$, and $^{3}J_{C'\text{-}C\gamma}$ couplings) were measured using quantitative *J* correlation spectroscopy (36). Stereospecific assignment of valine and leucine methyl groups were derived from an analysis of cross-peak multiplet patterns in a high resolution $^{1}H\text{-}^{13}C$ correlation spectrum recorded on a fractionally (10%) ^{13}C-labeled sample of IIB^{mtl} (37).

Residual dipolar couplings were measured in a liquid crystalline medium of 5% C12E5 polyethylene glycol/hexanol (with a surfactant to alcohol ratio of 0.68) (38) in 40 mM sodium phosphate buffer, pH 7.1, 25 mM dithiothreitol, and 7% D_2O at 30 °C. $^{1}D_{NH}$, $^{1}D_{NC'}$, and $^{1}D_{HNC}$ dipolar couplings were obtained from the difference between the *J* couplings measured in the aligned and isotropic (water) media. $^{1}J_{NH}$ couplings were measured from a two-dimensional in-phase/anti-phase $^{1}H\text{-}^{15}N$ heteronuclear single quantum coherence spectrum (HSQC) (39), and $^{1}J_{NC'}$ and $^{2}J_{HNC'}$ couplings were measured from a carbonyl coupled $^{1}H\text{-}^{15}N$ HSQC (40).

Structure Calculations—Interproton distance restraints derived from the multidimensional NOE spectra were classified into the distance ranges 1.8–2.9 Å, 1.8–3.5 Å, 1.8–5 Å, and 1.8–6 Å corresponding

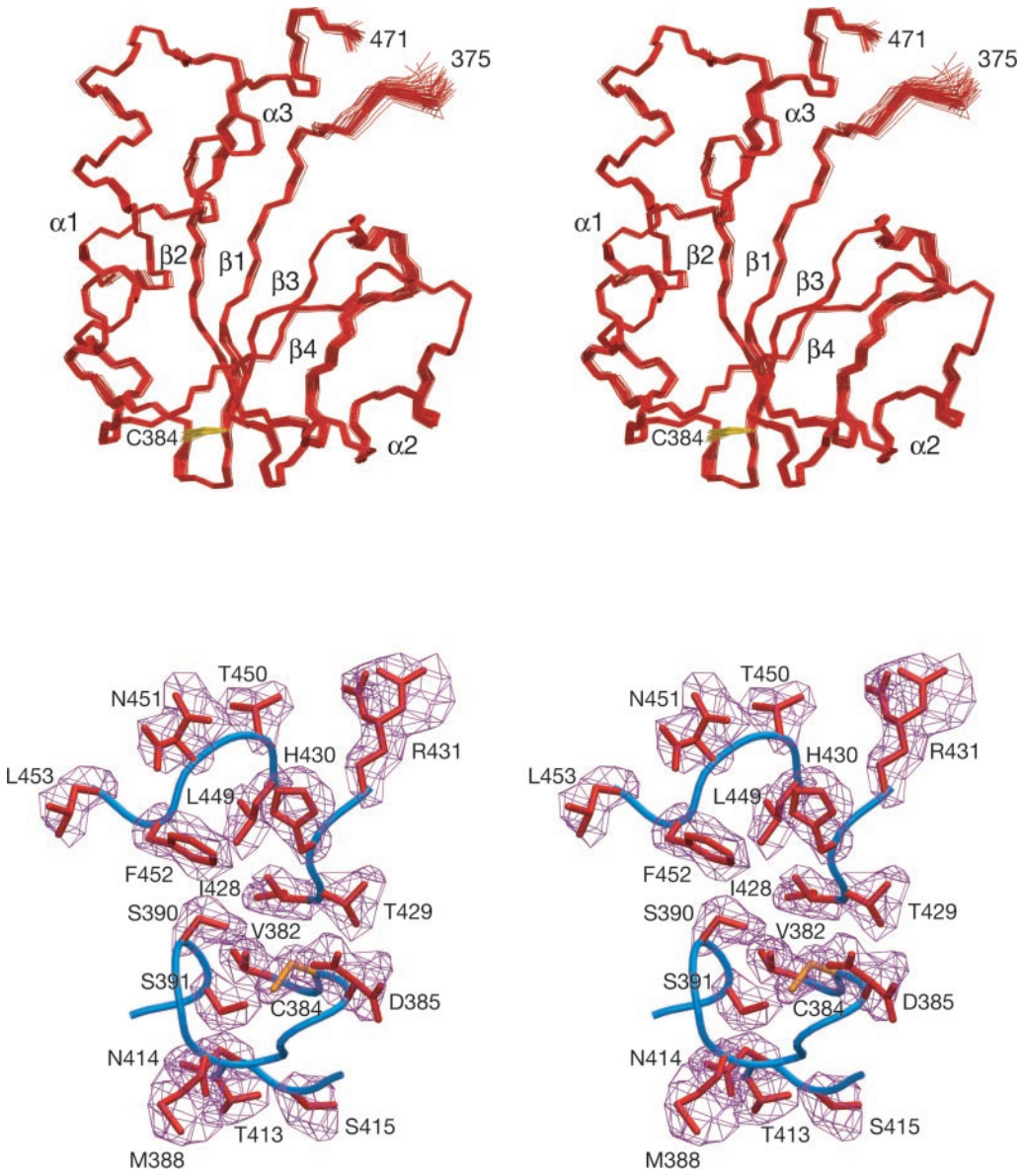

FIG. 1. **The structure of IIB^{Mtl}.** *A*, stereo view of a superposition of the backbone of the final 80 simulated annealing structures. Only residues 375–471 are shown, because residues 367–374 and 472–489 are disordered in solution. *B*, isosurface of the reweighted atomic density map (*purple*) for selected side chains drawn at a value of 20% maximum (53) calculated from the 80 simulated annealing structures. The backbone of the restrained regularized mean structure is shown as a *blue tube*, and side-chain coordinates within the atomic density map are shown in *red*. Note that some side chains are in multiple rotamer conformations (*e.g.* Asp-385, Arg-431, and Asn-451). Single-letter amino acid abbreviations are used with position numbers.

to strong, medium, weak, and very weak NOEs, with an additional 0.5 Å added to the upper distance bounds of NOEs involving methyl groups; non-stereospecifically assigned protons were represented by a $(\Sigma r^{-6})^{-1/6}$ sum (35). Backbone ϕ/ψ torsion angle restraints were derived from backbone chemical shifts using the program TALOS (41). Side-chain torsion angle restraints were derived from ^{3}J heteronuclear couplings coupled with analysis of a short mixing time (τ_m = 35 ms), three-dimensional, ^{13}C-separated NOE spectrum and a three-dimensional ^{15}N-separated rotating frame Overhauser effect spectrum.

The structures were calculated using well established procedures from the experimental restraints (42–45) by simulated annealing in torsion angle space (46) using the program Xplor-NIH (47). The non-bonded contacts in the target function were represented by a quartic van der Waals repulsion term (35, 42) supplemented by torsion angle (48) and hydrogen-bonding (49) data base potentials of mean force and a radius of gyration restraint to ensure optimal packing (50). Structure figures were generated with the programs VMD-XPLOR (51) and RIBBONS (52). Reweighted atomic density probability maps were calculated from the ensemble of simulated annealing structures as described previously (53). Optimal structural superpositions to structurally related proteins were carried out using the program O (54).

RESULTS AND DISCUSSION

Structure Determination—The solution structure of IIB^{Mtl} (residues 366–489 of II^{Mtl}) was solved on the basis of 2227 experimental NMR restraints, including 200 residual dipolar couplings. The N- (residues 367–374) and C termini (residues 472–489) are disordered as judged by heteronuclear ^{15}N-^{1}H NOE data (values < 0.4), as well as the absence of any non-sequential interproton NOE interactions. A summary of the structural statistics for the ordered IIB^{Mtl} domain (residues 375–471) is provided in Table I, and a superposition of the backbone atoms for the ensemble of 80 simulated annealing structures and an atomic density probability map representation of selected side chains derived from the ensemble are shown in Fig. 1.

Overall Structure—A ribbon diagram of IIB^{Mtl} is shown in Fig. 2*A*. IIB^{Mtl} is an α/β protein comprising a central four-stranded parallel β-sheet (β1, residues 379–382; β2, residues

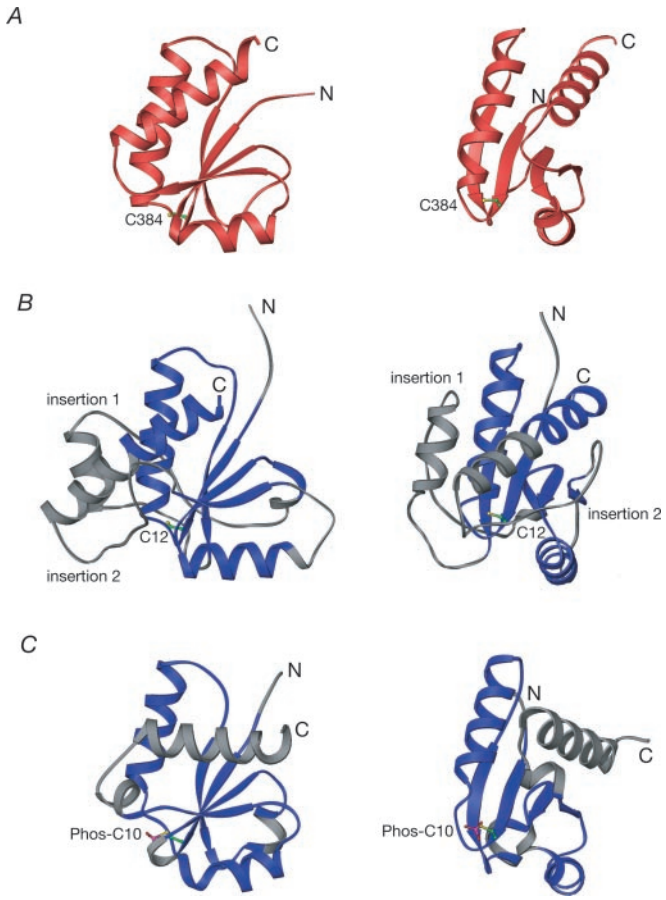

FIG. 2. **Comparison of the polypeptide folds of IIB^{Mtl}, BPTP, and IIB^{Chb}.** Ribbon diagrams showing two approximately orthogonal views of IIB^{Mtl} (*red*) (*A*), BPTP (*B*), and IIB^{Chb} (*C*). In *panels B* and *C* the regions of the polypeptide of BPTP and IIB^{Chb}, respectively, that superimpose on IIB^{Mtl} are shown in *blue* with the remainder in *gray*. Also shown in each case is the location of the active site cysteine. Single-letter amino acid abbreviations are used. For IIB^{Mtl} and BPTP (PDB accession code 1DG9) (64), the C_α atoms of 81 residues (sequence identity 12.3%) superimpose with an r.m.s. difference of 1.9 Å; for IIB^{Mtl} and IIB^{Chb} (PDB accession code 1H9C) (19), the C_α atoms of 74 residues (sequence identity 8.1%) superimpose with an r.m.s. difference of 2.0 Å. (The structural alignments are as follows: residues 378–411, 411–417, 422–429, 429–444, 444–450, and 455–466 of IIB^{Mtl} superimpose on residues 6–39, 40–46, 83–90, 91–103, 111–117, and 146–157, respectively of BPTP; residues 378–385, 386–408, 409–420, 421–430, and 431–451 of IIB^{Mtl} superimpose on residues 4–11, 11–33, 33–44, 48–57, and 61–81, respectively of IIB^{Chb}).

410–414; β3, residues 426–430, β4, residues 445–449) in a 1*x*, −2*x*, −1*x* topology, flanked on one face by helices α1 (residues 389–393) and α3 (residues 455–470) and on the other face by helix α2 (residues 431–440). Helices α1 and α3 are separated by a center-to-center distance of ~11 Å and oriented at an angle of ~45°. The overall architecture is characteristic of the dinucleotide binding (Rossmann) fold (55), with two right-handed βαβ motifs (β1-α1-β2 and β3-α2-β4) contributing to the central four-stranded β-sheet. The active site extends from the C terminus of strand β1 to the N terminus of helix α1 and comprises residues 384–391 with the active site cysteine located at position 384 (56, 57).

A stereo view of the active site of IIB^{Mtl} is shown in Fig. 3*A*. The geometry adopted by the active site loop (CDAGMGSS) is very similar to that of the phosphate binding P-loop (consensus sequence C*XXXXX*R(S/T)) of both the high (58–60) and low (61–64) molecular weight protein tyrosine phosphatases with C_α atomic r.m.s. differences of 0.6 Å (*versus* Protein Data Bank accession code 1A5Y; P-loop sequence CSAGIGRS) and 0.9 Å (*versus* Protein Data Bank accession code 1DG9; P-loop sequence CLGNICRS), respectively. The characteristic feature of these P-loops is the presence of a residue in the left-hand region of the Ramachandran plot at position i + 3 from the cysteine. As a consequence, the NH protons of residues 386, 387, 388, 390, and 391 are directed inwards, and the NH protons of Ala-386, Gly-387, and Met-388, as well as the hydroxyl group of Ser-391, are within hydrogen bonding distance of the sulfur atom (Fig. 3*A*). An equivalent set of interactions is seen in the protein tyrosine phosphatases (58–64). These interactions stabilize the thiolate state of Cys384, consistent with a low pK_a of ≤ 5.6 as determined by monitoring the ^{1}H-^{15}N cross-peak position of Ala-383 as a function of pH (which was easy to follow because of the absence of any chemical shift overlap).

The structure of the active site of IIB^{Mtl} suggests the following mechanism of phosphoryl transfer from IIA^{Mtl}. The phosphorylated His-554 of IIA^{Mtl} (26) undergoes a nucleophilic attack by the thiolate anion of Cys-384 of IIB^{Mtl}. The position of the side chain of Cys-384 is optimally positioned in a *g*+ conformation by the hydrogen bonding interactions described above. Modeling of a phosphorylated cysteine (by simply adding a phosphoryl group using standard geometry) indicates that the phosphoryl group can be readily stabilized by hydrogen bonding interactions with the backbone amides of Ala-386, Met-388, Gly-389, Ser-390, and Ser-391 and the hydroxyl group of Ser-390 (Fig. 3*B*). The hydroxyl group of Ser-390 may thus, in part, fulfill a similar structural role for the guanidino group of the equivalent conserved arginine residue in the protein tyrosine phosphatases. In addition, the helix dipole of helix α1 provides a favorable environment to accommodate the negatively charged phosphoryl group. In the subsequent transfer of the phosphoryl group from Cys-384 to mannitol bound to IIC^{Mtl}, Asp-385 may possibly act as a general base abstracting the hydroxyl proton from the terminal hydroxyl group of mannitol. The carboxylate group of Asp-385 may also simply serve to optimally position the mannitol via electrostatic interactions with one or more hydroxyl groups.

Both the low and high molecular weight eukaryotic tyrosine phosphatases have a conserved arginine residue at position i + 6 from the active site cysteine. It is thought that this arginine residue, in addition to its role in substrate binding through hydrogen bonding and electrostatic interactions with the phosphoryl group, may also act as an acid catalyst by neutralizing some of the negative charge surrounding the phosphate (65, 66). Although Ser-390 in IIB^{Mtl} cannot act as an acid catalyst, it is possible that in the context of the phosphorylated IIA^{Mtl}-IIB^{Mtl} complex this role may be taken over by Arg-538 of IIA^{Mtl}. The latter has been shown to be in a position to directly interact with the phosphoryl group in the context of the IIA^{Mtl}-HPr complex (25). Moreover, the absence of an acid catalyst in the active site of IIB^{Mtl} may serve a dual purpose, *i.e.* to inhibit nonspecific phosphorylation of IIB^{Mtl} as well as to inhibit hydrolysis of the phosphorylated IIB^{Mtl}.

Domain Organization of II^{Mtl}—In the case of II^{Mtl}, the three domains, C (at the N terminus), B, and A (at the C terminus), are present in a single polypeptide chain connected by flexible linkers. The transmembrane C domain extends from residues 1–337 (29), the cytoplasmic B domain from residues 375–461 (current study), and the cytoplasmic A domain from residues 494–637 (16). Thus, the C-B linker is 38 residues long, and the B-A linker is 33 residues in length. Intact II^{Mtl}, both *in vitro* and *in vivo*, is dimeric with the interface, being provided by the C domain (67–70). Complementation analysis of heterodimers containing an inactive partner indicates that phosphoryl transfer from the A domain to the B domain and from the B domain to the C domain is predominantly intramolecular rather than

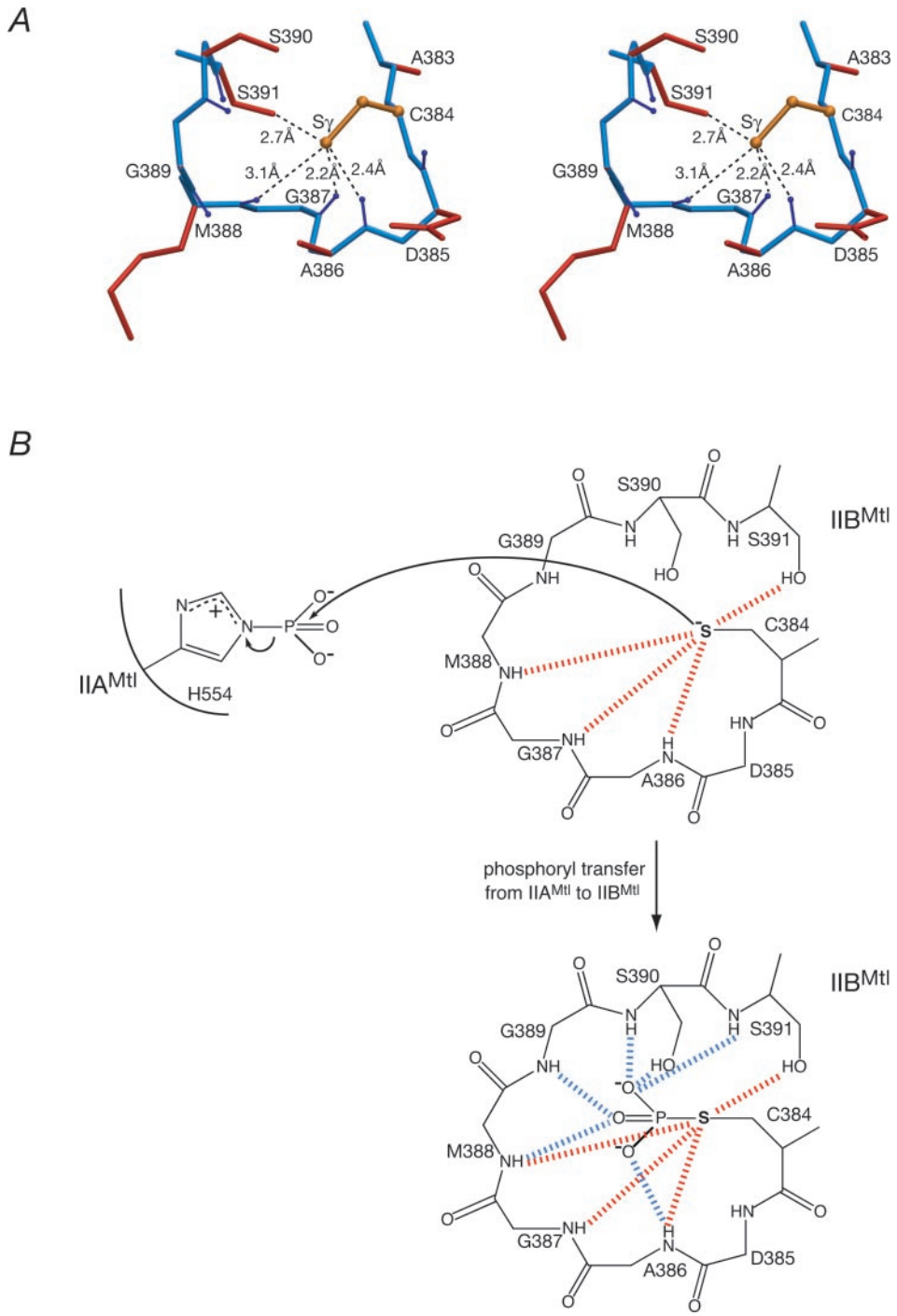

FIG. 3. **The active site of IIB^{Mtl}.** *A*, stereo view of the active site with the backbone in *light blue*, side chains in *red*, Cys-384 in *orange*, and the backbone NH bonds in *dark blue* with the H_N protons represented by small spheres. *B,* possible mechanism of phosphoryl transfer from IIA^{Mtl} to IIB^{Mtl}. Single-letter amino acid abbreviations are used with position numbers.

intermolecular (71). The N terminus of IIA^{Mtl} is located ~16 Å from the active site His-554 on the opposite face of the molecule (16). Likewise, the C terminus of IIB^{Mtl} is located ~26 Å from the active site Cys-384, again on the opposite face of the molecule (Fig. 2*A*). This sets a maximum C_α-C_α separation of ~40 Å between the C terminus of IIB^{Mtl} and the N terminus of IIA^{Mtl} and probably a minimum separation of ~ 30Å. The expected end-to-end distance for a 33-residue random coil is ~60 Å (given by $(C_n n l^2)^{1/2}$, where C_n = 7.5, the value of the characteristic ratio for n = 33 residues and l = 3.8 Å) (72). Thus, the structures of IIA^{Mtl} and IIB^{Mtl} are fully compatible with intramolecular phosphoryl transfer from His-554 of IIA^{Mtl} to Cys-384 of IIB^{Mtl}.

Strutural Similarity of IIB^{Mtl} to Bovine Protein Tyrosine Phosphatase and IIB^{Cellobiose}—A search of the Protein Data Bank data base using the program DALI (73) indicates that there are two proteins that have remarkably close structural similarity to IIB^{Mtl} despite the absence of any significant degree of sequence homology (Fig. 2). The first is bovine protein tyrosine phosphatase (BPTP), a member of the low molecular weight eukaryotic protein tyrosine phosphatases (61–64); the second is IIB^{Chb}, the cytoplasmic IIB component of II^{Chb}, the *N,N′*-diacetylchitobiose specific transporter of the lactose/chitobiose branch of the PTS (17–19). The similarities in the structures of IIB^{Chb} and the low molecular weight protein tyrosine phosphatases has been noted previously (17).

The ordered IIB^{Mtl} domain comprises 97 residues extending from Ser-375 to His-471 of II^{Mtl}. The C_α atoms of 81 of these 97 residues, comprising 84% of the entire IIB^{Mtl} domain, are superimposable on BPTP (Protein Data Bank accession code 1DG9) (64) with an atomic r.m.s difference of 1.9 Å; 74 residues of IIB^{Mtl} (*i.e.* 76% of the domain) are superimposable on IIB^{Chb} (PDB accession code 1H9C) (19) with an atomic r.m.s difference of 2 Å. The percentage sequence identities for the structurally aligned regions are 12.3% for IIB^{Mtl} *versus* BPTP and 8.1% for IIB^{Mtl} *versus* IIB^{Chb}.

IIB^{Mtl} and IIB^{Chb} are of comparable size (~100 residues) and catalyze analogous phosphoryl transfer reactions in the mannitol and *N,N′*-diacetylchitobiose branches of the PTS (4). Both catalyze nucleophilic attack at phosphorus using a cysteine as the nucleophile. Their topologies are virtually identical with only a few minor differences (Fig. 2, *A versus C*); the loops connecting strand β2 to β3 and strand β4 to helix α4 contain short (5–6 residue) helices in the case of IIB^{Chb}, and the orientation of the C-terminal helices in IIB^{Mtl} and IIB^{Chb} differ by 35°. In addition, the active site loop of IIB^{Chb} contains a single deletion relative to those of IIB^{Mtl} and the protein tyrosine phosphatases; specifically, in the structural superposition, residues i + 1 and i +2 from the active site cysteine are replaced by only a single residue in IIB^{Chb}.

In contrast, BPTP (157 residues and ~18 kDa) (61) is 50% larger than IIB^{Mtl}. Remarkably, essentially all of the structural elements present in IIB^{Mtl} are present in BPTP (Fig. 2, *A versus B*). BPTP, however, has two large insertions relative to IIB^{Mtl}, *i.e.* between strands β2 and β3 (36 residues) and between strand β4 and helix α4 (28 residues). Interestingly, a portion of the second insertion of BPTP is in close spatial proximity to the active site and provides the general base (Asp-129) that acts as the proton donor to the leaving group in the phosphorylation step (74). IIB^{Mtl}, on the other hand, may possibly use the aspartate residue (Asp-385) at position i + 1 from the active site cysteine in a similar manner for the phosphoryl transfer reaction from phosphorylated IIB^{Mtl} to IIC^{Mtl}-bound mannitol. The first insertion of BPTP contains a residue (His-72) that stabilizes the active site loop by hydrogen bonding to the asparagine residue located at position i + 3 from the active site cysteine (61). This residue is located in the left-handed helical region of the Ramachandran map and, as noted above, gives the active site P-loop its characteristic appearance. In the case of IIB^{Mtl} the equivalent residue is a glycine; consequently no additional stabilization is required because, in the case of glycine, the right- and left-handed α-helical regions of the Ramachandran map are equally favorable energetically.

Concluding Remarks—We have solved the structure of the cytoplasmic IIB^{Mtl} domain of the mannitol transporter II^{Mtl}. This structure sets the stage for future studies aimed at solving the structure of the complex between the A and B domains of II^{Mtl}. The structure reveals the presence of an active site P-loop whose conformation is essentially the same as that of the eukaryotic protein tyrosine phosphatases and accounts for the stabilization of the thiolate state of the active site cysteine and the stabilization of the phosphorylated intermediate. Remarkably, IIB^{Mtl} is structurally very similar to two other proteins, IIB^{Chb} and the eukaryotic low molecular weight class of eukaryotic protein tyrosine phosphatases, despite the absence of any significant sequence identity. All three proteins catalyze nucleophilic substitution at phosphorus using cysteine as the nucleophile. IIB^{Chb} is the analogous component to IIB^{Mtl} in the *N,N′*-diacetylchitobiose branch of the PTS, whereas BPTP plays a critical role in intracellular signal transduction pathways by controlling the level of the protein tyrosine phosphorylation of a large variety of target proteins involved in growth, cell proliferation, and differentiation. It is tempting to speculate that this commonality of structure represents convergent evolution toward a common fold that has been optimized for phosphoryl transfer involving a phosphorylated cysteine intermediate.

Acknowledgments—We thank David Williams, Junji Iwahara, and Al Mildvan for many useful discussions.

REFERENCES

1. Kundig, W., Ghosh, S., and Roseman, S. (1964) *Proc. Natl. Acad. Sci. U. S. A.* **52,** 1067–1074
2. Herzberg, O., and Klevit, R. E. (1994) *Curr. Opin. Struct. Biol.* **4,** 814–822
3. Tchieu, J. H., Norris, V., Edwards, J. S., and Saier, M. H. (2001) *J. Mol. Microbiol. Biotechnol.* **3,** 329–346
4. Robillard, G. T., and Broos, J. (1999) *Biochim. Biophys. Acta* **1422,** 73–104
5. Liao, D.-I., Silverton, E., Seok, Y.-J., Lee, B. R., Peterkofsky, A., and Davies, D. R. (1996) *Structure* **4,** 861–872
6. Garrett, D. S., Seok, Y.-J., Liao, D.-I., Peterkofsky, A., Gronenborn, A. M., and Clore, G. M. (1997) *Biochemistry* **36,** 2517–2530
7. Wittekind, M., Rajgopal, P., Cranchini, B. R., Reizer, J., Saier, M. H., and Klevit, R. E. (1992) *Protein Sci.* **1,** 1363–1376
8. Herzberg, O., Reddy, P., Sutrina, S., Saier, M. H., Reizer, J., and Kapadia, G. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 2499–2503
9. Jia, Z., Quail, J. W., Waygood, E. B., and Delbaere, L. T. (1993) *J. Biol. Chem.* **268,** 22490–22501
10. Kalbitzer, H. R., and Hengstenberg, W. (1993) *Eur. J. Biochem.* **216,** 205–214
11. van Nuland, N. A. J., Hangyi, I. W., van Shaik, R. C., Berendsen, H. J. C., van Gunsteren, W. F., Scheek, R. M., and Robillard, G. T. (1994) *J. Mol. Biol.* **237,** 544–559
12. Liao, D.-I., Kapadia, G., Reddy, P., Saier, M. H., Reizer, J., and Herzberg, O. (1991) *Biochemistry* **30,** 9538–9594
13. Worthylake, D., Meadow, N. D., Roseman, S., Liao, D.-I., Herzberg, O., and Remington, S. J. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 10383–10386
14. Nunn, R. S., Markovic-Housley, Z., Genovesio-Taverne, G., Flükiger, K., Rizkallah, P. J., Jansonius, J. N., Schirmer, T., and Erni, B. (1996) *J. Mol. Biol.* **259,** 502–511
15. Sliz, P., Engelmann, R., Hengstenberg, W., and Pai, E. F. (1997) *Structure* **5,** 775–788
16. Van Montfort, R. L., Pijning, T., Kalk, K. H., Hangyi, I. W., Kouwijzer, M. L. C. E., Robillard, G. T., and Dijkstra, B. W. (1998) *Structure* **6,** 377–388
17. Van Montfort, R. L., Pijning, T., Kalk, K. H., Reizer, J., Saier, M. J., Thunnissen, M. M., Robillard, G. T., and Dijkstra, B. W. (1997) *Structure* **5,** 217–225
18. Ab, E., Schuurman-Wolters, G. K., Reizer, J., Saier, M. H., Dijkstra, M., Scheek, R. M., and Robillard, G. T. (1997) *Protein Sci.* **6,** 304–314
19. Ab, E., Shuurman-Wolters, G. K., Nijlant, D., Dijkstra, K., Saier, M. H., Robillard, G. T., and Scheek, R. M. (2001) *J. Mol. Biol.* **308,** 993–1009
20. Schauder, S., Nunn, R. S., Lanz, R., Erni, B., and Schirmer, T. (1998) *J. Mol. Biol.* **276,** 591–602
21. Orniss, G. L., Erni, B., and Schirmer, T. (2003) *J. Mol. Biol.* **327,** 1111–1119
22. Cai, M., Williams, D. C., Wang, G., Lee, B. R., Peterkofsky, A., and Clore, G. M. (2003) *J. Biol. Chem.* **278,** 25191–25206
23. Garrett, D. S., Seok, Y.-J., Peterkofsky, A., Gronenborn, A. M., and Clore, G. M. (1999) *Nat. Struct. Biol.* **6,** 166–173
24. Wang, G., Louis, J. M., Sondej, M., Seok, Y.-J., Peterkofsky, A., and Clore, G. M. (2000) *EMBO J.* **19,** 5635–5649
25. Cornilescu, G., Lee, B. R., Cornilescu, C., Wang, G., Peterkofsky, A., and Clore, G. M. (2002) *J. Biol. Chem.* **277,** 42289–42298
26. Van Weeghel, R. P., Meyer, G. H., Keck, W., and Robillard, G. T. (1991) *Biochemistry* **30,** 1774–1779
27. Stephan, M. M., Khandekar, S. S., and Jacobson, G. R. (1989) *Biochemistry* **28,** 7941–7946
28. White, D. W., and Jacobson, G. R. (1990) *J. Bacteriol.* **172,** 1509–1515
29. Van Weeghel, R. P., Meyer, G. H., Pas, H. H., Keck, W., and Robillard, G. T. (1991) *Biochemistry* **30,** 9478–9485
30. Lolkema, J. S., Kuiper, H., ten Hoeve-Duurkens, R. H., and Robillard, G. T. (1993) *Biochemistry* **32,** 1396–1400
31. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfiefer, J., and Bax, A. (1995) *J. Biomol. NMR* **6,** 277–293
32. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) *J. Magn. Reson.* **95,** 214–220
33. Clore, G. M., and Gronenborn, A. M. (1991) *Science* **252,** 1390–1399
34. Bax, A., and Grzesiek, S. (1993) *Acc. Chem. Res.* **26,** 131–138
35. Clore, G. M., and Gronenborn, A. M. (1998) *Trends Biotechnol.* **16,** 22–34
36. Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R., and Zhu, G. (1994) *Methods Enzymol.* **239,** 79–105
37. Neri, D., Szypersrky, T., Otting, G., Senn, H., and Wüthrich, K. (1989) *Biochemistry* **28,** 7510–7516
38. Rückert, M., and Otting, G. (2000) *J. Am. Chem. Soc.* **122,** 7793–7797
39. Ottiger, M., Delaglio, F., and Bax, A. (1988) *J. Magn. Reson.* **131,** 373–378
40. Bewley, C. A., Gustafson, K. R., Boyd, M. R., Covell, D. G., Bax, A., Clore, G. M., and Gronenborn, A. M. (1988) *Nat. Struct. Biol.* **5,** 571–578
41. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* **13,** 289–302
42. Nilges, M., Gronenborn, A. M., Brünger, A. T., and Clore, G. M. (1988) *Protein Eng.* **2,** 27–38
43. Garrett, D. S., Kuszewski, J., Hancock, T. J., Lodi, P. J., Vuister, G. W., Gronenborn, A. M., and Clore, G. M. (1994) *J. Magn. Reson. B* **104,** 99–103
44. Kuszewski, J., Qin, J., Gronenborn, A. M., and Clore, G. M. (1995) *J. Magn. Reson. B* **106,** 92–96
45. Clore, G. M., Gronenborn, A. M., and Tjandra, N. (1988) *J. Magn. Reson.* **131,**

159–162
46. Schwieters, C. D., and Clore, G. M. (2001) *J. Magn. Reson.* **152,** 288–302
47. Schwieters, C. D., Kuszewski, J., Tjandra, N., and Clore, G. M. (2003) *J. Magn. Reson.* **160,** 65–73
48. Clore, G. M., and Kuszewski, J. (2002) *J. Am. Chem. Soc.* **124,** 2866–2867
49. Grishaev, A., and Bax, A. (2004) *J. Am. Chem. Soc.* **126,** 7281–7292
50. Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1999) *J. Am. Chem. Soc.* **121,** 2337–2338
51. Schwieters, C. D., and Clore, G. M. (2001) *J. Magn. Reson.* **149,** 239–244
52. Carson, M. (1991) *J. Appl. Crystallogr.* **24,** 958–961
53. Schwieters, C. D., and Clore, G. M. (2002) *J. Biomol. NMR* **23,** 221–225
54. Jones, T. A., Bergdoll, M., and Kjeldgarrd, M. (1990) in *Crystallographic and Modeling Methods in Molecular Design* (Bugg, C., and Ealick, S., eds), pp. 189–195, Springer-Verlag, Frankfurt
55. Rossman, M. G., Liljas, A., Branden, C.-I., and Banaszak, L. J. (1975) in *The Enzymes*, (Boyer P. D., ed) Vol. 11, pp. 61–112, Academic Press, New York
56. Pas, H. H., and Robillard, G. T. (1988) *Biochemistry* **27,** 5835–5839
57. Pas, H. H., Meyer, G. H., Kruizinga, W. H., Tamminga, K. S., van Weeghel, R. P., and Robillard, G. T. (1991) *J. Biol. Chem.* **266,** 6690–6692
58. Barford, D., Flint, A. J., and Tonks, N. K. (1994) *Science* **263,** 1397–1404
59. Stuckey, J. A., Schubert, H. L., Fauman, E. B., Zhang, Z.-Y., Dixon, J. E., and Saper, M. A. (1994) *Nature* **370,** 571–575
60. Pannifer, A. D. B., Flint, A. J., Tonks, N. K., and Barford, D. (1998) *J. Biol. Chem.* **273,** 10454–10462
61. Zhang, M., Van Etten, R. L., and Stauffacher, C. V. (1994) *Biochemistry* **33,** 11097–11105
62. Su, X.-D., Taddei, N., Stefani, M., Ramponi, G., and Nordlund, P. (1994) *Nature* **370,** 575–578
63. Logan, T., Zhou, M.-M., Nettesheim, D. G., Meadows, R. P., Van Etten, R. L., and Fesik, S. W. (1994) *Biochemistry* **33,** 11087–11096
64. Zhang, M., Zhou, M., Van Etten, R. L., and Stauffacher, C. V. (1997) *Biochemistry* **36,** 15–23
65. Cirri, P., Chiarugi, P., Camici, G., Manao, G., Raugei, G., Cappugi, G., and Ramponi, G. (1993) *Eur. J. Biochem.* **214,** 647–657
66. Davis, J. P., Zhou, M.-M., and Van Etten, R. L. (1994) *J. Biol. Chem.* **269,** 8734–8740
67. Pas, H. H., Ellory, J. C., and Robillard, G. T. (1987) *Biochemistry* **26,** 6689–6696
68. Khandekar, S. S., and Jacboson, G. R. (1989) *J. Cell. Biochem.* **39,** 207–216
69. Boer, H., ten Hoeve-Duurkens, R. H., and Robillard, G. T. (1996) *Biochemistry* **35,** 12901–12908
70. Koning, R. I., Keegstra, W., Oostergetel, G. T., Schuurman-Wolters, G., Robillard, G. T., and Brisson, A. (1999) *J. Mol. Biol.* **287,** 845–851
71. van Weeghel, R. P., van der Hoek, Y. Y., Pas, H. H., Elferink, M., Keck, W., and Robillard, G. T. (1991) *Biochemistry* **30,** 1768–1773
72. Cantor, C. R., and Schimmel, P. R. (1980) in *Biophysical Chemistry*, Part III, pp. 1008–1010, W.H. Freeman, San Francisco
73. Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233,** 123–138
74. Zhang, Z., Harms, E., and Van Etten, R. L. (1994) *J. Biol. Chem.* **269,** 25947–25950
75. Clore, G. M., and Garrett, D. S. (1999) *J. Am. Chem. Soc.* **121,** 9008–9012
76. Clore, G. M., Gronenborn, A. M., and Bax, A. (1998) *J. Magn. Reson.* **133,** 216–221
77. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
78. Vriend, G., and Sander, C. (1993) *J. Appl. Crystallogr.* **26,** 47–60