

## Examination of the Structure, Stability, and Catalytic Potential in the Engineered Phosphoryl Carrier Domain of Pyruvate Phosphate Dikinase<sup>†,‡</sup>

Ying Lin,<sup>§</sup> Jacqueline D. Lusin,<sup>||</sup> Dongmei Ye,<sup>§</sup> Debra Dunaway-Mariano,<sup>§</sup> and James B. Ames<sup>\*,||</sup>

Department of Chemistry, University of New Mexico, Albuquerque, New Mexico 87131, and Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, Maryland 20850

Received September 8, 2005; Revised Manuscript Received December 9, 2005

**ABSTRACT:** Pyruvate phosphate dikinase (PPDK) is a multidomain protein that catalyzes the interconversion of ATP, pyruvate, and phosphate with AMP, phosphoenolpyruvate (PEP), and pyrophosphate using its central domain to transport phosphoryl groups between two distant active sites. In this study, the mechanism by which the central domain moves between the two catalytic sites located on the N-terminal and C-terminal domains was probed by expressing this domain as an independent protein and measuring its structure, stability, and ability to catalyze the ATP/phosphate partial reaction in conjunction with the engineered N-terminal domain protein (residues 1–340 of the native PPDK). The encoding gene was engineered to express the central domain as residues 381–512 of the native PPDK. The central domain was purified and shown to be soluble, monomeric (13,438 Da), and stable ( $\Delta G = 4.3$  kcal/mol for unfolding in buffer at pH 7.0, 25 °C) and to possess native structure, as determined by multidimensional heteronuclear NMR analysis. The main chain structure of the central domain in solution aligns closely with that of the X-ray structure of native PPDK (the root-mean-square deviation is 2.2 Å). Single turnover reactions of [<sup>14</sup>C]-ATP and phosphate, carried out in the presence of equal concentrations of central domain and the N-terminal domain protein, did not produce the expected products, in contrast to efficient product formation observed for the N-terminal central domain construct (residues 1–553 of the native PPDK). These results are interpreted as evidence that the central domain, although solvent-compatible, must be tethered by the flexible linkers to the N-terminal domain for the productive domain–domain docking required for efficient catalysis.

Pyruvate phosphate dikinase (PPDK)<sup>11</sup> catalyzes the interconversion of ATP, P<sub>i</sub>, and pyruvate with AMP, PP<sub>i</sub>, and PEP (1). In C<sub>4</sub> plants the reaction produces PEP for carbon fixation, whereas in pathogenic protists it appears to function in the production of ATP (2) from scavenged PP<sub>i</sub>. In Figure 1 the three phosphoryl transfer steps (3) are depicted for the reaction proceeding in the PEP-forming direction. Phosphoryl transfer is mediated by a carrier histidine (His455 in *C. symbiosum* PPDK (4)), which first displaces AMP from the

β-phosphoryl of ATP to form a pyrophosphorylated histidine intermediate (E-His-PP) (4, 5). Next, P<sub>i</sub> attacks the γ-phosphoryl group of E-His-PP to form PP<sub>i</sub> and a phosphorylated histidine intermediate (E-His-P). In the final step, the E-His-P phosphorylates pyruvate to generate PEP.

The *C. symbiosum* PPDK is a homodimer of 96 kDa subunits. The X-ray crystal structure shows that each subunit is organized into three consecutive structural domains connected by helical linkers (6) (Figure 2A). Mg<sup>2+</sup>, ATP, and P<sub>i</sub> bind to the active site located at the concave surface of the N-terminal domain. The small central domain docks with this face of the N-terminal domain, desolvating its active site and positioning the carrier His455 for attack at the β-phosphoryl group of the bound ATP ligand. This structure corresponds to a conformer (henceforth referred to as conformer 1) competent to catalyze the first two phosphoryl transfer steps pictured in Figure 1. The third phosphoryl transfer (step 3 in Figure 1) occurs in the active site located at the concave surface of the C-terminal domain. This active site binds pyruvate and Mg<sup>2+</sup>. The central domain docks with this face of the C-terminal domain, desolvating the active site and positioning the phosphorylated His residue for phosphoryl transfer to the C(2)O of the pyruvate forming PEP and regenerating the unphosphorylated H455 (picture in Figure 2B). The enzyme conformer that catalyzes PEP formation is named conformer 2 and is shown in Figure 2A. Conformer 2 was originally modeled from conformer 1 by

<sup>†</sup> This work was supported by a Beckman Foundation Young Investigator Award to J.B.A. and NIH grants EY012347/NS045909 to J.B.A. and GM 36260 to D.D.-M.

<sup>‡</sup> Atomic coordinates have been deposited into the Protein Data Bank (2FM4).

\* To whom correspondence should be addressed. Tel: (240) 314–6120. Fax: (240) 314–6255. E-mail: james@carb.nist.gov.

<sup>§</sup> University of New Mexico.

<sup>||</sup> University of Maryland Biotechnology Institute.

<sup>1</sup> Abbreviations used are as follows: pyruvate phosphate dikinase (PPDK), adenosine 5-triphosphate (ATP), adenosine 5-diphosphate (ADP), adenosine 5'-monophosphate (AMP), heteronuclear single quantum coherence (HSQC), nuclear Overhauser effect (NOE), nuclear Overhauser effect spectroscopy (NOESY), total correlation spectroscopy (TOCSY), pyruvate (Pyr), phosphate (P<sub>i</sub>), phosphoenolpyruvate (PEP), pyrophosphate (PP<sub>i</sub>), PPDK pyrophosphorylated histidine intermediate (E-His-PP), PPDK phosphorylated histidine intermediate (E-His-P), the PPDK central domain construct of residues 381–512 (Cent-I), the PPDK central domain construct of residues 389–503 (Cent), PPDK central domain construct of residues 341–533 (Cent-L), the PPDK N-terminal domain construct of residues 1–340 (Tem-340), the PPDK N-terminal domain construct of residues 1–553 (Tem340-Cent-I).



the central domain might take. Specifically, it might travel through solvent or it might travel over a protein surface of the two “touching” terminal domains. Movement along the protein surface would be analogous to a ball sliding between opposing sides of a bowl. Although such a model is attractive by virtue of its simplicity, it requires the central domain to stay “His455-face-down” on the protein surface. Unlike the ball in the bowl, we cannot invoke gravity as the force to hold it to the surface. Furthermore, this role cannot be assigned to the domain–domain linkers.

The central domain has homologues in the domains of several other multidomain proteins (10–12). However, it has not been observed as a protein and consequently there was no evidence to support the proposal that it might alternate between terminal domains via solvent diffusion. In this paper we examine the central domain in a fully solvated environment to determine whether it is structurally and/or functionally dependent on the protein context of the holoenzyme. The central domain was expressed from an engineered PPDK gene and purified for NMR structure determination. Here we report on its native structure, stability, and catalytic potential.

## MATERIALS AND METHODS

**Materials.**  $^{15}\text{N}$ - $\text{NH}_4\text{Cl}$  and  $^{13}\text{C}_6$ -glucose were from Sigma-Aldrich, Nde I and BamH I restriction enzymes were purchased from Promega, pET 3a vector was purchased from Novagen, and Sephadex G-75 resin, Calibration Kit proteins, and Hiload 16/60 Superdex 200 column were purchased from Amersham Pharmacia Biotech. The [ $^{14}\text{C}$ ]ATP was from Perkin-Elmer. All other chemicals were from Sigma. A Beckman DU 7400 spectrophotometer was used for UV absorption measurements, and a Rainin Dynamax SD-200 pump system was used for HPLC separations. Scintillation counting was carried out with a Beckman LS 6500 multi-purpose scintillation counter. An Amersham Pharmacia P-500 FPLC pump system was employed in the determination of protein molecular size by gel filtration.

**Gene Engineering, Expression, and Purification of  $^{15}\text{N}$ - and  $^{13}\text{C}$ -Labeled Central Domain with Partial Linkers (Cent-I).** Mutant genes were prepared from the plasmid pA-CYC184-D12 (13) using PCR techniques analogous to those described in (14). Mutagenic primers, 18–22 base pairs in length, were synthesized by Invitrogen products. The protein construct named “Cent-I” included residues 381–512 of the recombinant *Clostridium symbiosum* PPDK. Nde I and BamH I restriction sites were added to the PPDK gene following residue 380 and residue 512 codon positions, respectively. The gene was digested with Nde I and BamH I restriction enzymes and subcloned into the similarly digested pET3a vector. The sequence of the cloned gene was confirmed by DNA sequencing carried out by the Department of Pathology at the University of New Mexico. The plasmid was used to transform competent *E. coli* JM101 cells, which were grown in LB media at 28 °C, with gentle shaking for ~12.5 h until the OD reached ~1.3. IPTG was added at a concentration of 0.4 mM, and after a 4 h induction period the cells were harvested by centrifugation. Cell lysis was carried out in buffer containing 20 mM imidazole, 2.5 mM EDTA, 7.5 mM KCl, 2 mM DTT, 1 mM 1,10-phenanthroline, 1 mM PMSF, 1 mM benzamidine, and 50 mg/mL

trypsin inhibitor (pH 6.8) using a French press (1300–1500 psi). The cell extract was fractionated by ammonium sulfate induced protein precipitation. The 50–80% ammonium sulfate protein precipitant was dialyzed at 4 °C against 4 L of 50 mM triethanolamine, 2.5 mM EDTA, and 1 mM DTT (pH 7.0) before loading onto a 90 × 2.5 cm DEAE-cellulose column equilibrated with 50 mM triethanolamine, 2.5 mM EDTA, and 1 mM DTT (pH 7.0). The column was eluted with a 2 L linear gradient of 0–0.3 M KCl in 50 mM triethanolamine, 2.5 mM EDTA, and 1 mM DTT (pH 7.0) at 4 °C. The eluant was analyzed by SDS-PAGE, and the Cent-I containing fractions (which eluted at 0.14–0.16 M KCl) were combined and concentrated using an Amicon with 3K membrane disk filter. The concentrate was chromatographed on a 120 × 2 cm Sephadex G-75 column equilibrated with 50 mM triethanolamine, 2.5 mM EDTA, 100 mM KCl, and 1 mM DTT (pH 7.0). The column was eluted with 50 mM triethanolamine, 2.5 mM EDTA, 100 mM KCl, and 1 mM DTT, at pH 7.0 and 4 °C. The eluant was analyzed by SDS-PAGE, and the Cent-I containing fractions were combined and concentrated to 50 mg/mL using a Centricon with 3K filter and stored in 50 mM triethanolamine, 2.5 mM EDTA, 100 mM KCl, and 1 mM DTT at pH 7.0 and 4 °C. The yield was 1.5 mg/g of wet cells. The  $^{15}\text{N}$ - and  $^{13}\text{C}$ -labeled Cent-I protein was made by adaptation of the isotope labeling method described by Sambrook and Ames (15, 16). Accordingly, transformed *E. coli* JM101 cells were grown in 1 L of M9 media containing 12.66 g of  $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ , 3.0 g of  $\text{KH}_2\text{PO}_4$ , 0.5 g of NaCl, 1.0 g of  $^{15}\text{N}$ - $\text{NH}_4\text{Cl}$ , 0.2 g of  $\text{MgSO}_4$ , 50  $\mu\text{L}$  of 1 M  $\text{CaCl}_2$ , 40 mL of 10%  $^{13}\text{C}_6$ -glucose, 1 g of thiamine, 1 g of pyridoxal, 1 g of niacinamide, and 0.1 g of riboflavin. Cells were grown with shaking at 28 °C for ~13 h until the OD reached ~0.7 and then induced with 142.5 mg/1.5 L of IPTG for ~10 h. The same procedure used to isolate the unlabeled Cent-I was employed for purification of the labeled protein. The yield of labeled protein was 1 mg/g of wet cells (or 3 mg/L of cell culture).

**Cent-I Molecular Weight Determination.** The theoretical molecular weight of Cent-I was calculated by using the EXPASY molecular biology server program to compute PI/MW. The molecular weight of Cent-I was determined by electrospray mass spectrometric analysis carried out by the mass spectral facility of the University of New Mexico and by FPLC gel filtration chromatographic analysis. Chromatography was carried out at 4 °C using a Hiload 16/60 Superdex 200 column (Amersham Pharmacia Biotech) equilibrated with 25 mM  $\text{K}^+$ Hepes (pH 7.5) and calibrated with molecular weight standards ribonuclease A (13.7 kDa), chymotrypsinogen A (25 kDa), ovalbumin (43 kDa), and albumin (67 kDa). The MWs of Cent-I were determined by extrapolation from the plot of elution volume vs log MW of the protein standards.

**Cent-I Denaturation.** Cent-I (37  $\mu\text{M}$ ) unfolding was induced by incubation in 50 mM triethanolamine, 0.1 mM EDTA, 100 mM KCl, and 1 mM DTT (pH 7.0, 25 °C) containing 0–6 M urea for 4 h. The equilibrium fluorescence spectra of the samples were measured at 25 °C with a Fluoromax-2 spectrofluorometer (JUBIN YVON-SPEX Instruments SA Inc.). The excitation wavelength was 278 nm. The procedures used to determine the denaturation curve, and from the curve to determine the  $\Delta G$ , for unfolding were adapted from (17). Accordingly, the percent decrease in

fluorescence intensity at  $\lambda_{\max}$  304 nm was determined for the conversion of native protein (0 M urea) to fully denatured protein (6 M urea) to be 23.5%. The mole fractions of native ( $f^N$ ) and denatured protein ( $f^D$ ) in the urea titration solutions were calculated from the observed fluorescence intensity ( $I_{\text{obs}}$ ), the fluorescence intensity of the native protein ( $I_N$ ) and the denatured protein ( $I_D$ ) using eq 1. A two-state system was assumed for which  $f^D + f^N = 1$ .

$$I_{\text{obs}} = f^N I_N + f^D I_D \quad (1)$$

The equilibrium constants ( $K_{\text{eq}}$ ) for the conversion of native to denatured protein titration solutions were calculated from the ratio  $f^D/(1 - f^D)$ . The Gibbs free energy of unfolding ( $\Delta G$ ) for the protein at each urea concentration was calculated from the corresponding  $K_d$  using eq 2, where  $R$  is the gas constant and  $T$  is the temperature. A linear plot of

$$\Delta G = -RT \ln K_{\text{eq}} \quad (2)$$

$\Delta G$  vs urea concentration was constructed, which allowed extrapolation to 0 M urea to define the  $\Delta G$  value for protein unfolding in 50 mM triethanolamine, 0.1 mM  $\text{Na}_2\text{EDTA}$ , 100 mM KCl, and 1 mM DTT (pH 7.0, 25 °C).

*Kinetic Determination of Catalysis of the E-His + ATP + P<sub>i</sub> to E-P + AMP + PP<sub>i</sub> Partial Reaction by the N-Terminal Domain (Tem 340), Central Domain (Cent-I), and Tem340-Cent-I Mixture Proteins.* Reactions were initiated by mixing 14  $\mu\text{L}$  of buffered enzyme/cofactors with 13  $\mu\text{L}$  of buffered substrate to generate solutions initially containing protein (40  $\mu\text{M}$  Tem340 protein, 40  $\mu\text{M}$  Cent-I protein, or a mixture (40  $\mu\text{M}$  each) of Tem340 and Cent-I proteins), 10  $\mu\text{M}$  [ $^{14}\text{C}$ ]ATP, 5 mM  $\text{MgCl}_2$ , 40 mM  $\text{NH}_4\text{Cl}$ , and 11 mM  $\text{P}_i$  in 50 mM  $\text{K}^+\text{Hepes}$  (pH 7.0, 25 °C). Tem340 is the N-terminal domain protein obtained by engineering the PPDK gene to express the N-terminal domain (residues 1–340). Details regarding its preparation and properties are provided in (18). Reactions were terminated after specified incubation periods by the addition of 193  $\mu\text{L}$  of 0.6 M HCl. Protein was removed from the quenched reactions by centrifugal size filtration using a 500  $\mu\text{L}$  filter (cutoff molecular weight 5 kDa) purchased from Pall Gelman Inc. The supernatant containing the unconsumed [ $^{14}\text{C}$ ]ATP and the formed [ $^{14}\text{C}$ ]AMP and [ $^{14}\text{C}$ ]ADP product were separated by HPLC using a Beckman Ultrasphere C18 reversed-phase analytical column and 0.02 M triethylamine, with 4% methanol eluant, pH 7.5. The ATP, ADP, and AMP fractions were analyzed for  $^{14}\text{C}$  content by liquid scintillation counting. Control reactions were carried out in which either the Tem340 or the Cent-I protein was excluded, or both of them were excluded. The [ $^{14}\text{C}$ ]AMP was 2.5%, 2.7%, and 4.3%, respectively, of the reactant [ $^{14}\text{C}$ ]ATP. The [ $^{14}\text{C}$ ]ADP in the reactant was 6.6%, 9.7%, and 6.9%, respectively. The percent conversion was adjusted using the background value and then converted to product concentration on the basis of the original reactant concentration.

*NMR Spectroscopy.* Samples for NMR analysis consisted of  $^{15}\text{N}$ -labeled or  $^{13}\text{C}$ -/ $^{15}\text{N}$ -labeled Cent-I (1.0 mM) in 0.3 mL of a 95%  $\text{H}_2\text{O}/5\%$  [ $^2\text{H}$ ] $\text{H}_2\text{O}$  solution containing 50 mM KCl and 20 mM  $\text{KH}_2\text{PO}_4$  at pH 7.0 by adding 5%  $\text{D}_2\text{O}$  for deuterium lock. All NMR experiments were performed at 25 °C on Bruker Avance 500 or 600 MHz spectrometers

Table 1: Structural Statistics for the Ensemble of 25 Calculated Structures of Cent-I

NOE restraints (total)	2286
intra ( $ i - j  = 0$ )	586
medium ( $1 \leq  i - j  \leq 4$ )	810
long ( $ i - j  > 4$ )	890
dihedral angle restraints ( $\phi$ and $\psi$ )	116
H bond restraints in $\beta$ -sheet regions	76
RMSD from ideal geometry	
bond length ( $\text{\AA}$ )	$0.0009 \pm 0.00004$
bond angle (deg)	$1.95 \pm 0.003$
Ramachandran plot (%)	
most favored region	75
allowed regions	23
disallowed regions	2
RMSD of atom position from av structure ( $\text{\AA}$ )	
$\beta$ -sheet regions (main-chain atoms)	$0.65 \pm 0.10$
$\beta$ -sheet regions (non-hydrogen atoms)	$1.19 \pm 0.13$

equipped with a four-channel interface and triple-resonance probe with triple-axis pulsed field gradients. The  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectra (see Figure 5) were recorded on a sample of  $^{15}\text{N}$ -labeled Cent-I (in 95%  $\text{H}_2\text{O}/5\%$   $^2\text{H}_2\text{O}$ ). The number of complex points and acquisition times were as follows: 256, 180 ms ( $^{15}\text{N}$  ( $F_1$ )); and, 512, 64 ms ( $^1\text{H}$  ( $F_2$ )). All triple-resonance experiments were performed, processed, and analyzed as described (19, 20) on a sample of  $^{13}\text{C}$ -/ $^{15}\text{N}$ -labeled Cent-I (in 95%  $\text{H}_2\text{O}/5\%$   $^2\text{H}_2\text{O}$ ) with the following number of complex points and acquisition times: HNCO { $^{15}\text{N}$  ( $F_1$ ) 32, 23.7 ms;  $^{13}\text{C}$  ( $F_2$ ) 64, 42.7 ms;  $^1\text{H}$  ( $F_3$ ) 512, 64 ms}; HNCACB { $^{15}\text{N}$  ( $F_1$ ) 32, 23.7 ms;  $^{13}\text{C}$  ( $F_2$ ) 48, 6.3 ms;  $^1\text{H}$  ( $F_3$ ) 512, 64 ms}; CBCACONNH { $^{15}\text{N}$  ( $F_1$ ) 32, 23.7 ms;  $^{13}\text{C}$  ( $F_2$ ) 48, 6.3 ms;  $^1\text{H}$  ( $F_3$ ) 512, 64 ms}; HBHACONNH { $^{15}\text{N}$  ( $F_1$ ) 32, 23.7 ms,  $^1\text{H}_{\text{ab}}$  ( $F_2$ ) 64 21 ms,  $^1\text{H}$  ( $F_3$ ) 512, 64 ms}.  $^{15}\text{N}$ -edited/ $^{13}\text{C}$ -edited NOESY-HSQC, TOCSY-HSQC, and HCCH-TOSCY experiments were performed as described previously (21, 22).

Triple-resonance and NOESY spectra measured above were analyzed to determine resonance assignments and secondary structure of Cent-I. The chemical shift index (see (23) for a detailed description of the chemical shift index),  $^3J_{\text{NH}\alpha}$  coupling constants, and nuclear Overhauser effect (NOE) connectivity patterns for each residue were analyzed and provided a measure of the overall secondary structure. Small  $^3J_{\text{NH}\alpha}$  coupling constants (<5 Hz), strong NOE connectivities (NN( $i, i+1$ ) and  $\alpha\text{N}(i, i+3)$ ), and positive chemical shift index are characteristic of residues in an  $\alpha$ -helix. Conversely, large  $^3J_{\text{NH}\alpha}$  coupling constants (>8 Hz), strong  $\alpha\text{N}(i, i+1)$  and weak NN( $i, i+1$ ) NOE connectivities, and negative chemical shift index are characteristic of residues in a  $\beta$ -strand.

*Structure Calculation.* Backbone and side-chain NMR resonances of Cent-I were assigned as described previously (20). Analysis of NOESY data determined over 2200 interproton distance relationships throughout the protein and estimates of  $\phi$  and  $\psi$  dihedral angles (19). The NMR-derived distances and dihedral angles then served as restraints (see Table 1) for calculating the three-dimensional structure of the protein using distance geometry and restrained molecular dynamics. Structure calculations were performed using the YASAP protocol within X-PLOR (24, 25), as described previously (26). A total of 2286 interproton distance constraints (including stereospecific NOE assignment of leucine

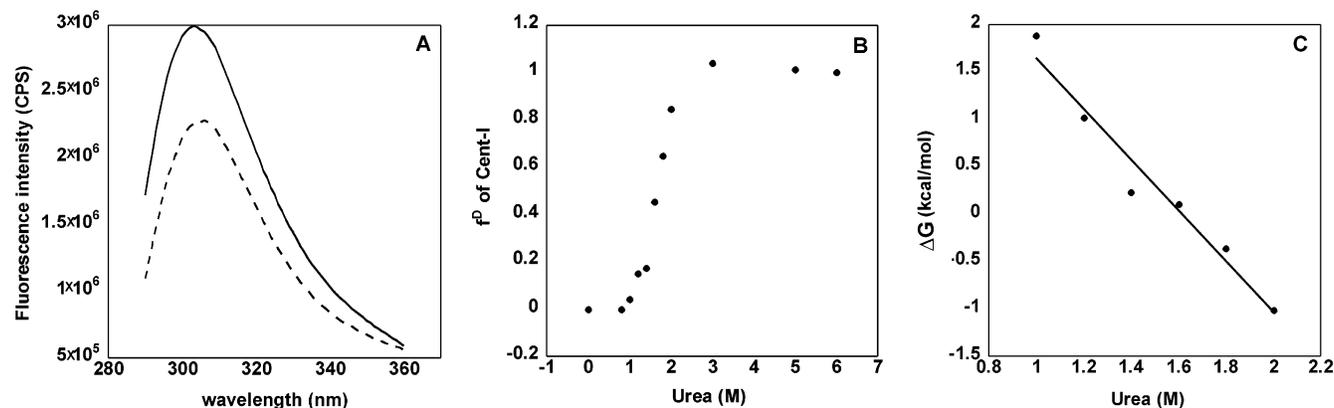


FIGURE 3: (A) Fluorescence spectrum of the native Cent-I protein in buffer (solid line) and unfolded Cent-I protein in buffer containing 6 M urea (dashed line). The intrinsic fluorescence emission spectrum was obtained by excitation at 278 nm. (B) Urea-induced denaturation of Cent-I at 25 °C, in 50 mM triethanolamine, 0.1 mM EDTA, 100 mM KCl, and 1 mM DTT buffer, at pH 7.0. The intrinsic fluorescence was monitored at 304 nm, as a function of urea concentration. (C) Transformation of the fluorescence data to give a linear plot of change in Gibbs free energy versus urea concentration around the midpoint of the unfolding transition. Linear extrapolation of this plot to 0 M urea results in the Gibbs free energy change of denaturation in the absence of denaturant.

and valine methyl groups (27)) were obtained as described (20) by analysis of  $^{13}\text{C}$ -edited and  $^{15}\text{N}$ -edited NOESY-HSQC spectra (100 ms mixing time) of  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled Cent-I. This represents about 18 NOEs per residue, consistent with a medium-resolution structure. In addition to the NOE-derived distance constraints, the following additional constraints were included in the structure calculation: 116 dihedral angle constraints ( $\phi$  and  $\psi$ ) and 76 distance constraints for 38 hydrogen bonds verified by identifying slowly exchanging amide protons in hydrogen–deuterium exchange experiments (15). Fifty independent structures were calculated, and the 25 structures of lowest energy were selected. The root-mean-square difference of the ensemble with respect to the energy-minimized mean structure was  $0.65 \pm 0.1$  Å for main chain atoms and  $1.19 \pm 0.13$  Å for all heavy atoms. The average total and experimental distance energies are 2469 and 159 kcal mol $^{-1}$ . The average root-mean-square (RMS) deviations from an idealized geometry for bonds and angles are 0.009 Å and  $1.95^\circ$ . None of the distance and angle constraints were violated by more than 0.40 Å or  $4^\circ$ , respectively.

## RESULTS

**Construction and Stability of the PPDK Central Domain Protein (Cent-I).** Three constructs of the central domain were prepared. Each differed in the length of the residual domain linker (see Figure 2B). The first construct (Cent) consisted of PPDK residues 389–503 and contained no linker region. H NMR analysis revealed that this construct did not have a stable, unique native conformation (data not shown). The second construct (Cent-L; residues 341–533) contained both full-sized linkers. H NMR analysis revealed that this construct did have a stable, unique native conformation (data not shown). However, SDS-PAGE chromatography of stored protein sample, carried out in conjunction with automated N-terminal sequence analysis, revealed that both the N-terminal and C-terminal linker regions were subject to proteolytic cleavage (data not shown). The third construct (Cent-I; residues 381–512) contained only the first eight residues of each linker. Cent-I proved superior in stability and yield and consequently was carried forward for structure determination. The mass spectral determination of molecular mass of 13,438 Da matched the theoretical weight (13,438

Da). The native molecular weight was estimated by size filtration chromatography to be ~19 kDa. The Cent-I protein is thus a monomer.

The stability of the Cent-I protein was examined using urea to unfold the protein and intrinsic protein fluorescence to monitor the solution equilibrium. The fluorescence emission spectra for the native protein in buffer and denatured protein in buffer containing 6 M urea are compared in Figure 3A. The urea titration, monitored at 304 nm, is shown in Figure 3B. A single inflection point was observed at 2 M urea, indicative of a two-state system. The  $\Delta G$  for unfolding is plotted as a function of urea concentration in Figure 3C. The  $\Delta G$  for unfolding in buffer was determined by extrapolation to be 4.3 kcal/mol. The central domain construct is thus a stable protein.

**NMR Structure of the PPDK Central Domain (Cent-I).** The  $^1\text{H}$ – $^{15}\text{N}$ -HSQC NMR spectrum of  $^{15}\text{N}$ -labeled Cent-I exhibited the expected number of well-resolved amide resonances (Figure 4), suggesting that the protein is stably folded. Pulsed field gradient NMR diffusion studies (28) revealed a hydrodynamic radius of 1.5 nm and confirmed that Cent-I is a protein monomer in solution under NMR conditions. More than 90% of the NMR resonances had very sharp and uniform peak intensities and were assigned as indicated in Figure 4. The assigned peaks in the spectrum represent main-chain and side-chain amide groups that serve as fingerprints of the overall conformation. Three-dimensional structures of Cent-I were calculated on the basis of NOE data, slowly exchanging amide NH groups, chemical shift analysis, and  $^3J_{\text{NH}\alpha}$  spin–spin coupling constants (see Materials and Methods). The analysis of chemical shift index (CSI) (29),  $^3J_{\text{NH}\alpha}$  (30), and short-range NOEs (31) is illustrated in Figure 5A and revealed at least 11  $\beta$ -strands and 3  $\alpha$ -helices. The overall topology of secondary structure elements consists of two antiparallel  $\beta$ -sheets (strands  $\beta 1$ – $\beta 11$ – $\beta 10$ – $\beta 2$  and  $\beta 7$ – $\beta 8$ – $\beta 9$ ) and a 4-stranded parallel  $\beta$ -sheet (strands  $\beta 3$ – $\beta 4$ – $\beta 5$ – $\beta 6$ ) flanked by 3 parallel  $\alpha$ -helices (shown schematically in Figure 5B). The final three-dimensional structures of Cent-I derived from the NMR data are superimposed and illustrated in Figure 6A. Table 1 summarizes the structural statistics calculated for 25 lowest energy conformers.

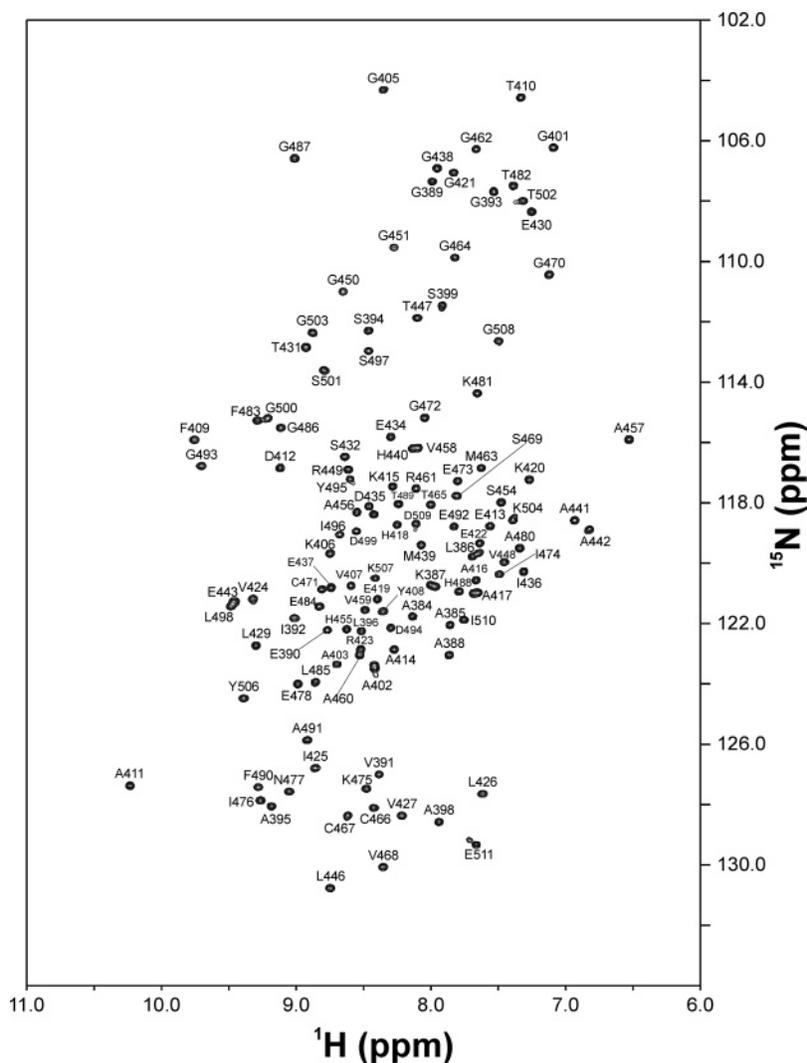


FIGURE 4:  $^1\text{H}$ – $^{15}\text{N}$  HSQC NMR spectrum of Cent-I. Sequence-specific assignments are indicated and deposited into the BioMagResBank Repository (BMRB accession no. 6932).

The NMR-derived Cent-I structures (Figure 6) revealed a three-dimensional fold very much like that seen previously for the central domain of the holoenzyme (6), except for minor differences in loop regions and near the N- and C-termini that are poorly defined by the NMR data. The Cent-I structure in solution contains a multilayered  $\beta/\beta/\alpha$  architecture with a base layer of three parallel  $\alpha$ -helices ( $\alpha 1$ – $\alpha 3$ ) that lay beneath a central four-stranded parallel  $\beta$ -sheet (strands  $\beta 3$ – $\beta 6$ ). A four-stranded antiparallel  $\beta$ -sheet (strands  $\beta 1$ – $\beta 11$ – $\beta 10$ – $\beta 2$ ) is positioned on top of the central parallel  $\beta$ -sheet in an orthogonal fashion, forming a  $\beta$ -sandwich structure that is spread apart on one side to accommodate a third three-stranded antiparallel  $\beta$ -sheet (strands  $\beta 7$ – $\beta 9$ ). The three-stranded  $\beta$ -sheet has an abundance of aromatic residues (Phe483, His488, Phe490) that interact with peripheral hydrophobic residues in the  $\beta$ -sandwich. The catalytic histidine (His455) located on helix  $\alpha 3$  is solvent exposed (Figure 6B) and represents the only catalytic residue from the central domain. Internal alanine residues (Ala414, Ala417, Ala442, Ala457, Ala460) located on one face of the helical base layer interact with opposing hydrophobic residues (Val424, Leu426, Ile445, Cys467) from the central parallel  $\beta$ -sheet (Figure 5B). Alternating hydrophobic resi-

dues (Val 407, Ile425, Val427, Leu429, Leu446, Val448, Cys466, Val468) on the opposite side of the central  $\beta$ -sheet interact with opposing hydrophobic residues (Ala 395, Ala403, Ile496, Leu498, Ile505) from the antiparallel four-stranded  $\beta$ -sheet.

The distribution of charged residues on the outer surface of Cent-I is shown in Figure 6C. One side of the surface is highly charged, whereas the other half appears uncharged with many exposed conserved hydrophobic residues (Val391, Leu396, Ala398, Ile436, Val448, Met452, Val458, Met463, Val468; see Figure 6C). The catalytic histidine is located on the uncharged surface and surrounded by the hydrophobic residues, Ile436 and Val458. X-ray structural analysis has shown that this face of the central domain alternates between the convex face of the N-terminal domain and the convex face of the C-terminal domain (6, 8, 32) (Figure 2A). It is significant that the exposure of the hydrophobic face of the central domain to solvent does not result in poor water solubility or in protein aggregation. This finding suggests that the nonpolar residues are not used to inhibit diffusion of the central domain from the protein surface. A more likely explanation, which is consistent with preliminary mutagenesis data, suggests that the nonpolar residues surrounding the catalytic His function to form a waterproof seal at the



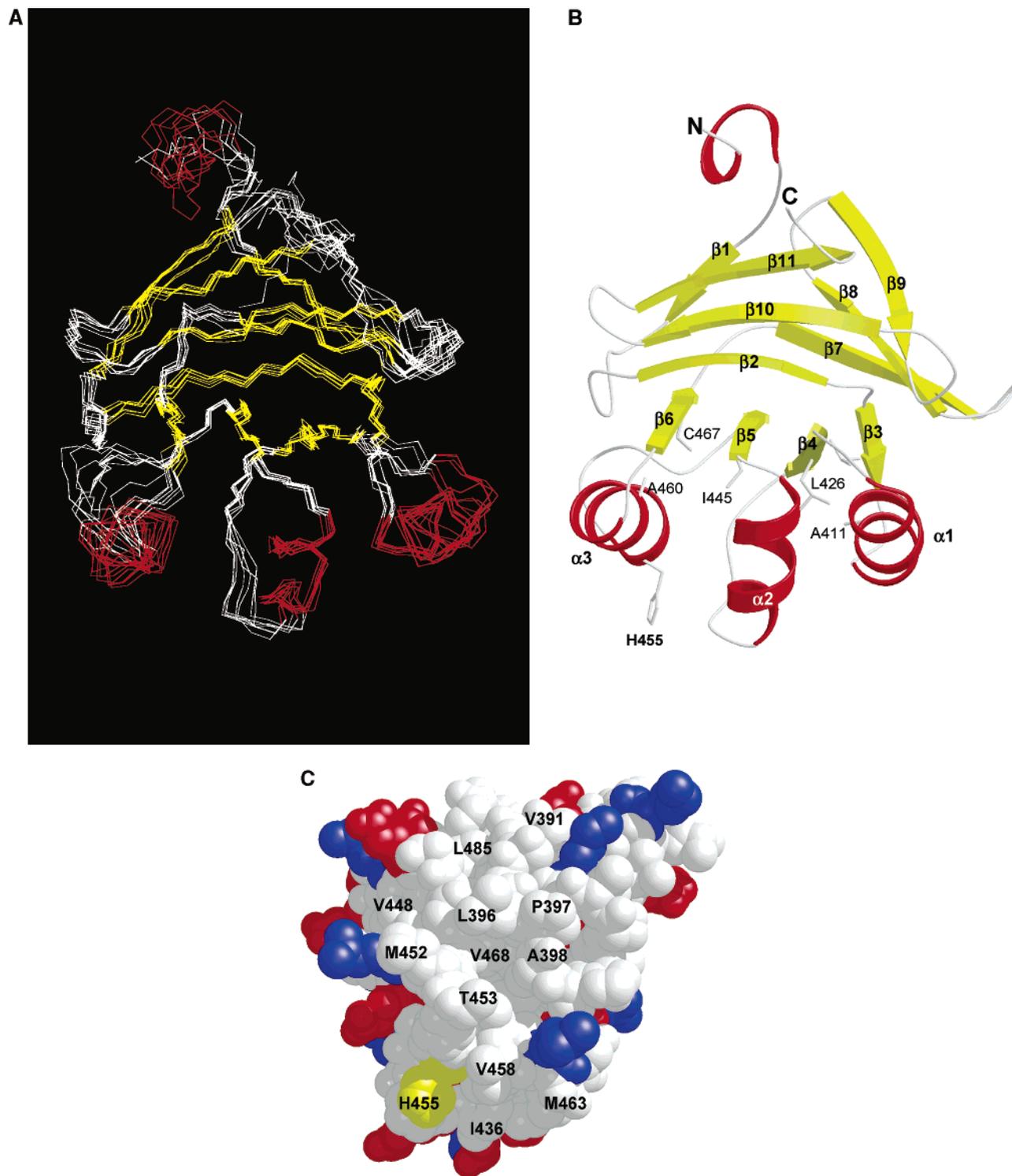


FIGURE 6: NMR structure of cent-I in solution. (A) Superposition of main-chain atoms of the 25 lowest energy calculated structures. (B) Ribbon representation of the energy-minimized average main-chain structure. Side-chain atoms of structurally important residues are indicated.  $\beta$ -Strand and  $\alpha$ -helical regions are highlighted in yellow and red. (C) Space-filling representation of Cent-I with acidic residues (Glu and Asp), basic residues (Lys and Arg), and His455 colored red, blue, and yellow, respectively. Atomic coordinates of cent-I have been deposited into the Protein Data Bank (accession no. 2FM4).

kinetic behavior is in contrast with that observed for the N-terminal domain–central domain protein, which does not catalyze ATP hydrolysis but rather the reaction  $\text{EHis} + \text{ATP} + \text{P}_i \rightarrow \text{EHis-P} + \text{AMP} + \text{P}_i$  (steps I and II of Figure 1) at a significant rate turnover rate of  $0.4 \text{ s}^{-1}$ .

Two control reactions were carried out to determine the source of the ATP hydrolase activity. Whereas incubation of Cent-I with  $^{14}\text{C}$ ATP and  $\text{P}_i$  did not lead to  $^{14}\text{C}$ AMP or

$^{14}\text{C}$ ADP formation (data not shown), incubation of Tem340 with  $^{14}\text{C}$ ATP and  $\text{P}_i$  did (see Figure 7B). Indeed, the time course for the reaction measured using Cent-I and Tem340 in a 1:1 ratio is identical with that measured with the Tem340 alone. Thus, the slow ATP hydrolysis is attributed to Tem340. The present studies show that Cent-I and Tem340 do not associate to catalyze the formation of  $^{14}\text{C}$ AMP from reaction with  $^{14}\text{C}$ ATP and  $\text{P}_i$ .

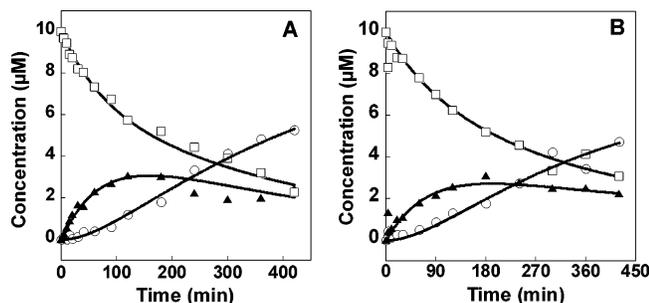


FIGURE 7: Reaction curves of 10  $\mu\text{M}$  [ $^{14}\text{C}$ ]ATP and 11 mM  $\text{P}_i$  with (A) 40  $\mu\text{M}$  Cent I and 40  $\mu\text{M}$  N-terminal domain (Tem340) and (B) 40  $\mu\text{M}$  Tem340 and  $\text{K}^+\text{Hepes}$  (pH 7.0, 25  $^\circ\text{C}$ ) containing 5 mM  $\text{MgCl}_2$  and 40 mM  $\text{NH}_4\text{Cl}$ : (○) [ $^{14}\text{C}$ ]AMP; (▲) [ $^{14}\text{C}$ ]ADP; (□) [ $^{14}\text{C}$ ]ATP. The simulated curves were generated using the

kinetic model  $\text{E}\cdot\text{ATP} \xrightleftharpoons[k_{-1}]{k_1} \text{E}\cdot\text{ADP} \xrightleftharpoons[k_{-2}]{k_2} \text{E}\cdot\text{AMP}$  the simulation program KINSIM (41), and the rate constants  $k_1 = 1.0 \times 10^{-4} \text{ s}^{-1}$ ,  $k_{-1} = 8.0 \times 10^{-5} \text{ s}^{-1}$ ,  $k_2 = 8.5 \times 10^{-5} \text{ s}^{-1}$ , and  $k_{-2} = 1.0 \times 10^{-7} \text{ s}^{-1}$  for part A and  $k_1 = 7.8 \times 10^{-5} \text{ s}^{-1}$ ,  $k_{-1} = 6.2 \times 10^{-4} \text{ s}^{-1}$ ,  $k_2 = 1.0 \times 10^{-4} \text{ s}^{-1}$ , and  $k_{-2} = 1.9 \times 10^{-5} \text{ s}^{-1}$  for part B.

## DISCUSSION

PPDK is a three-domain protein in which domains are connected by solvated linkers (Figure 2). This protein might have evolved through the process of gene duplication and fusion, leading to the linkage of three proteins, whose structures coevolved to support a novel function. The central domain is small and spherical. Its catalytic residue, His455, is responsible for the transport of the phosphoryl group between the active sites of the two terminal domains. A seemingly remarkable feature of this domain is the large cluster of nonpolar residues on its surface, forming a half-circle around the His455. The domain is essentially amphiphilic with the His455 face hydrophobic and the opposite face polar (Figure 6C). Nonpolar protein surfaces are unusual, and therefore it is not unreasonable to assume that the central domain evolved this surface to support its role in catalysis. Our original hypothesis was that this role is to ensure that the central domain remains in contact with the terminal domains as it alternates between active sites through a hydrophobic effect.

The finding that the central domain can exist in its native structure outside of the context of the holoPPDK is intriguing. Cent-I does not aggregate but remains monomeric. This indicates that its nonpolar face is compatible with solvent. Although this observation does not rule out the possibility that the nonpolar face contributes to the attraction of the central domain to the terminal domains, one cannot argue that the nonpolar surface is evidence that the central domain remains in contact with the protein surface during alternation. Instead, it appears that the attractive forces for the transfer of the central domain from one terminal domain to the other might be mediated through water molecules coating the protein surface.

The central domain consists of a scaffold of  $\beta/\beta/\alpha$  layers with four-stranded orthogonal  $\beta$ -sheets (Figure 6). It assumes the “phosphohistidine carrier domain” fold. The similar structure was found in the N-terminal domain of enzyme I of the PEP:sugar phosphotransferase system (Z-score is 10.3). It was also found in the glycerol dehydratase reactivase  $\alpha$  subunit (Z-score is 6.0), transferring receptor protein fragment (Z-score is 3.8), 3-isopropylmalate dehydratase small subunit

(Z-score is 3.4), pyruvate kinase (Z-score is 2.5), and aconitase complexed with sulfate (Z-score is 3.1) by a DALI structural similarity search. The histidine-containing phosphocarrier protein (HPr) is an independent protein having a function analogous to that of the central domain of PPDK (35, 36). However, its fold with segregated  $\alpha$  and  $\beta$  regions (37, 38) is different from that of PPDK. By checking the surface of HPr protein, it is also found to contain a large nonpolar face with a few hydrophilic residues surrounding its periphery. This face is used to interface HPr with other enzymes in the signal transduction pathway (39). It is a second example of a small, stable protein that has a hydrophobic face exposed to the solvent. The major distinction that can be made between the HPr and the central domain of PPDK is that the former docks with its receptor enzyme in an intermolecular process, whereas the central domain of PPDK docks in an intramolecular process.

The present study indicates that the N-terminal and central domains are not effective catalysts for the ATP/ $\text{P}_i$  partial reaction unless they are tethered by the N-terminal linker. The N-terminal domain construct (Tem340) used in this study has been shown in unpublished work (18) to be soluble, monomeric, and relatively stable ( $\Delta G = 2.8 \text{ kcal/mol}$ ) for unfolding in buffer at pH 7.0 (25  $^\circ\text{C}$ ), but the three-dimensional structure has not yet been determined. Therefore, we cannot rule out the possibility that the absence of collaboration between the engineered proteins is due to the non-native structure in Tem340. Regions of non-native or at least altered structure may also be present in Cent-I, as evidenced by its structurally disordered termini and the 2.2  $\text{\AA}$  RMSD from the X-ray structure which could be significant in catalysis. With this caveat in mind, we are still inclined to conclude that the binding forces between the two domains are not sufficient to populate a productive complex, which consists of Tem340 bound to  $\text{Mg}^{2+}$ , ATP,  $\text{P}_i$ , and Cent-I.

It follows that the domain–domain linkers play an essential role in optimizing domain–domain orientation and residence time. From previous studies (9) we know that the linkers do not direct domain–domain docking but rather function as well-designed leashes that restrict conformational space. The present studies have shown that the central domain is compatible with solvent and, therefore, that in the absence of strong surface attraction the central domain would dissociate from the terminal domains as it alternates between active sites. The PPDK domain–domain interface has a much smaller buried accessible area ( $\sim 170 \text{ \AA}^2$  originating from the N-terminal or central domain) than that of protein complexes for signal transduction (for instance,  $\sim 765 \text{ \AA}^2$  accessible surface area buried from HPr (39)). Thus, the “target” for the central domain might not be large enough to “hit” in the absence of the tethers. By fusing of the three domains, the efficiency of protein–protein docking is optimized to form a single enzyme catalyst, as opposed to a signal transduction system exemplified by the PTS.

## ACKNOWLEDGMENT

Dr. Brian Martin is gratefully acknowledged for carrying out N-terminal sequence analyses on the three central domain constructs. We also thank Professor Nick Pace for helpful advice regarding the measurement of the  $\Delta G$  for unfolding Cent-I.

## REFERENCES

- Wood, H. G., O'Brien, W. E., and Micheales, G. (1977) Properties of carboxytransphosphorylase; pyruvate, phosphate dikinase; pyrophosphate phosphofruktinase and pyrophosphate-acetate kinase and their roles in the metabolism of inorganic pyrophosphate, *Adv. Enzymol. Relat. Areas Mol. Biol.* 45, 85–155.
- Saavedra, E., Encalada, R., Pineda, E., Jasso-Chavez, R., and Moreno-Sanchez, R. (2005) Glycolysis in *Entamoeba histolytica*. Biochemical characterization of recombinant glycolytic enzymes and flux control analysis, *FEBS J.* 272, 1767–1783.
- Reeves, R. E., Menzies, R. A., and Hsu, D. S. (1968) The pyruvate-phosphate dikinase reaction. The fate of phosphate and the equilibrium, *J. Biol. Chem.* 243, 5486–5491.
- Wang, H. C., Ciskanik, L., Dunaway-Mariano, D., von der Saal, W., and Villafranca, J. J. (1988) Investigations of the partial reactions catalyzed by pyruvate phosphate dikinase, *Biochemistry* 27, 625–633.
- Mehl, A., Xu, Y., and Dunaway-Mariano, D. (1994) Energetics of pyruvate phosphate dikinase catalysis, *Biochemistry* 33, 1093–1102.
- Herzberg, O., Chen, C. C., Kapadia, G., McGuire, M., Carroll, L. J., Noh, S. J., and Dunaway-Mariano, D. (1996) Swiveling-domain mechanism for enzymatic phosphotransfer between remote reaction sites, *Proc. Natl. Acad. Sci. U.S.A.* 93, 2652–2657.
- Ye, D. (2001) Structure and Catalysis in *Clostridium Symbiosum* Pyruvate Dikinase. in *Chemistry*, University of New Mexico, Albuquerque, NM.
- Nakanishi, T., Nakatsu, T., Matsuoka, M., Sakata, K., and Kato, H. (2005) Crystal structures of pyruvate phosphate dikinase from maize revealed an alternative conformation in the swiveling-domain motion, *Biochemistry* 44, 1136–1144.
- Wei, M., Ye, D., and Dunaway-Mariano, D. (2001) Investigation of the role of the domain linkers in separate site catalysis by *Clostridium symbiosum* pyruvate phosphate dikinase, *Biochemistry* 40, 13466–13473.
- Ginsburg, A., and Peterkofsky, A. (2002) Enzyme I: the gateway to the bacterial phosphoenolpyruvate:sugar phosphotransferase system, *Arch. Biochem. Biophys.* 397, 273–278.
- Liao, D. I., Reiss, L., Turner, I., and Dotson, G. (2003) Structure of glycerol dehydratase reactivase: a new type of molecular chaperone, *Structure* 11, 109–119.
- Wang, J., and Chen, L. (2003) Domain motions in GroEL upon binding of an oligopeptide, *J. Mol. Biol.* 334, 489–499.
- Pocalyko, D. J., Carroll, L. J., Martin, B. M., Babbitt, P. C., and Dunaway-Mariano, D. (1990) Analysis of sequence homologies in plant and bacterial pyruvate phosphate dikinase, enzyme I of the bacterial phosphoenolpyruvate:sugar phosphotransferase system and other PEP-utilizing enzymes. Identification of potential catalytic and regulatory motifs, *Biochemistry* 29, 10757–10765.
- McGuire, M., Huang, K., Kapadia, G., Herzberg, O., and Dunaway-Mariano, D. (1998) Location of the phosphate binding site within *Clostridium symbiosum* pyruvate phosphate dikinase, *Biochemistry* 37, 13463–13474.
- Ames, J. B., Tanaka, T., Stryer, L., and Ikura, M. (1994) Secondary structure of myristoylated recoverin determined by three-dimensional heteronuclear NMR: implications for the calcium-myristoyl switch, *Biochemistry* 33, 10743–10753.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Pace, N. C., and Scholtz, J. M. (1996) *Measuring the Conformational Stability of a Protein*, 2nd ed., IRL Press, Oxford, U.K.
- Lin, Y. (2005) Mechanism of separate site catalysis of *clostridium symbiosum* pyruvate phosphate dikinase. in *Chemistry*, University of New Mexico, Albuquerque, NM.
- Clore, G. M., and Gronenborn, A. M. (1997) NMR structures of proteins and protein complexes beyond 20,000 M(r), *Nat. Struct. Biol.* 4, 849–853.
- Tanaka, T., Ames, J. B., Kainosho, M., Stryer, L., and Ikura, M. (1998) Differential isotope labeling strategy for determining the structure of myristoylated recoverin by NMR spectroscopy, *J. Biomol. NMR* 11, 135–152.
- Talluri, S., and Wagner, G. (1996) An optimized 3D NOESY-HSQC, *J. Magn. Reson. B* 112, 200–205.
- Muhandiram, D. R., Farrow, N. A., Xu, G., Smallcombe, S. H., and Kay, L. E. (1993) A gradient NOESY-HSQC Experiment for Recording NOESY Spectra of Proteins Dissolved in H<sub>2</sub>O, *J. Magn. Reson. B* 102, 317–321.
- Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy, *Biochemistry* 31, 1647–1651.
- Brünger, A. T. (1992) *X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
- Badger, J., Kumar, R. A., Yip, P., and Szalma, S. (1999) New features and enhancements in the X-PLOR computer program, *Proteins* 35, 25–33.
- Bagby, S., Harvey, T. S., Eagle, S. G., Inouye, S., and Ikura, M. (1994) NMR-derived three-dimensional solution structure of protein S complexed with calcium, *Structure* 2, 107–122.
- Neri, D., Szyperski, T., Otting, G., Senn, H., and Wuthrich, K. (1989) Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional <sup>13</sup>C labeling, *Biochemistry* 28, 7510–7516.
- Altieri, A. S., Hinton, D. P., and Byrd, R. A. (1995) Translational diffusion measured by NMR pulsed-field gradients, *J. Am. Chem. Soc.* 117, 7566–7567.
- Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between Nuclear Magnetic Resonance Chemical Shift and Protein Secondary Structure, *J. Mol. Biol.* 222, 311–333.
- Anglister, J., Grzesiek, S., Wang, A. C., Ren, H., Klee, C. B., and Bax, A. (1994) <sup>1</sup>H, <sup>13</sup>C, <sup>15</sup>N nuclear magnetic resonance backbone assignments and secondary structure of human calcineurin B, *Biochemistry* 33, 3540–3547.
- Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
- Cosenza, L. W., Bringaud, F., Baltz, T., and Vellieux, F. M. (2002) The 3.0 Å resolution crystal structure of glycosomal pyruvate phosphate dikinase from *Trypanosoma brucei*, *J. Mol. Biol.* 318, 1417–1432.
- Wei, M., Ye, D., Herzberg, O., and Dunaway-Mariano, D. (2000) Identification of domain-domain docking sites within *Clostridium symbiosum* pyruvate phosphate dikinase by amino acid replacement, *J. Biol. Chem.* 275, 41156–41165.
- Ye, D., Wei, M., McGuire, M., Huang, K., Kapadia, G., Herzberg, O., Martin, B. M., and Dunaway-Mariano, D. (2001) Investigation of the catalytic site within the ATP-grasp domain of *Clostridium symbiosum* pyruvate phosphate dikinase, *J. Biol. Chem.* 276, 37630–37639.
- Waygood, E. B. (1998) The structure and function of HPr, *Biochem. Cell. Biol.* 76, 359–367.
- Azuaga, A. I., Neira, J. L., and van Nuland, N. A. (2005) HPr as a model protein in structure, interaction, folding and stability studies, *Protein Pept. Lett.* 12, 123–137.
- Williams, D. C., Cai, M., Suh, J. Y., Peterkofsky, A., and Clore, G. M. (2005) Solution NMR structure of the 48-kDa IIAMannose-HPr complex of the *Escherichia coli* mannose phosphotransferase system, *J. Biol. Chem.* 280, 20775–20784.
- Maurer, T., Meier, S., Kachel, N., Munte, C. E., Hasenbein, S., Koch, B., Hengstenberg, W., and Kalbitzer, H. R. (2004) High-resolution structure of the histidine-containing phosphocarrier protein (HPr) from *Staphylococcus aureus* and characterization of its interaction with the bifunctional HPr kinase/phosphorylase, *J. Bacteriol.* 186, 5906–5918.
- Cornilescu, G., Lee, B. R., Cornilescu, C. C., Wang, G., Peterkofsky, A., and Clore, G. M. (2002) Solution structure of the phosphoryl transfer complex between the cytoplasmic A domain of the mannitol transporter IIMannitol and HPr of the *Escherichia coli* phosphotransferase system, *J. Biol. Chem.* 277, 42289–42298.
- Herzberg, O., Chen, C. C., Liu, S., Tempczyk, A., Howard, A., Wei, M., Ye, D., and Dunaway-Mariano, D. (2002) Pyruvate site of pyruvate phosphate dikinase: crystal structure of the enzyme-phosphonopyruvate complex, and mutant analysis, *Biochemistry* 41, 780–787.
- Barshop, B. A., Wrenn, R. F., and Frieden, C. (1983) Analysis of numerical methods for computer simulation of kinetic processes: development of KINSIM - a flexible, portable system, *Anal. Biochem.* 130, 134–145.