Newgard, C. B., Hwang, P. K., & Fletterick, R. J. (1989) *Crit. Rev. Biochem. Mol. Biol. 24*, 69–99.
Oikonomakos, N. G., Melpidou, A. E. & Johnson, L. N. (1985) *Biochim. Biophys. Acta 832*, 248–256.
Oikonomakos, N. G., Johnson, L. N., Acharya, K. R., Stuart, D. I., Barford, D., Hajdu, J., Varvill, K. M., Melpidou, A. E., Papageorgiou, T., Graves, D. J., & Palm, D. (1987) *Biochemistry 26*, 8381–8389.
Oikonomakos, N. G., Acharya, K. R., Stuart, D. I., Melpidou, A. E., McLaughlin, P. J., & Johnson, L. N. (1988) *Eur. J. Biochem. 173*, 569–578.
Praly, J. P., & Lemieux, R. U. (1987) *Can. J. Chem. 65*, 213–223.
Priestle, J. P. (1988) *J. Appl. Crystallogr. 21*, 572–576.
Sansom, M. S. P. (1983) D.Phil. Thesis, Oxford University, Oxford, U.K.
Sansom, M. S. P., Stuart, D. I., Acharya, K. R., Hajdu, J., McLaughlin, P. J., & Johnson, L. N. (1985) *J. Mol. Struct.: THEOCHEM 123*, 3–25.
Sprang, S. R., & Fletterick, R. J. (1979) *J. Mol. Biol. 131*, 523–551.
Sprang, S. R., Goldsmith, E. J., Fletterick, R. J., Withers, S. G., & Madsen, N. B. (1982) *Biochemistry 21*, 5364–5371.
Sprang, S. R., Acharya, K. R., Goldsmith, E. J., Stuart, D. I., Varvill, K., Fletterick, R. J., Madsen, N. B., & Johnson, L. N. (1988) *Nature 336*, 215–221.
Street, I. P., Armstrong, C. R., & Withers, S. G. (1986) *Biochemistry 25*, 6021–6027.
Street, I. P., Rupitz, K., & Withers, S. G. (1989) *Biochemistry 28*, 1581–1587.
Suguwara, Y., & Iwasaki, H. (1984) *Acta Crystallogr. C40*, 389–393.

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[†,‡]

David J. Pocalyko,[§] Lawrence J. Carroll,[§] Brian M. Martin,[⊥] Patricia C. Babbitt,[∥] and Debra Dunaway-Mariano*[,§]

Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742, Molecular Neurogenetics Unit, Clinical Neuroscience Branch, National Institute of Mental Health, Bethesda, Maryland 20892, and Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143

Received April 10, 1990; Revised Manuscript Received July 30, 1990

ABSTRACT: In this paper we report the amino acid sequence of pyruvate phosphate dikinase (PPDK) from *Bacteroides symbiosus* as determined from the nucleotide sequence of the PPDK gene. Comparison of the *B. symbiosus* PPDK amino acid sequence with that of the maize PPDK [Matsuoka, M., Ozeki, Y., Yamamoto, N., Hirano, H., Kamo-Murakami, Y., & Tanaka, Y. (1988) *J. Biol. Chem. 263*, 11080] revealed long stretches of homologous sequence (>70% identity), which contributed to an overall sequence identity of 53%. The circular dichrosim spectra, hydropathy profiles, and calculated secondary structural elements of the two dikinases suggest that they may have very similar tertiary structures as well. A comparison made between the amino acid sequence of the maize and *B. symbiosus* dikinase with other known protein sequences revealed homology, concentrated in three stretches of sequences, to a mechanistically related enzyme, enzyme I of the *Escherichia coli* PEP:sugar phosphotransferase system [Saffen, D. W., Presper, K. A., Doering, T. L., Roseman, S. (1987) *J. Biol. Chem. 262*, 16241]. It is proposed that (i) these three stretches of sequence constitute the site for PEP binding and catalysis and a possible site for the regulation of enzymatic activity and (ii) the conserved sequences exist in a third mechanistically related enzyme, PEP synthase.

Pyruvate phosphate dikinase (PPDK)[1] catalyzes the reversible phosphorylation of pyruvate and orthophosphate with the β- and γ-phosphoryl groups of a single molecule of ATP (Reeves et al., 1968; Evans & Wood, 1968):

$$\text{Ade } \overset{\beta\gamma}{\text{PPP}} + \text{P}_\text{i} + \text{pyruvate} \rightleftharpoons \text{Ade P} + \overset{\gamma}{\text{PP}_\text{i}} + \overset{\beta}{\text{PEP}}$$

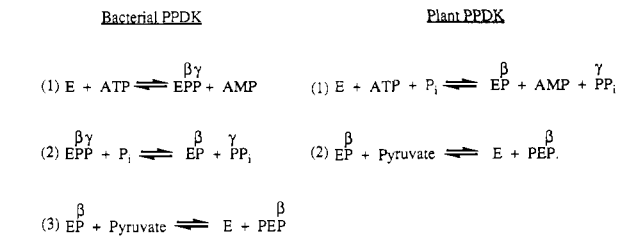
The enzyme has been found in a variety of unicellular organisms and in C_4 and some Crassulacean acid metabolism plants (Reeves, 1968; Reeves et al., 1968; Evans & Woods, 1968; Benzimam & Palgi, 1970; Hatch & Slack, 1968; Kluge & Osmond, 1971). In *Entamoeba histolytica* and *Bacteroides symbiosus* where pyruvate kinase is absent, PPDK functions in the direction of ATP synthesis. In *Propionibacterium shermanii*, *Acetobacter xylinum*, the photosynthetic bacterium, *Rhodospirillum rubrum*, and the C_4 and Crassulacean acid

[†] This work was supported by NIH Grants GM-36260 and GM-28688.
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J05295.
* To whom correspondence should be addressed.
[§] University of Maryland.
[⊥] National Institute of Mental Health.
[∥] University of California.

[1] Abbreviation ATP, adenosine 5′-triphosphate; PEP, phosphoenolpyruvate; PP_i, inorganic pyrophosphate; P_i, orthophosphate; EP, phosphoenzyme; EPP, pyrophosphoenzyme; ADP, adenosine 5′-diphosphate; AMP, adenosine 5′-monophosphate; PPDK, pyruvate phosphate dikinase; PTS, PEP:sugar phosphotransferase system; BRFP, bifunctional regulatory protein; HPLC, high-pressure liquid chromatography; CD, circular dichroism.

0006-2960/90/0429-10757$02.50/0

Scheme I: Mechanism Proposed for the Bacterial PPDK[a] and for the Plant PPDK[b,c]



[a] Milner et al, 1978; Milner & Wood, 1976. [b] Sugiyama, 1973; Andrews and Hatch, 1969; Jenkins and Hatch, 1985. [c] Where E represents the free enzyme, EPP represents the pyrophosphoryl enzyme intermediate and EP represents the phosphoryl enzyme intermediate.

metabolism plants PPDK is responsible for PEP production.

Studies of the structure, regulation, and mechanism of action of PPDK have focused primarily on the sugar cane and maize enzymes and on the dikinases isolated from the bacteria *P. shermanii* and *B. symbiosus*. Emerging from these studies are two fundamentally distinct profiles for plant PPDK versus bacterial PPDK. Studies of the mechanism of action of the two bacterial enzymes carried out by Wood and co-workers (Milner et al., 1978; Milner & Wood, 1976) and of the sugarcane and maize dikinases carried out by Sugiyama (1973) and by Hatch and co-workers (Andrews & Hatch, 1969; Jenkins & Hatch, 1985) have lead to the proposal of a three-step, Tri Uni Uni mechanism for the bacterial dikinase and a two-step, Bi Bi Uni Uni mechanism for the plant dikinase (see Scheme I).

The plant PPDK has also been distinguished from the *P. shermanii* and *B. symbiosus* dikinases on the basis of its in vivo regulation. In C_4 plants, PEP serves as the site of CO_2 fixation during photosynthesis and, therefore, PEP formation by PPDK is under strict light/dark control (Hague et al., 1983; Ashton & Hatch, 1983; Ashton et al., 1984; Burnell & Hatch, 1983; Edwards et al., 1985; Sheen & Bogorad, 1987; Hudspeth et al., 1986; Roeske et al., 1988). PPDK activity at the protein level is regulated by a phosphorylation/dephosphorylation mechanism that is mediated by a bifunctional regulatory protein (BFRP). BFRP-catalyzed phosphorylation of the regulatory threonine residue of PPDK with ADP leads to inactivation and BFRP-catalyzed dephosphorylation of the inactivated enzyme with P_i restores catalytic activity. No such activity control has been reported to exist for the dikinases of *P. shermanii* or *B. symbiosus*.

Ongoing studies in our laboratory are examining the structure and catalytic mechanism of PPDK. In this paper we report the amino acid sequence of the *B. symbiosus* PPDK as deduced from the sequence of the gene encoding this protein. A comparison is made between the amino acid sequence of the bacterial enzyme and the maize enzyme (Matsuoka et al., 1988) and the high degree of similarity observed is examined in relationship to the structure and catalytic functioning of these two distinct enzymes. Emerging from this analysis are interesting implications for PPDK catalysis and regulation that provide the basis for the recognition of a possible catalytic and regulatory motif in two other PEP-utilizing enzymes, *Escherichia coli* PEP synthase and enzyme I of the *E. coli* PEP: sugar phosphotransfer system (PTS).

Materials and Methods

Materials

Maize PPDK was a generous gift from Dr. Raymond Chollet, Department of Biochemistry, University of Nebraska. *E. coli* strain JM101 and vectors M13mp18 and M13mp19 were generous gifts from Dr. John Gerlt, Department of Chemistry, University of Maryland. *E. coli* strain JM83 was purchased from the American Type Culture Collection. Sequenase, trypsin, and buffers were obtained from United States Biochemical Corp. Restriction enzymes were purchased from Promega or Bethesda Research Laboratories. The preparation of pUC19 containing the 2.3-kbp *Hin*dIII fragment and pACYC184 containing the 3.6-kbp *Eco*RI fragment is described by Pocalyko et al. (1990). Oligonucleotide primers were prepared at the University of Maryland protein/DNA center with a Biosearch DNA synthesizer (Model 8750). Chemicals were purchased from Sigma Chemical Co. or Aldrich Chemical Co.

Methods

Preparation and Sequence Analysis of Tryptic Peptides. *B. symbiosus* PPDK (4 mg), purified according to Wang et al. (1988), was denatured in a 1-mL solution of 6 M urea and 0.4 M Tris·HCl (pH 8.0) at 37 °C over a period of 5 h. The protein solution was diluted with 3 mL of 0.4 M Tris·HCl (pH 8.0) and then made 2% (w/w) in trypsin. Subsequent additions of trypsin, each corresponding to a level of 2% (w/w), were made every 12 h over a 36-h period. The crude tryptic digest was fractionated by using three different chromatographic procedures. Peptides 1 and 2 were first separated on a C-18 reverse-phase HPLC (Alltech; 250 × 4.6 mm) using a linear gradient of 0.05% phosphoric acid in H_2O (A) and 0.05% phosphoric acid in CH_3CN (B) (0% B to 60% B in 60 min; column flow rate of 1 mL/min). All elution profiles were monitored at 220 nm with a UV detector. A mixture of peptides eluting between 36 and 38.2 mL was collected and then rechromatographed on a C-8 reverse-phase HPLC column (Alltech; 150 × 4 mm) using a linear gradient of 0.05% TFA in H_2O (A) and 0.05% TFA in CH_3CN (B) (0% B to 60% B in 60 min; column flow rate of 1 mL/min). Peptide 1 eluted at 19.0 mL and peptide 2 eluted at 19.6 mL. Peptides 3–6 were obtained by first chromatographing the crude tryptic digest on a C-8 reverse-phase HPLC column (Alltech; 150 × 4.6 mm) with a linear gradient of 0.05% TFA in H_2O (A) and 0.05% TFA in CH_3CN (B) (0% B to 60% B in 60 min; column flow rate of 1 mL/min). Peptides 3, 4, 5, and 6 eluted at 23.0, 30.8, 33.2, and 35.6 mL, respectively. The peptides were further purified by chromatographing them separately on the C-8 reverse-phase column with a more shallow gradient (0% B to 60% B in 120 min; column flow rate 1 mL/min). Peptide 3, 4, 5, and 6 eluted at 24.2% B, 30.4% B, 31.5% B, and 32.5% B, respectively. Peptide 7 was separated from the crude tryptic digest by using a C-18 reverse-phase HPLC column (Vydac; 250 × 4.6 mm) and a linear gradient of 0.05% TFA in H_2O (A) and 0.05% TFA in CH_3CN (B) (0% B to 60% B in 60 min; column flow rate of 1 mL/min). Peptide 7 eluted at 47 mL. The purified peptides and the N-terminal region of PPDK were sequenced by Edman's degradation with an Applied Biosystems 470A gas-phase protein sequencer. The sequences determined for the peptides are as follows: peptide 1, Val-Tyr-Phe-Thr-Ala-Asp-Glu-Ala-Lys; peptide 2, Phe-Ala-Tyr-Asp-Ser-Tyr-Arg; peptide 3, Val-Asp-Glu-Leu-His-Glu-Phe-Asn-Pro-Met-Met-Gly-His; peptide 4, Ser-Leu-Asp-Gln-Leu-Leu-His-Pro-Thr-Phe-Asn-Pro-Ala-Ala-Leu; peptide 5, Met-Asn-Asp-Ile-Pro-Gly-Asp-Trp-Gly-Thr-Ala-Val-Asn-Val-Gln-Thr-Met-Val; peptide 6, Glu-Glu-Thr-Gly-Ile-Asp-Ile-Val-Pro-Glu; peptide 7, Gln-Ile-Thr-Gln-Glu-Ile-Gln-Asp-Gln-Ile-Phe-Glu-Ala-Ile-Thr; N-terminal, Ala-Lys-Trp-Val-Tyr-Lys-Phe-Glu-Glu-Gly. In Table I (supplementary material) the quantities of the amino acids obtained during each cycle of the peptide sequencer are summarized.

DNA Sequencing. A single strand of a 2.3-kbp *Hin*dIII fragment containing ca. 1 kbp of the C-terminal region of the gene was sequenced first, followed by sequence analysis of both strands of the 3.6-kbp *Eco*RI fragment containing the entire PPDK coding region. The 2.3-kbp *Hin*dIII insert was excised from the cloning plasmid, pUC19, and cut into 1.3- and 1.0-kbp fragments with *Pst*I. The *Pst*I restriction fragments were subcloned into the cloning vectors M13mp18 and M13mp19. Nucleotide sequence was determined by the dideoxy chain termination method (Sanger et al., 1977) using a modified form of T7 DNA polymerase (Tabor & Richardson, 1987) and M13 universal and custom made primers. The 3.6-kbp *Eco*RI fragment was sequenced in the double-stranded plasmid pACYC184 by using the supercoil sequencing method (Chen & Seeburg, 1985; Hattori & Sakaki, 1986).

Primary and Secondary Structure Analysis. Hydropathy plot analysis of the PPDK protein sequence was conducted according to Hopp and Woods (1981) using the computer program MicroGenie (Beckman Instruments). The secondary structure was predicted by the method of Chou and Fasman (1978) using the computer program MicroGenie. The PPDK sequence was compared with protein sequences contained in the NBRF data base (release no. 23) by using the programs in the EuGene sequence package (Lawrence & Goldman, 1988; Pearson & Lipman, 1988). The alignments of the PPDK sequences and enzyme I of the bacterial PEP:sugar phosphotransferase system were generated by using the algorithm of Smith and Smith (1990). Statistical analysis of the alignment was performed by serially randomizing the enzyme I sequence 25 times and aligning these sequences with the PPDK sequences by using the same conditions as those used to generate the authentic alignment. CD spectra of the buffered solutions of *B. symbiosus* and maize PPDK (0.15 μM enzyme in 5 mM potassium phosphate, pH 6.9) were recorded at ambient temperature with a JASCO 500-C spectropolarimeter equipped with a microcell.

Results and Discussion

DNA and Protein Sequence Determination. Our first attempt to clone and express the *B. symbiosus* PPDK gene in *E. coli* resulted in the cloning of a 2.3-kbp *Hin*dIII fragment (in the plasmid, pUC19) but did not lead to expression. The 2.3-kbp fragment was cut into 1.3- and 1.0-kbp pieces with the restriction enzyme *Pst*I and these were subcloned into M13 vectors for sequencing. The 1.3-kbp piece was found to contain $\sim$1 kbp of the C-terminal region of the PPDK gene. From this sequence a probe was designed that ultimately led us to target a 3.6-kbp fragment generated in an *Eco*RI DNA digest for cloning. By using the low copy number plasmid pACYC184 as the vector we successfully cloned and expressed the PPDK gene in *E. coli*. However, because the level of expression of the PPDK gene, off its own promoter, turned out to be phenomenally high (>50% of the cell protein is PPDK)[2] standard (high-copy) sequencing vectors could not be cloned and therefore could not be used in sequencing the remainder of the PPDK gene. We instead elected to use the primer extension approach directly on the pACYC184 plasmid containing the 3.6-kbp *Eco*RI insert. The sequencing strategy used is illustrated in Figure 1. The sequences of the primers used and their location along the template strand are provided in Table II of the supplementary material.

The sequence of the *Eco*RI fragment shown in Figure 2 includes an open reading frame of 2520 nucleotides starting with the initiation codon ATG and ending with a termination codon TAG at position 2592. This region encodes a protein of 840 amino acids. The N-terminal region of PPDK, the seven PPDK tryptic peptides, and the active-site peptide (Goss et al., 1980) all coincided with the amino acid sequence derived from the gene sequence. The N-terminal Met, however, seems to have been removed from the enzyme following translation. On the basis of the sequence, the subunit molecular weight of the bacterial PPDK is 93 126, which compares closely with the reported subunit weight of 94 000 that was measured by using SDS–PAGE techniques (Goss et al., 1980).

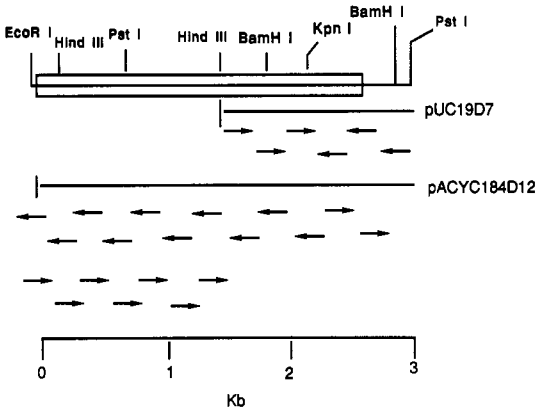


FIGURE 1: DNA sequencing scheme for the 1.3-kbp *Hin*dIII/*Pst*I fragment in M13 and the 3.6-kbp *Eco*RI fragment in pACYC184. The boxed region shows the coding region for PPDK. See the Methods section for details.

Included in the *Eco*RI fragment is a 70 base pair stretch upstream from the start codon that provides for the expression of the PPDK gene in transformed *E. coli*. Within this 70 base pair segment is a 12 base pair stretch of purines that contains the Shine-Dalgarno sequence. As indicated in Figure 2, regions that show similarity in sequence to the -35 and -10 promoter consensus sequences in *E. coli* can be observed in the remaining segment upstream from the ribosome binding site.

Sequence Homology and Secondary Structure Comparison with the Maize PPDK. PPDK from *B. symbiosus* was found to be slightly smaller than the PPDK from maize (840 versus 876 amino acids per subunit; 93 126 Da versus 95 353 Da per subunit). Although the two proteins differ in quaternary structure [the bacterial enzyme is α_2 (Milner et al., 1975) and the plant is α_4 (Sugiyama, 1973)], their primary structure is quite similar. The 53% overall sequence identity is not distributed uniformly throughout the protein but derives from long stretches of high similarity [viz., *B. symbiosus* PPDK resides 10 → 30 (75%), 86 → 117 (81%), 214 → 250 (72%), 396 → 412 (81%), 425 → 462 (81%), 553 → 568 (93%), and 741 → 786 (89%)] separated by intervals of comparatively low similarity (see Figure 3). The high degree of sequence similarity observed between the *B. symbiosus* and maize pyruvate phosphate dikinases indicates that their structural genes evolved from a common ancestral gene.

Whether the bacterial and plant dikinases are homologous in tertiary structure will not be precisely known until their crystal structures become available. However, at the present, a comparison of the circular dichroism spectra, hydropathy profiles, and calculated secondary structures for the two enzymes can be made to provide a rough indication of structural homology. The far-ultraviolet CD spectra of the two proteins (measured from 190 to 250 nm) are almost identical. Both spectra display a negative Cotton effect from 200 to 250 nm with the λ_{max} occurring at 220 nm ($[\theta] = -2.7 \times 10^4$ deg·

[2] Studies that are designed to test the basis for this high level of expression are in progress. The results from these investigations will be reported at a later date.

```
                                   GAA TTC TCA ATC CTT TGC TCA TCG CAG CAT ATC AAT GTT AAC ACA TAA ACT TTA GGA GGA AGA AAA CTT    69
                                           -35                               -10                              S.D.

ATG GCA AAA TGG GTT TAT AAG TTC GAA GAA GGC AAT GCA TCT ATG AGA AAC CTT CTT GGA GGC AAA GGC TGC AAC CTT GCA GAG ATG ACC   159
Met Ala Lys Trp Val Tyr Lys Phe Glu Glu Gly Asn Ala Ser Met Arg Asn Leu Leu Gly Gly Lys Gly Cys Asn Leu Ala Glu Met Thr

ATC TTA GGA ATG CCG ATT CCA CAG GGC TTT ACT GTA ACA ACA GAA GCT TGT ACA GAG TAC TAC AAC AGT GGA AAA CAG ATC ACA CAG GAA   249
Ile Leu Gly Met Pro Ile Pro Gln Gly Phe Thr Val Thr Thr Glu Ala Cys Thr Glu Tyr Tyr Asn Ser Gly Lys Gln Ile Thr Gln Glu

ATT CAG GAT CAG ATT TTC GAA GCT ATC ACA TGG TTA GAG GAA CTG AAC GGC AAG AAG TTC GGC GAC ACT GAA GAT CCG TTA TTA GTA TCT   339
Ile Gln Asp Gln Ile Phe Glu Ala Ile Thr Trp Leu Glu Glu Leu Asn Gly Lys Lys Phe Gly Asp Thr Glu Asp Pro Leu Leu Val Ser

GTA CGT TCC GCG GCC CGC GCA TCC ATG CCG GGT ATG ATG GAT ACC ATC CTG AAC CTT GGT TTA AAC GAC GTT GCA GTA GAG GGC TTT GCA   429
Val Arg Ser Ala Ala Arg Ala Ser Met Pro Gly Met Met Asp Thr Ile Leu Asn Leu Gly Leu Asn Asp Val Ala Val Glu Gly Phe Ala

AAG AAA ACG GGA AAT CCA AGA TTT GCA TAT GAT TCT TAC AGA AGA TTT ATC CAG ATG TAT TCC GAC GTA GTT ATG GAA GTT CCG AAG TCC   519
Lys Lys Thr Gly Asn Pro Arg Phe Ala Tyr Asp Ser Tyr Arg Arg Phe Ile Gln Met Tyr Ser Asp Val Val Met Glu Val Pro Lys Ser

CAT TTC GAG AAA ATC ATC GAT GCG ATG AAA GAA GAA AAG GGC GTT CAC TTC GAT ACA GAC CTG ACT GCC GAT GAT TTA AAA GAG CTG GCT   609
His Phe Glu Lys Ile Ile Asp Ala Met Lys Glu Glu Lys Gly Val His Phe Asp Thr Asp Leu Thr Ala Asp Asp Leu Lys Glu Leu Ala

GAG AAG TTC AAA GCT GTT TAC AAA GAG GCT ATG AAC GGC GAA GAG TTC CCA CAG GAG CCG AAG GAT CAG TTA ATG GGC GCT GTT AAA GCA   699
Glu Lys Phe Lys Ala Val Tyr Lys Glu Ala Met Asn Gly Glu Glu Phe Pro Gln Glu Pro Lys Asp Gln Leu Met Gly Ala Val Lys Ala

GTT TTC CGT TCC TGG GAC AAC CCT CGT GCA ATC GTA TAC CGC CGT ATG AAC GAT ATC CCT GGA GAC TGG GGT ACT GCA GTT AAC GTT CAG   789
Val Phe Arg Ser Trp Asp Asn Pro Arg Ala Ile Val Tyr Arg Arg Met Asn Asp Ile Pro Gly Asp Trp Gly Thr Ala Val Asn Val Gln

ACC ATG GTA TTT GGT AAC AAG GGC GAG ACC AGC GGT ACA GGC GTT GCC TTC ACA CGT AAC CCA TCC ACA GGT GAA AAA GGC ATC TAC GGT   879
Thr Met Val Phe Gly Asn Lys Gly Glu Thr Ser Gly Thr Gly Val Ala Phe Thr Arg Asn Pro Ser Thr Gly Glu Lys Gly Ile Tyr Gly

GAG TAC CTG ATC AAT GCA CAG GGC GAG GAC GTA GTT GCA GGT GTC CGC ACA CCA CAG CCT ATC ACC CAG TTA GAG AAC GAT ATG CCT GAC   969
Glu Tyr Leu Ile Asn Ala Gln Gly Glu Asp Val Val Ala Gly Val Arg Thr Pro Gln Pro Ile Thr Gln Leu Glu Asn Asp Met Pro Asp

TGC TAC AAG CAG TTC ATG GAT CTG GCC ATG AAG CTG GAG AAA CAT TTC CGT GAC ATG CAG GAT ATG GAG TTC ACA ATC GAG GAA GGT AAA  1059
Cys Tyr Lys Gln Phe Met Asp Leu Ala Met Lys Leu Glu Lys His Phe Arg Asp Met Gln Asp Met Glu Phe Thr Ile Glu Glu Gly Lys

TTA TAC TTC TTA CAG ACA CGT AAC GGC AAG AGA ACA GCT CCG GCT GCT CTT CAG ATT GCC TGC GAT TTA GTA GAC GAA GGC ATG ATC ACA  1149
Leu Tyr Phe Leu Gln Thr Arg Asn Gly Lys Arg Thr Ala Pro Ala Ala Leu Gln Ile Ala Cys Asp Leu Val Asp Glu Gly Met Ile Thr

GAG GAA GAG GCT GTT GTA AGA ATC GAA GCA AAA TCT CTT GAT CAG TTA CTT CAC CCG ACC TTC AAC CCG GCT GCT TTA AAG GCC GGC GAA  1239
Glu Glu Glu Ala Val Val Arg Ile Glu Ala Lys Ser Leu Asp Gln Leu Leu His Pro Thr Phe Asn Pro Ala Ala Leu Lys Ala Gly Glu

GTA ATC GGT TCC GCT CTT CCG GCA TCT CCT GGC GCA GCA GCA GGT AAA GTA TAC TTC ACC GCT GAT GAG GCT AAG GCT GCC CAC GAG AAG  1329
Val Ile Gly Ser Ala Leu Pro Ala Ser Pro Gly Ala Ala Ala Gly Lys Val Tyr Phe Thr Ala Asp Glu Ala Lys Ala Ala His Glu Lys

GGT GAG AGA GTT ATC CTT GTT CGT CTT GAG ACA TCT CCG GAA GAT ATC GAA GGT ATG CAT GCA GCC GAA GGT ATC CTG ACA GTG CGC GGC  1419
Gly Glu Arg Val Ile Leu Val Arg Leu Glu Thr Ser Pro Glu Asp Ile Glu Gly Met His Ala Ala Glu Gly Ile Leu Thr Val Arg GLY

GGT ATG ACA AGC CAT GCA GCC GTA GTT GCA CGT GGT ATG GGA ACA TGC TGC GTA TCC GGA TGC GGT GAG ATC AAG ATC AAC GAA GAA GCT  1509
GLY MET THR SER HIS ALA ALA VAL VAL ALA ARG Gly Met Gly Thr Cys Cys Val Ser Gly Cys Gly Glu Ile Lys Ile Asn Glu Glu Ala
         *       *

AAG ACA TTC GAA CTT GGC GGA CAC ACA TTT GCA GAG GGA GAT TAC ATC TCC TTA GAT GGT TCC ACA GGT AAG ATT TAC AAG GGC GAC ATC  1599
Lys Thr Phe Glu Leu Gly Gly His Thr Phe Ala Glu Gly Asp Tyr Ile Ser Leu Asp Gly Ser Thr Gly Lys Ile Tyr Lys Gly Asp Ile

GAG ACT CAG GAA CGT TCC GTA AGC GGA AGC TTC GAG CGT ATC ATG GTA TGG GCT GAC AAG TTC AGA ACA TTA AAG GTT CGT ACA AAT GCC  1689
Glu Thr Gln Glu Arg Ser Val Ser Gly Ser Phe Glu Arg Ile Met Val Trp Ala Asp Lys Phe Arg Thr Leu Lys Val Arg Thr Asn Ala

GAC ACA CCG GAA GAT ACA CTC AAT GCC GTT AAA CTG GGT GCA GAG GGC ATC GGT CTT TGC CGT ACA GAG CAT ATG TTC TTC GAG GCT GAC  1779
Asp Thr Pro Glu Asp Thr Leu Asn Ala Val Lys Leu Gly Ala Glu Gly Ile Gly Leu Cys Arg Thr Glu His Met Phe Phe Glu Ala Asp

AGA ATC ATG AAG ATC AGA AAG ATG ATC CTT TCC GAT TCA GTG GAA GCA AGA GAA GAG GCT CTG AAC GAA TTA ATC CCG TTC CAG AAG GGC  1869
Arg Ile Met Lys Ile Arg Lys Met Ile Leu Ser Asp Ser Val Glu Ala Arg Glu Glu Ala Leu Asn Glu Leu Ile Pro Phe Gln Lys Gly

GAT TTC AAG GCT ATG TAC AAA GCT CTG GAA GGC AGG CCA ATG ACG GTT CGC TAC CTG GAT CCG CCG CTG CAT GAG TTC GTT CCT CAT ACA  1959
Asp Phe Lys Ala Met Tyr Lys Ala Leu Glu Gly Arg Pro Met Thr Val Arg Tyr Leu Asp Pro Pro Leu His Glu Phe Val Pro His Thr

GAA GAG GAG CAG GCT GAA CTG GCT AAG AAC ATG GGC CTT ACT TTA GCA GAA GTA AAA GCA AAA GTT GAC GAA TTA CAC GAG TTC AAC CCA  2049
Glu Glu Glu Gln Ala Glu Leu Ala Lys Asn Met Gly Leu Thr Leu Ala Glu Val Lys Ala Lys Val Asp Glu Leu His Glu Phe Asn Pro

ATG ATG GGC CAT CGT GGC TGC CGT CTT GCA GTT ACC TAT CCG GAA ATT GCA AAG ATG CAG ACA AGA GCC GTT ATG GAA GCT GCT ATC GAA  2139
Met Met Gly His Arg Gly Cys Arg Leu Ala Val Thr Tyr Pro Glu Ile Ala Lys Met Gln Thr Arg Ala Val Met Glu Ala Ala Ile Glu

GTG AAG GAA GAG ACA GGA ATC GAT ATT GTT CCT GAG ATC ATG ATT CCG TTA GTT GGC GAG AAG AAA GAG CTT AAG TTC GTT AAG GAC GTA  2229
Val Lys Glu Glu Thr Gly Ile Asp Ile Val Pro Glu Ile Met Ile Pro Leu Val Gly Glu Lys Lys Glu Leu Lys Phe Val Lys Asp Val

GTT GTG GAA GTA GCT GAG CAG GTT AAG AAA GAG AAA GGT TCC GAT ATG CAG TAC CAC ATC GGT ACC ATG ATC GAA ATT CCT CGT GCA GCT  2319
Val Val Glu Val Ala Glu Gln Val Lys Lys Glu Lys Gly Ser Asp Met Gln Tyr His Ile Gly Thr Met Ile Glu Ile Pro Arg Ala Ala

CTC ACA GCA GAT GCC ATC GCT GAG GAA GCA GAG TTC TTC TCC TTC GGT ACA AAC GAC TTA ACA CAG ATG ACA TTC GGC TTC TCC CGT GAC  2409
Leu Thr Ala Asp Ala Ile Ala Glu Glu Ala Glu Phe Phe Ser Phe Gly Thr Asn Asp Leu Thr Gln Met Thr Phe Gly Phe Ser Arg Asp

GAC GCC GGC AAG TTC CTG GAT TCC TAC TAT AAA GCA AAA ATT TAT GAG TCC GAT CCA TTC GCA AGA CTT GAC CAG ACA GGC GTT GGC CAG  2499
Asp Ala Gly Lys Phe Leu Asp Ser Tyr Tyr Lys Ala Lys Ile Tyr Glu Ser Asp Pro Phe Ala Arg Leu Asp Gln Thr Gly Val Gly Gln

TTA GTA GAG ATG GCA GTT AAG AAA GGC CGT CAG ACA CGT CCG GGC CTT AAG TGC GGC ATC TGC GGC GAG CAC GGC GAG ATC CTT CTT CCG  2589
Leu Val Glu Met Ala Val Lys Lys Gly Arg Gln Thr Arg Pro Gly Leu Lys Cys Gly Ile Cys Gly Glu His Gly Glu Ile Leu Leu Pro

TAG AGT TCT GCC ACA AAG TAG GCC TGA ACT ATG TTT CCT GCT CAC CAT TCC GTG TGC CGA TCG CAC GTC TGG CAG CAG CAC AGG CAG CAT  2679
End

TAA ATA ATA AAT AAT TAG CAA TAA ATA AAG CGA AAG CCC CAG AGT TGC AGC TCT GGG GCT TTC TAT ATG CTT AG                       2753
```

FIGURE 2: Nucleotide sequence and deduced amino acid sequence of *B. symbiosus* PPDK. The N-terminal (10 residues) region and the regions corresponding to the seven tryptic peptides (sequenced by automated Edman degradation) are underlined as is the active-site peptide containing the catalytic histidine and upstream threonine residue. In the 5′-flanking region of the nucleotide sequence upstream from the ATG codon, the -35 and -10 regions of the putative promoter site and the Shine-Dalgarno sequence (SD) for the ribosome binding site are also indicated.

```
PPDK(M)     maasvsraicvqkpg skctrdreatsfarr svaaprpphakaagv irsdsgagrgqhcsp lravvdaapiqttkK rVfhFgkgksEGNkt    90
PPDK(B)     --------------- --------------- --------------- --------------- ------------maK rVykF----eEGNas    14
Enz. I      --------------- --------------- --------------- --------------- --------------- ---------------     0

PPDK(M)     MkeLLGGKGaNLAEM asiGlsvPpGFTVsT EACqqYqdaGcalpa glwaeIvdglqWvEE ymGatlGDpqrPLLl SVRSgAavSMPGMMD   180
PPDK(B)     MrnLLGGKGcNLAEM tilGmpiPqGFTVtT EACteYynsGkqitq eiqdqIfeaitWlEE lnGkkfGDtedPLLv SVRSaAraSMPGMMD   104
Enz. I      --------------- --------------- --------------- --------------- --------------- ---------------     0

PPDK(M)     TvLNLGLNDevaaGl AaKsGe-RFAYDSfR RFldMfgnVVMdiPr SlFEeklehMKEsKG lknDTDLTAsDLKEL vgqyKeVYlsA-kGE   268
PPDK(B)     TiLNLGLNDvaveGf AkKtGnpRFAYDSyR RFiqMysdVVMevPk ShFEkiidaMKEeKG vhfDTDLTAdDLKEL aekfKaVYkeAmnGE   194
Enz. I      --------------- --------------- --------------- --------------- --------------- ---------------     0

PPDK(M)     pFPsdPKkQLelAVl AVFnSWesPRAkkYR siNqItGlrGTAVNV QcMVFGNmGnTSGTG VLFTRNPnTGEKKlY GEfLVNAQGEDVVAG   358
PPDK(B)     eFPqePKdQLmgAVk AVFrSWdnPRAivYR rmNdIpGdwGTAVNV QtMVFGNkGeTSGTG VaFTRNPsTGEKgiY GEyLiNAQGEDVVAG   284
                                                                    *   *    *          *       *    *    *
Enz. I      --------------- --------------- --------------m isgilaspGiafGka lLlkedeividrKki sadqVdqevErflsG    46

PPDK(M)     i-RTPedLdamKNlM PqaYdElvEncnILE sHykeMQDiEFTvqE nrLwmLQcRtGKRTg ksAvkIAvDmVnEGl veprSAikmvEpghL   447
PPDK(B)     v-RTPqpitqleNdM PdcYkqfmdlamkLE kHfrdMQDmEFTiEE gkLyfLQtRnGKRTa paAlqIAcDlVdEGm iteeeAvvriEaksL   373
                   *   *         *  *   * *  *     * *   *    *        *      *   *   * **      **        *
Enz. I      rakasaqLetiKtka getfgE--EkeaIfE gHimlleDeEleqEi iaLi-------Kdkh mtAdaaAhe-ViEG- --qaSAleelddeyL   123

PPDK(M)     DQLLHPqFenpsayk dqVIatgLPASPGAA vGqVvFTAedAeAwH sqGkaaILVRaETSP EDvgGMHAAvGILTe RGGMTSHAAVVARww   537
PPDK(B)     DQLLHPtFnpaalka geVIGsaLPASPGAA aGkVyFTADeAkAaH ekGervILVRlETSP EDieGMHAAeGILTv RGGMTSHAAVVARgm   463
                                *                  **              *   **             *  *   ** ***    **
Enz. I      keraadvrdigkrll rnilGlkiidlsaiq devilvaAD------ -------LtpsETaq lnlkkv---lGfiTd aGGrTSHtsimARsl   197

PPDK(M)     GkCCVSGCsgIrvNd aeKlvtIGsHVlrEG ewlSLNGSTGevilG kqplsppalS-GdLg tfMaWvDdvRkLKVl ANADTPdDaLtArnN   626
PPDK(B)     GtCCVSGCGeIkiNe eaKtfelGgHtfaEG dyiSLdGSTGKiykG dietqersvS-Gsfe riMvWaDkfRtLKVr tNADTPeDtLnAvkl   552
                  * *             *   *          *    *                  *               *  **  *  *   *  *
Enz. I      elpaivGtGsvtsqv knddylIldaVnnqv yvnptNevidKmrav qeqvasekaelakLk dlpaitldghqveVc ANigTvrDvegAerN   287

PPDK(M)     GAqGIGLCRTEHMFF asDeRIkavRqMIma ptlElRqqALdrLlt yQrsDFefifrAmdG lPvTiRlLDhPsyEF LPegniEdivsELca   716
PPDK(B)     GAEGIGLCRTEHMFF eaD-RImkiRkMIls dsvEaReeALneLip fQkgDFKAVykAleG rPmTVRyLDpPlhEF vPht--EeeqaELak   639
            **** ** ***  *    *                     **       *    ***  *  *     **  *       **
Enz. I      GAEGvGLyRTEflFm drD------------ --------ALpt-ee eQfaayKAVaeAcgs qaviVRtmDiggdke LPym-----------   345

PPDK(M)     etGanqedalAriek LsEvNPMLGfRGCRL gisYPElteMQaRAI fEAAIamtnq-Gvqv fPEIMvPLVGtpqEL ghqvtlIrqVAEkVf   805
PPDK(B)     nmGltlaevkAkvde LhEfNPMmGhRGCRL AvtYPEIakMQtRAv mEAAIevkeetGidi vPEIMiPLVGekkEL kfvkdvvveVAEqVk   729
                              * ** ** *  ●  *    **   * ***   *                **  *      *        *
Enz. I      -------------nf pkEeNPfLGwRaiRi AmdrrEIlrdQlRAI lrAsafgklr----- ---IMfPmiisveEv ralrkeIeiykqelr   414

PPDK(M)     anv-GktigYkvGTM IEIPRAALvADeIAE qAEFFSFGTNDLTQM TFGySR-DDvGKFip vhlaqgILqhDPFev LDQRGVGeLVkfAtE   893
PPDK(B)     kEk-GsdmqYhIGTM IEIPRAALtADaIAE EAEFFSFGTNDLTQM TFGfSR-DDaGKFLd syykakIyesDPFar LDQtGVGqLVemAvk   817
             *         ** *  * * **   *   *  *  *** *******  *    *  *    *         *              *       *
Enz. I      dEgkafdesieIGvM vEtPaAAtiArhlAk EvdFFSiGTNDLTQy TlavdRgnDmishLy qpmspsvLnlikqvi dashaeGkwtgmcgE   504

PPDK(M)     rGRkaRPnLKvGICG EHGgepSsvaffaka gldfvscspfrvpia rlaaaqvlv                                         947
PPDK(B)     kGRqtRPgLKcGICG EHGeillp                                                                          840
                 *  *     *       *
Enz. I      lagdeRatLlllgmG ldefsmSaisiprik kiirntnfedakvla eqalaqpttdelmtl vnkfieektic                       575
```

FIGURE 3: Linear alignment of the amino acid sequences of PPDK from maize (PPDK(M)), PPDK from *B. symbiosus* (PPDK(B)), and enzyme I from *E. coli* (ENZ. I) with amino acid numbering shown in the right-hand margin. Amino acid residues are capitalized at sequence positions at which any two of the sequences are identified. Positions at which either PPDK sequence matches the enzyme I sequence are indicated by an asterisk.

cm^2/dmol for the maize PPDK and $[\theta] = -2.7 \times 10^4$ deg·cm^2/dmol for the *B. symbiosus* PPDK) and a shoulder centered at 204 nm for the *B. symbiosis* enzyme ($[\theta] = -2.1 \times 10^4$ deg·cm^2/dmol) or at 208 nm for the maize enzyme ($[\theta] = -2.5 \times 10^4$ deg·cm^2/dmol). The hydropathy profiles of the two enzymes sequences are aligned in Figure 4 with some gaps incorporated to provide maximum sequence overlap. In general, the profiles match quite well, possibly reflecting a similar pattern of secondary structure arrangements in the two proteins. The few regions of dissimilar hydropathy that do exist are found within stretches of both high homology and low homology and, for the most part, result from the occurrence of two consecutive charged amino acids in one protein and two uncharged amino acids at the same position in the other protein. Analysis of secondary structure (data not shown) by the method of Chou and Fasman (1978) indicated that both sequences display similar profiles of alternating α-helix and β-structure throughout the sequence. Taken together these three predictors are consistent with the idea that the *B. symbiosus* and maize enzymes may have quite similar tertiary structure.

Comparison of the Active-Site Structure and Catalytic Functioning of the Pyruvate Phosphate Dikinase from B. symbiosus and Maize. Historically, the dikinase from C_4 plants has been viewed to be distinctly different from the bacterial dikinase. This view was based in part on the differences known to exist in their quaternary structure (Sugiyama, 1973; Milner et al., 1975), stability (Sugiyama, 1973; South & Reeves, 1975; Milner et al., 1975; Hatch & Slack, 1975), and regulatory control (Hague et al., 1983; Ashton & Hatch, 1983; Ashton et al., 1984; Burnell & Hatch, 1983; Edwards et al., 1985; Sheen & Bogorad, 1987; Hudspeth et al., 1986; Roeske et al., 1988; Burnell, 1984) and in part on perceived differences in their mode of catalysis. To elaborate on this last point, the bacterial enzyme had been reported (Milner et al., 1978; Milner & Wood, 1976) to proceed via the three-step, Tri Uni Uni mechanism shown in Scheme I, while the plant enzyme had been reported (Sugiyama, 1973,

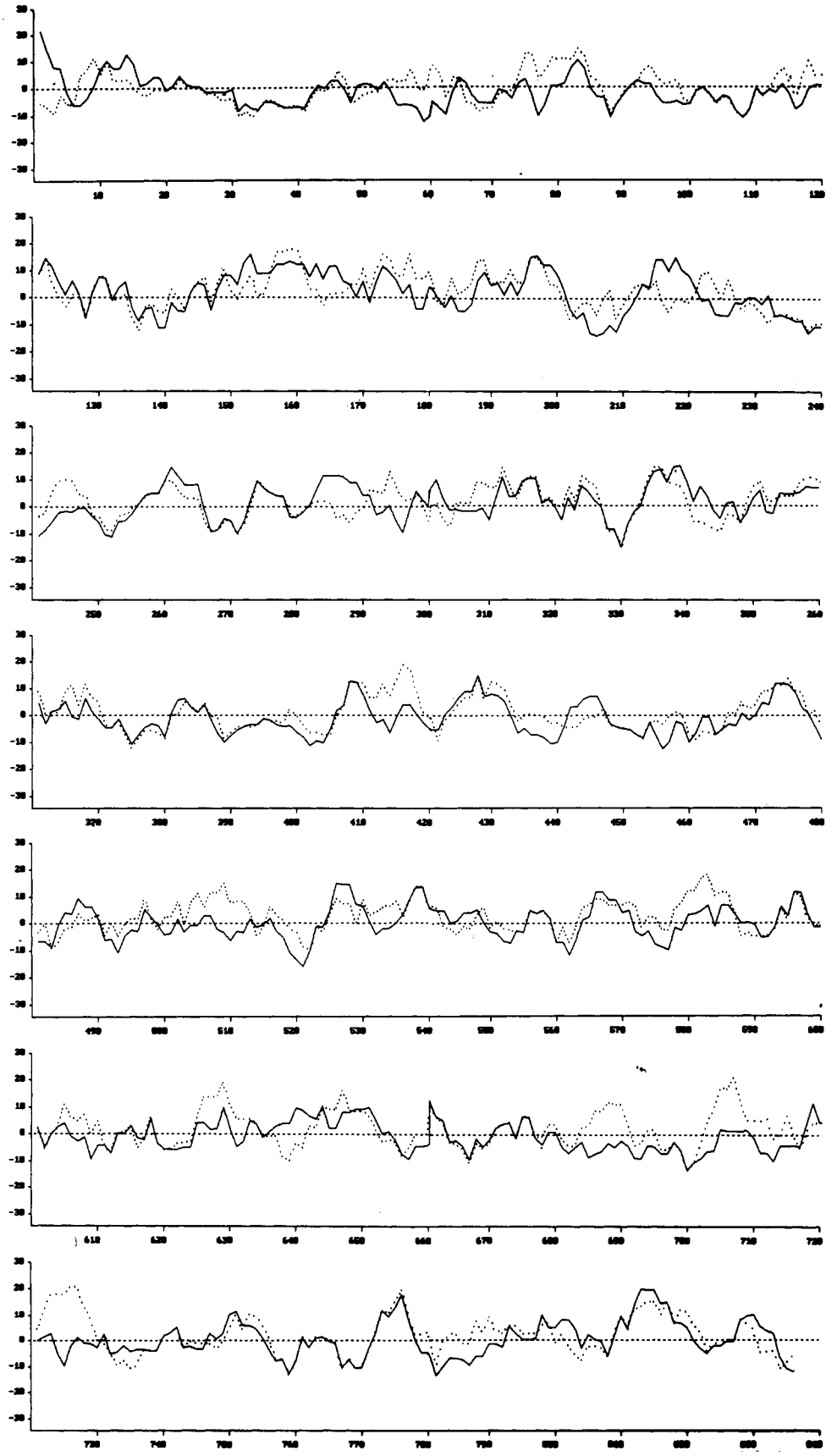


FIGURE 4: Comparison of hydropathy profiles of PPDK from *B. symbiosus* (···) and maize (—). Consecutive hydropathy averages are plotted for a four-residue window advancing from the N- to the C-terminals. Relative hydrophilicity (positive) and hydrophobicity (negative) were recorded in the range +30 to −30 hydropathy units. The two sequences were aligned by introducing gaps to maximize identities. The amino acid residue number in the hydropathy profile is based on the maize PPDK sequence.

Andrews & Hatch, 1969; Jenkins & Hatch, 1985) to proceed via the Bi Bi Uni Uni mechanism also depicted in Scheme I. The different kinetic mechanisms observed for the plant versus bacterial dikinases suggested that the chemical pathways leading to the respective phosphoenzyme intermediates (see Scheme I) may differ[3] or that in the case of the plant, but not

the bacterial enzyme, bound P_i is required for the alignment of catalytic groups that act on the first substrate, ATP.

Our own studies of PPDK lead us to quite different view of the relatedness of the plant and bacterial dikinases particularly with respect to their catalytic machinery. First, we have reexamined (Wang et al., 1988) the kinetic mechanism of the *B. symbiosus* PPDK (activated by Mg^{2+}) and found contrary to the earlier report (Milner et al., 1978) that this enzyme, like the plant dikinases (Sugiyama 1973; Andrews & Hatch, 1969; Jenkins & Hatch, 1985), proceeds via the Bi BI Uni Uni mechanism shown in Scheme I. Second, in most recent studies we have found that the phosphoenzyme intermediate of both the *B. symbiosus* (Carroll et al., 1989) and maize (Carroll et al., 1990) dikinase catalyzed reactions is formed via the pyrophosphoenzyme intermediate formed in turn from the E·ATP·P_i complex. Thus, operationally, the plant and bacterial dikinases are alike. Moreover, close examination of the amino acid sequences of the maize (Matsuoka et al., 1988) and *B. symbiosus* dikinases (Figure 3) reveals very high sequence identity in the region near the catalytic histidine [His 458 in the maize enzyme (Burnell, 1984; Roeske et al., 1988; Matsuoka et al., 1988) and His 455 in *B. symbiosus* (Goss et al., 1980)]. The catalytic histidine accepts the pyrophosphoryl group from ATP (Phillips & Wood, 1986; Carroll et al., 1989; Carroll et al., 1990) in the PEP-forming direction and the phosphoryl group from PEP (Spronk et al., 1976; Milner et al., 1978; Sugiyama, 1973; Andrews & Hatch, 1969; Carroll et al., 1989; Carroll et al., 1990) in the ATP-forming direction. The high degree of conservation of amino acid sequence observed near the catalytic histidine (amino acids 425–462 of the *B. symbiosus* enzyme) and in the regions that we believe make up the PEP binding site (see below) indicates, that the structures of the active sites of *B. symbiosus* and maize dikinase may be very similar.

Furthermore, in addition to the active-site histidine, the amino acid 425 → 462 region of the bacterial enzyme also contains a threonine residue (453), which is positioned two residues upstream from the catalytic histidine (at position 455). These residues are found in an identical arrangement in the plant PPDK (see Figure 3). In the plant enzyme the threonine residue serves as the site of regulation via phosphorylation by ADP (inactivation) and subsequent dephosphorylation by P_i (activation), both of which are catalyzed by a bifunctional regulatory protein (BFRP) (Edwards et al., 1985; Sheen & Bogorad, 1987; Hudspeth et al., 1986; Roeske et al., 1988). Communication between the histidine and threonine residues is evident from the reciprocal control based on phosphorylation states. Only when the histidine is phosphorylated is the threonine susceptible to phosphorylation by the BFRP and only when the histidine is not phosphorylated can the phosphothreonine be dephosphorylated. This seemingly fine-tuned active-site framework, which is present in both plant and bacterial enzymes, seems only in the plant enzyme to be exploited as a dual catalytic and regulatory center. An interesting question that arises is whether yet to be detected structural differences in the plant and bacterial enzyme stemming from the process of evolution are responsible for the functional regulatory motif existing in the plant enzyme but not the bacterial enzyme or whether a functional but unexploited (or, perhaps, undiscovered) regulatory center also exists in the bacterial enzyme.

Comparison of the Amino Acid Sequence of PPDK with Other PEP- and/or ATP-Utilizing Enzymes. A second point of interest that we have examined is whether the active site sequence ...Gly-Gly-Met-Thr-Ser-His-Ala-Ala-Val... found in the pyruvate phosphate dikinases of *B. symbiosus* and maize is present in other enzymes that are known to catalyze the transfer of a phosphoryl group from PEP via a phosphohistidine intermediate. A global search of reported protein sequences led to the identification of a mechanistically related enzyme, enzyme I of the bacterial PEP:sugar phosphotransferase system. Enzyme I is a protein phosphorylase that catalyzes the transfer of a phosphoryl group from PEP to its own active-site histidine and from there to the active-site histidine of HPr, the second enzyme of the phosphotransferase system (the phosphoryl group is ultimately transferred from HPr-P to enzyme II and, hence, to the transported sugar [for recent reviews on this system see Postma and Lengeler (1985) and Saier (1989)]). Enzyme I and PPDK thus share the capacity for binding PEP and catalyzing the transfer of the phosphoryl group from it to an active-site histidine. By comparing the primary structure of the PPDK from *B. symbiosus* (Figure 3) and maize (Matsuoka et al., 1988) with that of enzyme I from *E. coli* (Saffen et al., 1987), we discovered that an overall alignment of the three sequences can be generated. Statistical analysis of this alignment (Figure 3) shows that the alignment scores of enzyme I with both PPDK sequences are at least 10 standard deviations from the mean of analogous alignment scores generated by using serial radomizations of the enzyme I sequence in place of the authentic enzyme I sequence. These results suggest that it is "highly probable" that the proteins shown in Figure 3 are indeed homologous (Doolittle, 1986). In addition to the overall alignment, there are regions of the enzyme I sequence that exhibit particularly high sequence identity with the two PPDK sequences. The enzyme I region 184 → 189 (Gly-Gly-Arg-Thr-Ser-His) is of particular significance insofar as it contains the catalytic histidine[4] and it matches closely with the 450 → 455 region (Gly-Gly-Met-Thr-Ser-His) containing the catalytic histidine in the bacterial and maize PPDK's. The Gly-Gly-Arg-Thr-Ser-His active-site sequence is also found in enzyme I of *Streptococcus faecalis* (Alpert et al., 1985). Two other regions show very high sequence similarity to the PPDK sequences. They are enzyme I region 286 → 298 (10 of 12 residues identical between enzyme I and either PPDK sequence) and enzyme I region 448 → 460 (11 of 13 residues identical between enzyme I and either PPDK sequence). In view of the fact that the sequence homology between enzyme I and PPDK is highest in the vicinity of these three stretches of sequence, it is likely that these three regions are directly involved in PEP binding and the histidine-mediated phosphoryl transfer. In addition, there also exist the intriguing possibility (which to our knowledge has not been tested) that enzyme I activity in vivo might be under the control of a threonine phosphorylation/dephosphorylation mechanism.

The sequence homology observed between the plant and bacterial dikinases and enzyme I of the PTS led us to explore the possible relationship between these two enzymes and the bacterial enzyme PEP synthase. The reaction catalyzed by PEP synthase, ATP + pyruvate + H_2O → AMP + P_i + PEP, is analogous to the PPDK reaction and it, in a metabolic sense, links PEP synthase to the enzymes of the PTS [in fact, recent studies (Geerse et al., 1989; Chin et al., 1989) have shown that

[3] An alternate mechanistic pathway consists of the transfer of the γ-P of ATP to P_i to form E·ADP·PP_i followed by phosphorylation of the enzyme with the β-P of ADP to form EP·AMP·PP_i.

[4] The phosphorylated histidine has been experimentally identified in enzyme I of *S. faecalis* only (Alpert et al., 1985). The phosphorylated histidine of the *E. coli* enzyme I was identified by Saffen et al. (1987) on the basis of the sequence comparison.

the repressor of the fructose PTS, FruR, is required for the inactivation of PEP synthase at the level of transcription]. Common to PPDK, enzyme I, and PEP synthase (Narindrasorasak & Bridger, 1977) catalysis is the use of an active-site histidine residue to mediate phosphoryl transfer either from PEP to the enzyme or from the enzyme to pyruvate to generate PEP. As in the PPDK-catalyzed reaction, the β-P of ATP is, in the PEP synthase reaction, transferred to pyruvate with retention of configuration (Cook & Knowles, 1985). Whether or not the PEP synthase reaction actually proceeds via a pyrophosphoenzyme intermediate, as does the PPDK reaction, is currently under examination in our laboratory.

In view of the sequence homology that exists between PPDK and enzyme I and the chemical and, possibly, mechanistic similarities seen to exist between the reactions catalyzed by PPDK and PEP synthase, we expect to find the active-site segment and the two other conserved regions common to PPDK and enzyme I (see Figure 3) in PEP synthase. The PEP synthase gene has been cloned (Geerse et al., 1989) and the gene sequence is presently being determined in H.J. Hirsch's laboratory. Hirsch has communicated to us that while the full sequence has not yet been completed a segment reading ... Gly-Gly-Arg-Thr-Cys-His-Ala-Ala... has been found. If the His residue of this segment is indeed the catalytic histidine, this finding suggests that the catalytic/regulatory motif of PPDK and enzyme I does, in fact, also exist in PEP synthase.

Aside from enzyme I, the global search of reported protein sequences failed to identify statistically significant similarities between PPDK- and other PEP-utilizing enzymes (including pyruvate kinase, PEP carboxykinase, and EPSP synthase) or other enzymes that catalyze phosphoryl transfer via a phosphohistidine intermediate (e.g., HPr of the bacterial PTS). Hence, the homologous sequences that exist (see Figure 3) between PPDK, enzyme I, and, we predict, PEP synthase define a catalytic site specialized for PEP binding and phosphoryl transfer to a mediary histidine residue. This motif might also be eventually observed in the recently discovered enzyme, PEP phosphomutase (Bowman et al., 1988; Seidel et al., 1988; Hidaka et al., 1989). This enzyme, like PPDK, enzyme I, and PEP synthase, may use an active-site histidine residue to mediate phosphoryl group transfer, in this case, in the conversion of PEP to phosphonopyruvate.

Finally, the global search for reported sequences also failed to identify homology between PPDK and other ATP-utilizing enzymes, including members of the kinase family. A manual search for ATP binding site sequences (Serrano, 1988; Fry et al., 1985) within the PPDK structure did not locate one. The closest facsimile found is a Gly-X-X-Gly-X-Gly stretch (starting with residue 553 in the bacterial enzyme and 556 in the plant). This is a reverse ordering of the Gly-X-Gly-XX-Gly glycine triad conserved in the ATP binding domain throughout the protein kinase family (Hanks et al., 1988). Identification of the ATP binding site within the primary sequence of PPDK will have to await the results of affinity labeling experiments.

Acknowledgments

We are very grateful to Dr. H. J. Hirsch for communicating his sequence data on the *E. coli* PEP synthase prior to publication.

Supplementary Material Available

A table listing the quantities of the amino acids derived from the automated peptide sequencing of the *B. symbiosus* pyruvate phosphate dikinase N-terminal and seven tryptic peptides and a table of the oligonuleotide primers used in sequencing the *B. symbiosus* PPDK gene (2 pages). Ordering information is given on any current masthead page.

References

Alpert, C.-A., Frank, R., Stuber, K., Deutscher, J., & Hengstenberg, W. (1985) *Biochemistry 24*, 959.
Andrews, T. J., & Hatch, M. D. (1969) *Biochem. J. 114*, 117.
Ashton, A. R., & Hatch, M. D. (1983) *Biochem. Biophys. Res. Commun. 115*, 53.
Ashton, A. R., Burnell, J. N., & Hatch, M. D. (1984) *Arch. Biochem. Biophys. 230*, 492.
Benzimann, M., & Palgi, A. (1970) *J. Bacteriol. 104*, 211.
Bowman, E., McQueney, M., Barry, R. J., & Dunaway-Mariano, D. (1988) *J. Am. Chem. Soc. 110*, 5575.
Budde, R. J. A., Ernst, S. M., & Chollet, R. (1986) *Biochem. J. 236*, 579.
Burnell, J. N. (1984) *Biochem. Int. 9*, 683.
Burnell, J. N., & Hatch, M. D. (1983) *Biochem. Biophys. Res. Commun. 111*, 288.
Carroll, L. J., Mehl, A. F., & Dunaway-Mariano, D. (1989) *J. Am. Chem. Soc. 111*, 5965.
Carroll, L. J., Dunaway-Mariano, D., Smith, C. M., & Chollet, R. (1990) *FEBS Lett.* (in press).
Chen, E. J., & Seeburg, P. H. (1985) *DNA 4*, 165.
Chin, M. A., Feldheim, D. A., & Saier, M. H. (1989) *J. Biol. Chem. Soc. 171*, 2424.
Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. 47*, 45.
Cook, A. G., & Knowles, J. R. (1985) *Biochemistry 24*, 51.
Cooper, R. A., & Kornberg, H. L. (1974) *Enzymes (3rd Ed.)* 631.
Edwards, G. E., Nakamoto, H., Burnell, J. N., & Hatch, M. D. (1985) *Annu. Rev. Plant Physiol. 36*, 255.
Evans, H. J., & Woods, H. G. (1968) *Proc. Natl. Acad. Sci. U.S.A. 61*, 1448.
Fry, D. C., Kuby, S. A., & Mildvan, A. S. (1985) *Biochemistry 24*, 4680.
Geerse, R. H., van der Pluijm, J., & Postma, P. W. (1989) *Mol. Gen. Genet. 218*, 348.
Goss, N. H., Evans, C. T., & Wood, H. G. (1980) *Biochemistry 19*, 5805.
Hague, D. R., Uhler, M., & Collins, P. D. (1983) *Nucleic Acids Res. 11*, 4853.
Hanks, S. K., Quinn, A. M., & Hunter, T. (1988) *Science 241*, 42.
Hatch, M. D., & Slack, C. R. (1968) *Biochem. J. 106*, 141.
Hatch, M. D., & Slack, C. R. (1975) *Methods Enzymol. 42*, 212.
Hattori, M., & Sakaki, Y. (1986) *Anal. Biochem. 152*, 232.
Hidaka, T., Mori, M., Imai, S., Hara, O., Nagaoka, K., & Seto, H. (1989) *J. Antibiot. 42*, 491.
Hopp, T. P., & Woods, K. R. (1981) *P.N.A.S. 78*, 3824.
Hudspeth, R. L., Glackin, C. A., Bonner, J., & Grula, J. W. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2884.
Jenkins, C. L. D., & Hatch, M. D. (1985) *Arch. Biochem. Biophys. 239*, 53.
Kluge, M., & Osmond, C. B. (1971) *Naturwissenshaffen 58*, 414.
Lawrence, C. B., & Goldman, D. A. (1988) *CABIOS 4*, 25.
Matsuoka, M., Ozeki, Y., Yamamoto, N., Hirano, H., Kano-Murakami, Y., & Tanaka, Y. (1988) *J. Biol. Chem. 263*, 11080.
Milner, Y., & Wood, H. G. (1976) *J. Biol. Chem. 251*, 7920.
Milner, Y., Michaels, G., & Wood, H. G. (1975) *Methods Enzymol. 42*, 199.
Milner, Y., Michaelis, G., & Wood, H. G. (1978) *J. Biol. Chem. 253*, 878.

Milton, H., & Saier, M. H. (1989) *Microbiol. Rev. 53*, 109.
Narindrasorasak, S., & Bridger, W. A. (1977) *J. Biol. Chem. 252*, 3121.
Pearson, W. R., & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 185*, 244.
Phillips, N. F. B., & Wood, H. G. (1986) *Biochemistry 25*, 1644.
Pocalyko, D. J. (1990) Ph.D. Dissertation, University of Maryland.
Postma, P. W., & Lengeler, J. W. (1985) *Microbiol. Rev. 49*, 232.
Reeves, R. E. (1968) *J. Biol. Chem. 243*, 3203.
Reeves, R. E., Menzies, R. A., & Hsu, D. S. (1968) *J. Biol. Chem. 243*, 5486.
Roeske, C. A., Kutny, R. M., Budde, R. J. A., & Chollet, R. (1988) *J. Biol. Chem. 263*, 6683.
Saffen, D. W., Presper, K. A., Doering, T. L., & Roseman, S. (1987) *J. Biol. Chem. 262*, 16241.
Saier, M. H. (1989) *Microbiol. Rev. 53*, 109.
Sanger, F., Miklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463.
Seidel, H. M., Freeman, S., Seto, H., & Knowles, J. R. (1988) *Nature 335*, 457.
Serrano, R. (1988) *Biochim. Biophys. Acta 947*, 1.
Sheen, J. Y., & Bogorad, L. (1987) *J. Biol. Chem. 262*, 11726.
Smith, R. F., & Smith, T. F. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 118.
South, D. J., & Reeves, R. E. (1975) *Methods Enzymol. 42*, 187.
Spronk, A. M., Yoshida, H., & Wood, H. G. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 4415.
Sugiyama, T. (1973) *Biochemistry 12*, 2862.
Tabor, S., & Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4767.
Wang, H.-C., Ciskanik, L. Dunaway-Mariano, D., von der Saal, W., & Villafranca, J. J. (1988) *Biochemistry 27*, 625.

Synergistic Effects of Proton and Phenylalanine on the Regulation of Muscle Pyruvate Kinase†

Thomas G. Consler,‡ Michael J. Jennewein, Guang-Zuan Cai, and James C. Lee*,§
E. A. Doisy Department of Biochemistry and Molecular Biology, St. Louis University School of Medicine, St. Louis, Missouri 63104

Received June 13, 1990; Revised Manuscript Received August 17, 1990

ABSTRACT: Steady-state kinetic studies of muscle pyruvate kinase were conducted as a function of pH and phenylalanine concentrations. Results show that at a pH below 7.0, there is no observable effect of phenylalanine on the kinetic properties of muscle pyruvate kinase. When the results at a pH below 6.5 are used as the state for comparison, the kinetic results show that phenylalanine and proton exert a synergistic effect on the allosteric properties of the enzyme. A significantly greater change in Hill coefficients at high pH can be detected in the presence of phenylalanine than in its absence. To pinpoint the specific mechanism that leads to the synergistic effect, the kinetic data were resolved into the five equilibrium and two rate constants that characterize the basic two-state model. It can be shown that K_I^T, the binding constant of phenylalanine to the inactive T state, is strongly proton-linked. The affinity of phenylalanine for the T state increases with increasing pH. When the pH dependence of K_I^T was analyzed by the linked-function theory [Wyman, J. (1964) *Adv. Protein Chem. 19*, 224–285], it was shown that deprotonation favors phenylalanine binding to the T state. K_I^R (the binding constant of phenylalanine to the active R state), K_S^T (the binding constant of substrate to the T state), and L (the isomerization constant of the two states) not only are all weakly proton-linked but also it was shown that protonation favors the ligand–pyruvate kinase complex. K_S^R, the binding constant of substrate for the R state, shows no observable linkage to proton concentration. Thus, pH exhibits differential effects on these equilibrium constants both qualitatively and quantitatively. Knowing the proton linkage relationships, it is possible to conclude that the synergistic effect of phenylalanine and proton can be explained by the interplay among the strong proton-linked effect on the affinity of phenylalanine to the T state and the apparently weak or insignificant proton linkage in the other equilibrium parameters.

The basic molecular mechanism of regulation for muscle pyruvate kinase (PK)[1] has yet to be elucidated, although results from numerous studies have helped to establish a correlation between enzyme conformation and its function. It has been reported that PK undergoes a conformational change upon substrate binding (Kayne & Suelter, 1965; Mildvan & Cohn, 1965, 1966). These conformational transitions in PK can be affected by temperature and pH changes. These changes can be monitored by spectroscopic and hydrodynamic measurements (Kayne & Suelter, 1965, 1968). The structural transition(s) can be reversed by Phe, an allosteric inhibitor (Consler

† Supported by NIH Grants NS-14269 and DK-21489.
* Address correspondence to this author.
‡ Present address: Molecular Biology Institute, University of California, Los Angeles, CA 90024-1570.
§ Present address: Department of Human Biological Chemistry and Genetics, The University of Texas Medical Branch, Galveston, TX 77550.

[1] Abbreviations: PK, pyruvate kinase; TKM buffer, 50 mM Tris buffer that contains 72 mM KCl and 7.2 mM $MgSO_4$; PEP, phosphoenolpyruvate.

0006-2960/90/0429-10765$02.50/0