JOURNAL OF BACTERIOLOGY, May 1991, p. 3117–3127
0021-9193/91/103117-11$02.00/0

Vol. 173, No. 10

Nucleotide Sequence of the *Rhodobacter capsulatus fruK* Gene, Which Encodes Fructose-1-Phosphate Kinase: Evidence for a Kinase Superfamily Including Both Phosphofructokinases of *Escherichia coli*

LONG-FEI WU,[1] AIALA REIZER,[1] JONATHAN REIZER,[1] BING CAI,[2] JOHN M. TOMICH,[2] AND MILTON H. SAIER, JR.[1]*

Department of Biology, University of California at San Diego, La Jolla, California 92093-0116,[1] and The Medical Genetics Division, Children's Hospital of Los Angeles, Los Angeles, California 90027[2]

Received 13 August 1990/Accepted 11 March 1991

The *fruK* gene encoding fructose-1-phosphate kinase (FruK), located within the fructose (*fru*)-catabolic operon of *Rhodobacter capsulatus*, was sequenced. FruK of *R. capsulatus* (316 amino acids; molecular weight = 31,232) is the same size as and is homologous to FruK of *Escherichia coli*, phosphofructokinase B (PfkB) of *E. coli*, phosphotagatokinase of *Staphylococcus aureus*, and ribokinase of *E. coli*. These proteins therefore make up a family of homologous proteins, termed the PfkB family. A phylogenetic tree for this new family was constructed. Sequence comparisons plus chemical inactivation studies suggested the lack of involvement of specific residues in catalysis. Although the *Rhodobacter* FruK differed markedly from the other enzymes within the PfkB family with respect to amino acid composition, these enzymes exhibited similar predicted secondary structural features. A large internal segment of the *Rhodobacter* FruK was found to be similar in sequence to the domain bearing the sugar bisphosphate-binding region of the large subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase of plants and bacteria. Proteins of the PfkB family did not exhibit statistically significant sequence identity with PfkA of *E. coli*. PfkA, however, is homologous to other prokaryotic and eukaryotic ATP- and PP_i-dependent Pfks (the PfkA family). These eukaryotic, ATP-dependent enzymes each consist of a homotetramer (mammalian) or a heterooctamer (yeasts), with each subunit containing an internal duplication of the size of the entire PfkA protein of *E. coli*. In some of these enzymes, additional domains are present. A phylogenetic tree was constructed for the PfkA family and revealed that the bacterial enzymes closely resemble the N-terminal domains of the eukaryotic enzyme subunits whereas the C-terminal domains have diverged more extensively. The PP_i-dependent Pfk of potato is only distantly related to the ATP-dependent enzymes. On the basis of their similar functions, sizes, predicted secondary structures, and sequences, we suggest that the PfkA and PfkB families share a common evolutionary origin.

In earlier papers we presented the nucleotide sequences of two of the three structural genes within the *Rhodobacter capsulatus* fructose (*fru*) operon. These genes include *fruB(HI)*, encoding a multiphosphoryl transfer protein of the phosphoenolpyruvate:fructose phosphotransferase system (PTS) (46, 50, 61), and *fruA*, encoding the fructose permease, i.e., the fructose-specific enzyme II of the PTS (60). These studies revealed some unusual features of the *Rhodobacter* PTS proteins, which together catalyze the concomitant uptake and phosphorylation of fructose to yield fructose-1-phosphate (fructose-1-P). The further metabolism of fructose requires the ATP-dependent phosphorylation of fructose-1-P catalyzed by fructose-1-P kinase (FruK; ATP: D-fructose-1-phosphate-6-phosphotransferase, EC 2.7.1.56). This enzyme has been purified from *Escherichia coli* (7), and after this paper was submitted for publication, the nucleotide sequence of its structural gene in *E. coli* was published and the homology of the encoded protein with other carbohydrate kinases was noted (43). In this paper we report the complete sequence of the *fruK* gene encoding FruK of *R. capsulatus*.

Two phosphofructokinases (Pfks) have been identified in *E. coli*. The major Pfk (PfkA) is a tetramer consisting of four identical subunits (subunit molecular weight = 35,000) (58). The enzyme shows cooperative kinetics with respect to its substrate, fructose-6-phosphate, and is subject to allosteric control by phosphoenolpyruvate (6). The crystal structures of PfkA and of the Pfk from *Bacillus stearothermophilus* have been determined (19, 26). These two enzymes, very similar in sequence and structure, are homologous to two internally repeated domains in both the α- and β-subunits of the ATP-dependent octameric yeast Pfk, the two internally repeated domains of the single type of subunit of the mammalian ATP-dependent Pfks, and the two dissimilar subunits (α and β) of the potato PP_i-dependent Pfk (8, 25, 45). The family of these homologous proteins is here designated the PfkA family. The minor Pfk of *E. coli* (PfkB) is a homodimer (subunit molecular weight of about 33,000) which can associate to a less active homotetramer. It does not show cooperative kinetics or inhibition by phosphoenolpyruvate (4, 31). It lacks immunological cross-reactivity with PfkA (30) and has not been noted to show sequence similarity with members of the PfkA family (4, 25).

In this report, FruK of *R. capsulatus* is shown to be homologous to FruK and PfkB of *E. coli* as well as other bacterial kinases (the PfkB family). The predicted secondary structures of FruK and PfkB are shown to be similar to the known secondary structure of PfkA. Phylogenetic trees for

* Corresponding author.

both the PfkA and PfkB families are constructed. The reported results lead us to suggest that the PfkA and the PfkB families share a common origin but diverged from each other long before the genes encoding the individual members of either of these families diverged.

MATERIALS AND METHODS

Materials. Restriction endonucleases, T4 DNA ligase, and T4 DNA polymerase were obtained from Bethesda Research Laboratories, Gaithersburg, Md., or Boehringer Mannheim Biochemicals, Inc. The T7 polymerase sequencing kit was from Pharmacia. The deoxyadenosine 5′-([α-^{35}S]thio)triphosphate (1,335 Ci/mmol) was purchased from the New England Nuclear Corp., Boston, Mass. Isopropyl-β-D-thiogalactopyranoside, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal), and lysozyme were purchased from Boehringer Mannheim Biochemicals. Acrylamide, bisacrylamide, and *N,N,N′,N′*-tetramethylethylenediamine were from Bio-Rad Laboratories, and agarose was from Bethesda Research Laboratories. Deoxynucleotides and dideoxynucleotides were obtained from Pharmacia and P-L Biochemicals, Inc. All other chemicals and enzymes used were of the highest quality available commercially.

Bacterial strains, plasmids, and phage. *E. coli* TG1 was provided by T. Gibson of the Laboratory of Molecular Biology, Medical Research Council, Cambridge, United Kingdom (9), and strain XL1-Blue was obtained from Stratagene. *R. capsulatus* 37b4 was as described previously (13). The pBluescript SK(+) and KS(+) plasmids as well as helper phage VCS-M13 were obtained from Stratagene.

Synthesis of oligodeoxynucleotides. Oligodeoxynucleotides were synthesized from β-cyanoethyl phosphoramidite precursors on a model 380B Applied Biosystems DNA synthesizer. The oligonucleotides were synthesized as trityl-on derivatives which were deblocked at 58°C for 12 h in concentrated ammonium hydroxide. These solutions were then applied to Applied Biosystems OPC oligonucleotide purification columns. The oligonucleotides were eluted as described by the manufacturer. These solutions were dried under vacuum.

Growth media and selection conditions. Transformants were selected on Luria-Bertani (LB) plates (Difco Laboratories) (38) containing ampicillin (50 μg/ml). For the subcloning of fragments into pBluescript SK(+) or KS(+), LB plates containing 1 mM isopropylthio-β-galactoside and 30 μg of X-Gal per ml were used.

DNA procedures. *E. coli* XL1-Blue and TG1 carrying the pBluescript plasmids were grown at 37°C in LB medium containing ampicillin (50 μg/ml). Single-stranded DNA from the Bluescript plasmids was prepared with helper phage VCS-M13 as described by the Stratagene manual. Small- and large-scale preparation of plasmid DNAs was accomplished by the method of Birnboim (5). Competent cells of *E. coli* TG1 and XL1-Blue were prepared by the $CaCl_2$ method (37).

Nucleotide sequencing was conducted with the T7 sequencing kit of Pharmacia with Deaza ^{35}Sequencing Mixes as described previously in order to overcome GC compression due to the high GC content of *Rhodobacter* DNA (66% overall; 70% for the *fru* operon). We reported previously the location of the *fruK* gene on an *Eco*RI-*Pst*I fragment of 2.0 kb (61). This fragment was subcloned into plasmids Bluescript SK(+) and KS(+). Universal and synthesized primers were used for the sequencing. The entire fragment was sequenced in both directions at least twice, and all sequences were overlapping in each direction.

Computer analyses. The Staden programs were used for statistical analyses and structural prediction studies (57). For homology screening, the Fasta program (44) was used with the combined NBRF (PIR protein sequence data base release 25.0 plus preliminary entries new release 43.0), a translation of GenBank DNA libraries (release 65.0), and EMBL sequence data base (release 19.0). Multialignment was conducted by employing the Newat program (20). The significance of homology was calculated by using the ALIGN program (28) as well as the Los Alamos program (28). The Genetics Computer Group Sequence Analysis Software Package (GCG Package) was used for calculation of codon usage and determination of open reading frames. These programs are available through the University of California, San Diego, Computer Center. Phylogenetic tree construction was described by Doolittle and Feng (17, 20).

Extract preparation and enzyme inactivation studies. Cells of *R. capsulatus* were grown in complex medium, and extracts were prepared as described previously (13). FruK was assayed by using a spectrophotometric assay (13). Treatment with *N*-ethylmaleimide, phenylmethylsulfonyl fluoride (PMSF), and diethylpyrocarbonate (DEPC) was conducted in 20 mM Tris-HCl buffer, pH 7.5, with the reagent concentrations at 30 mM for *N*-ethylmaleimide and at 2, 5, and 10 mM for PMSF and DEPC. Reagent solutions were prepared fresh immediately before use. Exposure of the enzyme of interest (*R. capsulatus* PTS or FruK) was carried out at 37°C for 1 h (59). Subsequently, excess dithiothreitol (for *N*-ethylmaleimide inactivation), serine (for PMSF inactivation), or histidine (for DEPC inactivation) was added, and incubation was continued for 30 min. The fructose PTS of *R. capsulatus* was assayed as described previously (13).

RESULTS

Nucleotide sequence of the *fruK* gene encoding FruK of *R. capsulatus*. Figure 1 presents the nucleotide sequence of the *R. capsulatus fruK* gene and flanking regions. Preceding the start codon of the *fruK* gene is a potential Shine-Dalgarno sequence, GGGGG, located 5 bp upstream of the ATG initiation codon (+1), within the end of the *fruB(HI)* gene which overlaps the *fruK* gene by 4 bp. Interestingly, in *E. coli* and *Salmonella typhimurium* there is a 1-bp overlap between the *fruB(MH)* and *fruK* genes. Such overlapping genes suggest translational coupling between *fruB* and *fruK*. The *fruK* open reading frame in *R. capsulatus* is 951 bp long, corresponding to 317 codons (316 amino acids in the protein; molecular weight = 31,232). The molecular weight is substantially smaller than that determined experimentally for FruK of *R. capsulatus* (39 kDa [61]) but is essentially the same as that estimated for the *E. coli* FruK by sodium dodecyl sulfate (SDS) gel electrophoresis (molecular weight ≅ 30,000 [7]). The *Rhodobacter fruK* open reading frame is present within the 1,020-bp chromosomal fragment, the sequence of which is shown in Fig. 1.

Alignment of FruK from *R. capsulatus* with other bacterial kinases: the PfkB family. Figure 2 shows alignment of the *R. capsulatus* FruK protein with four bacterial kinases. These enzymes include FruK of *E. coli* (43), PfkB of *E. coli* (EC 2.7.1.11) (12), phosphotagatokinase of *Staphylococcus aureus* (the *lacC* gene product) (48, 57a), and ribokinase of *E. coli* (EC 2.7.1.15) (27). The consensus sequence is presented below the aligned sequences, and residues conserved in four or five of the proteins are indicated by asterisks or exclamation marks, respectively, above the aligned sequences.

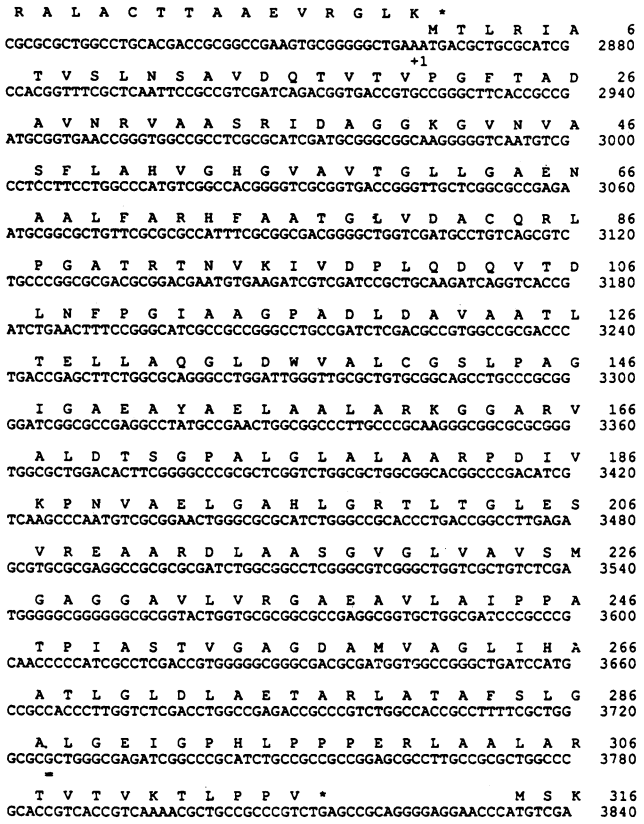

FIG. 1. Nucleotide sequence of a 1,020-bp segment encompassing the *fruK* gene of *R. capsulatus* and the deduced aminoacyl sequence of the 316 residues of FruK. The nucleotide sequence starts at bp 2821, numbered as described previously (61). Aminoacyl numbering starts at the ATG codon (+1) which is believed to serve as the initiation codon. Note the 4-bp overlap between the *fruK* gene and the preceding *fruB(HI)* gene. The intercistronic region between *fruK* and the following *fruA* gene (60) is 18 bp. The doubly underlined residue (corresponding to bp 3725) represents the start of the segment published previously (60).

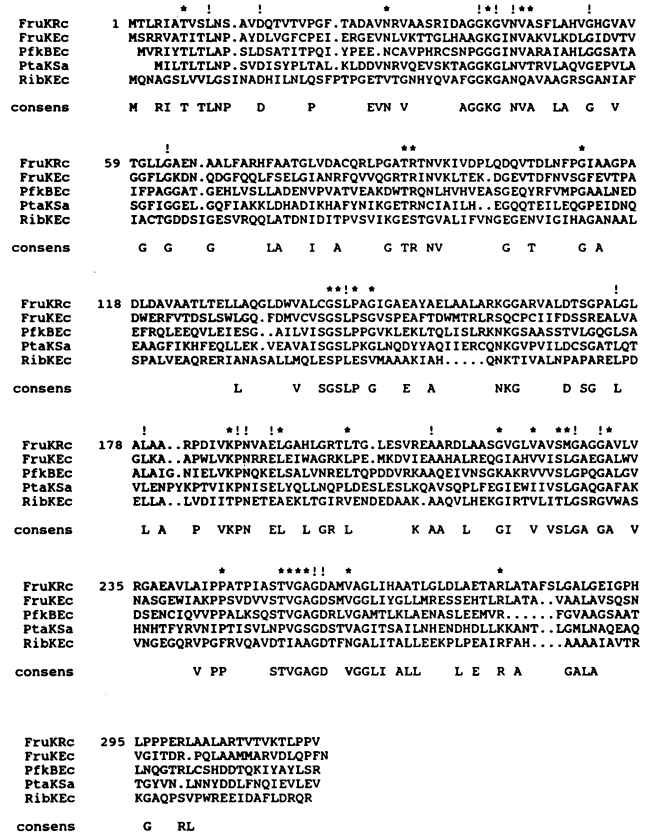

FIG. 2. Multiple alignment of FruK of *R. capsulatus* (FruKRc) with the same enzyme of *E. coli* (FruKEc), PfkB of *E. coli* (PfkBEc), phosphotagatokinase of *S. aureus* (PtaKSa), and ribokinase of *E. coli* (RibKEc). An exclamation mark above the aligned sequences indicates a residue conserved in all five sequences. An asterisk indicates a residue conserved in four of the five sequences. The consensus sequence (consens) is provided below the aligned sequences. The statistical analyses of the alignments are summarized in Table 1. The Fasta (44) and Bestfit (39) programs were used to give optimal alignment.

The statistical analyses of the aligned sequences are summarized in Table 1. The percent identities and the comparison scores (expressed in standard deviations) are sufficient to establish that these five proteins are homologous (16). A phylogenetic tree for these proteins is shown in Fig. 3. We refer to these proteins as the PfkB family. It can be seen that among these proteins, the two FruKs are most similar, that the phosphotagatokinase and PfkB are more distant from the two FruKs, and that ribokinase is more distant from the other proteins than the latter proteins are from each other.

Comparative secondary structural analyses, hydropathy, and amphiphilicity of *R. capsulatus* FruK and *E. coli* PfkB. Comparisons of secondary structural predictions for FruK of *R. capsulatus* and PfkB of *E. coli* are shown in Fig. 4. As can be seen, the plots for these two proteins are very similar for predicted random coil, turns, β-structure, and α-helix. Hydropathy plots (32) of these proteins reveal that both are strongly hydrophilic. They may both consist of alternating stretches of α- and β-structure as does PfkA, as determined from the X-ray crystallographic structure. Three extended regions of striking amphiphilicity (18) are also revealed (Fig. 4). Corresponding plots of the other homologous sequences depicted in Fig. 2 were similar (data not shown). The degrees of sequence identity observed for members of this family support the conclusion that they share essential structural features throughout their lengths.

Comparison of the amino acid compositions of FruK of *R. capsulatus* and homologous bacterial kinases. Table 2 presents the amino acid composition of FruK of *R. capsulatus* and compares it with those of FruK of *E. coli*, phosphotagatokinase of *S. aureus*, ribokinase of *E. coli*, and the two Pfks of *E. coli* (PfkB and PfkA). As described above, all of these proteins except PfkA of *E. coli* are sufficiently similar in sequence to establish that they are homologous, and as noted in Table 2, they are all of a very similar size. Despite these facts, the amino acid composition of FruK of *R. capsulatus* differs drastically from those of the other proteins. The amino acids are listed in Table 2 in the order of their frequency in FruK of *R. capsulatus*. All of the predominant amino acids in FruK of *R. capsulatus* (above the first space in Table 2) show increased frequencies relative to the other proteins, whereas almost all of the amino acids present in FruK of *R. capsulatus* in smaller amounts (below the second space in Table 2) show decreased frequencies rela-

TABLE 1. Statistical analyses[a] of the aligned sequences which compose the PfkB family

Protein	FruKEc		PfkBEc		PtaKSa		RibKEc	
	%I	CS	%I	CS	%I	CS	%I	CS
FruKRc	36 (316)	51	27 (299)	18	30 (233)	34	53 (15)	0.3
FruKEc			29 (283)	29	25 (307)	33	33 (127)	14
PfkBEc					24 (303)	25	21 (292)	11
PtaKSa							20 (243)	6

[a] Abbreviations used for the different proteins are the same as those described in the legend to Fig. 2. %I, percent identity; CS, comparison score given in SD determined by using the RDF2 program and 50 shuffles (44). Values in parentheses with the percent identities give the numbers of residues in the segment compared.

tive to those in the homologous proteins. Thus, for example, alanine, a predominant amino acid in all of these proteins, is about twice as abundant in FruK of *R. capsulatus* as in the other proteins. L, G, V, T, and P also show increased frequencies in FruK of *R. capsulatus* relative to the other proteins. By contrast, E, S, N, Q, K, C, and Y generally show increased frequencies of occurrence in the homologous proteins relative to FruK of *R. capsulatus*.

Nonessential residues in FruK. *R. capsulatus* FruK contains only 2 cysteyl, 1 tyrosyl, and 1 tryptophanyl residues, and none of these residues are conserved in *E. coli* FruK or PfkB (Fig. 2). A high concentration of *N*-ethylmaleimide did not inhibit *R. capsulatus* FruK activity, suggesting that no essential cysteyl residue is present in this enzyme. By contrast, the PTS activity of the same extract was completely abolished by this treatment. The PfkB family proteins contain four to eight histidyl residues, but none of them are conserved in *R. capsulatus* FruK and in *E. coli* FruK and PfkB. FruK of *R. capsulatus* was treated with DEPC as described in Materials and Methods. The enzyme was inactivated to the extent of only 25% by 2 mM DEPC, although it was largely inactivated by 10 mM DEPC.

FruK of *E. coli* has been shown to be inhibited by PMSF, a reagent specific for activated seryl residues (7). Employing the procedure outlined by Buschmeier et al. (7) (see Materials and Methods), we found that FruK in extracts of *R. capsulatus* was only weakly inactivated by 2 mM PMSF but was largely inactivated by 5 mM PMSF. FruK of *R. capsulatus* contains 11 seryl residues. Four of these are observed in the consensus sequence (Fig. 2), but none are conserved in all of the homologous kinases.

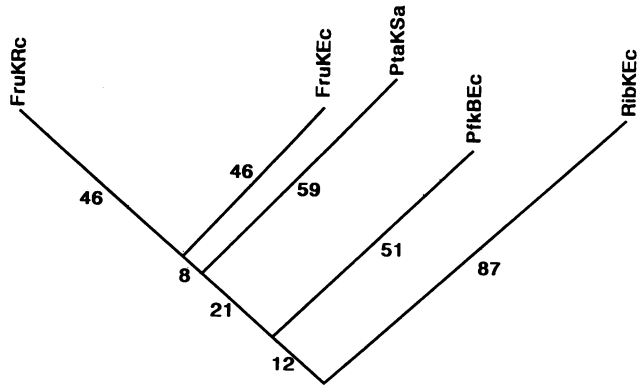

FIG. 3. Phylogenetic tree of homologous members of the PfkB family. Relative evolutionary distances are provided next to the branches. Abbreviations for the five enzymes are as indicated in the legend to Fig. 2. The tree was constructed as described previously (17, 20).

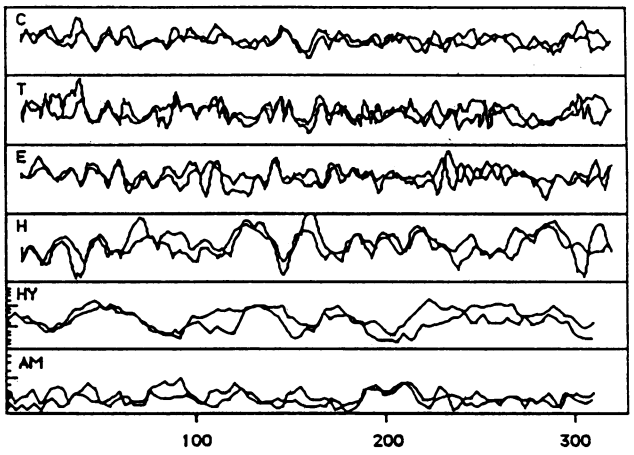

FIG. 4. Comparison of secondary structural predictions for FruK of *R. capsulatus* and PfkB of *E. coli*. C, coil; T, turn; E, extended β-structure; H, α-helix; HY, hydropathy; AM, amphiphilicity. The Analysep program from the Staden package was used to estimate secondary structure (57). Aminoacyl residue number is plotted on the *x* axis.

Multiple alignment of FruK with Rubisco large subunits from plants and bacteria. Figure 5 shows the alignment of a 127-aminoacyl residue segment of FruK from *R. capsulatus* with five different ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco) large subunits. Among the 21 totally

TABLE 2. Comparison of sizes and amino acid compositions of proteins in this study[a]

Amino acid	No. of each amino acid composing:					
	FruKRc	FruKEc	PfkBEc	PfkAEc	PtaKSa	RibKEc
A	66	30	37	27	22	49
L	42	30	34	24	37	26
G	36	29	27	38	24	27
V	32	33	28	25	23	25
T	23	16	16	15	16	17
P	18	13	15	9	13	14
R	16	18	14	21	5	15
D	15	17	10	23	14	13
E	11	19	22	22	23	23
S	11	21	24	14	16	15
I	10	15	15	28	25	24
N	7	11	14	8	22	17
F	6	12	4	9	9	8
H	6	4	6	7	8	6
Q	5	11	15	5	20	14
K	5	12	13	13	17	9
M	3	8	4	13	2	3
C	2	4	4	7	3	1
Y	1	2	5	11	9	1
W	1	7	1	1	1	2
Total	316	312	308	320	310	309
MW	31,232	33,708	33,853	32,290	32,388	34,758

[a] Abbreviations for the proteins are described in the legend to Fig. 2. PfkAEc, PfkA of *E. coli*.

```
        *^ *       * *    "          !     '     * **! ! ! "

FRUK  85 RLPGATRTNV KIVDPLQDQVTDLNFPGIAAGPADLDAVAATLTELLAQGLDWVALCGSL
ALFC 319 RMSGGDHIHAGTVVGKLEGE RDITLGFVDLLRDDFIEKDRSRGIFFTQ  DWVSLPGVL
PEHC 319 RMSGGDHIHAGTVVGKLEGE REMTLGFVDLLRDDFIEKDRARGIFFTQ  DWVSMPGVI
PEAC 326 RLSGGDHIHAGTVVGKLEGE REITLGFVDLLRDDYIKKDRSRGIYFTQ  DWVSLPGVI
TOBC 319 RMSGGDHIHSGTVVGKLEGE RDITLGFVDLLRDDFVEQDRSRGIYFTQ  DWVSLPGVL
AEUT 321 RLAGVDHMHTGTAVGKLEGD PLTVQGYYNVCRDAYTQTDLTRGLFFDQ  DWASLRKVM
                                                          - -

        *  *       !*   *  ' *  *  !*     !  ^  **        *

FRUK 144 PAGIGAEAYAELAALARKGGARVALDTSGPALGLALAARPDIVKPNVAELGAHLGRTLTG
ALFC 376 PVASGGIHVWHMPALTEIFGDDSVLQFGGGTLGHPWGNAPGAVANRVA          LEA
PEHC 376 PVASGGIHVWHMPALTEIFGDDSVLQFGGGTLGHPWGNAPGAAANRVA          LEA
PEAC 383 PVASGGIHVWHMPALTEIFGDDSVLQFGGGTLGHPWGNAPGAVANRVA          LEA
TOBC 376 PEASGGIHVWHMPALTEIFGDDSVLQFGGGTLGHPWGNAPGAVANRVA          LEA
AEUT 378 PVASGGIHAGQMHQLIHLFGDDVVLQFGGGTIGHPQGIQAGATANRVA          LEA
         -          -   -       - -          -                   -

            *   **!!  *

FRUK 204 LESVREAARDLAASG 218
ALFC 427 CVQARNEGRDLAREG 441
PEHC 427 CVQARNEGRDLAREG 441
PEAC 434 CVQARNEGRDLAREG 448
TOBC 427 CVKARNEGRDLAQEG 441
AEUT 429 MVLARNEGRDILNEG 443
         -          -
```

FIG. 5. Multiple alignment of FruK of *R. capsulatus* (FruK residues 85 to 318) with corresponding regions of five sequenced Rubisco large subunits. The five sources of Rubisco are alfalfa chloroplast (ALFC) (1), petunia hybrida chloroplast (PEHC) (2), pea chloroplast (PEAC) (62), tobacco chloroplast (TOBC) (56), and *Alcaligenes eutrophus* (AEUT) (3). Residue numbers within each protein are indicated at the beginning of each line and at the end of the segments. The following symbols indicate identity of the residue in FruK with the five Rubisco proteins: *, all five depicted proteins; !, four of the proteins; ^, three of the proteins; ", two of the proteins; and ', just one of the proteins. Double underlines indicate residues in FruK of *R. capsulatus* which align with all residues in the Rubisco large subunits as well as with those in PfkB of *E. coli* (Fig. 2).

TABLE 3. Statistical analyses of comparisons between the aligned sequence FruK of *R. capsulatus* and those of Rubisco large subunits

First enzyme and region compared	Second enzyme and region compared	% Identity	SD[a]	No. of gaps
FruK, 85–218	AlfC, 319–441	27.1	8.5	4
FruK, 85–218	PehC, 319–441	23.8	6.7	4
FruK, 85–218	PeaC, 326–448	26.2	7.5	4
FruK, 85–218	TobC, 319–441	27.1	6.6	4
FruK, 85–218	AeuT, 321–443	18.7	5.0	4
FruK, 85–218	RspR, 313–430	13.8	1.5	9
AlfC, 319–441	PehC, 319–441	95.1	63.5	0
AlfC, 319–441	PeaC, 326–441	95.1	56.6	0
AlfC, 319–441	TobC, 319–441	94.3	65.8	0
AlfC, 319–441	AeuT, 321–443	61.0	37.7	0
PehC, 319–441	PeaC, 326–441	93.5	53.1	0
PehC, 319–441	TobC, 319–441	89.4	59.2	0
PehC, 319–441	AeuT, 321–443	60.2	39.0	0
PeaC, 326–448	TobC, 319–441	91.1	60.2	0
PeaC, 326–448	AeuT, 321–443	61.0	41.2	0
TobC, 319–441	AeuT, 321–443	60.2	37.3	0
RspR, 313–430	AlfC, 319–441	31.7	14.8	4
RspR, 313–430	PehC, 319–441	31.7	11.4	4
RspR, 313–430	PeaC, 326–448	32.5	11.7	4
RspR, 313–430	TobC, 319–441	31.7	11.7	4
RspR, 313–430	AeuT, 321–443	32.5	11.9	4

[a] Los Alamos comparison scores are reported. ALIGN comparison scores were similar. For example, values for the ALIGN comparison of FruK with AlfC, PehC, and PeaC were 7.4, 6.8, and 7.2 SD, respectively, higher than those obtained with 1,000 comparisons of these randomized sequences. Abbreviations are the same as those described in the legend to Fig. 5. RspR, Rubisco from *R. rubrum*.

conserved residues, L335 and G464 of alfalfa Rubisco aligned with L100 and G172 of FruK. In the former enzyme, these residues are involved in ribulose-1,5-bisphosphate binding (29). As revealed by the statistical analyses reported in Table 3, the degrees of sequence similarity suggest that these proteins may be homologous. For example, FruK is 27% identical with Rubisco from alfalfa throughout a segment encompassing residues 319 to 441 of the latter enzyme. The Los Alamos comparison score (28) for this segment was 8.5 standard deviations (SD) higher than that obtained with 100 comparisons of randomized sequences of these proteins. The probability of obtaining such a score by chance is less than 10^{-16}. The corresponding value with the ALIGN program was 7.4 SD higher than that obtained with 1,000 comparisons of these randomized sequences. Thus, the results suggest but do not prove that a domain in FruK is homologous to a domain in alfalfa Rubisco. Other Rubisco proteins were also similar to FruK, although the comparison scores were somewhat lower. The sequence of Rubisco from *Rhodospirillum rubrum* was most divergent from that of FruK, but sequence comparisons among the different Rubisco large subunits revealed that all of these proteins are homologous (Table 3). The residues doubly underlined in Fig. 5 are those which are also aligned with PfkB in Fig. 2. The results show that FruK, PfkB, and the various Rubisco large subunits all possess domains which exhibit sequence similarity. These domains might be involved in sugar bisphosphate binding, as demonstrated in Rubisco by X-ray crystallography (29).

Codon preference for the *fru* operon of *R. capsulatus*. Figure 6 shows the codon preference plots for the three reading frames within the *fru* operon, based on the codon usage for *Rhodobacter* spp. (60a). It can be seen that the first gene, *fruB(HI)*, is in the A reading frame, whereas both *fruK* and *fruA* are in the C reading frame. Rare codons are utilized very infrequently in all three genes, suggesting (i) that the open reading frames *fruB(HI)*, *fruK*, and *fruA* are correctly determined and (ii) that the entire operon can be expressed at a high level. Preceding the *fruB(HI)* gene, a partial open reading frame (*orfX*) which shows a diminished frequency of rare codon usage relative to extragenic DNA, but an increased frequency of rare codon usage relative to the three recognized open reading frames of the *fru* operon is revealed. This 387-bp stretch encodes the C-terminal 128 amino acids of a putative protein which terminates within the *fruB(HI)* gene with an overlap of 4 bp. This is the same degree of overlap observed between the *fruB(HI)* and *fruK* genes (Fig. 1). Since no predominant protein bands were observed following expression of the 7.4-kb fragment, which starts 1 kb upstream from the beginning of the sequenced part of *orfX* in the T7 promoter/RNA polymerase system (61), we presume, in agreement with the increased frequency of rare codons within this open reading frame, that the product of this gene either is not expressed at a high level or is not expressed at all. A computer search of the data base for homology with the encoded C-terminal 128 residues of the putative protein did not yield significant positive results.

Alignment of Pfks of prokaryotes and eukaryotes: the PfkA family. Figure 7 shows alignment of three bacterial Pfks (from *E. coli* [25], *B. stearothermophilus* [21], and *Spiroplasma citri* [10]) with homologous domains of a number of eukaryotic Pfks. The statistical analyses of the sequences aligned in Fig. 7 are summarized in Table 4. The percent identities and the comparison scores (in SD) are sufficient to establish that all of these protein segments are homologous. We refer to these proteins and homologous protein segments as the PfkA family. The four sequenced mammalian en-

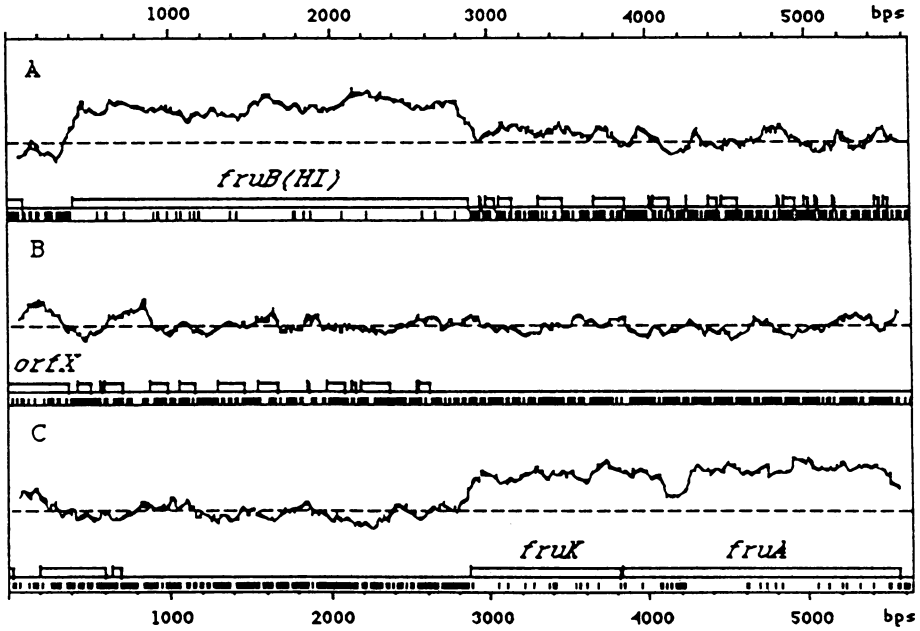

FIG. 6. Codon preference plots of the *fru* operon of *R. capsulatus* generated employing the program of Gribskov et al. (23) but based on *Rhodobacter* codon usage. The program identifies efficiently translated genes as peaks above open reading frames (open boxes below peaks) that contain few rare codons (vertical lines below the open reading frames). The window size and rare codon thresholds were set at 50 and 0.1, respectively. The number of the base pairs in the DNA fragment analyzed is indicated on the *x* axis.

zymes (from human muscle [55], human liver [34], rabbit muscle [33], and mouse liver [22]) are homotetramers of polypeptide chains about twice the size of the bacterial enzymes (each 780 residues). They consist of two adjacent segments which are homologous to each other and the bacterial Pfks, and they probably arose by tandem intragenic duplication (Table 5). The larger yeast enzyme is a heterooctamer consisting of two large, dissimilar polypeptide chains, α and β (987 and 959 residues, respectively), each of which, like the mammalian enzyme, contains an internal repeat of the bacterial Pfk equivalent and is present in four copies per enzyme (Fig. 7 and Table 5) (24). Examination of the aligned sequences (Fig. 7) reveals that the N-terminal segments are much more similar to the bacterial enzymes than are the homologous C-terminal segments, which, however, are closely related to the other C-terminal segments. This observation suggests that the duplication event, giving rise to the eukaryotic Pfk subunit from a prokaryotic Pfk-like protein, occurred before the divergence of the yeasts from mammals.

A phylogenetic tree for these ATP-dependent bacterial Pfks and the eukaryotic Pfk segments as well as the potato PP_i-dependent Pfk is shown in Fig. 8. It can be seen that all of the N-terminal sequences of the eukaryotic, ATP-dependent Pfks cluster together with the bacterial proteins, whereas the C-terminal sequences form a cluster which is quite distant (in relative, apparent, evolutionary time) from the N-terminal segments. Even more distant from the bacterial sequences are homologous segments in each of the two subunits of the potato PP_i-dependent Pfk (Fig. 8). These two subunits (α and β) exhibited 42% overall identity with each other throughout most of their regions of overlap (8). The percent identities observed between the α- and β-subunits of the PP_i-dependent plant enzyme and *E. coli* PfkA were 25% in a 148-residue overlap and 29% in a 222-residue overlap, respectively (comparison scores of 11 and 12 SD, respectively). Corresponding values when these two subunits (α and β) were compared with the *B. stearothermophilus* Pfk were 27% identity in 202 overlapping residues (comparison score of 14 SD) and 30% identity in 208 overlapping residues (comparison score of 16 SD), respectively. Thus, the plant enzyme is clearly homologous to members of the PfkA family. It more closely resembled the gram-positive bacterial enzyme than the gram-negative bacterial enzyme. The potato PP_i-dependent Pfk exhibited greater sequence similarity to the bacterial enzymes than to the other eukaryotic enzymes (data not shown).

Sequence similarity between the *R. capsulatus* FruK and the two yeast Pfk subunits. Members of the PfkA family exhibited regions of sequence similarity with members of the PfkB family. For example, FruK of *R. capsulatus* exhibited 17.2% identity with the yeast Pfk β-subunit. The regions of overlap included residues 4 to 302 in FruK of *R. capsulatus* and residues 668 to 959 in the yeast Pfk β-subunit. Three gaps were introduced to achieve this degree of identity. The percent similarity for this alignment, taking into account both identities and semiconservative substitutions, was 37.1%. The comparison score for this alignment was 3.9 SD. The probability of getting this comparison score by chance is less than 5×10^{-5}. These values, while insufficient to establish homology, demonstrate a degree of sequence similarity which is not likely to have arisen by chance. Divergent evolution from a common ancestor would provide a reasonable explanation for this sequence similarity, but the possibility of convergent evolution cannot be ruled out.

DISCUSSION

In early reports (49, 54), a fructose-specific PTS was detected and characterized in the photosynthetic bacteria *Rhodospirillum rubrum* and *Rhodobacter spheroides*. The system was shown to consist of two membrane proteins, one

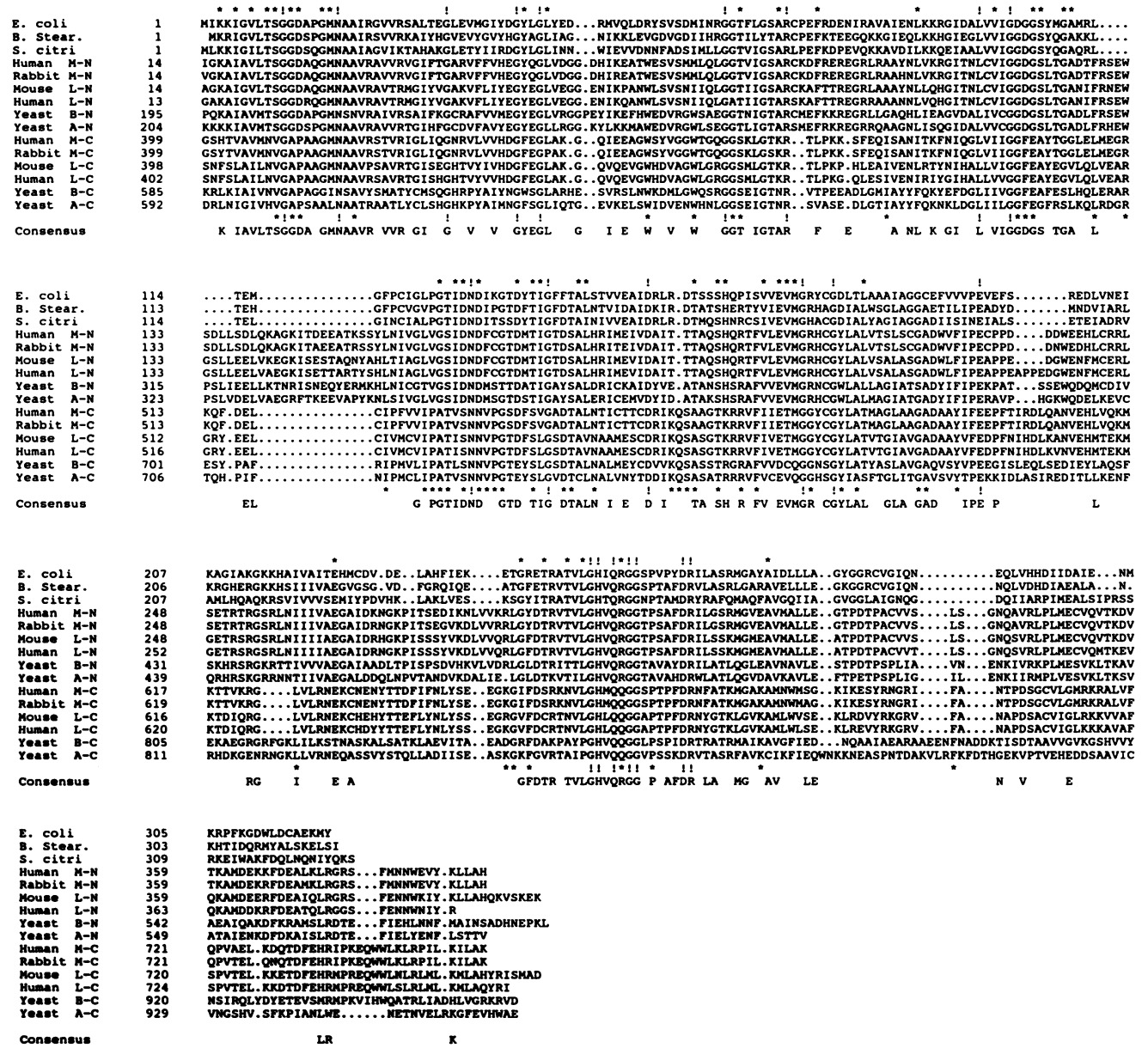

FIG. 7. Multiple alignment of the PfkA family of ATP-dependent Pfks of bacteria, yeasts, and mammals. The species are as indicated. M and L, muscle and liver enzymes of mammals, respectively; B and A, β- and α-subunit of the yeast enzyme, respectively; N and C, N-terminal and the C-terminal domains, respectively, of the eukaryotic protein subunits. An exclamation mark above and below the aligned sequences indicates a residue conserved in all sequences. An asterisk above the aligned sequences indicates a residue conserved in the bacterial enzymes and the eukaryotic N-terminal segments. An asterisk below the aligned sequences indicates a residue conserved among the eukaryotic C-terminal segments. The consensus sequence (Consensus) is provided below the aligned sequences. The statistical analyses of the alignments are summarized in Table 4. The Fasta (44) and Bestfit (39) programs were used to give the optimal alignment.

which could be released from the membrane as a high-molecular-weight soluble protein, the other which was an integral constituent of the membrane. The biochemistry of the former protein has been examined (35, 36). As shown in our previous reports (60, 61), these two proteins are a multiphosphoryl transfer protein and a fructose-specific enzyme II of the PTS. They act in concert to produce cytoplasmic fructose-1-P from extracellular fructose. This sugar phosphate ester is then further phosphorylated in the reaction catalyzed by FruK to give the common intermediate of glycolysis, fructose-1,6-bisphosphate. In species of *Rhodospirillum*, *Rhodobacter*, *Thiocapsa*, *Thiocystis*, *Pseudomonas*, *Alcaligenes*, and *Fusobacterium*, all of which possess a fructose-specific PTS (15), fructose is probably the only sugar metabolized via the PTS and glycolysis, because 6-Pfk is lacking.

Biochemical experiments have shown that in *R. rubrum*, *R. spheroides*, and *R. capsulatus*, a multiphosphoryl transfer protein, a fructose-specific enzyme II, and FruK are coordinately induced 3-, 10-, and >100-fold, respectively, by inclusion of fructose in the growth medium (13, 49, 54). These observations led to the prediction that the three

TABLE 4. Statistical analyses of the aligned sequences which compose the ATP-dependent Pfk members of the PfkA family[a]

Enzyme	*E. coli*		*B. stearothermophilus*		Human				Rabbit				Mouse				Human				Yeast							
					M-N		M-C		M-N		M-C		L-N		L-C		L-N		L-C		A-N		A-C		B-N		B-C	
	%I	CS	%I	CS	%I	CS	%I	CS	%I	CS	%I	CS	%I	CS	%I	CS	%I	CS	%I	CS	%I	CS	%I	CS	%I	CS	%I	CS
S. citri	48 (314)	57	46 (319)	89	34 (214)	19	29 (281)	31	34 (214)	27	29 (281)	38	33 (226)	23	30 (274)	26	31 (208)	24	30 (274)	31	31 (236)	26	25 (218)	17	35 (218)	19	27 (184)	19
E. coli			57 (302)	75	36 (203)	41	30 (288)	33	36 (203)	28	30 (288)	36	37 (197)	22	32 (281)	31	40 (172)	26	31 (281)	47	33 (202)	26	26 (170)	12	32 (208)	33	29 (172)	17
B. stearothermophilus					39 (244)	32	35 (288)	44	38 (244)	32	33 (285)	43	44 (203)	35	33 (285)	27	43 (196)	33	33 (292)	54	35 (233)	31	29 (246)	33	35 (202)	27	26 (178)	21
Human																												
M-N							35 (121)	22	96 (377)	168	35 (119)	20	77 (376)	290	42 (72)	17	74 (376)	156	43 (72)	21	53 (371)	125	25 (192)	16	52 (372)	82	29 (171)	21
M-C									35 (121)	24	98 (356)	263	26 (244)	26	66 (352)	204	32 (178)	19	66 (353)	133	37 (108)	18	37 (290)	57	36 (110)	19	35 (364)	82
Rabbit																												
M-N											35 (119)	14	77 (376)	131	42 (72)	11	75 (376)	157	43 (72)	15	54 (371)	120	24 (192)	13	52 (372)	87	28 (171)	27
M-C													25 (244)	30	65 (355)	110	32 (178)	26	65 (355)	156	35 (108)	19	36 (290)	69	34 (110)	20	35 (364)	56
Mouse																												
L-N															32 (123)	11	92 (377)	227	32 (124)	11	52 (371)	139	30 (115)	13	50 (384)	87	22 (266)	24
L-C																	24 (177)	13	94 (357)	195	41 (101)	26	38 (294)	58	34 (103)	19	38 (366)	78
Human																												
L-N																			23 (177)	19	52 (374)	154	30 (115)	15	51 (375)	107	26 (184)	21
L-C																					41 (101)	14	39 (296)	116	36 (103)	21	38 (368)	67
Yeast																												
A-N																							24 (159)	9	65 (375)	132	26 (179)	22
A-C																									31 (74)	10	48 (368)	146
B-N																											22 (256)	13

[a] The abbreviations used are those described in the legend to Fig. 7. %I, percent identity (values in parentheses give the number of residues in the segment compared); comparison score in SD calculated by using the RDF2 program and 50 shuffles (44).

TABLE 5. Size, subunit, and domain compositions of Pfks and selected related enzymes[a]

Enzyme	Approximate no. of aminoacyl residues/chain	No. of polypeptide chains/ enzyme	No. of identified homologous domains/ enzyme
R. capsulatus, FruK	316	1	1
E. coli PfkB	308	2	2
E. coli PfkA	320	4	4
B. stearothermophilus Pfk	319	4	4
S. citri Pfk	327	4	4
Mammalian Pfks	780 (N and C)	4	8
Yeast Pfk, α	987 (N and C)	4 ⎫ 8	16
Yeast Pfk, β	959 (N and C)	4 ⎭	
Potato P_2 Pfk, α	616	2 ⎫ 4	4
Potato P_2 Pfk, β	513	2 ⎭	

[a] Mammalian Pfk, the four mammalian Pfks for which sequence data are available (human muscle, human liver, rabbit muscle, and mouse liver); Yeast Pfk, *Saccharomyces cerevisiae* Pfk; Potato P_2 Pfk, potato PP_i-dependent Pfk; α and β, α- and β-subunits, respectively, of the yeast and potato enzymes; N and C, N-terminal and C-terminal homologous domains, respectively, of the mammalian enzymes and of the two subunits of the yeast enzyme.

fructose enzymes are encoded within a single operon in each of these organisms, a conclusion established for *R. capsulatus* by the sequence analyses reported here. The *fru* operon of *R. capsulatus* may possess a compact structure with a 4-nucleotide overlap between the *fruB(HI)* and *fruK* genes and an 18-nucleotide intercistronic region between *fruK* and *fruA*. These structural features resemble those in the *fru* operons of *E. coli* in which a 1-nucleotide overlap characterizes the chain terminating codon of *fruB(MH)* and the initiation codon of *fruK* and a 16-nucleotide intercistronic region separates *fruK* from *fruA* (43).

Determination of the *R. capsulatus fruK* nucleotide sequence, deduction of the aminoacyl sequence of FruK, and sequence comparisons with proteins in the data base revealed that FruK is homologous to several bacterial kinases, including PfkB of *E. coli*. A phylogenetic tree relating these proteins is shown in Fig. 3. Surprisingly, we also found that an extended segment of the large subunit of Rubisco exhibited a significant degree of sequence similarity (27.1% identity over a 122-residue stretch with four gaps and a comparison score of 8.5 SD) with FruK of *R. capsulatus* (Fig. 5; Table 3). Since this region encompasses the sugar bisphosphate-binding site of Rubisco, it is possible, although unlikely, that this sequence similarity arose by convergent evolution.

The sequence similarities between FruK of *R. capsulatus* and the homologous kinases from other bacteria were not reflected in the amino acid compositions of these proteins. The *R. capsulatus* protein showed a markedly different composition, with the predominant amino acids (Table 2, shown above the spaces; particularly A, L, and G) showing increased occurrence in the *R. capsulatus* FruK relative to the other proteins and the minor amino acids (Table 2, below the spaces; particularly E, S, N, Q, K, C, and Y) showing decreased occurrence in *R. capsulatus* FruK relative to that of the other kinases. Surprisingly, none of the C, S, Y, W, or H residues were fully conserved among the homologous proteins, suggesting that none of these residues play an essential catalytic role in the mechanism of action of these

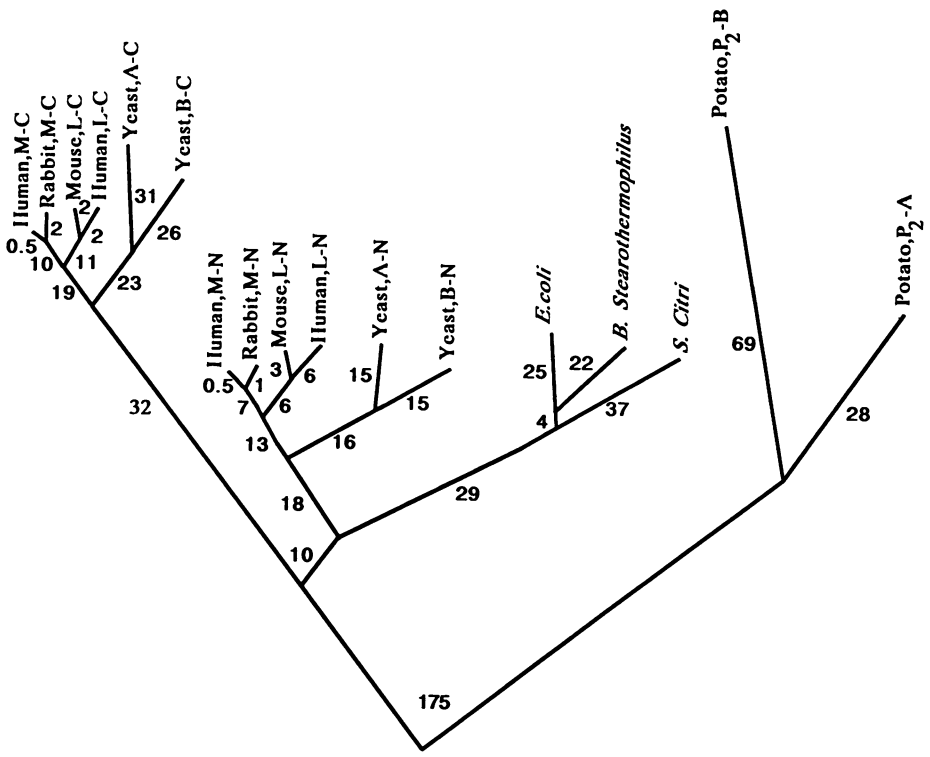

FIG. 8. Phylogenetic tree of homologous members of the PfkA family. Relative evolutionary distances are provided adjacent to the branches. Abbreviations are the same as indicated in the legend to Fig. 7, except that the PP_i-dependent Pfk subunits α and β of the potato are designated Potato, P_2-A, and P_2-B, respectively. The tree was constructed as described by Doolittle and Feng (17, 20).

functionally related enzymes. In fact, FruK was not sensitive to inhibition by the cysteyl-specific reagent, *N*-ethylmaleimide, and was inhibited only at high concentrations of the histidyl-specific reagent, DEPC, and the activated serine reagent, PMSF. Since the monomeric *Rhodobacter* protein (60a) contains just one W and one Y, this protein may be good material for fluorescence studies.

Secondary structural predictions (Fig. 4) led to the suggestion that members of the PfkB family of homologous proteins are all structurally similar. Predicted secondary structural elements (coil, β-turns, β-strands, and α-helix) as well as hydropathy and amphiphilicity analyses were nearly superimposable for FruK of *R. capsulatus* and PfkB of *E. coli* despite their considerable compositional differences. Both proteins appear to consist of alternate β- and α-structures, as has been demonstrated for PfkA (25). On the basis of their primary structural similarities, the tertiary structures of the homologous PfkB family proteins are probably nearly superimposable (16, 52, 53).

It is interesting to note that PfkA, the predominant Pfk in *E. coli*, is a nondissociable tetramer, whereas PfkB is a dimer which can associate to a less-active tetramer (see reference 51 for a review). Their subunits are of about the same size, and their genes are located at opposite positions on the circular *E. coli* chromosome (at 87 and 38 min, respectively). Despite these facts, PfkA does not exhibit significant sequence identity with PfkB, but PfkA is homologous to sequenced Pfks from other bacteria and eukaryotes (8, 25, 45). It has been proposed that the primordial bacterial chromosome was once one-half of its present size, and a chromosomal duplication event occurred, giving rise to a genome with genes of similar structure and function on opposite sides of the chromosome (47). If this hypothesis is correct, then this chromosomal duplication must have occurred before the divergence of prokaryotes from eukaryotes, and the subunits of the currently recognized eukaryotic Pfks (which presumably arose by intragenic duplication) must have been derived from the precursor of the present-day prokaryotic PfkA. This hypothesis presupposes also that divergence within the PfkB family occurred after this duplication. It is interesting to note that PfkA and PfkB of *E. coli* differ not only with respect to their sequences and subunit association properties but also with respect to their regulatory properties. Moreover, while the synthesis of PfkA is induced in the presence of glucose, that of PfkB has been reported to be constitutive (11). Thus, PfkA may function primarily in anaerobic glycolysis, whereas PfkB may function under conditions of aerobic metabolism. In view of the structural similarity of PfkB and FruK (Fig. 4), it is interesting to note that the *pfkB* gene in *E. coli* and the *fru* operon in *S. typhimurium* share the same unusual putative −10 region (CAGACT) and have similar −35 regions (40–42a). The functional and evolutionary significance of these observations has yet to be established.

The PfkA family is large and diverse, with representatives from gram-positive and gram-negative prokaryotes as well as lower and higher eukaryotes. The Pfk subunits of eukaryotes are minimally twice the size of the prokaryotic enzymes, and as documented here, they may have arisen by intragenic duplication. The N-terminal domains of these eukaryotic enzymes exhibit a higher degree of similarity to the prokaryotic enzymes than do the homologous C-terminal domains. It is therefore reasonable to propose that the N-terminal domains retained their original, primary function, namely,

catalysis, while the C-terminal domains evolved more rapidly, giving rise to allosteric regulatory domains. In this regard it is interesting to note that substrates and products of the Pfk-catalyzed reactions also serve as allosteric effectors. Thus, the substrate-binding residues have been largely retained in the N- and C-terminal domains of the eukaryotic enzymes (26, 46a). The eukaryotic enzymes have also acquired sensitivity to regulation by other metabolites, such as Krebs cycle intermediates (51). This fact is consistent with their increased degree of structural complexity relative to that of the prokaryotic enzymes. The property of cooperativity, characteristic of all members of the PfkA family, is consistent with their tetrameric structures, and these structures differ from those of the PfkB family members which do not exhibit the property of cooperativity and which can be monomeric or dimeric.

Although PfkA and PfkB of *E. coli* were not found to exhibit a statistically significant degree of sequence identity, we nevertheless propose that these proteins, and therefore all members of the PfkA and PfkB families, shared a common origin. The evidence is as follows. (i) PfkA and PfkB catalyze the same reaction, and all members of both families catalyze essentially the same reaction with various carbohydrate substrates. (ii) PfkA and PfkB are essentially the same size, a size that is common to all PfkB family members, prokaryotic PfkA family members, and homologous domains of eukaryotic Pfks (Table 2) (46a). (iii) The predicted secondary structures of PfkB and FruK (Fig. 4) correspond to the known secondary structures of PfkA and of the *B. stearothermophilus* Pfk (19, 26), and the degrees of sequence similarity within either the PfkA or the PfkB family imply conservation of secondary and tertiary structural features (16). (iv) FruK of *R. capsulatus*, a member of the PfkB family, exhibited regions of sequence similarity both with the PP_i-dependent Pfk of potatoes (8; unpublished data), which is clearly homologous to members of the PfkA family, and with the yeast Pfk (see Results). While the extent of similarity was insufficient to establish homology, the results are suggestive of a common origin. (v) As noted above, the *pfkA* and *pfkB* genes map to opposite sides of the *E. coli* chromosome, which is consistent with an origin resulting from a chromosomal duplication event (47). These observations, while compelling, must nevertheless be considered preliminary. Substantiation for the proposed common origin for the PfkA and PfkB families may be forthcoming when sequences of additional homologous proteins or the X-ray crystallographic structure of a member of the PfkB family becomes available.

ACKNOWLEDGMENTS

We thank Michael C. O'Neill for conducting computer analyses of the promoters of the *fru* operons of both *R. capsulatus* and *S. typhimurium*, Michael Baker and DaFei Feng for valuable suggestions concerning the computer analyses, and Mary Beth Hiller for assistance in preparation of the manuscript.

This work was supported by Public Health Service grants 5RO1AI21702 and 2RO1AI14176 from the National Institute of Allergy and Infectious Diseases.

REFERENCES

1. **Aldrich, J., B. Cherney, E. Merlin, and J. Palmer.** 1986. Sequence of the *rbcL* gene for the large subunit of ribulose bisphosphate carboxylase-oxygenase from petunia. Nucleic Acids Res. **14**:9534.
2. **Aldrich, J., B. Cherney, E. Merlin, and J. Palmer.** 1986. Sequence of the *rbcL* gene for the large subunit of ribulose bisphosphate carboxylase-oxygenase from alfalfa. Nucleic Acids Res. **14**:9535.
3. **Andersen, K., and J. Caton.** 1987. Sequence analysis of the *Alcaligenes eutrophus* chromosomally encoded ribulose bisphosphate carboxylase large and small subunit genes and their gene products. J. Bacteriol. **169**:4547–4558.
4. **Babul, J.** 1978. Phosphofructokinases from *Escherichia coli*. Purification and characterization of the nonallosteric isozyme. J. Biol. Chem. **253**:4350–4355.
5. **Birnboim, H. C.** 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol. **100**:243–255.
6. **Blangy, D., H. Buc, and J. Monod.** 1968. Kinetics of the allosteric interactions of phosphofructokinase from *Escherichia coli*. J. Mol. Biol. **31**:13–35.
7. **Buschmeier, B., W. Hengstenberg, and J. Deutscher.** 1985. Purification and properties of 1-phosphofructokinase from *Escherichia coli*. FEMS Microbiol. Lett. **29**:231–235.
8. **Carlisle, S. M., S. D. Blakeley, S. M. Hemmingsen, S. J. Trevanion, T. Hiyoshi, N. J. Kruger, and D. T. Dennis.** 1990. Pyrophosphate-dependent phosphofructokinase. Conservation of protein sequence between the α- and β-subunits and with the ATP-dependent phosphofructokinase. J. Biol. Chem. **265**: 18366–18371.
9. **Carter, P., H. Bedouelle, and G. Winter.** 1985. Improved oligonucleotide site-directed mutagenesis using M13 vectors. Nucleic Acids Res. **13**:4431–4443.
10. **Chevalier, C., C. Saillard, and J. M. Bové.** 1990. Organization and nucleotide sequences of the *Spiroplasma citri* genes for ribosomal protein S2, elongation factor Ts, spiralin, phosphofructokinase, pyruvate kinase, and an unidentified protein. J. Bacteriol. **172**:2693–2703.
11. **Daldal, F.** 1983. Molecular cloning of the gene for phosphofructokinase-2 of *Escherichia coli* and the nature of a mutation, *pfkB1*, causing a high level of the enzyme. J. Mol. Biol. **168**:285–305.
12. **Daldal, F.** 1984. Nucleotide sequence of gene *pfkB* encoding the minor phosphofructokinase of *Escherichia coli* K-12. Gene **28**:337–342.
13. **Daniels, G. A., G. Drews, and M. H. Saier, Jr.** 1988. Properties of a Tn*5* insertion mutant defective in the structural gene *fruA* of the fructose-specific phosphotransferase system of *Rhodobacter capsulatus* and the cloning of the *fru* regulon. J. Bacteriol. **170**:1698–1703.
14. **Dayhoff, M. O., W. C. Barker, and L. T. Hunt.** 1983. Establishing homologies in protein sequences. Methods Enzymol. **91**: 524–545.
15. **Dills, S. S., A. Apperson, M. R. Schmidt, and M. H. Saier, Jr.** 1980. Carbohydrate transport in bacteria. Microbiol. Rev. **44**: 385–418.
16. **Doolittle, R. F.** 1986. Of URFs and ORFs. A primer on how to analyze derived amino acid sequences. University Science Books, Mill Hill, Calif.
17. **Doolittle, R. F., and D.-F. Feng.** 1990. A nearest neighbor procedure for relating progressively aligned amino acid sequences. Methods Enzymol. **183**:659–669.
18. **Eisenberg, D.** 1984. 3-Dimensional structure of membrane and surface proteins. Annu. Rev. Biochem. **53**:595–623.
19. **Evans, P. R., G. W. Farrants, and P. J. Hudson.** 1981. Phosphofructokinase: structure and control. Philos. Trans. R. Soc. London B **293**:53–62.
20. **Feng, D. F., and R. F. Doolittle.** 1990. Progressive alignment and phylogenetic tree construction of protein sequences. Methods Enzymol. **183**:375–387.
21. **French, B. A., and S. H. Chang.** 1987. Nucleotide sequence of the phosphofructokinase gene from *Bacillus stearothermophilus* and comparison with the homologous *Escherichia coli* gene. Gene **54**:65–71.
22. **Gehnrich, S. C., N. Gekakis, and H. S. Sul.** 1988. Liver (B-type) phosphofructokinase mRNA. Cloning, structure, and expression. J. Biol. Chem. **263**:11755–11759.
23. **Gribskov, M., J. Devereur, and R. R. Burgess.** 1984. The codon preference plot: graphic analysis of protein coding sequences and prediction of gene expression. Nucleic Acids Res. **12**:539–547.

24. **Heinisch, J., R. G. Ritzel, R. C. von Borstel, A. Aguilera, R. Rodicio, and F. K. Zimmermann.** 1989. The phosphofructokinase genes of yeast evolved from two duplication events. Gene **78**:309–321.
25. **Hellinga, H. W., and P. R. Evans.** 1985. Nucleotide sequence and high-level expression of the major *Escherichia coli* phosphofructokinase. Eur. J. Biochem. **149**:363–373.
26. **Hellinga, H. W., and P. R. Evans.** 1987. Mutations in the active site of *Escherichia coli* phosphofructokinase. Nature (London) **327**:437–439.
27. **Hope, J. N., A. W. Bell, M. A. Hermodson, and J. M. Groarke.** 1986. Ribokinase from *Escherichia coli* K12. Nucleotide sequence and overexpression of the *rbsK* gene and purification of ribokinase. J. Biol. Chem. **261**:7663–7668.
28. **Kanehisa, M.** 1982. Los Alamos sequence analysis package for nucleic acids and proteins. Nucleic Acids Res. **10**:183–196.
29. **Knaff, D. B.** 1989. Structure and regulation of ribulose-1,5-bisphosphate carboxylase/oxygenase. Trends Biochem. Sci. **161**:159–160.
30. **Kotlarz, D., and H. Buc.** 1977. Two *Escherichia coli* fructose-6-phosphate kinases. Preparative purification, oligomeric structure and immunological studies. Biochim. Biophys. Acta **484**: 35–48.
31. **Kotlarz, D., and H. Buc.** 1981. Regulatory properties of phosphofructokinase 2 from *Escherichia coli*. Eur. J. Biochem. **117**:569–574.
32. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
33. **Lee, C.-P., M.-C. Kao, B. A. French, S. D. Putney, and S. H. Chang.** 1987. The rabbit muscle phosphofructokinase gene. Implications for protein structure, function, and tissue specificity. J. Biol. Chem. **262**:4195–4199.
34. **Levanon, D., E. Danciger, N. Dafni, Y. Bernstein, A. Elson, W. Moens, M. Brandeis, and Y. Groner.** 1989. The primary structure of human liver type phosphofructokinase and its comparison with other types of PFK. DNA **8**:733–743.
35. **Lolkema, J. S., and G. T. Robillard.** 1985. The coupling between transport and phosphorylation in inside-out vesicles. Eur. J. Biochem. **147**:69–75.
36. **Lolkema, J. S., R. H. ten Hoeve-Duurkens, and G. T. Robillard.** 1986. Energetics of the phosphoryl group from phosphoenolpyruvate to fructose. Eur. J. Biochem. **154**:387–393.
37. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
38. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
39. **Needleman, S. B., and C. D. Wunsch.** 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. **48**:443–453.
40. **O'Neill, M. C.** 1989. Consensus methods for finding and ranking DNA binding sites. Application to *Escherichia coli* promoters. J. Mol. Biol. **207**:301–310.
41. **O'Neill, M. C.** 1989. *Escherichia coli* promoters. I. Consensus as it relates to spacing class, specificity, repeat substructure, and three-dimensional organization. J. Biol. Chem. **264**:5522–5530.
42. **O'Neill, M. C., and F. Chiafari.** 1989. *Escherichia coli* promoters. II. A spacing class-dependent promoter search protocol. J. Biol. Chem. **264**:5531–5534.

42a. **O'Neill, M. C., and M. H. Saier, Jr.** Unpublished data.

43. **Orchard, L. M. D., and H. L. Kornberg.** 1990. Sequence similarities between the gene specifying 1-phosphofructokinase (*fruK*), genes specifying other kinases in *Escherichia coli* K12, and *lacC* of *Staphylococcus aureus*. Proc. R. Soc. London B **242**:87–90.
44. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85**:2444–2448.
45. **Poorman, R. A., A. Randolph, R. G. Kemp, and R. L. Heinrikson.** 1984. Evolution of phosphofructokinase—gene duplication and creation of new effector sites. Nature (London) **309**:467–469.
46. **Postma, P. W., and J. W. Lengeler.** 1985. Phosphoenolpyruvate: carbohydrate phosphotransferase system of bacteria. Microbiol. Rev. **49**:232–269.

46a. **Reizer, J., A. Reizer, and M. H. Saier, Jr.** Unpublished data.

47. **Riley, M., and A. Anilionis.** 1978. Evolution of the bacterial genome. Annu. Rev. Microbiol. **32**:519–560.
48. **Rosey, E. L., and G. C. Stewart.** 1989. The nucleotide sequence of the *lacC* and *lacD* genes of *Staphylococcus aureus*. Nucleic Acids Res. **17**:3980.
49. **Saier, M. H., Jr.** 1977. Bacterial phosphoenolpyruvate:sugar phosphotransferase systems: structural, functional, and evolutionary interrelationships. Bacteriol. Rev. **41**:856–871.
50. **Saier, M. H., Jr.** 1985. Mechanisms and regulation of carbohydrate transport in bacteria. Academic Press, Inc., New York.
51. **Saier, M. H., Jr.** 1987. Enzymes in metabolic pathways, p. 131–134. Harper & Row, Publishers, Inc., New York.
52. **Saier, M. H., Jr.** 1990. Evolution of permease diversity and energy coupling mechanisms: an introduction to a forum on coupling of energy to transmembrane solute translocation in bacteria. Res. Microbiol. **141**:281–286.
53. **Saier, M. H., Jr.** Mechanisms of carbohydrate transport in bacteria and comparisons with those in eukaryotes. *In* P. L. Yeagle (ed.), The structure and function of biological membranes, in press. Telford Press, West Caldwell, N.J.
54. **Saier, M. H., Jr., B. U. Feucht, and S. Roseman.** 1971. Phosphoenolpyruvate-dependent fructose phosphorylation in photosynthetic bacteria. J. Biol. Chem. **246**:7819–7821.
55. **Sharma, P. M., G. R. Reddy, S. Vora, B. M. Babior, and A. McLachlan.** 1989. Cloning and expression of a human muscle phosphofructokinase cDNA. Gene **77**:177–183.
56. **Shinozaki, K., and M. Sugiura.** 1982. The nucleotide sequence of the tobacco chloroplast gene for the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. Gene **20**:91–102.
57. **Staden, R.** 1986. The current status and portability of our sequence handling software. Nucleic Acids Res. **14**:217–231.

57a. **Stewart, G. C.** Personal communication.

58. **Thornburgh, B. N., L. L. Wu, and C. C. Griffin.** 1977. Phosphofructokinase from *Escherichia coli*: further evidence for identical subunits. Can. J. Biochem. **56**:836–838.
59. **White, D. W., and G. R. Jacobson.** 1990. Molecular cloning of the C-terminal domain of *Escherichia coli* D-mannitol permease: expression, phosphorylation, and complementation with C-terminal permease deletion proteins. J. Bacteriol. **172**:1509–1515.
60. **Wu, L.-F., and M. H. Saier, Jr.** 1990. Nucleotide sequence of the *fruA* gene encoding the fructose permease of the phosphotransferase system from *Rhodobacter capsulatus* and analyses of the deduced protein sequence. J. Bacteriol. **172**:7167–7178.

60a. **Wu, L. F., and M. H. Saier, Jr.** Unpublished data.

61. **Wu, L.-F., J. M. Tomich, and M. H. Saier, Jr.** 1990. Structure and evolution of a multidomain, multiphosphoryl transfer protein: nucleotide sequence of the *fruB(HI)* gene in *Rhodobacter capsulatus* and comparisons with homologous genes from other organisms. J. Mol. Biol. **213**:687–703.
62. **Zurawski, G., P. R. Whitfeld, and W. Bottomley.** 1986. Sequence of the gene for the large subunit of ribulose 1,5-bisphosphate carboxylase from pea chloroplasts. Nucleic Acids Res. **14**:3975–3976.