Cloning, sequencing and expression of the pyrophosphate-dependent phosphofructo-1-kinase from *Naegleria fowleri*

Kathleen L. WESSBERG,* Sara SKOLNICK,* Jun XU,* Francine MARCIANO-CABRAL† and Robert G. KEMP*‡

*Department of Biological Chemistry, Chicago Medical School, North Chicago, IL 60064, U.S.A.
and †Department of Microbiology and Immunology, Medical College of Virginia, Richmond, VA 23298, U.S.A.

The cDNA for the PP_i-dependent phosphofructo-1-kinase has been cloned and sequenced from a cDNA library prepared from the free-living amoeba *Naegleria fowleri*. The coding sequence of the cDNA consists of 1311 bases which translates into 437 amino acids with a molecular mass of 48095 Da. Comparison of the sequence with those of the previously described sequences of PP_i-dependent phosphofructokinases from *Propionibacterium freudenreichii* and potato tuber revealed amino acid identities of 23 and 28% respectively and high conservation in those regions assumed to be part of the active site. The reading frame was cloned into an expression vector, which was transformed into *Escherichia coli*. Extracts of the transformed cells contained PP_i-dependent phosphofructokinase activity that could be purified to homogeneity. The activity was lost on incubation with the chaotropic agent, KSCN, and recovered by subsequent incubation with AMP. These properties are consistent with those described by Mertens, De Jonckheere and Van Schaftingen [Biochem. J. (1993) **292**, 797–803] for the enzyme prepared from *Naegleria* and support the idea that the cloned cDNA coded for the complete native enzyme. No nucleotide-binding motif or evidence for a nucleotide-binding site characteristic of the ATP-dependent phosphofructokinases could be found within the primary structure.

INTRODUCTION

Phosphofructokinase (PFK) is established as a key regulatory enzyme in glycolysis. The most broadly distributed form of the enzyme utilizes ATP as a phosphoryl donor and is subject to allosteric control. However, a second type of PFK that utilizes PP_i as a phosphoryl donor has been found in a variety of different parasitic protists, including *Entamoeba histolytica* (Reeves et al., 1974), *Trichomonas vaginalis* (Mertens et al., 1989), *Giardia lambia* (Mertens, 1990), *Toxoplasma gondii* (Peng and Mansour, 1992) and the free-living amoeba *Naegleria fowleri* (Mertens et al., 1993). A PP_i-dependent PFK is found also in higher plants where it is subject to regulatory control by fructose 2,6-bisphosphate. Whereas plants have both ATP- and PP_i-dependent PFKs, the protists appear to have one or the other. The relationship of the protozoan PP_i-dependent PFKs to the other types of PFK is particularly interesting. The pattern of occurrence in protozoa of ATP-dependent and PP_i-dependent PFKs appears to be somewhat random relative to the proposed phylogenetic relationship among these organisms (Sogin, 1989). Although it is found in the amitochondrial protozoa such as *Giardia* and *Trichomonas*, it has been found also in several protozoa with aerobic pathways such as *Toxoplasma* (Peng and Mansour, 1992) and *Naegleria* (Mertens et al., 1993). The random pattern could reflect a type of convergence in which the PP_i-PFKs of the protists each arose from mutation of an ATP-PFK rather the presence of two genes, one of which is silent or deleted from a particular lineage.

Until recently the PP_i-PFKs of bacteria and protists were thought not to have allosteric properties, but Mertens et al. (1993) characterized the PP_i-PFK of *Naegleria fowleri* and found stimulation by micromolar concentrations of AMP. Although stimulation by adenine nucleotides is reminiscent of the properties of the ATP-dependent PFKs, the mechanism was unique. AMP leads to the association of inactive monomers into active tetramers. Still, the presence in this organism of a binding site for AMP that is not present in more primitive protists suggests a close relationship to the ATP-dependent PFKs.

In the current paper, we deduce the primary structure of *N. fowleri* PFK from its cDNA sequence and compare that structure with those of other ATP- and PP_i-dependent PFKs. This represents the first published sequence of a protozoan PFK. The cloned enzyme has been expressed, purified to homogeneity, and examined for the regulatory action of AMP.

EXPERIMENTAL

Oligonucleotides

The oligonucleotides used for PCR amplification were synthesized by phosphoramidite chemistry on a Biosearch 8700 DNA synthesizer, deprotected and purified on a Sep-Pak column. Those oligonucleotides used for DNA sequencing were synthesized, deprotected, taken to dryness and used without further purification.

PCR cloning of *Naegleria* PFK

An initial amplification of a short segment of the PP_i-PFK gene was accomplished using degenerate primers and genomic DNA

Abbreviations used: ATP–PFK, ATP:fructose 6-phosphate 1-phosphotransferase (EC 2.7.1.11); PP_i-PFK, pyrophosphate:fructose 6-phosphate 1-phosphotransferase (EC 2.7.1.90); Fru 6-P, fructose 6-phosphate; IPTG, isopropyl thiogalactoside.

‡ To whom correspondence should be addressed.

The sequence data in this paper will appear in the EMBL/GenBank/DDBJ Nucleotide Sequence Databases under the accession number U11733.

of *N. fowleri* that was generously provided by Dr. E. Mertens of the Catholic University of Louvain. Primers ACTATT/CGGTGGTGAT/CGAT/CAC and A/GTCA/GTTA/GTCAATAGTCTTT/AGG were employed. Polymerization was carried out with *Taq* polymerase using the standard conditions provided by the supplier (Perkin–Elmer–Cetus) except for the inclusion of 6 mM $MgCl_2$ and 5% dimethyl sulphoxide. Cycling temperatures of 94 °C, 45 °C and 67 °C were used for 1 min each for 30 cycles. The most prominent product had a size of just over 100 bp as indicated by PAGE. The band was excised and run into a 1% agarose gel. The band was cut out, digested with agarose, and used directly for a second round of amplifications using the identical conditions as the first set of cycles. A single band was seen on PAGE. This band was removed and again run into agarose. The DNA band was cut from the gel, the gel digested with agarose, and the DNA precipitated with ethanol. The fragment was ligated into the AT cloning vector, PCRII (Invitrogen), which was subsequently transformed into competent INV cells (Invitrogen), which were in turn plated on blue/white selective medium. White colonies were selected and screened for the presence of an appropriately sized insert. A single colony containing an insert was grown for isolation of plasmid and subsequent sequencing.

Preparation of the cDNA library

N. fowleri LEE (ATCC 30894) was isolated from a fatal human case of primary amoebic meningoencephalitis and cultured axenically in Nelson medium (Martinez et al., 1973; Marciano-Cabral, 1988). Trophozoites were harvested by centrifugation, and total cellular RNA was obtained using guanidinium thiocyanate. Polyadenylated RNA was isolated using an oligo(DT)-cellulose column. The BRL (Gaithersburg, MD, U.S.A.) Superscript Plasmid System was employed for cDNA synthesis and cloning according to the manufacturer's instructions. First- and second-strand synthesis of cDNA was performed using reverse transcriptase and *E. coli* DNA polymerase respectively (Gubler and Hoffman, 1983). A directional cDNA library was constructed with *Not*I adapters at the 3′ end and *Sal*I adapters on the 5′ end, to enable directional cloning. The cDNA was subjected to size fractionation using a Sephacryl S-500 column. cDNA (longer than 500 bp) was ligated into predigested pSPORT 1 plasmid vector and introduced into DH10B cells by electroporation. Blue/white screening of colonies on 5-bromo-4-chloro-3-indolyl β-D-galactoside plates was performed to determine the frequency of pSPORT plasmid transformants.

PCR amplification of the PP_i-PFK in two fragments

Total plasmid DNA was isolated from the cDNA library and subjected to PCR amplification in two separate reactions. One reaction, which was designed to amplify the 5′ third of the cDNA, employed as primers the M13 sense primer, CAGGAAACAGCTATGAC, and an antisense 39-mer designed from the sequence of the initial 102 bp amplification product with the sequence TAACGTGACAACGTGAATTTCATTATTAGCCTTAGC. The second reaction, designed to amplify the 3′ two-thirds of the cDNA, contained the M13 antisense primer, GTAAAACGACGGCCAG, and a sense 39-mer also based on the first PCR-amplified product and having the sequence GCATTCAGTAGTATGAGTAATATTGCTAAGGCTGCTAAT. The PCR program consisted of 30 s at 94 °C, 1 min at 54 °C, and 1.5 min at 72 °C for 35 cycles. After electrophoresis of the PCR reaction mixture on agarose, the bands corresponding to approx. 0.6 and 1.5 kb were excised. The DNA was purified using the Geneclean protocol (BIO 101 Inc.) and cloned into the AT cloning vector PCRII. Clones were chosen and verified for the presence of the appropriate-sized insert by restriction digestion. Several clones were used for sequencing.

Sequencing

Double-stranded sequencing of each of the two PCR fragments that had been cloned into PCRII was accomplished using the Sequenase protocol (USB) utilizing alkaline denaturation of the DNA. Initially, the PCR primers as well as forward and reverse vector sequencing primers were used. On the basis of the sequence obtained, additional primers were synthesized to continue 'walking' through the sequence. The entire coding sequence was determined in both directions. Several areas of uncertainty were verified by sequencing the full-length reading frame described below.

Cloning of a full-length reading frame and expression

To obtain a full-length reading frame, PCR primers were designed on the basis of the determined 5′- and 3′-terminal sequences. The 5′-primer was designed to insert an *Nde*I restriction site coincident with the initiation codon. The antisense primer introduced a *Bam*HI restriction site in the 3′ non-coding region. The 5′-primer was CATCGCVCATATGTTGAGTTCATCCCA and the reverse primer was TAATACGGATCCACACCCACACATAC. PCR was performed using the cDNA library as template for 30 cycles of 20 s at 94 °C, 40 s at 55 °C and 1 min at 72 °C. A single band of approx. 1.4 kb was detected on electrophoresis; this was purified and cloned into the AT cloning vector pGEM. The vector was transformed into competent INFαF′ cells (Invitrogen), which were plated and grown. Colonies were selected and grown in order to find one or more containing the vector with the insert. The selected plasmid was purified, digested with *Nde*I and *Bam*HI, and the insert purified and cloned into the *Nde*I–*Bam*HI site of the pET3a vector, which was transformed subsequently into Bal21 cells. Colonies were isolated for growth and analysis of enzyme activity in extracts.

Purification of enzyme activity

The purification procedure was a modification of the method described for the preparation of the enzyme from *Propionibacterium* (Ladror et al., 1991). *E. coli* Bal21 cells bearing the pET vector were grown at 37 °C in Luria–Bertani media with 0.5 mM isopropyl thiogalactoside (IPTG), to an absorbance of 1.8–2.0 at 600 nm and were harvested by centrifugation at 6000 *g*. The cells were frozen and stored at −20 °C. Extraction buffer (0.2 M potassium phosphate, pH 6.8, containing 1 mM dithiothreitol and 0.1 mM EDTA) (2 vol.) were added to the frozen sediment, which was stirred until the cells were suspended. The suspension was subjected to sonication with a microprobe at maximum setting for three 20 s periods with intervening chilling in an ice-bath. The suspension was centrifuged for 10 min at 8000 *g*. The supernant was saved and the sediment was resuspended in extraction buffer and subjected to four cycles of sonication as described above. The suspension was centrifuged and the two extracts were combined. Using a 10% aqueous solution, polyethyleneimine was added to a final concentration of 0.5%, and the solution was stirred slowly at 4 °C for 30 min. The precipitate was removed by centrifugation. Solid $(NH_4)_2SO_4$ (43 g/100 ml) was added to the supernatant solution slowly with stirring. The suspension was further stirred in an ice-bath for 30 min. The precipitate was collected by centrifugation at 8000 *g*

for 15 min and dissolved in column buffer consisting of 50 mM Tes (potassium salt) (pH 7.0), 0.01 mM EDTA and 0.1 mM AMP. The dissolved sediment was dialysed in the cold against three changes of 2 litres each of the column buffer. The dialysate was applied at room temperature to a column of phosphocellulose (Whatman P-11) previously equilibrated with column buffer. The column was washed with 3–4 column vol. of column buffer, and the enzyme was eluted with column buffer plus 0.5 mM fructose 1,6-bisphosphate. Enzyme eluted by the sugar bisphosphate was more than 70% pure with a recovery of about 40%. Fractions containing enzyme activity was concentrated using a Centricon membrane (Amicon). The concentrated enzyme was loaded on to a Pharmacia Mono Q HR 5/5 column previously equilibrated with 20 mM Tris/HCl, pH 8.1, containing 0.1 mM AMP. The enzyme was eluted with a gradient of NaCl from 100 to 500 mM; it appeared as a single peak at approx. 190 mM NaCl. SDS/PAGE indicated a single component. The overall yield was in the range of 25%.

Assay of enzyme activity

Enzyme activity for PP_i-PFK was assayed at 30 °C and at pH 7.2 in a Gilford Response spectrophotometer. The assay medium contained 50 mM Tes buffer, 0.1 mM EDTA, 3 mM $MgCl_2$, 1 mM sodium pyrophosphate, 1.5 mM fructose 6-phosphate, 0.2 mM NADH and a mixture of auxiliary enzymes consisting of 0.6 unit of aldolase and 0.3 unit each of triosephosphate isomerase and glycerophosphate dehydrogenase. To detect total activity, 1 mM AMP was included in the assays. Blank reactions were measured by omitting either fructose 6-phosphate or PP_i from the reaction mixture containing the cell extracts. Reactions were initiated by the addition of either PP_i or fructose 6-phosphate. All reagents and assay auxiliary enzymes were purchased from Sigma.

RESULTS

Cloning strategy

At present, the sequences of PP_i-PFK from only two species are available to search for conserved stretches of common sequence: *Propionibacterium freudenreichii* (Ladror et al., 1991) and *Solanum tuberosum* (Carlisle et al., 1990). Two sequences that are almost completely conserved in the PP_i-PFK from *Propionibacterium* and potato are positions 117–123 (TIGGDDT) and 147–153 (PKTIDND) of *Propionibacterium*. The analogous sequences of ATP-dependent PFKs have been shown to be part of the active site (Shirakihara and Evans, 1988). Furthermore, in *Propionibacterium*, aspartate residues 151 and 153 have been implicated by site-directed mutagenesis experiments to be involved in catalysis and probably play a similar role to the aspartate residues shown to be in the active site of the ATP-dependent PFK (Green et al., 1993). On the basis of these observations, it was hypothesized that these two regions are crucial to catalytic activity of all PFKs and will be present in all PP_i-dependent PFKs. Degenerate oligomers were synthesized for each of these two areas, eliminating some of the possible degeneracy by taking into consideration the codon bias indicated by the sequences of *Naegleria* tubulin and actin that are available in the GenBank. Using the oligonucleotide primers described in the Experimental section, PCR was used to generate an approx. 100 bp fragment (after a number of variations in annealing temperature, Mg^{2+} concentration, dimethyl sulphoxide concentration, etc.). The insert was sequenced and found to be approximately the same size as the corresponding sequence of *Propionibacterium* PFK and to have some sequence similarity.

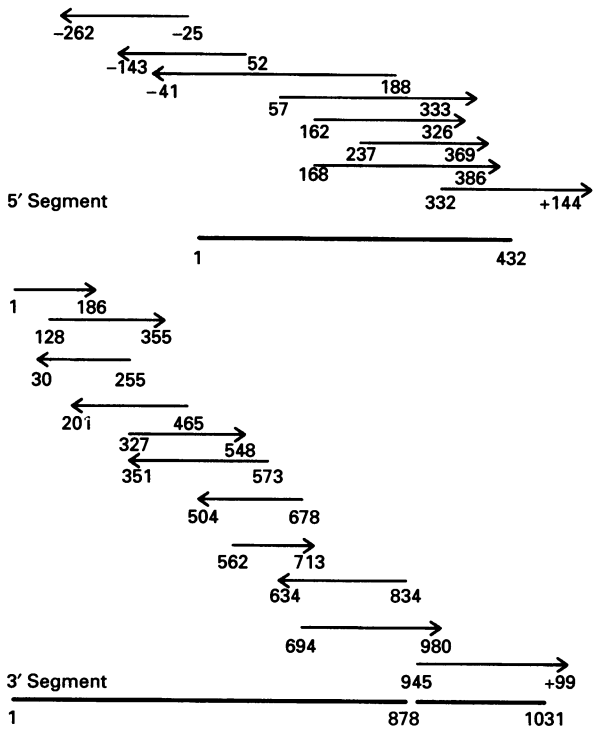

Figure 1 Sequencing strategy for the two overlapping fragments of the cDNA for the PP_i–PFK of *Naegleria*

Both forward and reverse PCR primers were designed on the basis of the unique DNA sequence determined from the initial PCR reaction. These two primers were used along with the forward and reverse primers of the cDNA cloning vector in two separate amplification reactions of the total cDNA library. This technique results in the cloning of the entire cDNA coding region for *Naegleria* PFK in two fragments, each starting with the segment of DNA originally amplified and terminating within the cDNA cloning vector, pSPORT.

The sequence

The two PCR fragments described above were ligated into PCRII, which was subsequently transformed into the INV host strain. Several white colonies from each plate representing the 3′ and 5′ halves of the cDNA of *Naegleria* were selected for sequencing. Figure 1 describes the sequencing strategy for each of the two overlapping fragments of the cDNA. The entire coding frame was sequenced in both directions from two different clones. The sequence of the total insert is described in Figure 2. The coding sequence consists of 1311 bases which translates into 437 amino acids. This predicts a molecular size of 48096 Da, which agrees reasonably well with the 51 kDa estimated by Mertens et al. (1993) from SDS/PAGE. The stop codon was identified as UAG, which differs from the translation termination codon of the tubulins which have been shown to overlap the polyadenylation signal in a very short 3′ untranslated region for those genes (Clark, 1990). The 3′ untranslated region of the PFK cDNA is considerably longer than that of the tubulin products and consists of 135 bases up to the poly(A) tail. The polyadenylation signal is located 14 bases from the tail and is ACUAAA, which is one of the rarer variants of typical signals found in higher eukaryotes (Wickens, 1990). Clark (1990) has

```
      ccacgcgtccgaacacactccttcttcatcgcgtc
  1   M  L  S  S  S  H  L  P  T  T  I  V  T  P  K  N  V  P  T  L  G  V  L  V  G  G  G  P  A  P  G  I   32
  1   ATGTTGAGTTCATCCCATCTACCTACGACCATTGTAACTCCAAAGAATGTTCCCACATTGGGTGTTCTTGTGGGTGGAGGTCCTGCTCCAGGTATT  96

 33   N  G  V  I  G  A  V  T  I  E  A  I  N  N  G  Y  R  V  L  G  F  L  E  G  F  Q  N  L  I  L  Q  D   64
 97   AATGGTGTCATTGGTGCAGTCACTATTGAAGCCATTAACAATGGTTATCGTGTGTTGGGATTCTTGGAAGGTTTTCAAAATCTCATCTTGCAAGAT  192

 65   D  S  K  I  V  E  L  T  I  D  S  V  S  R  I  H  F  E  G  G  S  I  L  K  T  S  R  A  N  P  T  K   96
193   GATTCTAAAATTGTTGAATTAACCATTGATTCTGTTTCAAGAATTCACTTTGAAGGTGGAAGTATTTTAAAAACCTCACGTGCCAATCCAACCAAG  288

 97   K  Q  E  D  L  Q  K  V  V  K  Q  L  Q  K  F  N  V  S  L  L  V  T  I  G  G  D  D  T  A  F  S  S  128
289   AAACAAGAGGATTTACAAAAAGTGGTGAAACAATTACAAAAGTTTAATGTTTCCTTATTGGTTACTATTGGAGGTGATGATACCGCATTCTCTAGT  384

129   M  S  V  A  K  A  A  N  N  E  I  H  V  C  H  V  P  K  T  I  D  N  D  L  P  L  P  Y  G  I  P  T  160
385   ATGAGTGTTGCTAAGGCTGCTAATAATGAAATTCATGTGTGCCATGTTCCTAAAACGATCGATAACGATCTTCCACTTCCGTATGGAATTCCTACT  480

161   F  G  Y  E  T  A  R  E  F  G  A  N  V  V  R  N  L  M  T  D  A  S  T  A  S  R  Y  F  I  V  V  A  192
481   TTTGGTTATGAAACAGCACGTGAATTTGGCGCCAATGTTGTGCGTAATCTCATGACTGATGCTTCGACTGCAAGTCGTTACTTTATTGTTGTGGCC  576

193   M  G  R  Q  A  G  H  L  A  L  G  I  G  K  S  A  G  S  H  L  T  L  I  P  E  E  F  L  P  T  T  D  224
577   ATGGGTCGTCAGGCTGGTCACTTGGCATTAGGTATTGGAAAATCTGCTGGCTCACATTTGACCTTAATTCCTGAAGAATTTTTACCAACCACCGAT  672

225   S  T  E  P  E  V  T  F  S  R  I  C  D  M  I  E  A  S  I  I  K  R  L  Y  T  S  K  K  D  H  G  V  256
673   AGTACTGAACCTGAAGTGACCTTTTCTCGAATTTGCGACATGATTGAAGCAAGTATTATTAAGAGATTATACACTTCCAAGAAAGATCATGGTGTA  768

257   I  V  L  A  E  G  L  L  E  Y  M  S  T  D  E  L  K  Q  A  F  G  S  S  L  K  Y  D  A  H  D  H  I  288
769   ATTGTATTGGCAGAAGGTCTCTTGGAATATATGAGTACCGATGAATTGAAACAAGCCTTTGGTTCCTCATTGAAATACGATGCTCATGATCATATC  864

289   M  L  A  E  L  D  F  G  R  L  V  R  D  E  M  R  E  R  M  N  R  R  G  L  K  I  A  F  T  E  K  N  320
865   ATGTTGGCCGAATTGGACTTTGGTCGATTGGTGCGCGATGAAATGAGAGAGCGCATGAATCGTCGTGGACTCAAAATTGCCTTCACTGAAAAAAAT  960

321   L  G  Y  E  L  R  C  A  P  P  N  A  F  D  R  E  Y  T  R  D  L  G  N  G  A  V  R  Y  L  L  N  G  352
961   TTGGGTTATGAATTGAGATGCGCCCCTCCCAATGCCTTTGATCGTGAATATACTCGGGATTTGGGTAATGGTGCAGTGAGATACTTGTTGAATGGA  1056

353   G  N  G  A  L  I  T  V  Q  G  V  K  M  V  P  L  S  F  D  D  L  K  D  P  R  T  G  K  T  R  T  R  384
1057  GGAAACGGAGCCTTAATTACTGTTCAAGGAGTGAAAATGGTACCTCTCAGTTTTGATGATTTGAAAGATCCTCGCACAGGAAAGACTCGTACCCGT  1152

385   Q  V  D  V  S  S  E  G  F  Q  V  A  K  R  Y  M  I  R  L  E  K  K  D  F  E  K  E  E  T  L  K  G  416
1153  CAAGTTGATGTTTCCAGTGAGGGTTTCCAAGTGGCCAAACGATACATGATCCGATTGGAGAAGAAGGACTTTGAGAAGGAAGAGACTTTGAAGGGT  1248

417   L  A  A  T  A  K  C  S  V  E  D  F  I  K  Q  F  K  Y  L  V  Q                                    437
1249  TTGGCTGCCACTGCAAAATGTAGTGTTGAAGATTTCATCAAACAATTCAAGTATTTGGTGCAA tagagtcagtagtagtatatgagcatatatg
      tatgtgtgggtgtgggtgtgtattataatattgtgagtatagattctattgatcgaagaggaatggccatccccacccctaaatacactaaacgctt
      tgccttgtaaaaaaaaaaaaaaaa
```

Figure 2 cDNA sequence and derived amino acid sequence of the PP_i-PFK of *Naegleria*

Table 1 Codon usage of the PP_i-PFK gene from *N. fowleri*

	Second position												
	U			C			A			G			
First position	Amino acid	No.	%	Amino acid	No.	%	Amino acid	No.	%	Amino acid	No.	%	Third position
U	F	13	3.0	S	5	1.1	Y	7	1.6	C	1	0.2	U
	F	6	1.4	S	5	1.1	Y	5	1.1	C	2	0.5	C
	L	11	2.5	S	5	1.1	*	0	0	*	0	0	A
	L	25	5.7	S	1	0.2	*	1	0.2	W	0	0	G
C	L	3	0.7	P	9	2.1	H	5	1.1	R	11	2.5	U
	L	5	1.1	P	2	0.5	H	3	0.7	R	3	0.7	C
	L	1	0.2	P	5	1.1	Q	12	2.7	R	4	0.9	A
	L	0	0	P	1	0.2	Q	1	0.2	R	1	0.2	G
A	I	22	5.0	T	14	3.2	N	15	3.4	S	11	2.5	U
	I	5	1.1	T	11	2.5	N	3	0.7	S	0	0	C
	I	0	0	T	3	0.7	K	17	3.9	R	5	1.1	A
	M	11	2.5	T	3	0.7	K	12	2.7	R	0	0	G
G	V	14	3.2	A	9	2.1	D	22	5.0	G	23	5.3	U
	V	3	0.7	A	12	2.7	D	3	0.7	G	2	0.5	C
	V	4	0.9	A	9	2.1	E	22	5.0	G	12	2.7	A
	V	12	2.7	A	0	0	E	6	1.4	G	0	0.0	G

reported that in previous published sequences from *Naegleria* the most common polyadenylation signal, AAUAAA, was found. The 5′ untranslated region consists of 35 bases if one assumes that a full-length transcript has been cloned. There is no obvious analogue of the Shine–Delgarno sequence upstream of the start codon.

Codon usage

The codon usage of *N. fowleri* PFK is described in Table 1. The AT/GC ratio was 59:41 and the base usage in the third position showed an AT preference of 67%. This is consistent with previous reports of high AT content in *Naegleria* (Mandel, 1967). Five of the seven codons not seen in *Naegleria* PFK DNA, CTG, AGG, ATA, GCG and GGG, are not found or are present at very low levels in *E. coli* coding sequences. Note the *Naegleria* PFK does not contain tryptophan.

Expression and purification

Because PP_i-dependent PFK is not present in *E. coli*, no activity was detected in assays of extracts of cells that do not contain the plasmid with insert. When Bal21 cells containing the pET3a with the full-length *Naegleria* PFK insert were grown in Luria–Bertani medium, low enzyme activity was detected in extracts (approx. 1.0 unit/g of cells). If, however, IPTG was added to the culture as described in the Experimental section, approx. 9.8 units of activity per g of cells was present in cell extracts. The enzyme was purified to homogeneity using a modification of the method described by Ladror et al. (1991) for the PP_i-dependent enzyme from *Propionibacterium* (see the Experimental section). The overall yield was approx. 25% and a single band corresponding to a mass of about 50 kDa was observed on SDS/PAGE. Kinetic studies of the purified enzyme indicated a K_m for PP_i of 14 μM and a K_m of 48 μM for fructose 6-phosphate, values reasonably close to those reported by Mertens et al. (1993).

AMP activation

AMP activation could be demonstrated in crude extracts or with purified enzyme. As described by Mertens et al. (1993), the enzyme can be inactivated by incubation in the presence of the chaotropic agent, KSCN. Purified PP_i-dependent PFK (5.2 units/ml) was dialysed against 20 mM Tris/HCl buffer, pH 8.1. The enzyme was diluted 1:5 with the Tris buffer, and KSCN was added to a final concentration of 0.75 mM. Incubation on ice led to the loss of 95% of the initial activity within 30 min. This inactivated preparation was diluted 1:5 in the Tris buffer containing 1 mM AMP. On incubation at 37 °C, activity was recovered in a time-dependent manner with activity reaching 96% of the original activity. These results are in agreement with the data obtained by Mertens et al. (1993) using enzyme purified from *Naegleria*.

DISCUSSION

The derived sequence for the *Naegleria* PP_i-dependent PFK shown in Figure 2 appears to represent the primary structure of the active enzyme. Cloning the full-length coding sequence from the ATG codon to past the stop codon into the expression vector followed by transformation into *E. coli* resulted in the presence of PP_i-dependent PFK activity in cell extracts. The production of enzyme activity was dependent on the presence of the vector and was induced by IPTG. When a sample of the extract was passed through a Superose 6 column equilibrated with extraction buffer, enzyme activity was eluted in the region expected for a protein with a molecular mass in the range of 200 kDa (results not shown). This indicates that the enzyme behaves as a tetramer and is compatible with the properties of the enzyme described by Mertens et al. (1993).

The amino acid sequence of PP_i-dependent PFK from *Naegleria* that was derived from the cDNA sequence can be readily aligned with other PP_i-PFKs as described in Figure 3. This alignment was achieved by an initial approximation using the Pile-Up program of the GCG package of structural analysis programs. Further adjustments were made by eye. This sequence represents the first reported sequence of a PP_i-PFK from protists and is, in fact, the first sequence of any PFK from this group of organisms. Although insufficient data are available for conclusions regarding phylogeny, *Naegleria* clearly stands between the bacterial and the potato enzymes in terms of relative similarity. The level of identity of the *Naegleria* enzyme with the other PP_i-dependent enzymes is relatively low: approx. 23% between *Naegleria* and *Propionibacterium* and 28% between *Naegleria* and the β-subunit of the potato enzyme. Although the relationship is distant, the enzymes are clearly homologous. This contention is supported by the conservation of residues established to be part of the active site of the PP_i-PFK of *Propionibacterium*. Striking sequence similarities among the PP_i-PFKs are seen in the regions of residues 38–50, 140–148, 167–175, 216–218, all of which align with active-site regions of the ATP-dependent PFK from *E. coli* (Figure 3). Furthermore, residues K-169, D-173, D-175, K-336 and R-349 in Figure 3 correspond to positions in the *Propionibacterium* enzyme that have been established to be part of the active site on the basis of site-directed mutagenesis studies (Green et al., 1993; Xu et al., 1994). Note that identity is complete among the three PP_i-dependent enzymes at these five positions.

The cloned and purified enzyme demonstrated activation by AMP that was similar to that seen for the enzyme isolated from the organism by Mertens et al. (1993). This indicates that the binding site for AMP is encoded in the indicated sequence and is not the consequence of an interaction with an additional regulatory protein or subunit in *Naegleria*. AMP activation of the *Naegleria* enzyme does not appear to be a typical allosteric phenomenon. Whereas ADP activation of the bacterial ATP-dependent PFK and AMP activation of mammalian PFK are allosteric events associated with conformational changes that enhance the affinity of the enzyme for substrate, the AMP activation of *Naegleria* appears to be an effect associated with the conversion of the enzyme from an inactive into an active conformation, possibly associated with the displacement of the equilibrium between inactive monomer and active tetramer (Mertens et al., 1993). Nonetheless, the phenomenon shows specificity for the nucleotide (Mertens et al., 1993) and thus requires the presence of a relatively specific binding site on the enzyme. One aspect of the motivation to sequence the enzyme was the possible existence of a binding site for AMP similar to the nucleotide-binding site of the ATP-dependent PFKs, which would indicate either a closer relationship between the enzyme families than previous data suggest or domain swapping that led to the acquisition of an AMP-binding domain. One can align the *Naegleria* enzyme and the ATP-dependent enzyme from *E. coli* to achieve a level of identity of about 25% (Figure 3). Such an alignment indicates a relatively high degree of identity in the fructose-6-phosphate-binding site and in the phosphate-binding regions of ATP (or PP_i). On the other hand, residues that are known to be important for binding of nucleotide to the allosteric site in *E. coli* PFK (Shirakihara and Evans, 1988), which are indicated in Figure 3 by a hash sign (#), do not align with amino acid residues of identical or similar properties in the *Naegleria* enzyme nor is there any similarity of primary sequence in those regions. It is obvious that domain swapping with the homologous ATP-PFK to generate an AMP-binding site has not occurred nor is there evidence that a PP_i-PFK ancestor retained the AMP-binding site characteristic of ATP–PFK. The other possibility regarding the evolution of the structural basis for the binding of AMP is that it may have evolved from that part of the

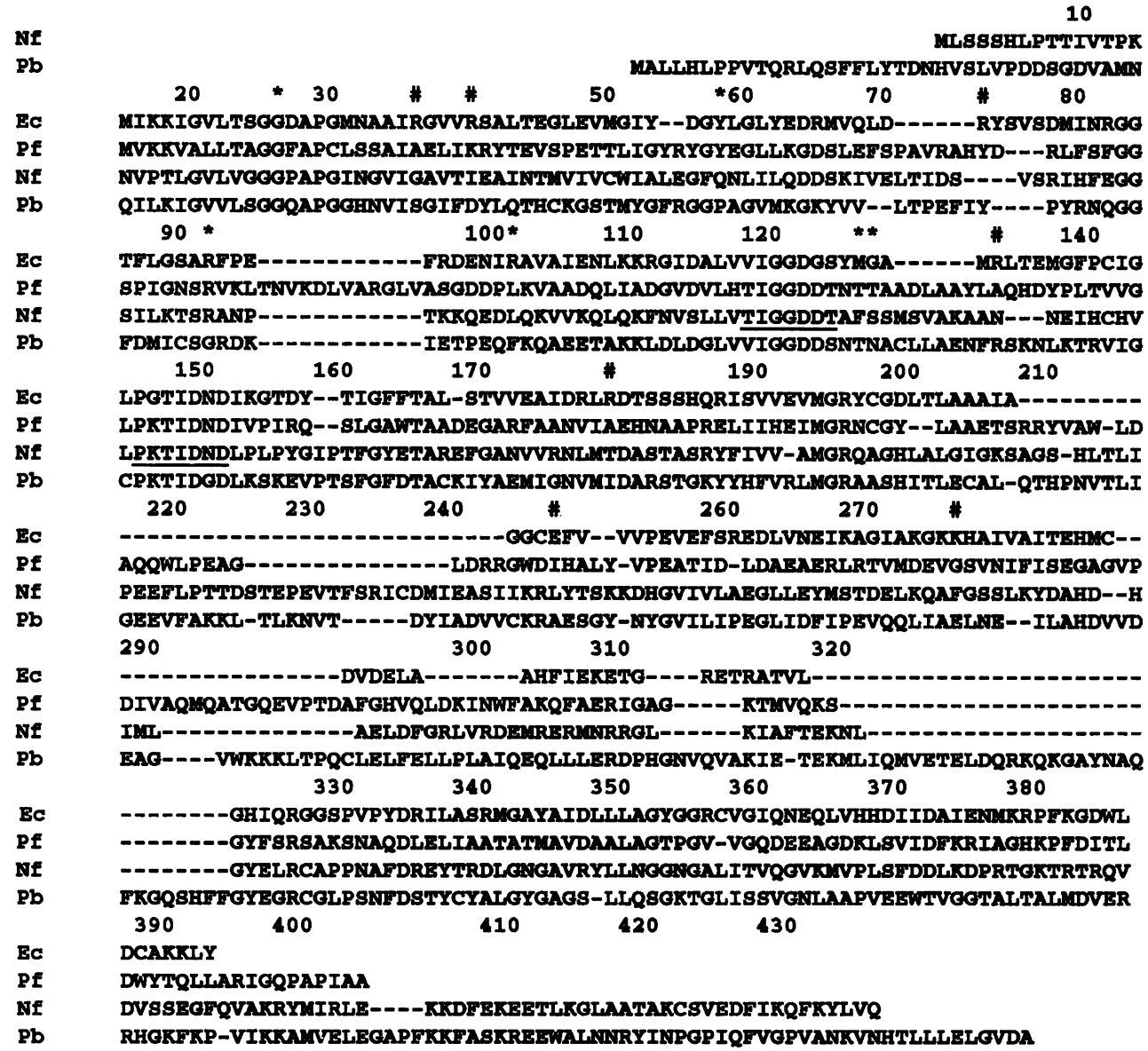

Figure 3 Sequence comparison among PFKs from *E. coli* (Ec), *P. freudenreichii* (Pf), *N. fowleri* (Nf) and the potato β-subunit (Pb)

The indicated numbering system is for the *Naegleria* enzyme. As discussed in the text, residues indicated by an asterisk (*) are those residues of *E. coli* ATP–PFK that have been implicated in the binding of the adenosine part of the catalytic ATP. Residues marked by a hash sign (#) are those residues of the ATP-dependent PFK that contribute to the binding of AMP (Shirakihara and Evans, 1988). The underlined residues starting at positions 120 and 147 represent the sequences corresponding to the PCR primers and are aligned with known active-site sequences of *E. coli* PFK (Shirakihara and Evans, 1988).

ATP-binding site of ATP-dependent PFK that bound the nucleoside and α-phosphate, i.e. the part of the site that does not include the terminal two phosphates that would comprise the PP_i-binding region. Those residues that appear to be important for nucleotide binding to the catalytic site of ATP-dependent PFK are indicated in Figure 3 by an asterisk. No sequence similarity is observed between these regions and the *Naegleria* enzyme indicating that the AMP-binding site is not likely to be a remnant of the ATP-binding site of ATP-dependent PFKs. A number of nucleotide-binding proteins contain similar glycine-rich sequences that are thought to constitute the nucleotide-binding site (Walker et al., 1982; Fry et al., 1986). A search of the primary sequence of the *Naegleria* PFK did not yield a stretch of residues that bore any similarity to these previously described motifs. The location of the AMP-binding site of the *Naegleria* enzyme remains to be established.

We express our thanks to Dr. Emmanuel Mertens for the gift of the *Naegleria* DNA that was used in the original PCR amplification. This work was supported by NIH grants DK19912 and the Thomas F. Jeffress and Kate Miller Jeffress Memorial Trust, Richmond, VA.

REFERENCES

Carlisle, S. M., Blakeley, S. D., Hemmingsen, S. M., Trevanion, S. J., Hiyoshi, T., Kruger, N. J. and Dennis, D. T. (1990) J. Biol. Chem. **265**, 18366–18377

Clark, C. G. (1990) J. Protozool. **37**, 2S–6S
Fry, D. C., Kuby, S. A. and Mildvan, A. S. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 907–911
Green, P. C., Tripathi, R. L. and Kemp, R. G. (1993) J. Biol. Chem. **268**, 5085–5088
Gubler, U. and Hoffman, B. J. (1983) Gene **25**, 263–269
Ladror, U. S., Gollapudi, L., Tripathi, R. L., Latshaw, S. P. and Kemp, R. G. (1991) J. Biol. Chem. **266**, 16550–16555
Mandel, M. (1967) in Chemical Zoology (Florkin, M. and Scheer, B. T., eds.), pp. 541–572, Academic Press, New York
Marciano-Cabral, F. (1988) Microbiol. Rev. **52**, 114–133
Martinez, A. J., Duma, R. J., Nelson, E. C. and Mretta, F. L. (1973) Lab Invest. **25**, 468–475
Mertens, E. (1990) Mol. Biochem. Parasitol. **40**, 147–150
Mertens, E., Van Schaftingen, E. and Müller, M. (1989) Mol. Biochem. Parasitol. **37**, 183–190
Mertens, E., De Jonckheere, J. and Van Schaftingen, E. (1993) Biochem. J. **292**, 797–803
Peng, Z.-Y. and Mansour, T. E. (1992) Mol. Biochem. Parasitol. **54**, 223–230
Reeves, R. E., South, D. J., Blytt, H. J. and Warren, L. G. (1974) J. Biol. Chem. **249**, 7737–7741
Shirakihara, Y. and Evans, P. R. (1988) J. Mol. Biol. **204**, 973–994
Sogin, M. L. (1989) Am. Zool. **29**, 487–499
Walker, J. E., Saraste, M., Runswick, M. J. and Gay, N. J. (1982) EMBO J. **1**, 945–951
Wickens, M. (1990) Trends Biochem. Sci. **15**, 277–281
Xu, J., Green, P. C. and Kemp, R. G. (1994) J. Biol. Chem. **269**, 15553–15557

Received 18 July 1994/24 November 1994; accepted 28 November 1994