The EMBO Journal vol.4 no.13B pp.3811-3817, 1985

Topogenesis of microbody enzymes: a sequence comparison of the genes for the glycosomal (microbody) and cytosolic phosphoglycerate kinases of *Trypanosoma brucei*

Klaas A.Osinga[1,4], Bart W.Swinkels[1], Wendy C.Gibson[1], Piet Borst[1], Gerrit H.Veeneman[2], Jacques H.Van Boom[2], Paul A.M.Michels[1,3] and Fred R.Opperdoes[3]

[1]Division of Molecular Biology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, [2]Organic Chemistry Laboratory, State University Leiden, Gorlaeus Laboratory, PO Box 9502, 2300 RA Leiden, The Netherlands, and [3]Research Unit for Tropical Diseases, International Institute of Cellular and Molecular Pathology, Avenue Hippocrate 74, B-1200 Brussels, Belgium

[4]Present address, Department of Yeast Genetics, Gist Brocades NV, Postbus 1, 2600 MA Delft, The Netherlands

Communicated by P.Borst

To determine how microbody enzymes enter microbodies, we are studying the genes for cytosolic and glycosomal (microbody) isoenzymes in *Trypanosoma brucei*. We have found three genes (A, B and C) coding for phosphoglycerate kinase (PGK) in a tandem array in *T. brucei*. Gene B codes for the cytosolic and gene C for the glycosomal isoenzyme. Genes B and C are 95% homologous, and the predicted protein sequences share ~45% amino acid homology with other eukaryote PGKs. The microbody isoenzyme differs from the cytosolic form and other PGKs in two respects: a high positive charge and a carboxy-terminal extension of 20 amino acids. Our results show that few alterations are required to redirect a protein from cytosol to microbody. From a comparison of our results with the unpublished data for three other glycosomal glycolytic enzymes we infer that the high positive charge represents the major topogenic signal for uptake of proteins into glycosomes.

Key words: glycosome/phosphoglycerate kinase/topogenesis/peroxisome biogenesis/gene conversion

Introduction

The existence of many compartments in eukaryotic cells requires an elaborate internal mail system to deliver newly synthesized proteins to the correct compartments and the address labels involved have been deciphered in outline for some organelles. We are interested in the delivery of proteins to a group of organelles collectively known as microbodies, which include peroxisomes, glyoxysomes and glycosomes. Proteins that enter microbodies are made on free polysomes and are usually not synthesized as larger precursors (reviewed by Kindl, 1982; Lazarow and Fujiki, 1985). Genes coding for microbody enzymes are just being cloned and analysed and, although *in vitro* import of nascent protein into microbodies has been observed, the lability of the organelle has thus far precluded more detailed studies of the import mechanism. Hence, little is known of the sequences involved in the import of these proteins or of the import mechanism itself.

We are also interested in the origin and evolution of microbodies. For a long time microbodies were considered to be a sub-compartment of the endoplasmatic reticulum, but this idea has wilted under the pressure of more rigorous biochemical experiments. All evidence now indicates that microbodies arise from pre-existing microbodies (see Kindl, 1982; Lazarow and Fujiki, 1985). Where the first microbody came from is a matter of speculation. It might have originated from an endosymbiont, or from another cellular compartment and have become independent later in evolution.

The flagellated protozoa of the order Kinetoplastida provide a nice system to study some of these questions. In trypanosomes the conversion of glucose to 3-phosphoglycerate takes place in a microbody-like organelle, called the glycosome to indicate this fact (Opperdoes and Borst, 1977). Glycosomes have been found in all representatives of the Kinetoplastida analysed, including *Crithidia* (Cannata *et al.*, 1982), *Leishmania* (Coombs *et al.*, 1982) and *Trypanosoma cruzi* (Taylor *et al.*, 1980). Initially there was some discussion as to whether glycosomes can be assigned full membership of the microbody family (de Duve, 1982), because the glycosomes of *Trypanosoma brucei* lack catalase, the classical marker enzyme for peroxisomes and glyoxysomes. This controversy now seems settled. First, catalase has been found in glycosomes of *Crithidia* (Muse and Roberts, 1973; Opperdoes *et al.*, 1977); second, the morphology of glycosomes is identical to that of peroxisomes, even to the semi-crystalline core (Opperdoes *et al.*, 1984); and third, glycosomes contain two enzyme systems now considered typical of peroxisomes: a β-oxidation system for fatty acids (Hart and Opperdoes, 1984) and the key enzymes for the biosynthesis of ether lipids (Opperdoes, 1984). The presence of additional enzymes in this organelle, like the glycolytic system and enzymes involved in glycerol-P and nucleotide metabolism and in CO_2 fixation (see Opperdoes, 1985), is not principally different from species-specific additions found in microbodies of some other organisms, such as a glyoxylate cycle or enzymes involved in methanol, methylamine or alkane degradation (reviewed by Tolbert, 1981; Kindl and Lazarow, 1982; Lord and Roberts, 1983; Veenhuis *et al.*, 1983; Trelease, 1984; Lazarow and Fujiki, 1985). We therefore think that glycosomes are typical microbodies and that information on their biogenesis and evolutionary origin is relevant to other eukaryotes as well.

The presence of a glycolytic system in the microbodies of the Trypanosomatidae represents a remarkable topogenic achievement and offers a unique opportunity to study microbody biogenesis. Glycolytic enzymes belong to the oldest housekeeping enzymes in nature and sequence information is already available for enzymes from pro- and eukaryotes. The sequences of these enzymes are highly conserved in evolution and therefore suitable for the study of ancient evolutionary relationships. Since these enzymes have always been maintained in the cytosol of bacteria and eukaryotes it is of interest to see how much change is required to allow import into microbodies. Finally, *T. brucei* contains a cytosolic isoenzyme for some of its glycosomal glycolytic enzymes (Opperdoes and Borst, 1977; Opperdoes *et al.*, 1981), notably phosphoglycerate kinase (PGK; EC 2.7.2.3) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH; EC

1.2.1.12), which should allow an instructive sequence comparison.

As part of a program to characterize fully the glycolytic enzymes of trypanosomes, we have cloned and sequenced the genes for PGK, GAPDH and triosephosphate isomerase (TIM; EC 5.3.1.1) and studied the expression of these genes. In this paper we present our results for PGK.

Results

Isolation of trypanosome PGK genes

Two PGK isoenzymes – one cytosolic and one glycosomal – have been detected in *T. brucei* and their properties are summarized in Table I. Each enzyme contains only a single polypeptide chain, as judged by SDS-polyacrylamide gel electrophoresis and gel filtration (O.Misset, personal communication). To identify the genes for these proteins we relied on the high degree of conservation of the genes for glycolytic enzymes and used cloned DNA complementary to the yeast PGK mRNA to screen trypanosome DNA. A summary of the trypanosome cDNA and genomic recombinant DNA clones studied in more detail is presented in Figure 1 together with a map of the genomic area containing the PGK genes. Three genes with homology to the yeast probe and to each other were found in tandem array. These were the only sequences that hybridized with cDNA clone pTcPGK8 in genomic DNA from *T. brucei* under the conditions employed (0.1 × SSC at 65°C).

Sequence of trypanosome PGK genes

To sequence the three PGK genes identified, segments of the genomic DNA clones were sub-cloned into M13 and sequenced by the chain termination method. We have used a serial sequencing approach, making new primers to extend the sequence just obtained, to avoid making many small sub-clones. The sequen-

Table I. Properties of the two PGK isoenzymes of *T. brucei*

	Cytosolic PGK	Glycosomal PGK	
Fraction in BF[a]	<0.1	>0.9	O.Misset, personal communication
Fraction in CF[b]	~0.9	~0.1	Opperdoes *et al.*, (1981)
Molecular mass (kd)	45.0 ± 1.2	47.0 ± 0.7	O.Misset, personal communication
Isoelectric point	6.2–6.7[c]	9.0	O.Misset, personal communication

[a]Bloodstream form trypanosomes.
[b]Cultured insect form trypanosomes.
[c]Multiple bands were observed.

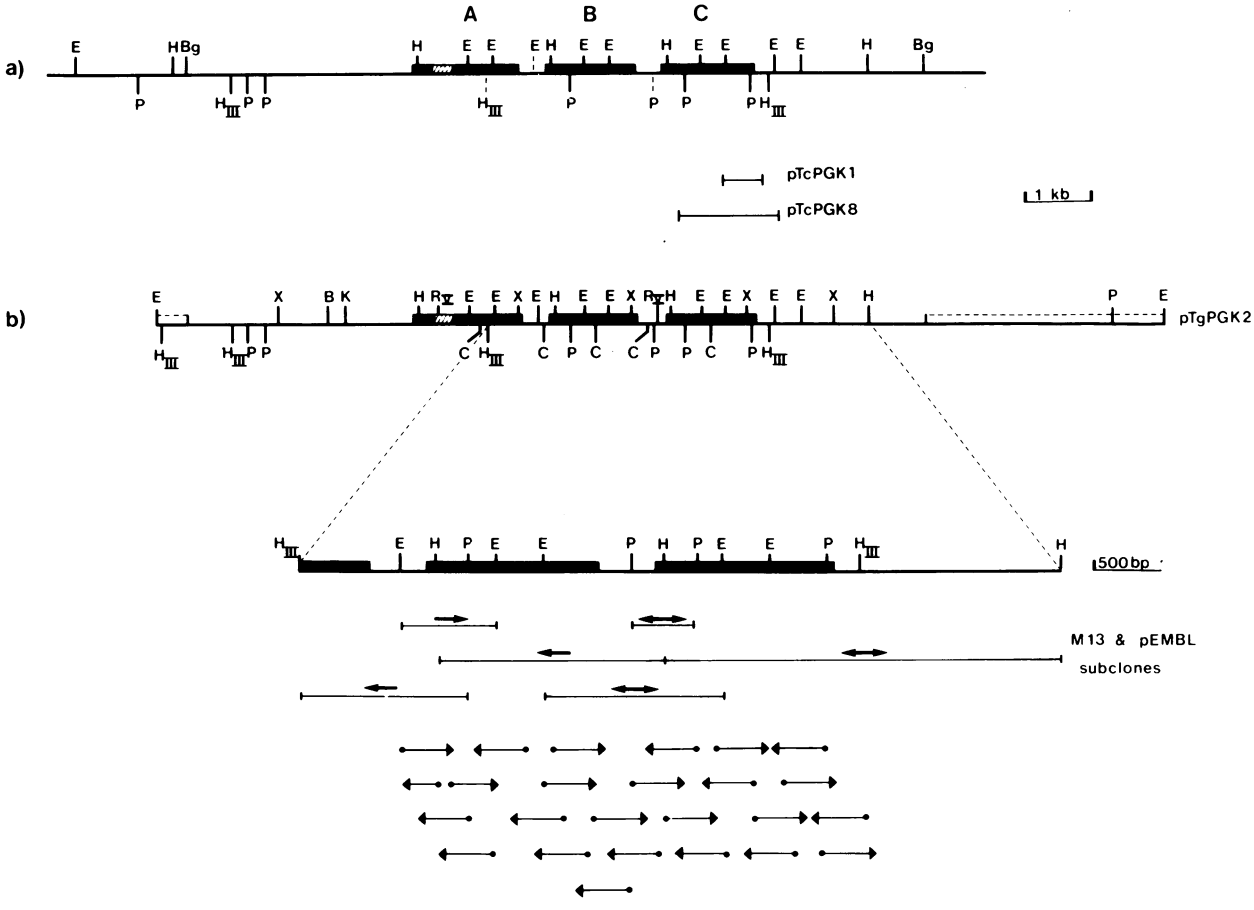

Fig. 1. (**a**) Genomic environment of the three tandemly linked PGK genes in *T. brucei* 427. The solid blocks indicate protein-coding sequences. The hatched box in gene A represents a 300-bp insertion which does not interrupt the reading frame of gene A. Broken lines indicate polymorphic restriction sites differing between the two allelic arrays in *T. brucei* 427. The PGK cDNA clones derived from bloodstream form RNA are shown below gene C. (**b**) Physical map of the genomic *Bgl*II fragment containing the tandem array of PGK genes inserted into pAT 153 (indicated by open box). Part of the map is enlarged to illustrate the sequencing strategy for genes B and C. Bars below map show the extent and orientation of M13 and pEMBL subclones used; vector arrows indicate sequence extended from each primer (dot). Restriction sites: B = *Bam*HI, Bg = *Blg*II, C = *Cla*I, E = *Eco*RI, H = *Hinc*II, HIII = *Hind*III, K = *Kpn*I, P = *Pst*I, RV = *Eco*RV, X = *Xho*I.

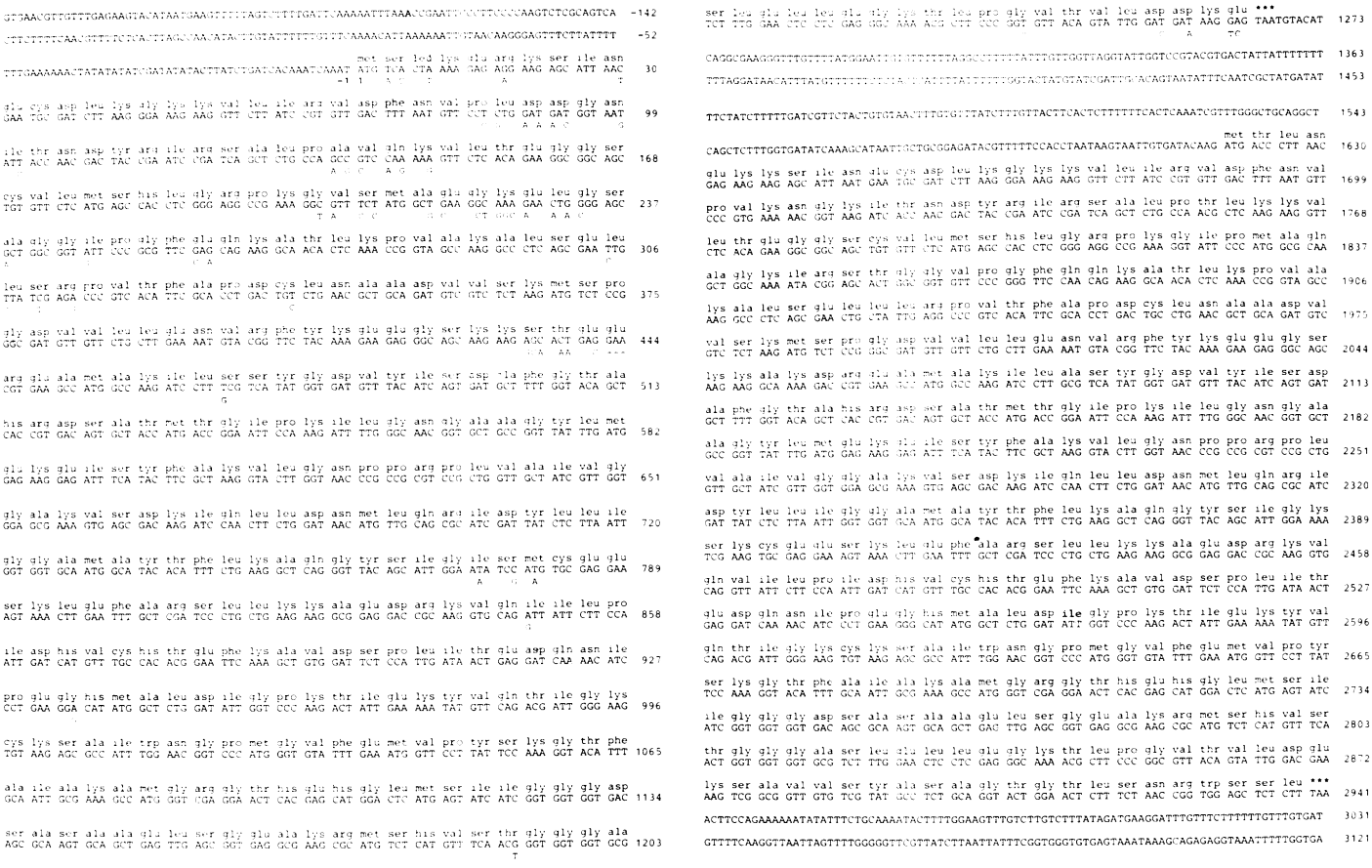

Fig. 2. Complete nucleotide and amino acid sequences of genes B and C. The coding region of gene B starts at position 1, and that of gene C at position 1618. Nucleotide substitutions in gene C relative to gene B are indicated below the gene B sequence.

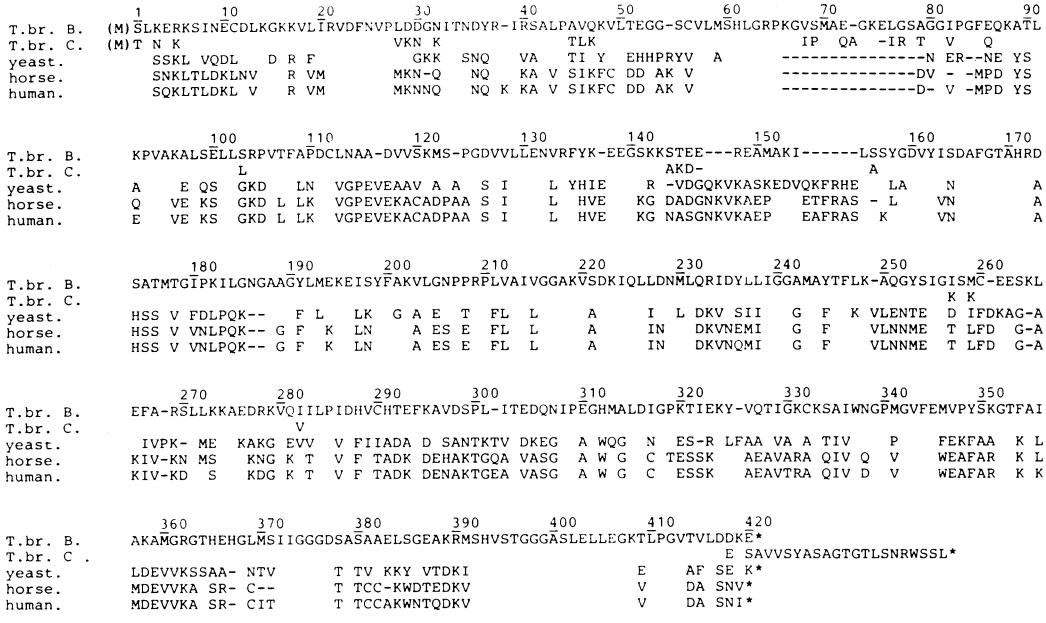

Fig. 3. Comparison of the amino acid sequences of trypanosome PGKs B (cytosolic) and C (glycosomal) and the PGKs of other organisms. The numbering refers to the trypanosome PGK B and only this sequence is shown in full. Deletions necessary to align the sequences are indicated by dashes. Amino acids and deletions are only shown in the other sequences where they differ from the trypanosome PGK B. Sequences are from: yeast (Watson *et al.*, 1982), horse (Banks *et al.*, 1979) and man (Michelson, 1983).

cing strategy is illustrated in Figure 1. Figure 2 shows the nucleotide and predicted amino acid sequences of genes B and C and also indicates the nucleotide differences between both genes; Figure 3 shows a comparison of the amino acid sequences of the proteins encoded by genes B and C and other eukaryotic PGKs (Watson *et al.*, 1982; Banks *et al.*, 1979; Michelson *et al.*, 1983). Gene A does not code for one of the major PGKs (see below) and its sequence will be presented in a later paper.

The sequence comparisons in Figures 2 and 3 show two main points. (i) There is a high degree of homology between the amino acid sequences deduced for trypanosome genes B and C and those of the other eukaryotic PGK genes. The sizes of the predicted proteins are similar, the amino acid homology is ~45% (see Table II) and, if conservative substitutions are included, the homology rises to 65%. No major insertions or deletions are required to line up the trypanosome and yeast genes, indicating that both trypanosome genes lack introns in the amino acid coding region. We conclude from these results that both genes B and C code for functional PGKs. (ii) The homology between genes B and C is very high, 95% at the nucleotide level and 93% at the amino acid level. Of the 65 substitutions in 1262 nucleotides, 38% were in the first position, 14% in the second position and 48% in the third position of the amino acid codon. The differences found in the amino acid sequence are of two types: gene C has

Table II. Percentage amino acid homology of PGKs calculated from Figure 3

	T. brucei B	*T. brucei* C	Human	Horse	Yeast
T. brucei B	–				
T. brucei C	93	–			
Human	44	44	–		
Horse	44	44	97	–	
Yeast	46	46	65	65	–

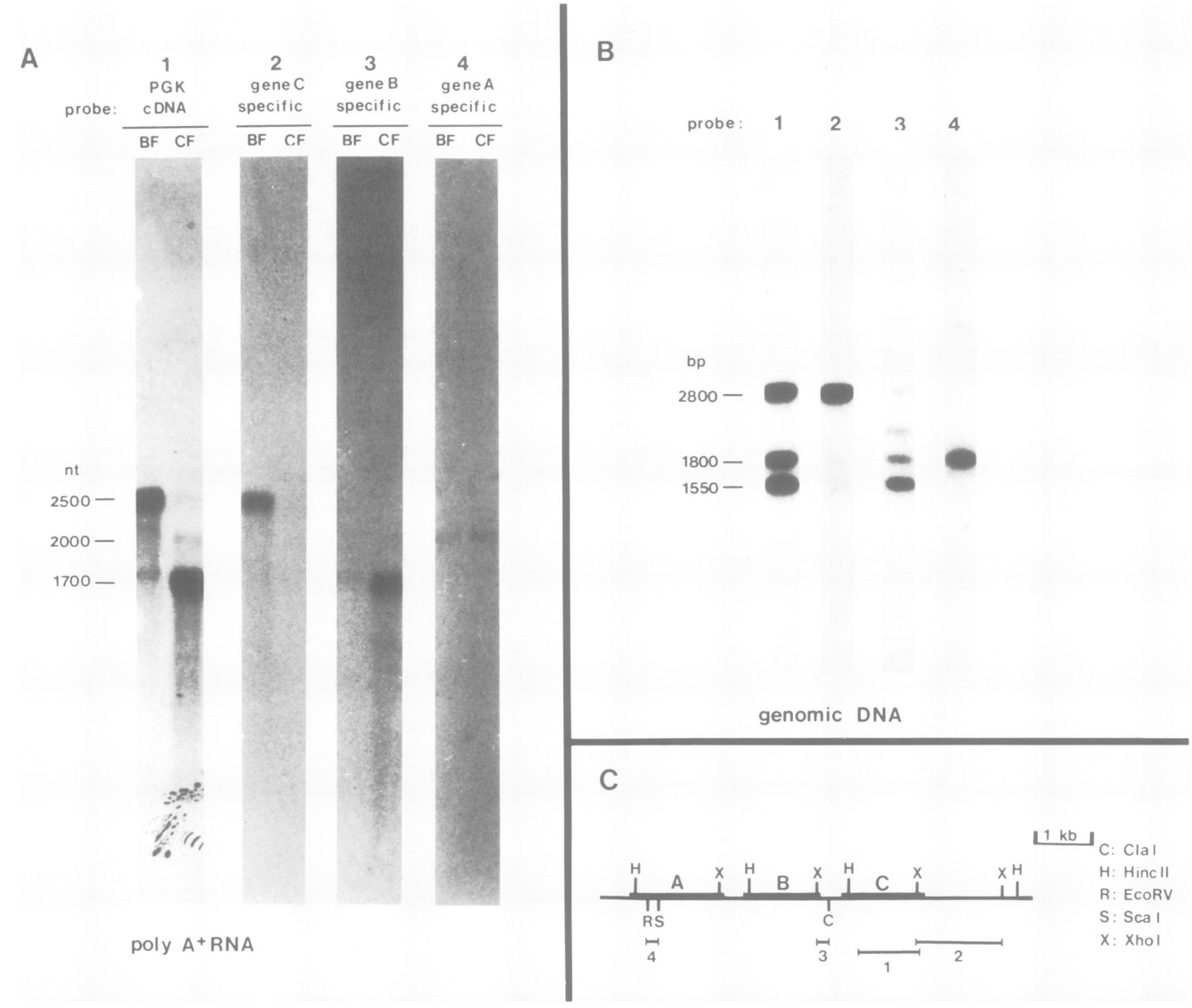

Fig. 4. (**A**) Poly(A)$^+$ RNA from bloodstream (BF) and culture (CF) form trypanosomes was size fractionated by agarose gel electrophoresis and transferred to a nitrocellulose filter. Identical strips of the filter were hybridized with ^{32}P-labelled PGK cDNA and gene-specific probes as indicated. Bands were visualized by autoradiography. (**B**) Nuclear DNA from *T. brucei* strain 427 was digested with the restriction enzyme *Hin*cII, size fractionated by agarose gel electrophoresis, transferred to nitrocellulose filters and identical strips were hybridized with the four probes used in **panel A**, to demonstrate their specificity. (**C**) Simplified map from Figure 1, indicating the location of each probe used.

Table III. PGK genes and transcripts in *T. brucei*

Gene	A	B	C
mRNA (nucleotides)	2000	1800	2500
Relative amount of mRNA in BF[a]	+	+	+++
Relative amount of mRNA in CF[b]	+	+++	+
Predicted mass of protein (daltons)	~55 000	44 914	46 890

[a]BF = bloodstream form trypanosomes.
[b]CF = cultured insect form trypanosomes.

a C-terminal extension of 20 amino acids not present in gene B or in any of the other eukaryotic PGKs; and in a total of 29 amino acid substitutions, gene C has gained 12 positive charges over gene B. We return to these striking differences below.

Which gene codes for the glycosomal PGK?

There are three PGK genes in *T. brucei*, but only two major PGK proteins, a cytosolic and a glycosomal one. To assign each protein to one of the three genes, we have made use of the differential production of the enzymes in the life cycle of *T. brucei*. During its life cycle, *T. brucei* shuttles back and forth between its insect vector and mammalian host. In the laboratory these life cycle forms are represented by *in vitro* cultured insect form trypanosomes (culture form; CF) and by rodent bloodstream form trypanosomes (BF). As indicated in Table I, glycosomal PGK is mainly produced in bloodstream trypanosomes which derive their energy only from glycolysis, while cytosolic PGK is the predominant enzyme in culture form trypanosomes. Figure 4A panel 1 shows an RNA blot with poly(A)$^+$ RNA from bloodstream and culture form trypanosomes hybridized with a PGK cDNA probe; the major PGK transcript in bloodstream trypanosomes is 2500 nucleotides in size, and that in culture form trypanosomes 1800 nucleotides. In addition, there is weak hybridization to a 2000-nucleotide transcript present in equal amounts in both types of RNA. Essentially similar results were obtained with total RNA instead of poly(A)$^+$ RNA (data not shown).

To match these RNAs with their corresponding genes, we hybridized identical lanes from the same blot with three gene-specific probes, the positions of which are indicated in Figure 4C. The gene A-specific probe corresponds to the 300-bp insert in gene A not present in genes B and C; the gene B-specific probe is located in the region between coding regions B and C; the gene C-specific probe is derived from the 3′-untranslated region of the C mRNA. The specificity of the probes is demonstrated in Figure 4B; their hybridization with poly(A)$^+$ RNA in Figure 4A. It is clear that the minor transcript of 2000 nucleotides comes from gene A; the 1800-nucleotide RNA from gene B; and the 2500-nucleotide RNA from gene C. We infer that gene B codes for cytosolic PGK, gene C for the glycosomal enzyme and gene A for an as yet undetected third enzyme species.

This conclusion tallies with the protein data: the glycosomal enzyme has a higher isoelectric point than the cytosolic enzyme (see Table I), in agreement with the higher net positive charge of enzyme C. It is also possible that the slightly lower mobility of the glycosomal enzyme on SDS gels (see Table I) is due to the C-terminal extension predicted for enzyme C. As we do not know whether the C-terminal extension predicted for protein C is present in the mature enzyme, and as mobility in SDS gels is not a perfect measure of mass, the correlation between the predicted size differences of proteins B and C and the observed mobility differences between cytosolic and glycosomal PGK may be coincidental.

We originally thought that gene A might be a pseudogene, but sequence analysis showed a complete open reading frame with an insertion of 100 amino acids relative to gene B (see Figure 1), and a correspondingly larger message as shown in Figure 4A. The predicted size of the protein encoded by gene A is 55 kd, larger than both cytosolic and glycosomal PGKs (see Table III); the mRNA level does not vary during the life cycle of *T. brucei*. Detailed results of gene A will be presented in a later paper when we have sorted out whether the predicted 55-kd protein exists and what its function is.

In conclusion, our data indicate that gene B codes for cytosolic PGK and gene C for glycosomal PGK. We note that the relative activities of these enzymes in bloodstream and culture trypanosomes, given in Table I, corresponds to the mRNA levels for each enzyme (see Tables I and III), indicating that the concentration of these enzymes is mainly determined by mRNA levels.

Discussion

We have identified and sequenced the genes for glycosomal and cytosolic PGKs in *T. brucei*. The most striking result is the high homology between these genes. This is not merely due to stringent evolutionary constraint, because conservation of eukaryotic PGKs is far from complete, as can be seen in Table II. Homology between yeast and *T. brucei* PGK is only 46%.

It seems unlikely that the high homology between genes B and C only reflects their recent evolution from a common ancestor. Glycosomes containing glycolytic enzymes are found in all kinetoplastid flagellates analysed (see Introduction); hence, the duplication of PGK genes probably occurred before the different kinetoplastid genera evolved from a common ancestor. Branching cannot have been a very recent event, because the genes for three conserved mitochondrial proteins, encoded in mitochondrial DNA, are only 75–83% homologous in *T. brucei* and *Leishmania tarentolae*, another kinetoplastid flagellate (see Benne, 1985). This contrasts with the 95% homology found for the two PGK genes of *T. brucei*. The high homology is therefore most simply explained by gene conversions, partially erasing the differences that have arisen since the duplication of the ancestral gene. Gene conversions could also be responsible for the very unequal distribution of the nucleotide differences over the length of the genes, with stretches of 301 bp (positions 471–772) and 252 bp (positions 938–1190) without any change, and 82% of all differences concentrated in the first 35% of the genes (see Figure 2).

Our results establish that few modifications are required to convert a cytosolic enzyme into a microbody enzyme. Only two types of difference are found between cytosolic and glycosomal PGK. (i) The glycosomal PGK has a C-terminal extension of 20 amino acids that is lacking in the cytosolic enzyme and in the yeast, horse and human PGKs. This 20-amino acid extension has a moderately hydrophobic character (eight amino acids are non-polar and only one amino acid is charged) and is rich in hydroxyl amino acids, having five serines, two threonines and one tyrosine. (ii) The glycosomal PGK has a strikingly higher net positive charge than the cytosolic isoenzyme and the yeast, horse and human PGKs.

Either of these differences or a combination of both must determine the difference in the routing of the cytosolic and the glycosomal PGK. Essential signal sequences for uptake by glycosomes should also be present in other glycosomal proteins. The amino acid sequences of triosephosphate-isomerase (TIM), glyceraldehyde phosphate-dehydrogenase (GAPDH; our unpublished results) and aldolase (Clayton, 1985) have recently been deduc-

ed from the gene sequences. TIM and aldolase do not have cytosolic isoenzymes in *T. brucei* and the gene for the cytosolic GAPDH has not yet been sequenced. None of the three deduced glycosomal proteins has a significant C-terminal extension. We also do not see an analogous sequence internally in positions where the *T. brucei* enzymes differ from their counterparts in other eukaryotes. However, trypanosome TIM does have a high ratio of basic to acidic residues and has the highest net positive charge seen thus far for any TIM sequenced. The same holds for the glycosomal GAPDH and aldolase of *T. brucei*. Moreover, native trypanosome TIM has an isoelectric point >pH 10 and the isoelectric points of the glycosomal and cytosolic GAPDH have been found to be >pH 9 and neutral, respectively (O.Misset, personal communication). This shows that a high positive charge is the only obvious common feature of the glycosomal enzymes analysed thus far. It is therefore likely that this feature is required for the uptake of the enzymes into the glycosome. The alternative is that a positive charge is required for function in the organelle and that we have missed the real topogenic signals. We consider this unlikely. To be informative, topogenic signals must have a minimal length, as actually found for the topogenic signals that specify protein delivery to organelles other than microbodies (see Borst, 1986). It is also unlikely that scattered secondary modifications of the protein chain, like glycosylation, might together represent the topogenic signal and that such scattered modifications might be overlooked in comparing protein sequences. Such modifications have not been found as part of the topogenic signal for proteins made on free ribosomes, like microbody proteins. Moreover, no such modifications are detected in microbody enzymes in eukaryotes other than trypanosomes (Lazarow and Fujiki, 1985).

Some of the extra positive charges on glycosomal PGK are clustered. This is accentuated if one inspects the three-dimensional structure of PGK, and a similar situation is found in the glycosomal TIM and GAPDH molecules (W.G.J.Hol, personal communication). Clustering might be caused by the way in which the positive charges were introduced into the glycosomal enzymes in evolution. If this occurred by segmental gene conversion, clustering would be the consequence (see Borst, 1986). There may also be only a few locations in glycolytic enzymes where positive charges can be freely inserted. Finally, a localized cluster of basic amino acids might function as the signal for microbody uptake. The other extra basic residues might have an adjuvant role in uptake or be required for function of the enzyme in the organelle. Construction of chimaeric PGK genes and analysis of the uptake of the product of such genes by glycosomes will be required to choose between these alternatives and to define precisely the topogenic signals for delivery to the glycosome.

There are two other examples in the literature of microbody enzymes that resemble isoenzymes present in other cellular compartments: the malate dehydrogenase of cucumber glyoxysomes and mitochondria and the carnitine acetyl transferase of fungal peroxisomes and mitochondria. The glyoxysomal malate dehydrogenase from cucumber has a much higher isoelectric point than the mitochondrial enzyme (Hock, 1984), but the peroxisomal and mitochondrial carnitine acetyl transferases of *Candida* hardly differ in isoelectric point (Ueda *et al.*, 1984). The peroxisomal enzyme enoyl-CoA:hydratase-3-hydroxyacyl-CoA dehydrogenase of rat liver (Osumi *et al.*, 1985) has a high net positive charge. There is no obvious preponderance of basic amino acids, however, in two enzymes from *Hansenula*, methanol oxidase and dihydroxy acetone synthase (Ledeboer *et al.*, 1985; Janowicz *et al.*, 1985). Whether a high net positive charge is a characteristic of microbody enzymes or precursors in eukaryotes other than trypanosomes is therefore far from clear.

Materials and methods

Isolation of trypanosome PGK genes

A *T. brucei* strain 427 cDNA library (De Lange *et al.*, 1984) was screened with a yeast PGK probe, isolated from plasmid pB1 [this plasmid was kindly provided by Dr. A.J.Kingsman, Department of Biochemistry, Oxford University, UK (unpublished results)]. Colony hybridization of filters were performed using standard procedures (Maniatis *et al.*, 1982), at relaxed stringency [5 × SSC (SSC = 0.15 M NaCl; 0.015 M sodium citrate; pH 7.0) at 58°C]. Post-hybridizational washes were carried out for 3 h at 55°C with several changes of 5 × SSC, 0.1% SDS. Under these conditions, yeast PGK hybridizes specifically with trypanosome DNA (not shown).

Genomic *T. brucei* 427 PGK clones were isolated from a *Bgl*II clone bank (Kooter and Borst, 1984), using the isolated insert of pTcPGK8 (see Figure 1) as probe. Two types of genomic clones were obtained, representing two allelic clusters of PGK genes which differed in some restriction sites (see Figure 1) (Gibson *et al.*, 1985).

DNA sequence analysis

DNA sequence analysis was performed using the dideoxy method (Sanger *et al.*, 1977) with the modifications described by Biggin *et al.* (1983) except that ^{32}P was used as label. Genomic fragments from pTgPGK2 (see Figure 1) were subcloned in M13 vectors (Yanish-Perron *et al.*, 1985). The *Hind*III-*Pst*I fragment spanning the 3′ end of gene A and a 5′ part of gene B (see Figure 1), which repeatedly underwent deletions in M13 vectors, was cloned in the pEMBL 9 vector (Dente *et al.*, 1983). Cloning and preparation of template DNA were performed using standard protocols (Maniatis *et al.*, 1982; Messing, 1983).

Oligonucleotides, used for sequencing, were prepared by a phosphotriester approach (Marugg *et al.*, 1984), using a fully automatic synthesizer (Biosearch, SAM1).

Isolation and blotting analysis of DNA and RNA

Total RNA was prepared from *T. brucei* 427 cultured insect or bloodstream form trypanosomes by LiCl precipitation (Auffray and Rougeon, 1980), and the poly(A)$^+$ fractions were purified by oligo(dT)-cellulose chromatography (Hoeijmakers *et al.*, 1980). RNA blots were prepared using glyoxylated RNA according to the procedure of Thomas (1980). Nuclear DNA from *T. brucei* 427 was isolated, digested with *Hin*cII, size-fractionated and blotted as described previously (Gibson *et al.*, 1985). Identical strips from the RNA and DNA blots were hybridized with the nick-translated ^{32}P-labelled (Rigby *et al.*, 1977) gene-specific probes indicated in Figure 4, essentially as described by Jeffreys and Flavell (1977), with the addition of 10% (w/v) Dextran sulphate (Wahl *et al.*, 1979). Post-hybridizational washes were to 0.1 × SSC at 65°C.

Acknowledgements

We thank Dr. A.J.Kingsman, Oxford University, UK for providing us with a yeast PGK clone. We are grateful to our colleagues for their helpful advice, to Dr. W.G.J.Hol (University of Groningen) for information on the three-dimensional structure of glycolytic enzymes, to Dr. O.Misset (International Institute of Cellular and Molecular Pathology, Brussels) for sharing unpublished results and to Helga Woudt for preparing the manuscript. The work in Amsterdam was supported by The Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO) and a fellowship of the Wellcome Trust to W.C.G. The work in Brussels was supported by the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases and the EEC Science and Technology for Development Programme, contract TSD-M-021.

References

Auffray,C. and Rougeon,F. (1980) *Eur. J. Biochem.*, **107**, 303-314.
Banks,R.D., Blake,C.C.F., Evans,P.R., Haser,R., Rice,D.W., Hardly,G.W., Merret,M. and Phillips,A.W. (1979) *Nature*, **279**, 773-777.
Benne,R. (1985) *Trends Genet.*, **1**, 117-121.
Biggin,M.D., Gibson,T.J. and Hong,G.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3963-3965.
Borst,P. (1986) *Biochim. Biophys. Acta*, in press.
Cannata,J.J.B., Valle,E., Docampo,R. and Cazzulo,J.J. (1982) *Mol. Biochem. Parasitol.*, **6**, 151-160.
Clayton,C.E. (1985) *EMBO J.*, **4**, 2997-3003.
Coombs,G.H., Craft,J.A. and Hart,D.T. (1982) *Mol. Biochem. Parasitol.*, **5**, 199-211.
De Duve,C. (1982) *Ann. N.Y. Acad. Sci.*, **386**, 1-4.

De Lange,T., Michels,P.A.M., Veerman,H.J.G., Cornelissen,A.W.C.A. and Borst,P. (1984) *Nucleic Acids Res.*, **12**, 3777-3790.
Dente,L., Cesarini,G. and Cortese,R. (1983) *Nucleic Acids Res.*, **11**, 1645-1655.
Gibson,W.C., Osinga,K.A., Michels,P.A.M. and Borst,P. (1985) *Mol. Biochem. Parasitol.*, **16**, 231-242.
Hart,D.T. and Opperdoes,F.R. (1984) *Mol. Biochem. Parasitol.*, **13**, 159-172.
Hock,B. (1984) *Physiol. Veg.*, **22**, 333-339.
Hoeijmakers,J.H.J., Borst,P., Van den Burg,J., Weissmann,C. and Cross,G.A.M. (1980) *Gene*, **8**, 391-417.
Janowicz,Z., Eckart,M.R., Drewke,C., Roggenkamp,R.O., Hollenberg,C.P., Maat,J., Ledeboer,A.M., Visser,C. and Verrips,C.T. (1985) *Nucleic Acids Res.*, **13**, 3043-3062.
Jeffreys,A.J. and Flavell,R.A. (1977) *Cell*, **12**, 429-439.
Kooter,J.M. and Borst,P. (1984) *Nucleic Acids. Res.*, **12**, 9457-9472.
Kindl,H. (1982) *Int. Rev. Cytol.*, **12**, 247-307.
Kindl,H. and Lazarow,P.B. (eds.) (1982) *Ann. N.Y. Acad. Sci.*, **386**, 1-550.
Lazarow,P.B. and Fujiki,Y. (1985) *Annu. Rev. Cell. Biol.*, **1**, 489-530.
Ledeboer,A.M., Edens,L., Maat,J., Visser,C., Bos,J.W., Verrips,C.T., Janowicz,Z., Eckart,M., Roggenkamp,R.O. and Hollenberg,C.P. (1985) *Nucleic Acids Res.*, **13**, 3063-3082.
Lord,J.M. and Roberts,L.M. (1983) *Int. Rev. Cytol.* (Suppl.), **15**, 115-156.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.
Marugg,J.E., Piel,N., McLaughlin,L.W., Tromp,M., Veeneman,G.H., Van der Marel,G.A. and Van Boom,J.H. (1984) *Nucleic Acids Res.*, **12**, 8639-8651.
Messing,J. (1983) *Methods Enzymol.*, **101**, 20-79.
Michelson,A.M., Markham,A.F. and Orkin,S.H. (1983) *Proc.Natl. Acad. Sci. USA*, **80**, 472-476.
Muse,K.E. and Roberts,J.F. (1973) *Protoplasma*, **78**, 343-348.
Opperdoes,F.R. (1984) *FEBS Lett.*, **169**, 35-39.
Opperdoes,F.R. (1985) *Br. Med. Bull.*, **41**, 130-136.
Opperdoes,F.R. and Borst,P. (1977) *FEBS Lett.*, **80**, 360-364.
Opperdoes,F.R., Markos,A. and Steiger,R.F. (1981) *Mol. Biochem. Parasitol.*, **4**, 291-309.
Opperdoes,F.R., Borst,P., Bakker,S. and Leene,W. (1977) *Eur. J. Biochem.*, **76**, 29-39.
Opperdoes,F.R., Baudhuin,P., Coopens,I., De Roe,C., Edwards,S.W., Weyers, P.J. and Misset,O. (1984) *J. Cell. Biol.*, **98**, 1178-1184.
Osumi,T., Ishii,N., Hijikata,M., Kamijo,K., Hisashi,O., Furuta,S., Miyazawa,S., Kondo,K., Inoue,K., Kagamiyama,H. and Hashimoto,T. (1985) *J. Biol. Chem.*, **260**, 8905-8910.
Rigby,P.W.J., Dieckmann,M., Rhodes,C. and Berg,P. (1977) *J. Mol. Biol.*, **113**, 237-251.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.
Taylor,M.B., Berghausen,H., Heyworth,P., Messenger,N., Rees,L.J. and Gutteridge,W. (1980) *Int. J. Biochem.*, **11**, 117-120.
Thomas,P.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201-5205.
Tolbert,N.E. (1981) *Annu. Rev. Biochem.*, **50**, 133-157.
Trelease,R.N. (1984) *Annu. Rev. Plant Physiol.*, **35**, 321-347.
Ueda,M., Tanaka,A. and Fukui,S. (1984) *Eur. J. Biochem.*, **138**, 445-449.
Veenhuis,M., Van Dijken,J.P. and Harder,W. (1983) *Adv. Microb. Physiol.*, **24**, 1-82.
Wahl,G.M., Stern,M. and Stark,P.G.R. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 3683-3687.
Watson,H.L., Walker,N.P.C., Shaw,P.J., Bryant,T.N., Wendell,P.L., Fothergill,L.A., Perkins,R.E., Conroy,S.C., Dobson,M.J., Tuite,M.F., Kingsman,A.J. and Kingsman,S.M. (1982) *EMBO J.*, **1**, 1635-1640.
Yanish-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103-119.

Received on 9 October 1985; revised on 28 October 1985