Sequence differences between human muscle and liver cDNAs for UDPglucose pyrophosphorylase and kinetic properties of the recombinant enzymes expressed in *Escherichia coli*

Ronald G. DUGGLEBY¹, Yu Chyi CHAO², Joyce G. HUANG², Hwei-Ling PENG³ and Hwan-You CHANG²

¹ Centre for Protein Structure, Function and Engineering, Department of Biochemistry, University of Queensland, Brisbane, Australia

² Department of Molecular and Cellular Biology, Chang-Gung Medical College, Taiwan, Republic of China

³ Department of Microbiology and Immunology, Chang-Gung Medical College, Taiwan, Republic of China

(Received 25 July/5 October 1995) - EJB 95 1234/4

UDP-Glc pyrophosphorylase (EC 2.7.7.9) catalyses the interconversion of MgUTP plus Glc1*P* and UDP-Glc plus MgPP_i. Complementation of an *Escherichia coli* strain lacking this activity has allowed isolation of cDNA encoding this enzyme from a human muscle library. Two forms were identified and the nucleotide sequence of each was determined; they were found to differ only in the 5' region and we suggest that these arise from the use of a different first exon in the two transcripts. These nucleotide sequences are different from that of the cDNA which was isolated previously from a human liver library [Peng, H.-L. & Chang, H.-Y. (1993) *FEBS Lett.* 329, 153–158] and it is proposed that these liver and muscle forms are derived from different genes. The cDNA for muscle form I, muscle form II, the liver form, and the liver form fused to part of the *lacZ* gene were expressed in *Escherichia coli* and the kinetic properties of each enzyme were characterised. Muscle form I and the LacZ/liver fusion enzyme exhibit Michaelis-Menten kinetics towards all substrates while muscle form II has a sigmoidal dependence of rate upon the concentration of MgPP_i. The liver form shows Michaelis-Menten kinetics towards MgUTP. For the remaining three substrates, complex kinetics were observed involving a combination of sigmoidicity at low substrate concentration and partial inhibition at high substrate concentration.

Keywords: cDNA sequence; enzyme kinetics; *galU*; human isoenzymes; UDPglucose pyrophosphory-lase.

UDP-Glc is an important intermediate in mammalian carbohydrate interconversions. In muscle, its primary fate is conversion to glycogen while in lactating mammary gland it is converted to UDPgalactose, a precursor of lactose. In liver, glycogen is also a major product of UDP-Glc metabolism but smaller amounts are converted to UDPglucuronate, which then reacts with xenobiotic and endobiotic compounds to aid in their solubilisation and excretion.

UDP-Glc is formed in the reaction catalysed by UDP-Glc pyrophosphorylase (UDP-Glc PPase):

 $MgUTP + Glc1P \leftrightarrow UDP-Glc + MgPP_i$.

The enzyme has been studied from several mammalian tissues including human liver (Knop and Hansen, 1970) and erythro-

Correspondence to R. G. Duggleby, Centre for Protein Structure, Function and Engineering, Department of Biochemistry, University of Queensland, Brisbane, Australia 4072

Fax: +617 3365 4699.

Abbreviations. UDP-Glc PPase, UDPglucose pyrophosphorylase; Glc1P, glucose 1-phosphate; Glc6P, glucose 6-phosphate; Glc(1,6) P_2 , glucose 1,6-bisphosphate.

Enzymes. UDPglucose pyrophosphorylase, UTP: α -D-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9); phosphoglucomutase, α -D-glucose 1,6-bisphosphate: α -D-glucose-1-phosphate phosphotransferase (EC 5.4.2.2; formerly 2.7.5.1); glucose-6-phosphate dehydrogenase, D-glucose-6-phosphate:NADP⁺ oxidoreductase (EC 1.1.1.49).

Note. The nucleotide sequence and deduced amino acid sequence data published here have been submitted to the GenBank database and are available under the accession number U27460.

cytes (Tsuboi et al., 1969) although the bulk of the published literature (see review of Turnquist and Hansen, 1973) concerns the bovine liver enzyme.

In view of its diverse metabolic roles, tissue-specific isoforms of UDP-Glc PPase would not be surprising. Aksamit and Ebner (1972) reported that multiple chromatographic forms of the enzyme are present in bovine mammary tissue although it was not determined whether these arise from variable oligomerisation or post-translational modification rather than from different genes. Shows et al. (1978) demonstrated two electrophoretic forms in human cells and showed that these are encoded by genes on chromosomes 1 and 2. However, the molecular differences between these two genes, and the properties of the resulting enzymes, have not been investigated. Here we describe two distinct forms isolated as clones from a human muscle cDNA library. These clones, and a liver form that has been isolated previously (Peng and Chang, 1993), result in enzymes with somewhat different kinetic properties when expressed in Escherichia coli.

MATERIALS AND METHODS

cDNA library screening. A human skeletal muscle cDNA library, constructed in the *Eco*RI site of the λ ZAP vector, was purchased from Stratagene (La Jolla, CA). The λ library was converted into the phagemid form by *in vivo* excision as recommended by the supplier. The *E. coli* mutant strain HP069 (MC4100, *galU F1ac*) was used for the selection of clones with

the Gal⁺ phenotype. MacConkey-galactose agar was prepared with MacConkey agar base (Difco, Detroit MI) supplemented with 0.2% galactose and 50 µg/ml ampicillin.

Recombinant DNA techniques and DNA sequencing. All restriction enzymes were obtained from Boehringer Mannheim. The cDNA was subcloned into M13 vectors and the sequence determined with the Sequenase kit (Amersham) with either universal M13 primer or synthetic oligonucleotides. The sequence was confirmed in both strands, and the question of compressions was resolved by sequencing with dITP. Nucleotide sequences were analysed with the DNASTAR program using a Macintosh LCII computer. The expression vector pTrc99a was obtained from Pharmacia P-L Biochemicals.

Overexpression of the human liver form of UDP-Glc PPase in E. coli. The human cDNA clone (pHC308) encoding the liver form of UDP-Glc PPase that was reported previously (Peng and Chang, 1993) did not yield a high level of expression in E. coli, presumably due to the presence of a long 5' noncoding region. To improve expression, the entire coding sequence was synthesised by PCR using the high-fidelity, thermostable enzyme Pfu DNA polymerase (Stratagene). Two synthetic oligonucleotides, 5'-ATATAACCCGACCATGGCGAGATTT-3' and 5'-CTTATCTTAGGGATCCAAGAT-3' were used as the upstream and downstream primer, respectively. An NcoI site and a BamHI site (underlined) were incorporated in these primers, which facilitated subcloning of the amplified DNA fragment. The reaction mixture contained 0.1 µg pHC308, 0.1 nmol of each primer, 250 μ M of the four dNTPs, 1.5 mM MgCl₂ and 5 units Pfu DNA polymerase. After 32 cycles of 94°C (1 min), 55°C (1 min) and 72°C (2 min) the amplified DNA fragment was isolated from SeaPlaque low-melting temperature agarose gel (FMC Corp., Rockland ME). This was then digested with NcoI and BamHI, and ligated into the corresponding restriction endonuclease sites of the expression vector pTrc99a. E. coli JM109 harbouring the resulting plasmid (pHC404) was used as the source of the human liver form of UDP-Glc PPase.

Enzyme preparation. The enzyme was partially purified from E. coli clones containing plasmid-encoded UDP-Glc PPase. In a typical preparation, cells were harvested from 2-1 cultures in Luria-Bertani medium after overnight induction with 0.2 mM isopropyl β -D-thiogalactopyranoside at 37 °C, and suspended in 60 ml ice-cold buffer A (50 mM Tris/HCl pH 7.8 containing 10 mM 2-mercaptoethanol and 10 mM MgCl₂). Buffer A was used throughout the purification and the temperature was maintained at 0-4°C. Cells were disrupted by sonication, the crude extract clarified by centrifugation, and nucleic acids removed by slow addition of 0.083 vol. 20% (mass/vol.) streptomycin sulfate. After centrifugation, the enzyme was precipitated from the supernatant by slow addition of 326 g/l solid ammonium sulfate, dissolved in 6 ml buffer A and dialysed twice against 1 l buffer A. Half of this solution was diluted with an equal volume of water and applied to a column $(2.1 \times 14.5 \text{ cm})$ of DEAE-cellulose (Whatman DE52) that had been equilibrated with buffer A. After washing with 10 ml half-strength buffer A and 90 ml full-strength buffer A, the column was developed with a 180-ml linear gradient of 0-0.25 M NaCl in buffer A at a flow rate of 3 ml/min; the enzyme eluted at approximately 0.1 M NaCl. The most active fractions were pooled, precipitated by slow addition of 436 g/l solid ammonium sulfate, and dissolved in 3 ml buffer A. When the enzyme was not to be used within a few days, an equal volume of glycerol was added and the solution stored at -20 °C.

Enzyme assays. Tetrasodium pyrophosphate was purchased from Nacalai Tesque Inc. (Kyoto, Japan) and triethanolamine was from Ajax Chemicals (Auburn, Australia). Other substrates, coenzymes, coupling enzymes and buffers were obtained from

Sigma. The kinetic properties were assessed for each of the substrates using assays based on those described by Tsuboi et al. (1969). Assays were conducted at 25°C in 0.1 M triethanolamine/HCl pH 8.0 (unless otherwise stated), in the presence of 1 mM dithiothreitol. In what is conventionally designated the forward direction (using Glc1P and MgUTP as substrates), the reaction was monitored at 340 nm by addition of NAD+ (2 mM) and UDP-Glc dehydrogenase (0.25 U/ml). In the reverse reaction (UDP-Glc and MgPP_i as substrates), the reaction was monitored at 340 nm by addition of phosphoglucomutase (5 U/ml) and its cofactor Glc(1,6)P2 (2 µM), Glc6P dehydrogenase (5 U/ ml) and NADP⁺ (1 mM). For assays in both directions, MgCl₂ was added to a total concentration calculated so as to give a free magnesium ion concentration of 1 mM, using the stability constants cited by Aksamit and Ebner (1972) for the Mg complexes of the reactants.

Data analysis. Kinetic data were analysed using the nonlinear regression computer program DNRPEASY that is an adaptation of the DNRP53 program described elsewhere (Duggleby, 1984). The following equations were fitted to the data; where the results from several experiments were pooled, the combined analysis was performed as described by Duggleby (1990). Michaelis-Menten kinetics:

$$v = V_{\rm m}[S]/(K_{\rm m} + [S]).$$
 (1)

Two-substrate kinetics:

$$v = V_{\rm m}[A][B]/([A][B] + K_{\rm a}[B] + K_{\rm b}[A] + K_{\rm ia}K_{\rm b}).$$
(2)

Competitive inhibition:

$$v = V_{\rm m}[S]/(K_{\rm m}(1 + [I]/K_{\rm i}) + [S]).$$
 (3)

Hill equation:

$$v = V_{\rm m} / \{1 + (S_{0.5} / [S])^h\}.$$
(4)

Modified Monod-Wyman-Changeux model:

$$v = V_{\rm m}([S]/K_{\rm s})(r + [S]/K_{\rm s})/\{L + (1 + [S]/K_{\rm s})^2\}.$$
 (5)

RESULTS

Cloning of human cDNA encoding skeletal muscle UDP-Glc PPase. We have shown previously that an *E. coli* UDP-Glc-PPase-deficient mutant can be used as a tool for the isolation by complementation of cDNA clones encoding the enzyme (Peng and Chang, 1993). Taking advantage of this simple selection method, we have extended our investigation to explore the presence of different isoforms of the enzyme in skeletal muscle. Five phagemid clones that are capable of transforming *E. coli* HP069 into the GalU⁺ phenotype were isolated from a human skeletal muscle cDNA library. The UDP-Glc PPase activity in cell lysates of these recombinant clones was determined and ranged over 2-7 U/mg protein, compared with 0.02 U/mg in *E. coli* JM109 (GalU⁺) and no detectable activity in the GalU⁻ host *E. coli* HP069. These results strongly suggest that these phagemid clones encode a UDP-Glc PPase.

Sequence analysis. DNA was prepared from these clones and their restriction enzyme sites were mapped. Two distinct groups were noted on the basis of the presence or absence of an Asp718 site. The nucleotide sequence of a representative phagemid of each of these groups, designated pHC377 (muscle form I) and pHC379 (muscle form II), was determined. Fig. 1A shows the sequence of pHC377, which comprises 1823 nucleotides. Fig. 1B compares the first 42 nucleotides of pHC379 with the corresponding region (nucleotides 70–111) of pHC377; beyond nucleotide 92, pHC379 has an identical sequence.

Α

1				~~~~~						-			CC	CCCC	ATT1	CCC2	AGCT	SCTA	AGGA	AGA
30	GGA	AGGTZ	ACCTO	Freco	STGCA	ACGC/	AGACO	GGA	feee	TGGC	GAA	90000	AGGA	ACTGA	1GAAA	ACCA	4GATO		JUAA!	AGCA
100	M	D D D D D D D D D D D D D D D D D D D	Q CNN	U CD	G	A	5 The the	CNC.	ድ ጥጥሮ	2	CDA	CTTC	፲ ስጥጥ	CCC	CAA	GAG	СТА	CDD	ፈ ፈጥጥ	о тст
21	AIG V	ICI V	UMA V	GAI	T. T.	F	r r	T	T	TT CAA	UAA T	a la	<u>د</u>	c000	слл н	E	T T	E	н	τCI Ψ
169	CTTC	ADC:	AAC	<u>د م</u>	ርጥል	222	מממ	בידים	CTC	200	272	GCA	a DTT	тса	TA'	GAA	ጥጥጥ	GAG	CAC	ACC
41	ĸ	K	D	T.	D	G	F	R	ĸ	T,	F	н	R	F	L	0	E	K	G	P
229	AAA	AAA	GAC	CTG	GAT	GGA	TTT	CGG	AAG	CTA	TTT	CAT	AGA	TTT	TTG	CÂA	GAA	AAG	GGG	CCT
61	S	v	D	W	G	K	I	0	R	P	P	Е	D	S	I	Q	Р	Y	Е	к
289	TCT	GTG	GAT	TGG	GGA	AAA	ATC	CAG	AGA	CCC	CCT	GAA	GAT	TCG	ATT	CAA	CCC	TAT	GAA	AAG
81	I	к	А	R	G	\mathbf{L}	Ρ	D	Ν	I	S	S	v	\mathbf{L}	N	Κ	L	v	V	v
349	ATA	AAG	GCC	AGG	GGC	CTG	CCT	GAT	AAT	ATA	TCT	TCC	GTG	TTG	AAC	AAA	CTA	GTG	GTG	GTG
101	к	г	N	G	G	L	G	т	S	М	G	С	К	G	Ρ	K	S	\mathbf{L}	Ι	G
409	AAA	CTC	AAT	GGT	GGT	TTG	GGA	ACC	AGC	ATG	GGÇ	TGC	AAA	GGC	CCT	AAA	AGT	CTG	ATT	GGT
121	v	R	N	Е	N	т	F	\mathbf{r}	Þ	Ľ	т	v	Q	Q	I	E	Н	\mathbf{L}	N	K
469	GTG	AGG	AAT	GAG	AAT	ACC	TTT	CTG	GAT	CTG	ACT	GTT	CAG	CAA	ATT	GAA	CAT	TTG	AAC	AAA
141	т	Y	N	т	D	v	P	L	v	L	М	N	S	F	N	Т	D	Е	D	Т
529	ACC	TAC	AAT	ACA	GAT	GTC	CCT	CTT	GTT	ATT	ATG	AAC	TCT	TTT	AAC	ACG	GA'T	GAA	GAT	ACC
161	K	K	1	L	Q	K	Y	N	Н	C	R	V	ĸ	1 Dma	Y mag	T	E'	N	Q	S
589	AAA	AAA	ATA	CTA	CAG	AAG	TAC	AAT	CAT	TGT	CGT	GTG	AAA	ATC	TAC	ACT	TTC	AAT	CAA	AGC
101	R RCC	I D D C	r	K NCC	1 1 1 1 1 1	אן. דע ג	707	CDD	5 	ատոր	R CCC	r ccm	v د ۲۰۰۰ م	A CCD	770	CNC	CTTC	0 00 00	1	5 17 C A
201	AGG	TAC	N	AGG	F	MAI	MAA W	GAA	u u	TIA	CGG	ц ц	GIA	n	T	V	GIG N	101	F	v
709	222	GDD	חמ	202	GAA	GCT	TGG	TAT	CCT	CCA	GGT	CAT	GGT	GAT	A ጥጥ	TAC	GCC	AGT	TTC	TAC
221	N	S	G	T.	T.	D	Ϋ́.	F	T	G	E	G	ĸ	F.	Ŷ	Т	F	v	s	N
769	AAC	TCT	GGA	TTG	CTT	GAT	ACC	TTT	ATA	GGA	GAA	GGC	AAA	GAG	TAT	ATT	TTT	GTG	TCT	AAC
241	I	D	N	L	G	A	т	v	D	L	Y	I	L	N	Н	L	I	N	P	P
829	עיוויע	~~~		~ ~ ~	aam	000	707	CTC	CAT	CTC	mam	D THE	CTD	חת מ	ርስም	ሮሞአ	አምሮ	A A C	CCA	CCC
	AIA	GAT	AAT	CTG	GGL	GUU	ACA	GTG	GUT	CIG	TWT	ALL	CII	AAI	CAI	CIA	ALC.	AAC	CON	
261	N	GAT	AAT K	R	C	E	F	V	M	E	V	Ť	N	K	T	R	A	D	V	K
261 889	N AAT	GAT G GGA	K AAA	R CGC	C C TGT	E GAA	F TTT	V GTC	M ATG	E GAA	V GTC	T ACA	N AAT	K AAA	T ACA	R CGT	A GCA	D GAT	V GTA	K AAG
261 889 281	N AAT G	GAT G GGA G	K AAA T	R CGC L	C TGT T	E GAA Q	F TTT Y	V GTC E	M ATG G	E GAA K	V GTC L	T ACA R	N AAT L	K AAA V	T ACA E	R CGT I	A GCA A	D GAT Q	V GTA V	K AAG P
261 889 281 949	N AAT G GGC	GAT GGA GGA GGG	AAT K AAA T ACA	R CGC L CTC	C TGT T ACT	GAA Q CAA	F TTT Y TAT	V GTC E GAA	M ATG G GGC	E GAA K AAA	V GTC L CTG	T ACA R AGA	N AAT L CTG	K AAA V GTG	T ACA E GAA	R CGT I ATT	A GCA A GCT	D GAT Q CAA	V GTA V GTG	K AAG P CCA
261 889 281 949 301	N AAT G GGC K	GAT GGA GGA GGG A	K AAA T ACA H	R CGC L CTC V	C TGT T ACT D	E GAA Q CAA E	F TTT Y TAT F	V GTC E GAA K	M ATG G GGC S	E GAA K AAA V	V GTC L CTG S	T ACA R AGA K	N AAT L CTG F	K AAA V GTG K	T ACA E GAA I	R CGT I ATT F	A GCA A GCT N	D GAT Q CAA T	V GTA V GTG N	K AAG P CCA N
261 889 281 949 301 1009	N AAT G GGC K AAA	GAT G GGA G G G G G G G C A	K AAA T ACA H CAT	R CGC L CTC V GTT	C TGT T ACT D GAC	E GAA Q CAA E GAG	F TTT Y TAT F TTC	V GTC E GAA K AAG	M ATG G GGC S TCT	E GAA K AAA V GTA	V GTC L CTG S TCA	T ACA R AGA K AAG	N AAT L CTG F TTC	K AAA V GTG K AAA	T ACA E GAA I ATA	R CGT I ATT F TTT	A GCA A GCT N AAT	D GAT Q CAA T ACA	V GTA V GTG N AAC	K AAG P CCA N AAC
261 889 281 949 301 1009 321	N AAT G GGC K AAA L	GAT GGA GGA GGG A GCA W	K AAA T ACA H CAT I	CTG R CGC L CTC V GTT S	C TGT T ACT D GAC L	E GAA Q CAA E GAG A	F TTT Y TAT F TTC A	V GTC E GAA K AAG V	M ATG GGC S TCT K	E GAA K AAA V GTA R	V GTC L CTG S TCA L	T ACA R AGA K AAG Q	N AAT L CTG F TTC E	K AAA V GTG K AAA Q	T ACA E GAA I ATA N	R CGT I ATT F TTT A	A GCA A GCT N AAT I	D GAT Q CAA T ACA D	V GTA V GTG N AAC M	K AAG P CCA N AAC E
261 889 281 949 301 1009 321 1069	N AAT GGC K AAA L CTA	GAT GGA GGA GGG A GCA W TGG	AAT K AAA T ACA H CAT I ATT	CTG R CGC L CTC V GTT S TCT	C TGT T ACT D GAC L CTT	GCC GAA Q CAA E GAG A GCA	F TTT Y TAT F TTC A GCA	GIG V GTC E GAA K AAG V GTT	M ATG GGC S TCT K AAA	E GAA K AAA V GTA R AGA	V GTC L CTG S TCA L CTG	T ACA R AGA K AAG Q CAG	N AAT L CTG F TTC E GAG	K AAA V GTG K AAA Q CAA	T ACA E GAA I ATA N AAT	R CGT I ATT F TTT A GCC	A GCA A GCT N AAT I ATT	D GAT Q CAA T ACA D GAC	V GTA V GTG N AAC M ATG	K AAG P CCA N AAC E GAA
261 889 281 949 301 1009 321 1069 341	N AAT G GGC K AAA L CTA I I	GAT GGA GGA GGG A GCA W TGG I	K AAA T ACA H CAT I ATT V	CTG R CGC L CTC V GTT S TCT N	C TGT T ACT D GAC L CTT A	GAA Q CAA E GAG A GCA K	F TTT Y TAT F TTC A GCA T	V GTC E GAA K AAG V GTT L	M ATG GGC S TCT K AAA D	E GAA K AAA V GTA R AGA G	V GTC L CTG S TCA L CTG G	T ACA R AGA K AAG Q CAG L	N AAT L CTG F TTC E GAG N	K AAA V GTG K AAA Q CAA V	T ACA E GAA I ATA N AAT I	R CGT I ATT F TTT A GCC Q	A GCA A GCT N AAT I ATT L	D GAT Q CAA T ACA D GAC E	V GTA V GTG N AAC M ATG T	K AAG P CCA N AAC E GAA A
261 889 281 949 301 1009 321 1069 341 1129	N AAT G GGC K AAA L CTA I ATC	GAT GGA GGA GGG A GCA W TGG I ATT	K AAA T ACA H CAT I ATT V GTG	CTG R CGC L CTC V GTT S TCT N AAT	C TGT T ACT D GAC L CTT A GCA	GAA Q CAA E GAG A GCA K AAG	F TTT Y TAT F TTC A GCA T ACT	V GTC E GAA K AAG V GTT L TTG	M ATG GGC S TCT K AAA D GAT	E GAA K AAA V GTA R AGA G GGA	V GTC L CTG S TCA L CTG G GGC	T ACA R AGA K AAG Q CAG L CTG	N AAT L CTG F TTC E GAG N AAT	K AAA V GTG K AAA Q CAA V GTC	T ACA E GAA I ATA N AAT I ATT	R CGT I ATT F TTT A GCC Q CAA	ATC A GCA A GCT N AAT I ATT L TTA	D GAT Q CAA T ACA D GAC E GAA	V GTA V GTG N AAC M ATG T ACT	K AAG P CCA N AAC E GAA A GCA
261 889 281 949 301 1009 321 1069 341 1129 361	N AAT G GGC K AAA L CTA I ATC V	GAT GGA GGA GGG A GCA W TGG I ATT G	AAT K AAA T ACA H CAT I ATT V GTG A	CTG R CGC L CTC V GTT S TCT N AAT A	C TGT T ACT D GAC L CTT A GCA I ATC	GCC E GAA Q CAA E GAG A GCA K AAG K AAG	F TTT Y TAT F TTC A GCA T ACT S	U GTC E GAA K AAG V GTT L TTG F	M ATG GGC S TCT K AAA D GAT E	E GAA K AAA V GTA R AGA GGA N N	V GTC L CTG S TCA L CTG G GGC S TCT	T ACA R AGA K AAG Q CAG L CTG L	N AAT L CTG F TTC E GAG N AAT G	K AAA V GTG K AAA Q CAA V GTC I	T ACA E GAA I ATA N AAT I ATT N	R CGT I ATT F TTT A GCC Q CAA V GTC	A GCA A GCT N AAT I ATT L TTA P CCA	D GAT Q CAA T ACA D GAC E GAA R	V GTA V GTG N AAC M ATG T ACT S	K AAG P CCA N AAC E GAA A GCA R CCT
261 889 281 949 301 1009 321 1069 341 1129 361 1189 381	N AAT GGCC K AAA L CTA I ATC V GTA	GAT G GGA GGG A GCA W TGG I ATT G GGG	AAT K AAA T ACA H CAT I ATT V GTG A GCT P	CTG R CGC L CTC V GTT S TCT N AAT A GCC	C TGT T ACT D GAC L CTT A GCA I ATC	GCC E GAA Q CAA E GAG A GCA K AAG K AAA T	F TTT Y TAT F TTC A GCA T ACT S AGC	GTG E GAA K AAG V GTT L TTG F TTT	M ATG GGC S TCT K AAA D GAT E GAG	E GAA K AAA V GTA R AGA GGA N AAT	V GTC L CTG S TCA L CTG G GGC S TCT L	T ACA R AGA K AAG Q CAG L CTG L CTG L	N AAT L CTG F TTC E GAG N AAT G GGT	K AAA V GTG K AAA Q CAA V GTC I ATT M	T ACA E GAA I ATA N AAT I ATT N AAT	R CGT I ATT F TTT A GCC Q CAA V GTG	ATC A GCA A GCT N AAT I ATT L TTA P CCA	D GAT Q CAA T ACA D GAC E GAA R AGG	V GTA V GTG N AAC M ATG T ACT S AGC	K AAG P CCA N AAC E GAA A GCA R CGT
261 889 281 949 301 1009 321 1069 341 1129 361 1189 381 1249	N AAT GGC K AAA L CTA I ATC V GTA F	GAT GGA GGG GGG A GGG A TGG GGG L CTG	AAT K AAA T ACA H CAT I ATT V GTG A GCT P CCT	CTG R CGC L CTC V GTT S TCT N AAT A GCC V GTC	C TGT T ACT D GAC L CTT A GCA I ATC K AAA	GAA Q CAA E GAG A GCA K AAG K AAA T ACC	F TTT Y TAT F TTC A GCA T ACT S AGC T ACA	U GTC E GAA K AAG V GTT L TTG F TTT S TCA	M ATG GGC S TCT K AAA D GAT E GAG D GAT	E GAA K AAA V GTA R AGA G GGA N AAT L CTC	V GTC L CTG S TCA L CTG GGC S TCT L TTG	T ACA R AGA K AAG Q CAG L CTG L CTG L CTG	N AAT L CTG F TTC E GAG N AAT G GTG V GTG	K AAA V GTG K AAA Q CAA V GTC I ATT M ATG	T ACA E GAA I ATA N AAT I AAT S TCA	R CGT I ATT F TTT A GCC Q CAA V GTG N AAC	A GCA A GCT N AAT I ATT L TTA P CCA L CTC	D GAT Q CAA T ACA D GAC E GAA R AGG Y TAT	V GTA V GTG N AAC M ATG T ACT S AGC S AGT	K AAG P CCA N AAC E GAA A GCA R CGT L CTT
2251 889 281 949 301 1009 321 1069 341 1129 361 1189 381 1249 401	N AAT GGC K AAA L CTA I ATC V GTA F TTT N	GAT G GGA G GGG A GCA W TGG I ATT G GGG L CTG A	AAT K AAA T ACA H CAT I ATT V GTG A GCT P CCT G	CTG R CGC L CTC V GTT S TCT N AAT A GCC V GTC S	C TGT T ACT D GAC L CTT A GCA I ATC K AAA I	GCC E GAA CAA E GAG A GCA K AAG K AAA T ACC T	F TTT Y TAT F TTC A GCA T ACT S AGC T ACA M	V GTC E GAA K AAG V GTT L TTG F TTT S TCA S	M ATG GGC S TCT K AAA D GAT E GAG D GAT E	E GAA K AAA V GTA R AGA GGA N AAT L CTC K	V GTC L CTG S TCA L CTG G GGC S TCT L TTG R	T ACA R AGA K AAG Q CAG L CTG L CTG L CTG E	N AAT L CTG F TTC E GAG GAG AAT GGT V GTG F	K AAA V GTG K AAA Q CAA V GTC I ATT M ATG P	T ACA E GAA I ATA N AAT I AAT S TCA T	R CGT I ATT F TTT A GCC Q CAA V GTG N AAC V	A GCA GCT N AAT I ATT L TTA P CCA L CTC P	D GAT Q CAA T ACA D GAC E GAA R AGG Y TAT I,	V GTA V GTG N AAC M ATG T ACT S AGC S AGT V	K AAG P CCA N AAC E GAA A GCA R CGT L CTT K
2261 889 281 949 301 1009 321 1069 341 1129 361 1189 381 1249 381 1249 1309	N AAA G GGC K AAA L CTA I ATC V GTA F TTN N AAT	GAT GGA GGG A GCA W TGG I ATT GGG CTG A GCA	AAT K AAA T ACA H CAT I ATT V GTG A GCT P CCT GGA	CTG R CGC L CTC V GTT S TCT N AAT A GCC V GTC S TCT	GGT C TGT T ACT D GAC L CTT A GCA I ATC K AAA L CTG	GCC E GAA CAA E GAG A GCA K AAG K AAA T ACC T ACA	F TTT Y TAT F TTC A GCA T ACT S AGC T ACA M ATG	V GTC E GAA K AAG GTT L TTG F TTT S TCA S AGT	M ATG GGC S TCT K AAA D GAT E GAG GAT E GAA	E GAA K AAA V GTA R AGA GGA N AAT L CTC K AAG	V GTC L CTG S TCA L CTG GGC S TCT L TTG R CGG	T ACA R AGA K AAG CAG CAG CAG CTG CTG CTG CTG CTG CTG CTG	N AAT L CTG F TTC E GAG GAG GGT GGT GTG F TTT	K AAA V GTG K AAA Q CAA V GTC I ATT M ATG P CCT	T ACA E GAA I ATA N AAT I AAT S TCA T ACA	R CGT I ATT F TTT A GCC Q CAA V GTG N AAC V GTG	A GCA GCT N AAT I ATT L TTA P CCA L CTC P CCC	D GAT Q CAA T ACA D GAC E GAA R AGG Y TAT L TTG	V GTA V GTG N AAC M ATG T ACT S AGC S AGT V GTT	K AAG P CCA N AAC E GAA C CA R CGT L CTT K AAA
2251 889 281 949 301 1009 321 1069 341 1129 361 1189 381 1249 401 1309 421	N AAA GGCC K AAA L CTA I ATC V GTA F TTT N AAT L	GAT G GGA GGG A GGG I ATT GGG L CTG A GCA GCA	K AAA T ACA H CAT I ATT V GTG A GCT P CCT GGA S	CTG R CGC L CTC V GTT S TCT AAT A GCC V GTC S TCT S	C TGT T ACT D GAC L CTT A GCA I ATC K AAA L CTG F	GCC E GAA CAA E GAG A GCA K AAA T ACC T ACA T	F TTT Y TAT F TTC A GCA T ACT S AGC T ACA K	GIG V GTC E GAA K AAG V GTT L TTG F TTT S TCA S AGT V	M ATG GGC S TCT K AAA D GAT E GAG D GAT E GAA O	E GAA K AAA V GTA R AGA GGA N AAT L CTC K AAG D	V GTC L CTG S TCA L CTG G GCC S TCT L TTG R CGG Y	T ACA R AGA K AAG Q CAG L CTG CTG CTA L CTG GAA L	N AAT L CTG F TTC E GAG N AAT GGT V GTG F TTT R	K AAA V GTG K AAA Q CAA V GTC I ATT ATG P CCT R	T ACA E GAA I ATA N AAT I AAT S TCA T ACA F	R CGT I ATT F TTT A GCC Q CAA V GTG N AAC V GTG E	ATC. A GCA A GCT N AAT I ATT L TTA P CCA L CTC P CCC S	D GAT Q CAA T ACA D GAC E GAA R AGG Y TAT L TTG I	V GTA V GTG N AAC M ATG T ACT S AGC S AGT V GTT P	K AAG P CCA N AAC E GAA C CA R CGT L CTT K AAA D
2251 889 281 949 301 1009 321 1069 341 1129 361 1189 381 1249 401 1309 421 1369	N AAT G GGC K AAA L CTA CTA CTA F TTT N AAT CTA L TTA	GAT G GGA GGG A GGG A GGG A GGG C T G G G C C A G G G G G G G G C A G G G G	K AAA T ACA H CAT I ATT V GTG A GCT P CCT GGA S AGT	CTG R CGC L CTC V GTT S TCT AAT A GCC V GTC S TCT S TCT	C TGT T ACT D GAC L CTT A GCA I ATC K AAA L CTG F TTT	GCC E GAA CAA E GAG A GCA K AAG K AAA T ACC T ACC	F TTT Y TAT F TTC A GCA T ACT S ACT S ACT ACA K AAG	GTC E GAA K AAG V GTT TTG F TTT S CA S AGT V GTT	M ATG GGC S TCT K AAA D GAT GAT GAT GAA Q CAA	E GAA K AAA GTA GGA GGA L CTC K AAG GAT	V GTC L CTG S TCA L CTG G G G C TCT TTG R CGG Y TAT	T ACA R AGA K AAG CAG CAG CAG CTG CTG CTG CTG CTA CTA	N AAT L CTG F TTC E GAG GGT GGT GTG F TTT R AGA	K AAA V GTG K AAA Q CAA V GTC I ATT ATG CCT R AGA	T ACA E GAA I ATA N AAT I AAT S TCA T ACA F TTT	R CGT I F TTT A GCC Q CAA V GTG AAC V GTG E GAA	A GCA A GCT N AAT I ATT L TTA CCA CTC P CCC S AGT	D GAT Q CAA T ACA D GAC E GAA R AGG Y TAT L TTG I ATA	V GTA V GTG N ACT ACT ACT S AGT V GTT P CCA	K AAG P CCA N AAC E GAA A GCA R CGT L CTT K AAA D GAT
221 889 281 949 301 1009 321 1069 341 1129 361 1249 401 1269 421 1269 421	N AAT G GGC K AAA L CTA I ATC V GTA F TTT N AAT L TTA M	GAT G G G G G G G G G G G G G G G G G G	K AAA T ACA H CAT I ATT V GTG A GCT P CCT G GGA S AGT E	CTG R CGC L CTC V GTT S TCT N AAT A GCC V GTC S TCT S TCT L	C TGT T D GAC L CTT A GCA L CTT A GCA L CTT A CTC K AAA L CTG F TTT D	GCC E GAA CAA E GAG A GCA K AAA C AAA T ACC T ACG H	F TTT Y TAT F TTC A GCA T ACT S AGC T ACA M A AGC K AAG L	GTC GAA K AAG GTT L TTG F TTT S CA S GTT T CA S GTT T	M ATG G GGC S TCT K AAA D GAT E GAA GAA Q CAA V	E GAA K AAA V GTA R AGA GGA N AAT L CTC K AAG GAT S	GTC CTG CTG CTG CTG GGC CTG GGC S TCT L TTG R CGG Y TAT G	T ACA R AGA K AAG Q CAG L CTG CTG CTG CTG CTG CTA L CTA L CTA D	AAT L CTG F TTC GAG N AAT GGT F TT R AGA V	K AAA V GTG K AAA V GTG CAA V GTC I ATT M ATG P CCT R AGA T	TACA EGAA IATA AAT I AAT S TCA TACA F TTT F	R CGT I ATT F TTT A GCC Q CAA V GTG N AAV GTG E GAA G	A GCA A GCT N AAT I ATT L TTA P CCA L CTC P CCC S AGT K	D GAT Q CAA T ACA D GAC E GAA R AGG Y TAT L TTG I ATA N	V GTA V GTG N AAC M ATG T ACT S AGC S AGT V GTT P CCA V	K AAG P CCA AAC E GAA A GCA R CGT L CTT K AAA D GAT S
261 889 281 949 301 1009 321 1069 341 1129 361 1129 361 1129 381 1249 401 1309 421 1309 421 13441 1429	ATA N AAT GGC K AAA L CTA I ATC V GTA F TTT N AAT L TTA M ATG	GAT G GGA GGG GGG A GGA A TGG GGG L CTG GCA GCA GCA CTT	AAT K AAA T ACA H CAT I ATT V GTG A GCT P CCT G GA S AGT E GAA	CTG R CGC L CTC V GTT S TCT N AAT A GCC V GTC S TCT S TCT S TCT L TTG	GGT C TGT T AACT D GAC L CTT A GCA L CTT A ATC K AAA L CTG F TTT D GAT	GCC E GAA Q CAA E GAG A GCA K AAA T C ACA T ACA CAC	F TTT Y TAT F TTC A GCA T ACT S AGC T ACT S AGC T ACT S CTC CTC	GTC GTC E GAA K AAG GTT L TTG F TTT S CA S AGT V GTT T ACA	M ATG G GGC S TCT K AAA D GAT E GAG GAT E GAA Q CAA V GTT	E GAA E GAA K AAA V GTA R AGA G GGA N AAT L CTC K AAG D GAT S TCA	V GTC L CTG S TCA L CTG G GC S TCT L TTG R CGG Y TAT G GGA	T T ACA AGA K AGA CAG CAG CAG CTG CTG CTG CTG CTG CTA L CTG GAA L CTA D GAT	N AAT L CTG F TTC E GAG F TTC GAG F TTT R AGA V GTG	AAAI K AAAA V GTG K AAAA Q CAA V GTC I ATT M GTC I ATT ACA	T T ACA E GAA I ATA N AAT I AAT S AAT T CA T CA F TTT F TTT	R CGT I F TTT A GCC Q CAA V GTG GTG GGA GGA	AIC. A GCA A GCT N AAT I ATT L TTA P CCA L CTC P CCC S AGT K AAA	D GAT Q CAA T ACA D GAC E GAA R GAC R GAA R TAT L TTG I I ATA N AAT	V GTA V GTG GTG N AACC M ATG T ACT S AGC S AGC V GTT P CCA V GTT	K AAG P CCA N AAC E GAA CGT CTT K AAA D GAT S TCA
261 889 281 949 301 1009 321 1069 341 1129 361 1189 381 1249 401 1309 421 1369 441 1429 461	AIA N AAT GGCC K AAA L CTA I CTA F TTT N AAT L TTA M ATG L	GAT G GGA GGA GGG A GCA GGG A TGG A TGG A TGG CTG A GGCA G GCA CTT K	AAT K AAA T ACA H CAT I ATT V GTG A GCT P CCT G GGA S AGT E GGAA G	CTC R CGC L CTC V GTT S TCT N AAT A GCC V GTC S TCT S TCT S TCT L TTG T	GGT C TGT T ACT D GAC L CTT A GCA L CTT A ACT C TT T T T T T T T T T V V	GCC E GAA Q CAA E GAG A GCA K AAA T ACC T ACA T ACC T ACA I I	F TTT Y TAT F TTC A GCA T T ACT S AGC T ACT S AGC T ACA M ATG K AAG K CTC L I	V GTC E GAA K AAG V GTT L GTT TTG S TCA S TCA S AGT T T ACA I	M ATG G GGC S TCT K AAA D GAT E GAA C A Q CAA V GTT A	E GAA E GAA K AAA V GTA R AGA G GGA N AAT L CTC K AAG G GAT S TCA N	V GTC L CTG S TCA L CTG G GC S TCT L TTG R CGG C Y TAT GGA H	T T ACA R AGA K AAG Q CAG C CTG CTG CTG CTG CTG CTG CTG CTG GAA C T C T G GAT G	N AAT L CTG F TTC E GAG GGT F TTT C GTG GGT TTT R AGA V GTG D	AAAI K AAAA V GTG K AAAA Q CAAA V CCAA V CCAA V GTCC I AATT R AATG P CCT R AATA ATG R AATA R	T T ACA E GAA I ATA N AAT T AAT T AAT T T T T T T T T T	R CGT I ATT F TTT A GCCC Q GTG Q CAA V GTG GTG GTG GTG GTG GTG GTG CAA D	AICCA A GCA A GCT N AAT I AAT I AAT CTA P CCA CTC P CCCA S AGT K AAA I	D GAT Q CAA T ACA D GAC E GAA R AGG TAT L TTG I I TTG I N AAT P	V GTA V GTG N AAC T AAC S AGC S AGC S AGC V CCA V GTT P CCA V GTT	K AAG P CCA N AAC E GAA R CGT L CTT K AAA D GAT S TCA G
261 889 281 949 301 1009 321 1069 341 1129 361 1189 381 1249 401 1309 421 1369 441 1429	N AAT G GGC K AAA L CTA I CTA F TTT N AAT L TTA M G L TTA	GAT G GGA GGA GGG A GCA W TGG A TGG A TGG L CTG A GGCA G GCA CTT K AAG	AAT K AAA T ACA H CAT I ATT V GTG A GCT P CCT G GGA S AGT E GAA GGA	CTC R CGC L CTC V GTT S TCT N AAT A GCC V GTC S TCT S TCT S TCT L GTC S TCT S CGC C C C C C C C C C C C C C C C C C	GGT C TGT ACT D GAC L CTT A ACC L CTT A ACT C TT A CTT C TT T TT D C TTT D C T T T T T T T	GCC E GAA Q CAA E GAG A GCA K AAA T ACC T ACA T ACC T ACC I ATC	F TTT Y TAT F TTC A GCA T ACT S AGC T ACA M ATG K AAG L C TC I ATC	V GTC E GAA K AAG V GTT T TTT S TCA S AGT V GTT T ACA I ATT	M ATG G GGC S TCT K AAA D GAT E GAG GAT E GAA Q CAA V CAA V CAA V GTT A GCA	E GAA E GAA K AAA V GTA R AGA G GGA N AAT L CTC K AAG G GAT S TCA N AAT	V GTC L CTG S TCA L CTG G GC S TCT L TTG R CGG Y TAT G GA H CAT	T T ACA R AGA K AAG Q CAG CTG CTG CTG CTG CTG CTG CTG GGAT GGT	N AAT L CTG F TTC E GAG GGT F TTT R AGA V GTG GGT GGT G D GAC	AAAI K AAAA V GTG K AAAA Q CAAA V CCAA V GTCC I ATT M ATG P CCT R AGA T ACA R ACA	T T ACA E GAA I ATA N AAT S TCA T T T T T T T T T T T T T T T T T T	R CGT I ATT F TTT A GCC Q CAA V GTG GTG GTG GAA GAA D GAT	AIC. A GCA A GCT N AAT I AAT I CTC P CCA L CTC P CCC S AGT K AAA I AAC	D GAT Q CAA T ACA D GAC E GAA R GAC R GAC T T T T T T T T T T T T T T T T T T T	V GTA V GTG N AAC M AAC T AACT S AGC S AGC V CCT P CCT	K AAG P CCA N AACC E GAA A GCA R CGT L L CTT K AAA D GAT S TCA G GGA
261 889 281 949 301 1009 321 1069 341 1129 381 1249 401 1309 421 1309 421 1309 421 1369 441 1429 461	AIA N AAT G G G K AAA L C TA C TA C TA C TTA ATG TTA A TTA A	GAT G GGA GGA GGG A GGG A GGG A GGG CTG GGC CTG A GCA GCA CTT K AAG V	AAT K AAA T ACA H CAT I CAT GGG A GGG A GGG A GGA S GGA L	CTC R CGC L CTC V GTT S TCT N AAT A GCC V GTC S TCT S TCT S TCT L GTC S TCT S CGC C C C C C C C C C C C C C C C C C	GGT C TGT T ACT D GAC L CTT A ACT T CTT A GCA L CTG F TTT D GAT V V S GTT N	GCC E GAA Q CAA E GAG A GCA K AAA CAC T ACA T ACC T ACA T ACG H CAC I ATC K	F TTT Y TAT F TTC A GCA T ACT S AGC T ACT S AGC T ACA M ATG K AAG L C TC I I	V GTC E GAA K AAG V GTT TTG F TTT S CA S AGT V GTT T CA S AGT V U GTT T CA S AGT V V U U U U U U U U U U U U U U U U U	M ATG GGC S TCT K AAA D GAT E GAG GAT E GAA Q CAA V T A GCA S	E GAA K AAA V GTA R GGA GGA N AAAT L CTC K AAG GGAT S CTC K AAG D GAT S CTC K AAG G GAT S CTC K AAG G GAT S CTC S	V GTC L CTG S TCA L CTG G GGC S TCT L CTG R CGG G GGA H CAT N	T T ACA R AGA K AAG Q CAG CAG CAG CTG CTG CTG CTG CTG CTG CTG CTG CTG CT	N AAT L CTG F TTC E GAG GGT F TTT R AGA V GTG D GAC R	AAAI K AAAA V GTG K AAAA Q CAA V GTCC I ATT M ATG P CCT R AGA T ACA R AGA I	T ACA E GAA I ATA N AAT I ATT N AAT T CA T TTT F T T TT I ATT L	R CGT I ATT F TTT A GCC Q CAA V GTG GTG GAA GGAA D GAT D	AICCA A GCT A GCT N AATT I TTA P CCA L CTC P CCCCS AGT K AAAA I ATC H	D GAT Q CAA T ACA D GAC E GAA R AGG T ATA TAT L TTG I ATA N TAT P CCA *	V GTA V GTG N AACG S AGT V GTT P CCA V GTT P CCT	K AAG P CCA N AACC E GAA A GCA R CGT L CTT K AAA D GAT S CGA GGA
221 889 281 949 301 1009 321 1069 341 1129 381 1229 381 1249 421 1309 421 1309 421 1309 421 1309 421 1369 441 1429 461 1489	N AAAT G GGCC K AAAA L CTAA I CTAA F TTT N AAAT L TTA M ATG L A GCA	GAT G GGA GGA GGG A GGG A GGG A GGG CTG GGC CTG GGC CTT K AAG V GTA	AAT K AAA T ACA H CAT I ATT V GTG A GCT G G GCT G GGA S AGT E GAA G GGA L TTA	CTCC R CGC U CTC V GTT S TCT N AAT A GCC S TCT S TCT L TTG T ACG E GAG	GGT TGT ACT D GACC L CTT A A GCA L CTT A A CTG F TTT D GAT V GTT N AAC	GCCC E GAA Q CAA E GAG A GCA K AAA C ACC T ACCA T ACCA I ACCC I AACC K AAG	F TTT Y TAT F TTC A CCA T T ACCA M ATC ACA M ATG K AAG L CTC I ATC I ATA	V GTC E GAA K AAG V GTT TTG F TTT S CA S AGT V GTT T ACA ATT V GTG	M ATG GGC S TCT K AAA D GAT E GAA Q CAA V CAA V GTT A GCA S TCT	E GAA K AAA V GTA R AGA GAA CTC K GGA CTC K AAG CTC K AAG CTC K GGA CTC K GGA A AAT GGA CTC C C C C C C C C C C C C C C C C C	V GTC L CTG S TCA L CTG G GGC S TCT L TTG R CGG GGC Y TAT G GAT H CAT N AAC	T T ACA R AGA K AAG Q CAG CAG CAG CAG CAG CTG CTG CTG CTG GAA L CTA G GAT L CTT	N AAT L CTG F TTC E GAG GAG F TTT G GGT G GGT C TTT R AGA V GTG GTG G GCC R CGC	AAAI K AAAA V GTG K AAAA Q CAA V GTCC I AATT AATG P CCT R AGA T AAGA I ATC	T T ACA E GAA I ATA N AAT I AAT T AAT T T T T T T T T T T T T	R CGT I AATT F TTT A GCC Q CAA V GTG GTG GTG GGA GGA D GAC	A CCA A CCA CCCCA CCCA	D D D CAA T ACA D CAA T A CAA T A CAA C CAA T A C A C A C	V GTA V GTG GTG N AACC M AACC T AACC S AGT V GTT P CCA V GTT P CCT AATC	K AAG P CCA N AAC E GAA A CGT L CGT L CGT L CGT L CGT L CGT CGT CGT GGA GGA GGA GGA
261 889 281 949 301 1009 321 1069 341 1129 361 1189 381 1249 401 1309 421 1309 421 1309 421 1309 441 1429 461 1429 461 1429 461	N N AAT G G G C K AAA L C TA ATC C TTT N AAT L TTA M ATG L TTA A G CA AAT/A C TA A TA A TA C TA A TA C C C C C K C C C C K C C C K C C C C	GAT G GGA GGG GGG A GCA W TGG GGA A ATT G GGG CTT K A AGGC L CTT K A AGGC V CTG CTT K CTT CTT C CTT	AAT K AAA T ACA H CAT I ATT V GTG A GCT P CCT G GGA GGA GGA GGA L TTA CGGA	CTC CGC R CGC L CTC V GTT S TCT N AAT A GCC V GTC S TCT S TCT S TCT L TTG TCT TG ACG GAG CACT	GGT TGT TGT ACT D GAC L CTT A ACC CTT A ACC K AAA L CTG F TTT T T T T T T T T T T T T T T T T	E GAA Q CAA E GAG GAG GCA K AAAG T ACC T ACC T ACC I ACC I ACC K AAG CAAC	F F TTT Y TAT F TTC A C C A C C C C C C C C C C C C C C	V GTC E GAA K AAG V GTT L GTT T T S TCA S TCA S TCA S TCA S TCA S T CA S T T T T T T T T T T T T T T T T T T	M ATG G GGC S TCT K AAA D GAT E GAG GAT E GAG GAT V GTT A GCA S TCT TTCT	E GAA K AAA V GTA R AGA GGA CTC K AAGA GGAT S TCA N AAT GGA TACAJ	V GTC L CTG S TCA L CTG G GCC S TCT TTG R CGG G GCS TCT TTG R CGG G GA H CAT N AACC	T T ACA R AGA R AGA CAG CAG CCAG CCAG CCAG	N AAT L CTG F TTC E GAG GGT GGT GGT F TTT R AGA V GTG GGT D GAC R CGC CTCTC	AAAI K AAAA V GTG K AAAA Q CAA V CAA V GTC I M ATG P CCT R AGA T AGA T AGA I I ATC CTAG	T ACA E GAA I ATA N AAT I ATA N AAT T N AAT T T T T T T T T T T T T T T T T G GATT	R CGT I ATT F TTT A GCC Q CAA V GTG GCC V GTG GGA GGA D GAC D CAA CAA C CAA CAA CAA C CAA CAA CAA	A C CACTA	D D CAA T ACA D CAA T A CAA C A CAA C CAA T A CAA C CAA T A CAA C CAA T A CAA T C CAA T A CAA T C A CAA T A CAA T A CAA T A CAA T A CAA T A CAA T A CAA T A CAA T A CAA T C A CAA T C A CAA T C A CAA T C A CAA T C A CAA T C A CAA T C A CAA T C A C A	V GTA V GTG N AAC M ATG T S AGC V GTT P CCCA QGTT P CCT	K AAG P CCA N AAC E GAA A CGT L CTT K AAA CGT CTT K AAA D GAT S TCA G GGA CGA
261 889 281 949 301 1009 321 1069 341 1129 361 1189 381 1249 401 1369 421 1369 441 1429 461 1429 461 1429 461 1549 1610 1689	N AAAT G GGC K AAAA L CTA ATC V GTA TTT N AATC L TTA ATC L TTA A GCA AAATI ACTA	GAT G GGA GGG GGG A GCA G GGG GGG CTG CTG CTG K AAG GCA CTT K K AAG GCA CTT CTT K CTG CTT K CTT CTT C CTT C CTT C CTT C C C C	AAT K AAA T ACA H CAT I ATT V G G G A G G CAT C G G G A C CT E G G A C CT T E G G A L TTA C A CA T A CA T A CA T A CA T A CA T A CA A A A	R CGC L CGC V GTT S TCT N AAT A AAT A AGCC V GTC S TCT T C T T C T T C S C C C C C C C	GGT TGT TGT ACT D GAC L CTT A A GCA L CTG GTT TTT T D GAT V GTT N AACC CTG GAT CTGG C TA CTG C CTG C C C C C C C C C C C C C C C	GCC E GAA CAA E GAG GAG A GCA K AAA T C ACA T ACA T ACA T ACA T ACA T ACA T AAA A AAA A AAA C AACA T AAA A AAA A AAA A AAA A AAA A AAA A AAA A	F TTTT Y TAT F CA GCA T ACT S ACT S ACT T ACT S ACT T ACT S ACT I ATG L CTCC I I ATA S GCCA T ACT S ACC A ACT T C C A C A C A C A C A C A C	V GTC E GAA K AAG GTT L TTG F TTT S TCA S TTT S TTT S TTT S TTT V GTT T T CA T T T CA T T T T T T T T T T T	M ATG G GGC S TCT K AAA D GAT E GAG GAT E GAA Q GAT A GCA S TCT TTAA	E GAA K AAAA V GTA R AGA GGA N AAAT CTCC K K AAGG GGA N CTCC K AAGG GGA N AAAAA AAAAAAAAAAAAAAAAAAAA	V GTC L CTG S TCA L CTG GGC S TCT TTG GGC S TCT TTG R CGG GGA H CAT G GGA N AAC CATGAN	T T ACA R AGA R AGA CAG CAG CCAG CCAG CCAG	N AAT L CTG F GAG GGG F TTC C GGG GGT GGTG GGTG GGTG C GCCC C CCCC	AAAA K AAAA V GTG K AAAA Q CAA V GTC CAA V GTC I AATT R AATG P CCT R AGA T ACA R AATG C TAGC TAGC TAGC	T ACA E GAA I ATA N AAT I AATA N AAT S ATT T AAT T T T T T T T T T T T T T T G GATT	R CGT I ATT F TTT A GCC Q CAA V GTG CAA V GTG CAA V GTG GAT GGA C D GAC C CAA C CAA C CAA C CAA C CAA C CAA C CAA C C CAA C	A CCA A GCA A GCT N AAT I ATT I ATT L CTC C CCC S AGT K AAA A I ATC CCC C CCC S AGT C CCC C C C C C C C C C C C C C C C C	D GAT Q CAA T ACA D GAC E GAA R ACA C ACA C ACA T T T G GAC T ATA T T T G GAC T ATA T T T G GAC A C A A C A A C A A C A A C A A C A A C A A C A A C A A C A A C A C A A C C A C A C A C A C	V GTA V GTG N AACC M ATG T AACT S AGC S GTT P CCA V V CCT AACC S GTT CCT AACC	K AAG P CCA N AAC E GAA A CGT L CTT K AAAA D GAT TCA G GGA SAAA CTTT TCA G GGA

В

				М	S	R	F	v	Q	D	\mathbf{L}	S	K	А	М	12
pHC379	1	CCT	AAA	ATG	TCG	AGA	$\mathbf{T}\mathbf{T}\mathbf{T}$	GTA	CAA	GAT	\mathbf{CTT}	AGC	AAA	GCA	ATG	42
								- I			111					
pHC377	70	GAA	GCG	GGA	GGA	CTG	AGA	AAA	CCA	GAT	CTT	AGC	AAA	GCA	ATG	111
										1					М	1

С

рНСЗ79 M85693	CCCCCAATTTCCCAGCTGCTAAAGGAAGGAAGGAAGGTACCTGTGCGTGC	60 161
pHC379 M85693 R58578	AGGGCTGGGGAAGCGGGAGGACTGAGAAAACC-AGATCTTAGCAAAGCAA	119 221 30
pHC379 M85693 R58578	TGGTGCTTCTCAGTTCCAAGAAGTCATTCGGCAAGAGCTAGAATTATCTGTGAAGAAGGA N	179 281 90
pHC379 M85693 R58578	ACTAGAAAAAATACTCACCACAGCATCATCACATGAATTT-GAGCACACCAAAAAAGACC	238 341 146
рНС379 M85693	TGGATGGATTTCGGAAGCTATTTCATAGATTT G	270 373

Fig. 1. Nucleotide sequence and deduced amino acid sequence of human muscle UDP-Glc PPase. (A) The nucleotide sequence is numbered in the 5' to 3' direction, beginning at the first position of the EcoRI linker. The amino acid sequence is numbered from the first methionine at the putative translation initiation site. (B) The nucleotide sequences of the 5'-regions of the two muscle forms. (C) Alignment of the nucleotide sequences of pHC377 and two human expressed sequence tags. A dot indicates identity to pHC377 while a hyphen indicates a gap.

Human liver Bovine liver Human muscle-I Human muscle-II	MSRFVQDLSKAMSQDGASQFQEVILQELELSVKKELEKILTTATSHEYEH R P F F R S F F R S F F	50 50 39 50
Human liver Bovine liver Human muscle-I Human muscle-II	TKKDLDGFRKLYHRFLQEKGPSVDWGKIQRPPEDSIQPYEKIKARGLPDN F F F	100 100 89 100
Human liver Bovine liver Human muscle-I Human muscle-II	ISSVLNKLVVVKLNGGLGTSMGCKGPKSLIGVRNENTFLDLTVQQIEHLN V	150 150 139 150
Human liver Bovine liver Human muscle-I Human muscle-II	KSYNTDVPLVLMNSFNTDEDTKKILQKYNHCRVKIYTFNQSRYPRINKES • T • T • T • T	200 200 189 200
Human liver Bovine liver Human muscle-I Human muscle-II	LRPVAKDVSSSGESTEAWYPPGHGDIYASFYNSGLLDTFLEEGKEYIFVS •L••••N••Y•••N•••••••••••••••••••••••••	250 250 239 250
Human liver Bovine liver Human muscle-I Human muscle-II	NIDNLGATVDLYILNHLINPPNGKRCEFVMEVTNKTRADVKGGTLTQYEG	300 300 289 300
Human liver Bovine liver Human muscle-I Human muscle-II	KLRLVEIAQVPKAHVDEFKSVSKFKIFNTNNLWISLAAVKRLQEQNAIDM	350 350 339 350
Human liver Bovine liver Human muscle-I Human muscle-II	EIIVNPKTLDGGLNVIQLETAVGAAIKSFENSLGINVPRSRFLPVKTTSD A.A.	400 400 389 400
Human liver Bovine liver Human muscle-I Human muscle-II	LLLVMSNLYSLNAGSLTMSEKREFPTVPLVKLGSSFTKVQDYLRRFESIP	450 450 439 450
Human liver Bovine liver Human muscle-I Human muscle-II	DMLELDHLTVSGDVTFGKNVSLKGTVIIIANHGDRIDIPPGAVLENKIVS	500 500 489 500
Human liver Bovine liver Human muscle-I Human muscle-II	GNLRILDH	508 508 497 508

Fig. 2. Comparison of the deduced amino acid sequences of the human liver, bovine liver and two human muscle forms of UDP-Glc PPase. The human liver sequence (Peng and Chang, 1993) is shown in the first line; for the other proteins, a dot indicates identity to the human liver sequence.

pHC379 encodes a protein of 508 amino acid residues (Fig. 2) as in the liver form of the enzyme from which it differs at 10 positions. The overall nucleotide identity within the coding regions is 92.3%. The protein encoded by pHC377 is somewhat smaller (497 residues) as it uses as a translation initiation codon the second in-frame ATG normally found in the muscle II form.

The liver form that we identified previously (Peng and Chang, 1993) was isolated from a commercial cDNA library purchased from Stratagene. To confirm that this DNA is truly of human liver origin, we sought to identify this sequence in an independent human liver cDNA library, purchased from Clontech (cat. no. HL3006b). Approximately 10⁷ plaque-forming units of this library were subjected to PCR (32 cycles of 94°C, 1 min; 55°C, 1 min; and 72°C, 2 min) using the synthetic oligonucleotides primers 5'-CAGAAGCTTGGTACCCCC-3' and 5'-GATTAGATGATTAAGAAT-3'. These primers correspond to bases 719–736 and 874–857, respectively, of the muscle sequence (Fig. 1A). The amplified fragment was isolated from an agarose gel and its nucleotide sequence was determined as de-

scribed by Cullmann et al. (1993). Among the eight nucleotides between positions 760 and 840 of the muscle cDNA that are different from those of the liver form, two were not identifiable unambiguously while the remaining six matched the liver sequence.

Kinetic properties. These cDNA clones were expressed in *E. coli* and the enzymes were partially purified and characterised with respect to their kinetic properties. Before presenting these results, the assay conditions will be described briefly.

Assay conditions. There have been conflicting reports of the kinetic properties of UDP-Glc PPase with respect to PP_i, with some authors describing a sigmoidal dependence of rate upon concentration (Villar-Palasi and Larner, 1960; Knop and Hansen, 1970). With the bovine enzyme, assayed in 0.1 M Tris/HCl pH 8.0, we also found a sigmoidal relationship (Fig. 3, circles). This appears to be due entirely to a competition between the buffer and PP_i for magnesium ions. Changing to triethanolam-ine/HCl pH 8.0, and maintaining a constant concentration of free





Fig. 3. Effect of the buffer on the apparent kinetics towards PP_i of bovine liver UDP-Glc PPase. The rate was measured over a range of PP_i concentrations using 0.1 M Tris/HCl (\bullet) or 0.1 M triethanolamine/HCl (\blacksquare) at pH 8.0. For Tris/HCl, the line represents the best fit of Eqn (4) to the data. In the case of 0.1 M triethanolamine/HCl, the concentrations of PP_i and MgCl₂ were varied simultaneously so as to maintain free [Mg²⁺] at 1 mM; the line represents the best fit of Eqn (1) to the data.

magnesium ion while varying the concentration of MgPP_i, gave a saturation curve that is hyperbolic (Fig. 3, squares). As will be seen later, the human liver and muscle II forms of the enzyme show non-hyperbolic kinetics with respect to MgPP_i that we believe represent true properties of the enzyme rather than an artefact of the assay.

Muscle form I. The human muscle I form of the enzyme exhibited Michaelis-Menten kinetics with respect to each of the four substrates. UDP-Glc PPase is thought to have an ordered sequential mechanism (Tsuboi et al., 1969; Aksamit and Ebner, 1972). As expected for this mechanism, the human muscle I form yielded an intersecting pattern in double-reciprocal plots obtained by varying the concentrations of both UDP-Glc and MgPP_i (Fig. 4). From these data, Michaelis constants for each substrate and an inhibition constant for UDP-Glc were determined. In the forward direction, varying the concentration of each substrate (MgUTP and Glc1P) gave hyperbolic curves from which Michaelis constants were determined. The inhibition constant for MgUTP was determined by using it as a product inhibitor of the reverse reaction; this inhibition was competitive with UDP-Glc as the varied substrate. The kinetic constants are summarised in Table 1.

Fig. 4. Kinetic properties of the human muscle UDP-Glc PPase form I with respect to MgPP_i and UDP-Glc. The rate was measured over a range of MgPP_i concentrations with fixed UDP-Glc concentrations (μ M) of 300 (\oplus), 100 (\bigcirc), 60 (\blacksquare), 30 (\square) and 15 (\blacklozenge). The results are plotted in double reciprocal form and the lines represent the best fit of Eqn (2) to the data. A represents the absorbance at 340 nm.

Liver form. The kinetics of the human liver enzyme in the reverse reaction are illustrated in Fig. 5 with MgPP_i as the varied substrate. Double-reciprocal plots (Fig. 5A) are markedly curved, especially at low concentrations of UDP-Glc (diamonds). This curvature results from a combination of two effects. First, there is some substrate inhibition as can be seen in Fig. 5B; this is most readily apparent at an intermediate UDP-Glc concentration (squares) of 70 µM. This substrate inhibition is partial in the sense that the rate falls to a finite plateau rather than to zero as the concentration of MgPP_i is raised. The curvature seen in Fig. 5A is partly due to sigmoidicity in the rate versus concentration curve at low concentrations of MgPP_i, as shown in Fig. 5C. Although the effect is not large (described by a Hill coefficient of 1.415 ± 0.064), the data clearly do not follow a hyperbola as shown by the broken line in Fig. 5C that represents the best-fit of the Michaelis-Menten equation to this subset of the data.

A similar pattern to that described for MgPP_i emerges with UDP-Glc as the varied substrate; that is, a sigmoidal curve with partial substrate inhibition that becomes close to hyperbolic at high concentrations of MgPP_i.

The kinetics in the forward reaction have been less extensively characterised. For MgUTP, the kinetics appear to be hyperbolic with a K_m of 419 ± 32 µM which is approximately half

Table 1. Kinetic constants of three forms of human UDP-Glc PPase. The constants were obtained by fitting the appropriate equation to the data by nonlinear regression and are reported as the best fit value ± standard error.

Reaction	Constant	Value for							
		muscle form I	muscle form II	LacZ/liver form					
		μΜ							
Forward	$K_{ m m}({ m MgUTP})$ $K_{ m m}({ m Glc1}P)$ $K_{ m i}({ m MgUTP})$	917 ± 147 404 ± 51 979 ± 80	301 ± 15 207 ± 28 216 ± 21	563 ± 115 172 ± 10 643 ± 47					
Reverse	$K_m(\text{UDP-Glc})$ $K_m(\text{MgPP}_i)$ $K_i(\text{UDP-Glc})$	$\begin{array}{rrrr} 63 \pm & 6 \\ 384 \pm & 36 \\ 52 \pm & 8 \end{array}$	41 ± 2 sigmoidal not determined	$\begin{array}{rrrr} 49 \pm & 4 \\ 166 \pm & 13 \\ 13 \pm & 4 \end{array}$					



Fig. 5. Kinetic properties of the human liver UDP-Glc PPase with respect to MgPP₁ and UDP-Glc. (A, B) The rate was measured over a range of MgPP₁ concentrations with fixed UDP-Glc concentrations (μ M) of 300 (\odot), 70 (\blacksquare) and 20 (\diamondsuit). The results are plotted with (A) or without (B) double-reciprocal transformation; the lines represent individual best fits of Eqn (5) to each set of data. (C) The results obtained at 70 μ M UDP-Glc shown in greater detail; (——) the best fit of Eqn (4) to the data; (-—) best fit of Eqn (1) to the data. A represents the absorbance at 340 nm.



Fig. 6. Kinetic model for UDP-Glc PPase. The two-subunit Monod-Wyman-Changeux model was modified by allowing for different rate constants for catalysis when one (k_1) or both (k_2) active sites are occupied.

the value observed for the muscle form I under identical conditions. For Glc1P, the saturation curve is non-hyperbolic and is described by Eqn (5).

LacZ/Liver fusion enzyme. As noted earlier, one of the differences between muscle form I and the liver form is in the Nterminal region of the protein and it seems probable that the kinetic complexity of the liver form may arise, in part, from the way in which this region folds. This hypothesis was supported by experiments on a fusion protein in which the first nine residues of the liver sequence (MSRFVQDLS) are replaced with the first 39 residues of the pBluescript *lacZ* gene product (MTMIT-PSSKLTLTKGNKSWSSTAVAAALELVDPPGCRNS) plus the linker sequence ARG. The kinetic properties of this fusion protein are summarised in Table 1. Although these properties differ from those of muscle form I, no kinetic anomalies were observed and the enzyme followed Michaelis-Menten kinetics with respect to all substrates.

Muscle form II. To dissect further the relative contributions to the kinetic differences of the various sequence variations between the liver and muscle I forms, the properties of the muscle II forms were investigated (Table 1). This form was found to exhibit Michaelis-Menten kinetics with respect to all substrates except MgPP_i, for which the kinetics were sigmoidal. Thus it appears that this particular kinetic anomaly is associated with the N-terminal region while the non-hyperbolic kinetics of the liver form with respect to the other substrates are controlled by sequence differences in the remainder of the protein.

DISCUSSION

Complementation of an *E. coli* GalU⁻ mutant has proved to be a useful means of isolating cDNAs encoding UDP-Glc PPase. Because it relies on expression of active enzyme, any cDNA clones with major 5'-truncations are eliminated. Using this method we have isolated previously a clone from a human liver cDNA library and here we report the detection of two forms from a human muscle cDNA library.

Comparing these sequences with those contained in the Gen-Bank and EMBL databases revealed a number of human expressed sequence tags that matched well with sections of our clones, and two of these (M85693 from brain and R58578 from heart) are almost identical to pHC379 over a region spanning the first 270 bases (Fig. 1C); thus, there is evidence for the expression of the muscle II form in at least two other tissues.

The 5'-end of pHC379 is very similar to that of the liver form that we have described previously (Peng and Chang, 1993). By contrast, the cDNA for the two muscle forms of UDP-Glc PPase are totally different before nucleotide 93 but from that point onwards, the sequences are identical. This is unlikely to occur if the two cDNAs arise from different genes and we suggest that they may result from the use of a different first exon in these two transcripts. A similar situation has been reported (Putt et al., 1993) to occur in the preceding enzyme of UDP-Glc formation, phosphoglucomutase.

It is of interest that there is a closer relationship between the amino acid sequences of bovine liver and human muscle forms (6 differences) than between the bovine liver and human liver forms (13 differences). In addition, the 3' untranslated region of the human muscle cDNA, but not the human liver form, exhibits significant similarity to the corresponding region of the bovine liver cDNA. For example, the 15 nucleotides immediately following the termination codon of the two transcripts are identical. Together, these data strongly suggest that duplication of the human UDP-Glc PPase gene into the liver and muscle isoforms has occurred before the diversification of the evolutionary lines leading to the human and bovine species.

Human liver UDP-Glc PPase shows complex kinetics with respect to all substrates except MgUTP. A possible explanation of the kinetic behaviour is given by the model shown as Fig. 6. This is an adaptation of the two-subunit Monod-Wyman-Changeux allosteric model with the added complexity that the catalytic rate constant depends on whether one or both active sites are occupied. Although the enzyme is represented as a dimer in Fig. 6, we acknowledge that the active form of the enzyme is probably the octamer (Levine et al., 1969). The rate equation for this model is given as Eqn (5) where r is the ratio k_1/k_2 . We wish to emphasise that this is only one possible interpretation of the curves; there is not enough information at present to be confident that this model is correct, although we note that it does describe the individual curves extremely well.

Some authors have reported that the kinetics of mammalian UDP-Glc PPase with respect to PP_i follow a sigmoidal saturation curve (Villar-Palasi and Larner, 1960; Knop and Hansen, 1970). This may be due partly to the failure to appreciate that the true substrate is MgPP_i, coupled with the use of a buffer that can complex magnesium ion. However, Steelman and Ebner (1966) described non-hyperbolic kinetics for the bovine mammary gland enzyme that cannot be explained simply as an experimental artefact. Thus it appears that non-hyperbolic kinetics similar to that which we have observed for the human liver and muscle II forms may be an intrinsic property of the enzyme, at least in some circumstances.

The metabolic significance of the kinetic anomalies that we have observed is unclear, but they tend to make the enzyme more responsive to changes in substrate concentration. For example, at a UDP-Glc concentration of 70 μ M, the liver enzyme goes from 20% to 80% of maximum activity over a 6-fold concentration range of MgPP_i, in contrast to muscle form I which requires a 16-fold range. We are conscious that the properties that we have measured are those of human enzymes expressed in *E. coli* and that post-translational modifications that could affect these properties may be different. Thus, we are reluctant to extrapolate these findings to the situation that might occur in human tissues.

The cDNAs that we have isolated have been designated as the liver, muscle I and muscle II forms but we emphasise that this indicates only the libraries from which they were isolated. For example, the 'liver' form could conceivably be the predominant one in muscle; the fact that we did not detect it in the muscle cDNA library may simply result from the relatively small number of positive clones that were observed and characterised. In addition, this 'liver' transcript may be preferentially degraded and therefore would not be detected by complementation. It has been shown previously by Northern analysis (Peng and Chang, 1993) that the UDP-Glc PPase cDNA is expressed at high levels in muscle but the probe used in these experiments was not specific enough to identify which of the three forms was present. Measurements of the relative abundance of the three cDNAs in various tissues will help to resolve whether any of them can correctly be described as muscle or liver forms.

It is not yet known whether the 'muscle' and 'liver' forms of the cDNA are derived from two independent genes or different alleles of a single gene but the diversity of their nucleotide and amino acid sequences favours the former possibility. Support for this hypothesis comes from the observation that two different loci for UDP-Glc PPase have been demonstrated and assigned to chromosomes 1 and 2 (van Someren et al., 1974; Shows et al., 1978). We have recently isolated genomic DNA corresponding to the muscle form of the enzyme and its genetic locus is currently being determined.

This work was supported in part by National Science Council of Republic of China (NSC83-0412-B182-061) and Chang-Gung College of Medicine and Technology (CMRP459) to HYC.

REFERENCES

- Aksamit, R. R. & Ebner, K. E. (1972) Purification, properties and kinetic analysis of UDP-glucose pyrophosphorylase from bovine mammary tissue, *Biochim. Biophys. Acta* 268, 102–112.
- Cullmann, G., Hubscher, U. & Berchtold, M. W. (1993) A reliable protocol for dsDNA and PCR product sequencing, *BioTechniques* 14, 578-579.
- Duggleby, R. G. (1984) Regression analysis of nonlinear Arrhenius plots: an empirical model and a computer program, *Comput. Biol. Med.* 14, 447-455.
- Duggleby, R. G. (1990) Pooling and comparing estimates from several experiments of a Michaelis constant for an enzyme, *Anal. Biochem.* 189, 84–87.
- Knop, J. K. & Hansen, R. G. (1970) Uridine diphosphate glucose pyrophosphorylase. IV. crystallization and properties of the enzyme from human liver, J. Biol. Chem. 245, 2499-2504.
- Levine, S., Gillett, T. A., Hageman, E. & Hansen, R. G. (1969) Uridine diphosphate glucose pyrophosphorylase. II. Polymeric and subunit structure, J. Biol. Chem. 244, 5729-5734.
- Peng, H.-L. & Chang, H.-Y. (1993) Cloning of a human liver UDPglucose pyrophosphorylase cDNA by complementation of the bacterial galU mutation, FEBS Lett. 329, 153-158.
- Putt, W., Ives, J. H., Hollyoake, M., Hopkinson, D. A., Whitehouse, D. B. & Edwards, Y. H. (1993) Phosphoglucomutase 1: a gene with two promotors and a duplicated first exon, *Biochem. J.* 296, 417– 422.
- Shows, T. B., Brown, J. A., Goggin, A. P., Haley, L. L., Byers, M. G. & Eddy, R. L. (1978) Assignment of a molecular form of UDP glucose pyrophosphorylase (UGPP₂) to chromosome 2 in man, Cytogenet. Cell Genet. 22, 215–218.
- Steelman, V. S. & Ebner, K. E. (1966) The enzymes of lactose biosynthesis. I. Purification and properties of UDPG pyrophosphorylase from bovine mammary tissue, *Biochim. Biophys. Acta* 128, 92–99.
- Tsuboi, K. K., Fukunaga, K. & Petricciani, J. C. (1969) Purification and specific kinetic properties of erythrocyte uridine diphosphate glucose pyrophosphorylase, J. Biol. Chem. 244, 1008–1015.
- Turnquist, R. L. & Hansen, R. G. (1973) Uridine diphosphoryl glucose pyrophosphorylase, in *The enzymes* (Boyer, P. D., ed.) vol. 8, pp. 51–71, Academic Press, New York.
- van Someren, H., Beijersbergen van Henegouwen, H., Westerveld, A. & Bootsma, D. (1974) Synteny of the human loci for fumarate hydratase and UDPG pyrophosphorylase with chromosome 1 markers in somatic cell hybrids, *Cytogenet. Cell Genet.* 13, 551–557.
- Villar-Palasi, C. & Larner, J. (1960) Uridinediphosphate glucose pyrophosphorylase from skeletal muscle, Arch. Biochem. Biophys. 86, 61-66.