

# Molecular cloning and characterization of a glucose transporter-related gene from Trypanosoma brucei

Bayele, Henry K.

*Award date:*
1992

*Awarding institution:*
University of Bath

# MOLECULAR CLONING AND CHARACTERIZATION OF A GLUCOSE TRANSPORTER-RELATED GENE FROM TRYPANOSOMA BRUCEI

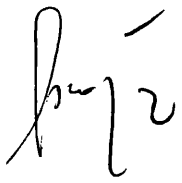
Submitted by **Henry K. Bayele**

for the degree of PhD of the

University of Bath, 1992

## COPYRIGHT





**Henry K. BAYELE**

**January 1992**





ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

**Dedicated to my parents**

**ACKNOWLEDGEMENTS**

I wish to thank my supervisors Dr. Robert Eisenthal, Dr. Paul Towner, and Dr. Geoffrey Holman for their support and encouragement, and for helping me find joy (and frustration!) in Biochemistry.

My special thanks go to the Wingate Foundation for a scholarship, and to the Sir Ernest Cassel Trust and the Diocese of Clifton for financial support. I should particularly thank Dr. Robert Eisenthal and Professor George Lunt in this regard, for pulling the strings when it mattered.

Dr. Adrian Wolstenholme has provided helpful hints and ideas, and "gleaned" the databases with my gene sequences when I needed him. Dr. Phil Harris helped to "change my first nappies" in gene cloning. My colleagues in the "Middle Lab" and "Transport House" were sometimes a "live-wire", especially Dr. Isabella Kozka and Ms Katina Constantinides. Mr. Andy Mayes has been very helpful, among other things, in processing this thesis. Thanks to Dr. Alison Fry for sharing information, and for mapping some genes.

Dr. Paul Hodgson of the Molecular Graphics Unit, Bath University, has been very helpful in sequence analysis, and I should thank him for providing some of the figures presented in this thesis.

My very sincere thanks go to all my friends, both within Biochemistry and without, for supporting me throughout these years- Rev. Brian McEvoy and especially Rev. Patrick Shanahan who has been very instrumental from the outset. To all who shared a smile and a word, thanks a million! I should say I would not have done much on my

own without the support and love of my family both here and abroad. Rebecca and Tony deserve my gratitude for understanding and appreciating my concept of Biochemistry-long and irregular working hours!

# CONTENTS

**CHAPTER 3**

CHAPTER 4

## SUMMARY

A lambda EMBL4 genomic clone was isolated from a trypanosome gene library using a 1500bp rat liver glucose transporter cDNA probe. Restriction mapping and Southern blot analysis identified an insert of about 2550bp. This was sequenced and found to have an open reading frame of 1731bp and encoding a protein with a predicted molecular weight of 63,346 daltons. An enriched plasmid library of trypanosome genomic fragments hybridizing to the cDNA, was also constructed and screened with the cDNA and with an oligonucleotide probe encompassing the putative glucose binding site of the GT superfamily. One of the clones isolated was identical in sequence to the lambda clone.

Analysis of the sequence showed that the gene contained a classical eukaryotic promoter core structure. The modular arrangement of the gene includes a TATA box and a CCAAT box, both of which are consistent with a RNA polII-transcribed gene. This is the first report of what appears to be a classical promoter in Kinetoplastid genes. The codon usage of this gene is consistent with that of trypanosome genes, and with that of a lowly or poorly expressed gene.

Run-off transcripts from the gene in pBluescript were used for in vitro translation in a rabbit reticulocyte lysate system. This yielded a protein of about 34-36kD on a SDS-PAGE fluorogram.

Northern blot analysis using the gene as probe on RNAs of the developmental forms of the parasite (bloodstream and insect/culture forms), indicated that it may be differentially expressed in the bloodstream form, a view that may be ascribed

to the location of a putative enhancer core upstream of the start of transcription. This motif is associated with developmentally regulated genes, serving as a hot-spot for gene conversion or rearrangement. This gene may therefore be "mobile".

At the amino acid, the protein bears little (global) homology to the glucose transporter superfamily. However, there is limited/segmental homology to the probes used to clone it, some of which correspond to conserved motifs in the superfamily. Hydrophobicity analysis indicated that it lacked the classical putative 12 membrane-spanning domains typical of the glucose transporters. This is attributed to a preponderance of charged groups in the sequence. The presence of potential N-linked glycosylation sites analogous to those in the GT family indicate that it may be a glycoprotein.

A run of regularly spaced arginine residues within the sequence was noted, and the implications for a channel or a coupled carrier are discussed within the context of similar observations in ligand- and voltage-gated channels.

On the basis of the diminished hydropathy profile and limited homology to the GTs, the cloned gene may not encode a glucose transporter. IF it does, it is substantially different from the rest of the family.

A comparison of the C-terminus of the protein with those of microbody/peroxisomal proteins identified a topogenic signal that is highly homologous, and may suggest that the protein is probably targeted to the glycosome. This may be credible on the basis of the similarities shared between this protein and other glycosomal proteins with respect to charges, insertions and deletions.

**ABBREVIATIONS**

| | |
|---|---|
| ATP | adenosine triphosphate |
| bp | base pair |
| BSA | bovine serum albumin |
| cpm | counts per minute |
| Ci | Curie |
| CTP | cytosine triphosphate |
| DEAE | diethylamino ethyl |
| DEPC | diethylpyrocarbonate |
| DMEM | Dulbecco's minimal essential medium |
| DNA | deoxyribonucleic acid |
| DNase | deoxyribonuclease |
| DTT | dithiothreitol |
| EDTA | ethylenediamine tetraacetic acid |
| EtBr | ethidium bromide |
| dNTP | deoxynucleoside triphosphate |
| GT | glucose transporter |
| GTP | guanosine triphosphate |
| $^{3}H$ | tritium |
| HCl | hydrochloric acid |
| HEPES | 4-(2-hydroxyethyl)-1-piperazine sulphonic acid |

| | |
|---|---|
| KAc | potassium acetate |
| KRPB | Krebs Ringer Phosphate Buffer |
| KRH | Krebs Ringer Hepes Buffer |
| $m^7G(5')ppp(5')G$ | $P^1$-5'-(7-methyl)-guanosine-$P^3$-5'-guanosine triphosphate |
| mM | millimolar |
| MOPS | 4-morpholine propane-sulphonic acid |
| mVAT | metacyclic variant antigen type |
| MEGA-10 | decanoyl-N-methylglucamide |
| PAGE | polyacrylamide gel electrophoresis |
| PARP | procyclic acidic repetitive protein |
| PBS | phosphate-buffered saline |
| PEG | polyethylene glycol |
| rRNA | ribosomal ribonucleic acid |
| RFLPs | restriction fragment length polymorphisms |
| RNA | ribonucleic acid |
| RNase | ribonuclease |
| SDS | sodium dodecyl sulphate |
| NAc | sodium acetate |
| SSC | standard saline citrate |
| TBE | tris-borate-EDTA |
| TEMED | N,N,N'N'-tetramethylethylenediamine |
| TE | tris-EDTA |

| | |
|---|---|
| Tglutrp | Trypanosome glucose transporter-related protein |
| Tris | tris(hydroxymethyl)aminomethane |
| TTP | thymidine triphosphate |
| vol | volume |
| VSG | variant surface glycoprotein |
| X-gal | 5-bromo-4-chloro-3-indolyl-B-galactopyranoside |

# CHAPTER 1

## Introduction

### 1.1 Overview of the trypanosomiases: aetiology and epidemiology

The two genera of the Trypanosomatidae which are of medical/economic importance to man are the Trypanosoma and the Leishmania. Within the Trypanosoma are the Stercoraria, which are transmitted through the excrement of reduviid bugs, and the Salivaria, which constitute the largest collection of trypanosomes and are transmitted prandially. T. cruzi, the most medically important Stercorarian, is the etiological agent of Chagas disease. Human infection is endemic throughout major regions of South and Central America between latitudes 40°N and 40°S (1). The chief vectors are the hemipterans Triatoma infestans, T. braziliensis and Rhodnius prolixus. The early stages of metacyclic T. cruzi infection are typified by conjunctivitis or a "chagoma" depending on their destination. This phase may be fatal due to the development of febrile paroxysms, lymphadenopathy and malaise. The subsequent phase is asymptomatic, irregular and protracted, while the chronic phase follows with cardiac and visceral involvement which invariably are enlarged. In the meantime, the host is immunocompromised as the parasite develops a predilection for macrophages.

Trypanosoma brucei, the agent of African trypanosomiasis or sleeping sickness, is a complex of various strains differing more in their virulence than in their life styles or morphology. The most significant difference between both forms of trypanosomiasis is the degree of severity. Whereas African trypanosomiasis is inevitably terminal in its prognosis, Chagas disease may remain chronic for a lifetime with no overt

symptoms despite occasional recrudescence of parasitaemia.

African trypanosomiasis is spread by the tsetse fly Glossina spp. Of this there are various strains that differ in the strain of trypanosome they transmit. The main vectors for T.b. rhodesiense, which causes the rhodesian or East African form of the disease, are G. morsitans, G. swynertonni and G. pallidipes. These inhabit the woodlands and savanna regions of East Africa. The gambian or West African form of the disease is caused by T.b. gambiense which is transmitted by G. fuscipes, G. palpalis and G. tachinoides. These inhabit riverine vegetation. The distribution of these vectors is inextricably linked to that of the two forms of the disease as depicted in the map below.

Needless to say, both diseases ravage large populations of the Third world and to date, are refractory to chemotherapy. They are two of the six major tropical diseases targeted by the WHO Special Programme for Research and Training in Tropical Diseases (TDR).

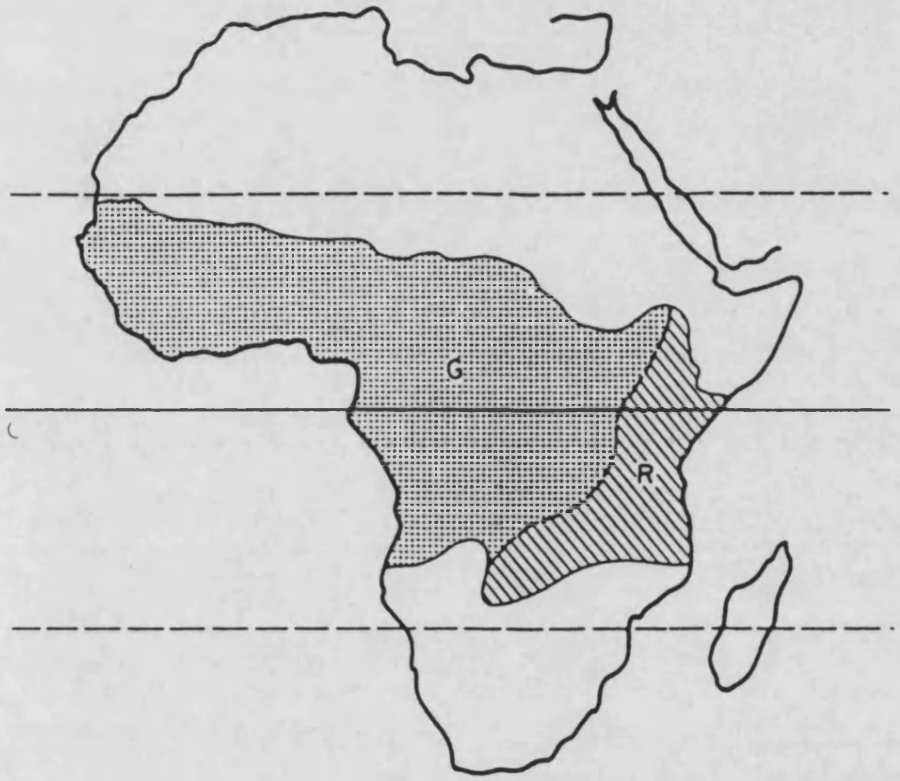



Fig. 1 Distribution of the African trypanosome and trypanosomiasis. The distribution of the Gambian (G) and Rhodesian (R) forms of the parasite and disease are as indicated.

## 1.2 Medical and economic importance

African sleeping sickness is rife in regions of great deprivation. The sociobiology of the tsetse fly is intimately linked to population movements in sub-Saharan Africa, and the economic geography of several countries has been affected by the disease. Its fatal nature means that not only human populations are at risk of elimination, but also large herds of cattle are lost through the animal form of the disease- nagana. Thousands of people and cattle are infected or at risk of infection each year. The loss of large reserves of fertile arable land as a result of population drift from tsetse fly infested zones, coupled with the debilitation of people and cattle, produces unmitigated misery in national economies. Malnutrition or under-nutrition as a result of the loss of cattle, a vital source of meat and milk, is an intractable sequel to trypanosome afflictions. The refractory nature of the disease, and the inability of national economies to deal with it due to the lack of proper surveillance, equipment or trained man-power, makes the plight of many real indeed. Even though large areas have been rendered tsetse-free, there remain residual foci of smoldering endemicity from which new cases may evolve sporadically.

The importance of this disease to mankind is reflected in concerted efforts worldwide to eliminate it at the level of the vector or through drug-targeting strategies in the parasite.

## 1.3 Basic biology of the African trypanosome

### 1.3.1 Identification and classification

The trypanosomes are dixenic parasites, infecting both vertebrates and invertebrates, and are characterized by flagella and kinetoplasts. In the bloodstream, the flagellum is the organ for motility but is modified for attachment in the insect vector. The characteristic morphological forms are the fusiform trypomastigotes, typically found in blood and tissues of their vertebrate hosts, and the epimastigotes in the invertebrate vector. They vary in shape and size from 15-100um in diameter. They have a typical unit plasma membrane approximately 2-4nm in width, which consists of dense osmiophilic layers on the periphery and separated by a clear layer (1). The flagellar pocket at the base of the flagellum is lined by the plasma membrane, and modified for food intake by pinocytosis.

Trypanosome classification has been based on a combination of phenotypic and genotypic features of the various strains. Similar structural features e.g the flagellum or the position of the kinetoplast may enable the identification of one strain with another by the cladistic or holistic approach. However, the subjective nature of this approach may be circumvented by applying molecular techniques that can delineate finer differences (2). It is therefore possible to separate trypanosomes into demes based on identical isoenzyme patterns or zymodemes (3), and similar restriction patterns of their kinetoplast DNA or schizodemes (4). Being the trypanosome equivalent of mitochondrial DNA in higher eukaryotes, the kinetoplast (mini-circle) DNA is suitable as a genotypic marker in determining closely related strains/demes, and for finger-printing individual strains.

Restriction profiles of the total genomic DNA may also identify polymorphs within a strain. This strategy may be made more sensitive if combined with DNA probes derived from a known strain. The polymerase chain reaction (PCR) is currently in vogue as a powerful tool to differentiate between single cells. It is also possible to use antibodies raised to key housekeeping proteins to classify trypanosomes. All these methods have their inherent shortfalls and may therefore be used to complement each other.

It is also conceptually possible to delineate inter-specific differences based on the sequences of small ribosomal RNA (rRNA) since they constitute the basic machinery for protein synthesis both in prokaryotes and eukaryotes. The primary and secondary structure, and the relative positions of the genes within the ribosomal cistron reflect phylogenetic distances between organisms (5,6). Comparative analysis based on the sequences of rRNAs indicated that protozoans in general diverged very early in evolution from a primordial ancestor (7).

| | M1 | M2 | M3 | M4 | M5 | M6 |
|---|---|---|---|---|---|---|
| T.cruzi vs T.brucei | 91.2 | 90.5 | 86.6 | 81.7 | - | 67.1 |
| T.cruzi vs Crithidia fasciculata | 94.0 | 91.9 | 89.5 | 89.0 | 89.3 | 70.9 |
| T.brucei vs C. fasciculata | 91.2 | 92.1 | 89.0 | 85.4 | - | 70.9 |

Fig. 2 Percentage homology between trypanosomal small ribosomal RNA subunits. Adapted from (8).

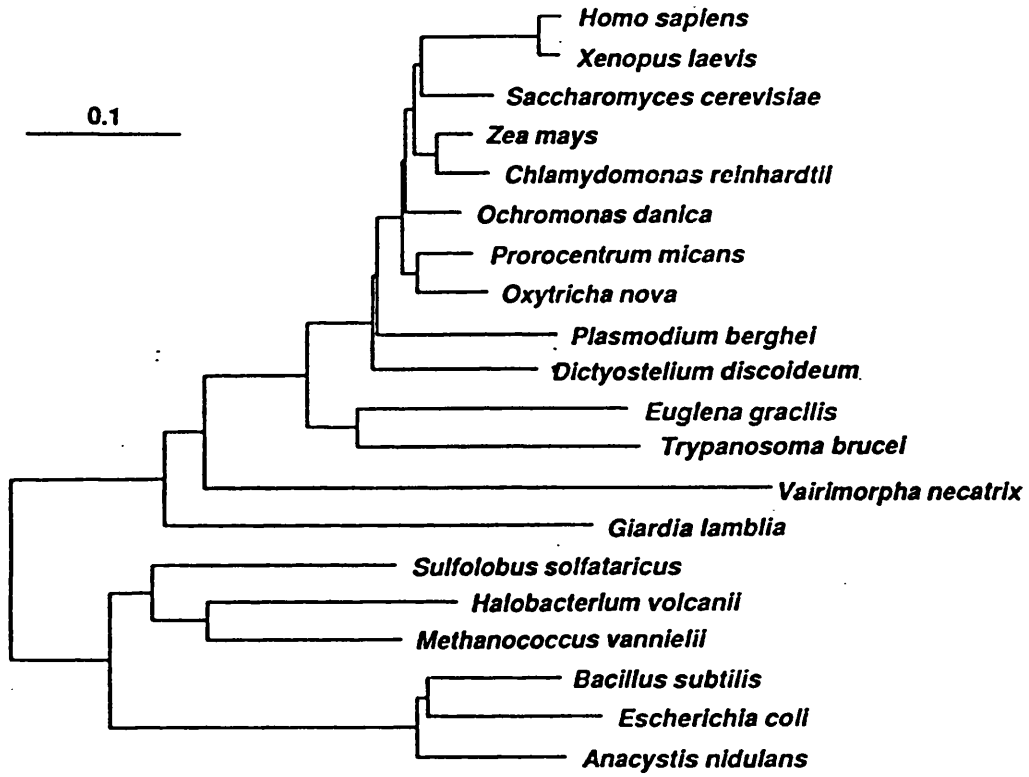



Fig. 3 Evolutionary relationship with other prokaryotes and eukaryotes inferred from identities in their small rRNA sequences. The bar represents 10 nucleotide substitutions per 100 residues. Reproduced with permission from Hyde, 1990 (9).

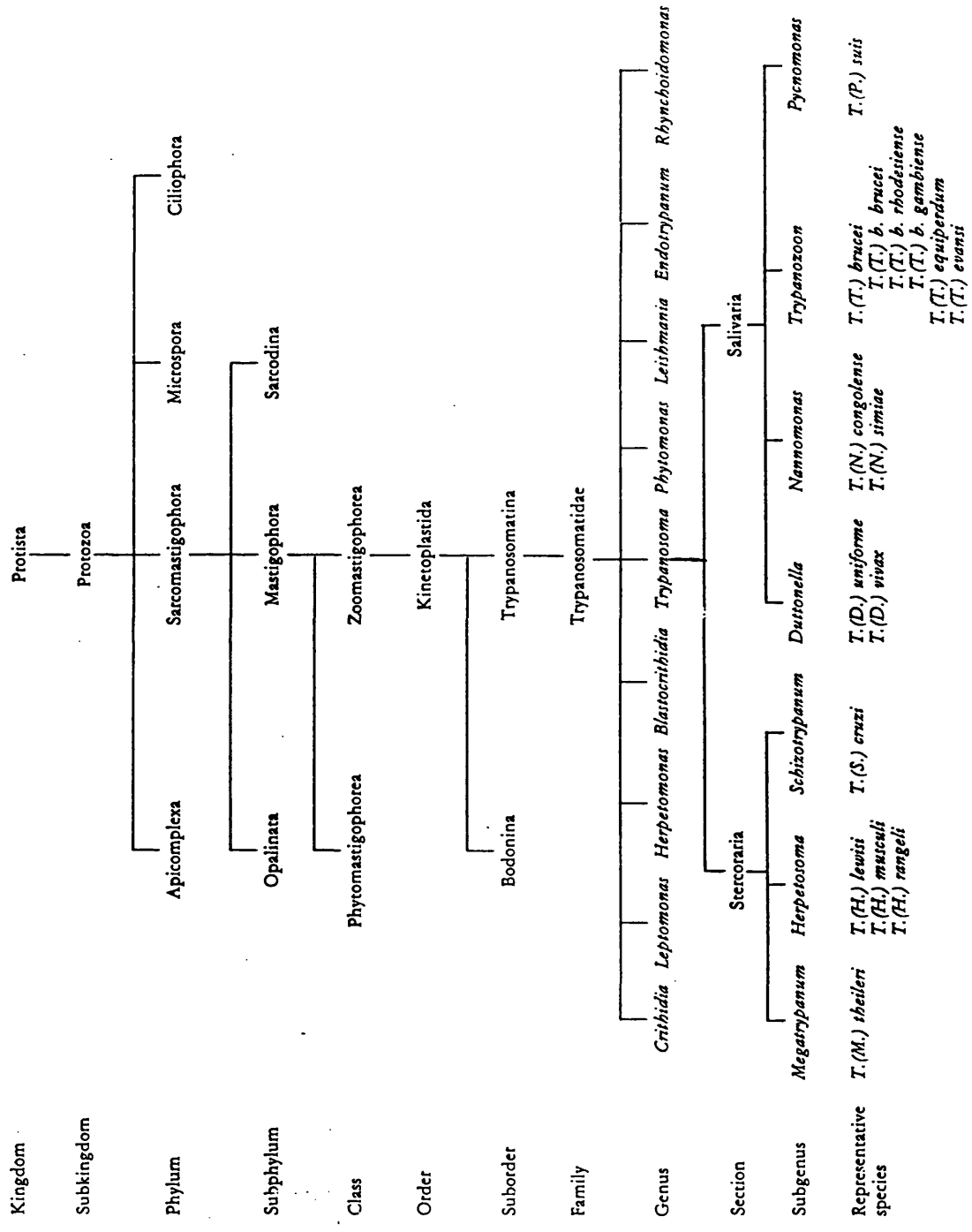



Fig. 4 Phylogenetic tree of the African trypanosome.

Adapted from (1).

### 1.3.2 Developmental patterns/cycles

Trypanosome development in the mammalian host begins shortly after a blood meal by the tsetse fly. In the immediate vicinity of the bite, the trypomastigotes proliferate causing an induration or trypanosomal chancre. From here, they are disseminated throughout the body via the lymphatic system. The first level of multiplication in the bloodstream causes a wave of parasitaemia that results in febrile paroxysms. Host antibody pressure results in a temporary remission of parasitaemia with short stumpy forms predominating (1,10). The surviving forms or heterotypes are long and slender and these reacquire a surface coat that confers immune protection for a further cycle of differentiation to intermediate and stumpy forms. These are the pre-adaptive forms (procyclics) for survival in the insect midgut after the next feed. These transform rapidly into epimastigotes and migrate to the salivary glands, attaching themselves to the microvilli of the gland. At 15-35 days after infection, they are transformed into the metacyclics and the fly is infective once again to the vertebrate host. Injected trypanosomes then again reproduce by binary fission to continue the cycle.

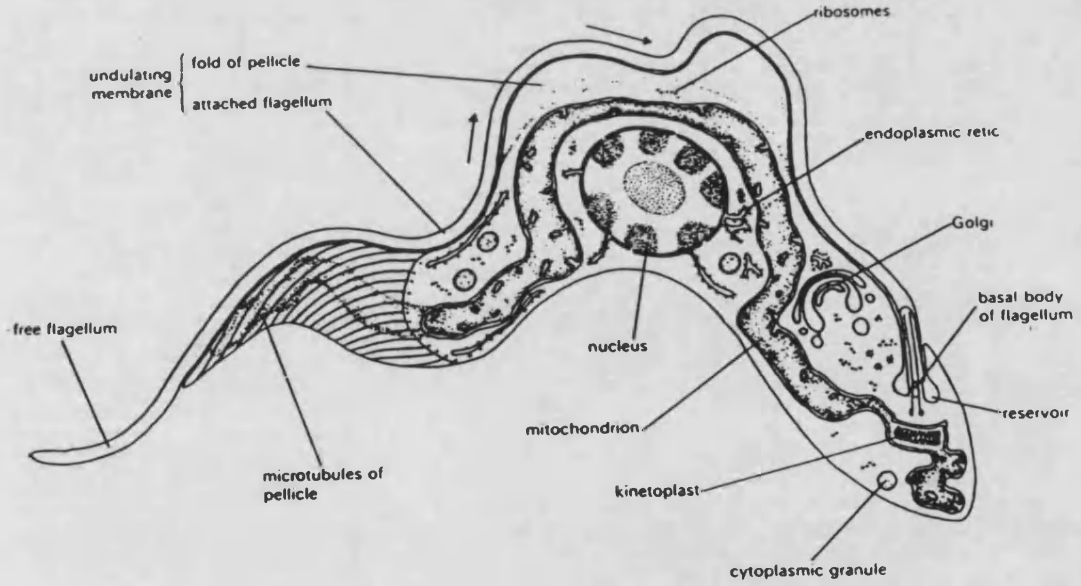



Fig. 5 Ultrastructure of the bloodstream form of the African trypanosome. Organelles are as labeled. The arrows indicate the direction of undulation or propulsion. Adapted from (1). Insert is a trypanosome among red blood cells.

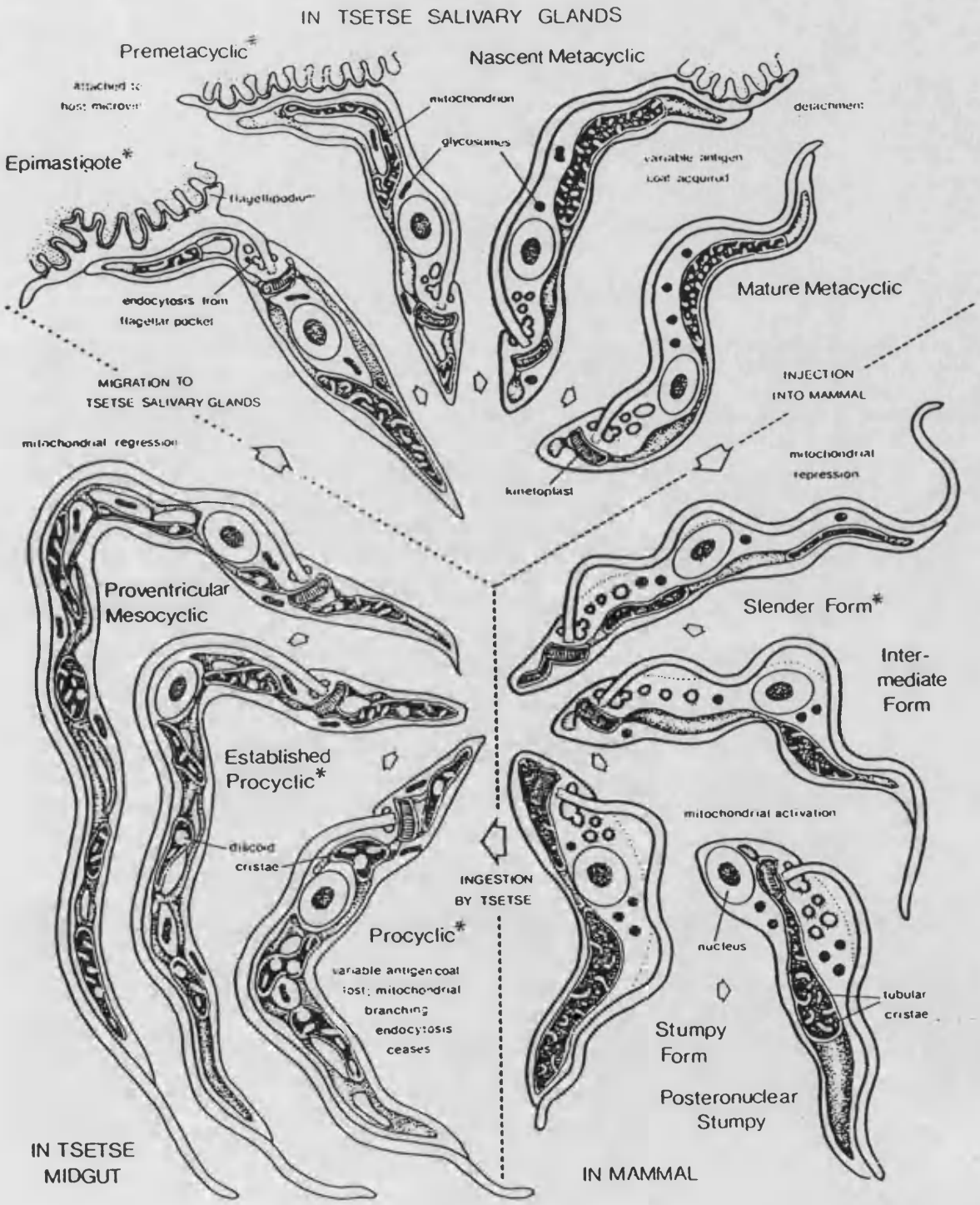



Fig. 6 Life cycle of the African trypanosome. Adapted from (10).

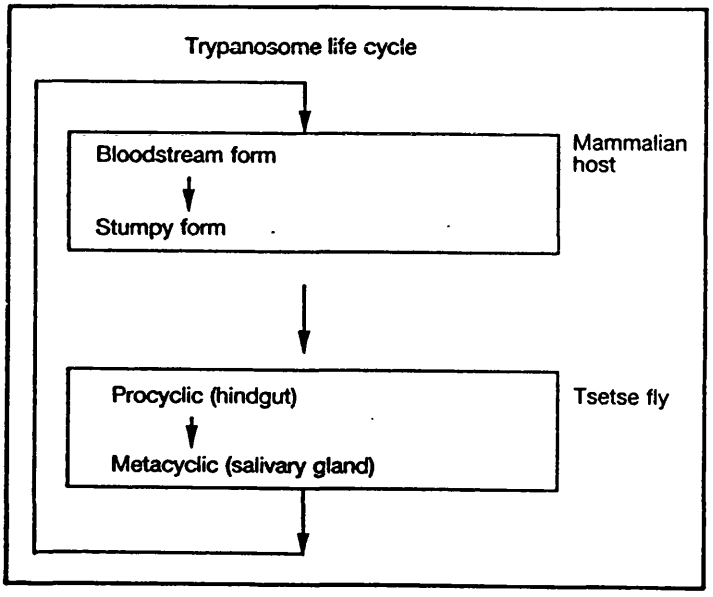



Fig. 7 Schematic of the trypanosome life-cycle. Adapted from (11).

### 1.3.3 Biochemical and molecular correlates of trypanosome biology

#### 1.3.3.1 Energy metabolism

The digenetic life cycle of the African trypanosome engenders a switching mechanism for amino acid and glucose dependence for its energy requirements, generating 2ATP and 2 pyruvate molecules per glucose molecule under aerobic conditions (12-16). The relatively inefficient energy production is compensated for by a high rate of glycolysis consequent upon a high rate of glucose transport (17), in order to support among other things, the rapid cell division which occurs in every 6-8 hours. The procyclic forms depend on proline for most of their energy requirements in the insect vector, which

has about 150mM amounts of this substrate in its haemolymph. These forms have a fully functional mitochondrion and a tricarboxylic acid (TCA) cycle, as shown by high levels of activity of proline, succinate, and a-ketoglutarate oxidases (10,18). Evidence from this accrues from the observation of an electromotive force (EMF) across the inner mitochondrial membrane of the transitional forms and the accumulation of rhodamine 123 (19). This EMF, which emanates from the NADH dehydrogenase site of oxidative phosphorylation, was inhibited by the classical inhibitors of the cytochrome mediated pathway of energy metabolism viz 2,4-dinitrophenol, rotenone, and salicylhydroxamic acid (SHAM) but not by antimycin A, cyanides or oligomycin. On this basis the existence of a functional $F_1F_0$ ATPase was possible. This was corroborated by the work of Williams et al (20), which illustrated a developmentally regulated expression of the ATPase. Western blot analysis using $F_0$ and $F_1$ antibodies showed elevated expression in the procyclics relative to the bloodstream forms. Additional evidence comes from the observation that there was a 30-fold lower steady state levels of the 9S and 12S mitochondrial ribosomal RNAs in the long-slender forms relative to the culture forms (21). Transcripts of cytochrome b and cytochrome oxidase subunits I and II were elevated 100-fold in stumpy forms compared to slender forms, heralding the onset of the development of the oxidative pathway of substrate phosphorylation (21). Further molecular evidence for the developmental regulation of energy metabolism comes from the observation that aldolase activity was decreased about 30-fold in the procyclics (22) and pyruvate kinase mRNA was 10X more abundant in the bloodstream forms (23). In the latter, the glycosomal form of phosphoglycerate kinase was 9X more abundant than

the cytoplasmic form whereas the converse was true in the procyclics (24) with an overall decrease in activity of about 2.5-fold (25).

The bloodstream forms, on the other hand, have a vestigial mitochondrion and a cyanide insensitive cytochrome system. They are therefore entirely dependent on substrate level phosphorylation/glycolysis for their energy supply. The metabolic processes involved in this are confined to the glycosomes, which are micro-organelles typical of the Kinetoplastids (13). They measure about 0.3um in diameter, and exist in groups of 200-300 per cell, thus constituting about 4% of cell volume and of protein content. A myriad of other functions have been ascribed to the glycosomes as indicated below. The first seven enzymes of the glycolytic pathway are contained within these organelles, in addition to the two involved in glycerol metabolism and the re-oxidation of NADH. These (seven) enzymes are tightly packed together, ensuring minimal transit time between metabolite intermediary states and the next and optimal local concentration of both enzymes and substrates, thus accounting for the high rates of energy metabolism or glycolytic flux in these organisms. This, by implication, means that none of these steps could possibly be the rate limiting step in glycolysis, as has been confirmed by pulse chase experiments (26).

| Metabolic pathway | Enzyme(s) | Organism | | | |
|---|---|---|---|---|---|
| | | Tb | Tc | L | Cl/Cf |
| Glycolysis | Hexokinase | + | + | + | +/ |
| | Phosphoglucose isomerase | + | | + | +/ |
| | Phosphofructokinase | + | | + | +/ |
| | Aldolase | + | | + | +/ |
| | Triosephosphate isomerase | + | | + | +/ |
| | Glyceraldehyde phosphate dehydrogenase | + | | + | +/ |
| | Phosphoglycerate kinase | + | | + | +/ |
| Glycerol metabolism | Glycerol 3-P dehydrogenase | + | | + | +/ |
| | Glycerol kinase | + | | + | |
| CO2 fixation | PEP carboxykinase | + | + | + | +/ |
| | Malate dehydrogenase | + | | | |
| Pyrimidine | Orotate phosphoribosyl transferase | + | + | + | /+ |
| | Orotidylate 5'-phosphate decarboxylase | + | + | + | /+ |
| Purine salvage | Hypoxanthine guanine phosphoribosyl transferase | + | + | + | |
| Ether lipid synthesis | DHAP acyltransferase | + | | + | |
| | Acyl/alkyl/DHAP reductase | + | | | |
| | Acyl CoA reductase | + | | | |
| Oxidation of fatty acids | Palmitoyl CoA synthetase | | | + | |
| | B-OH-butyrate CoA dehydrogenase | | | + | |

| | |
|---|---|
| Catalase | +/ |
| Phosphomannose isomerase | + |

Fig. 8 Integration and control of trypanosome metabolism: role of the glycosome. Tb (T. brucei); Tc (T. cruzi); Cl (Crithidia luciliae); L (Leishmania spp) and Cf (C.fascisculata). Adapted from (27).

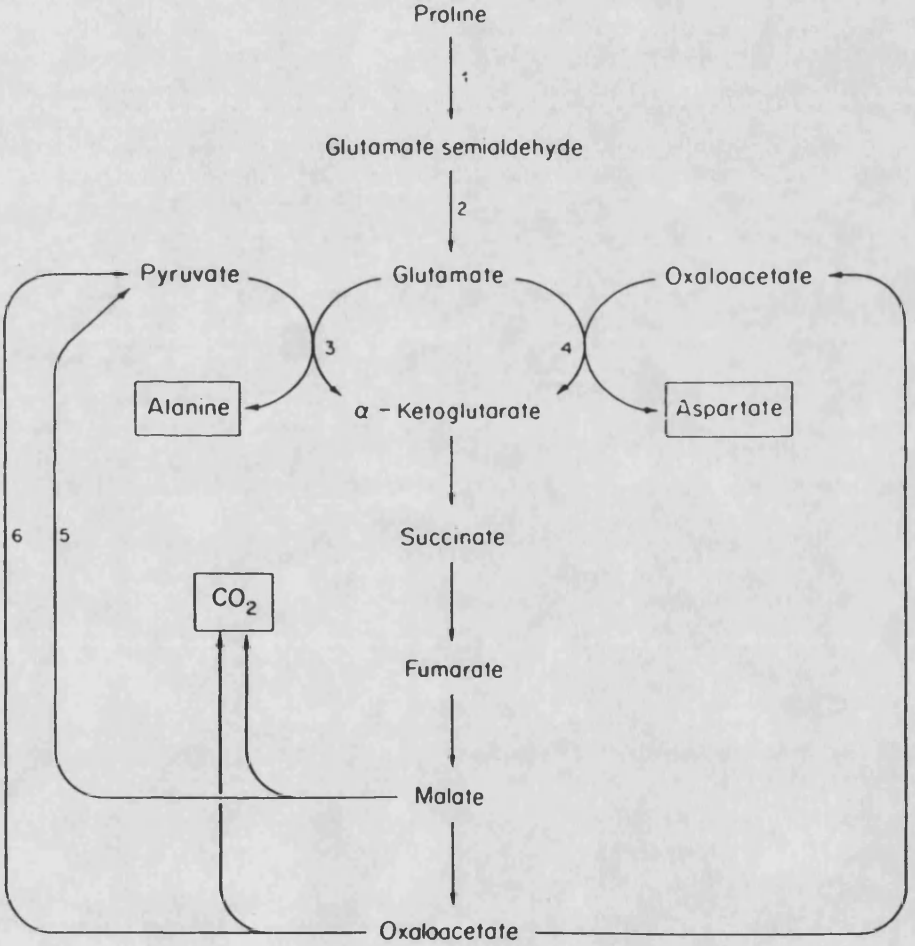



Fig. 9 Catabolism of proline in the culture forms of T. brucei. End products are boxed. Enzymes: 1, L-proline oxidase; 2, glutamate semi-aldehye dehydrogenase; 3, alanine aminotransferase; 4, aspartate aminotransferase; 5, malic enzyme; 6, pyruvate carboxylase. Adapted from (28).

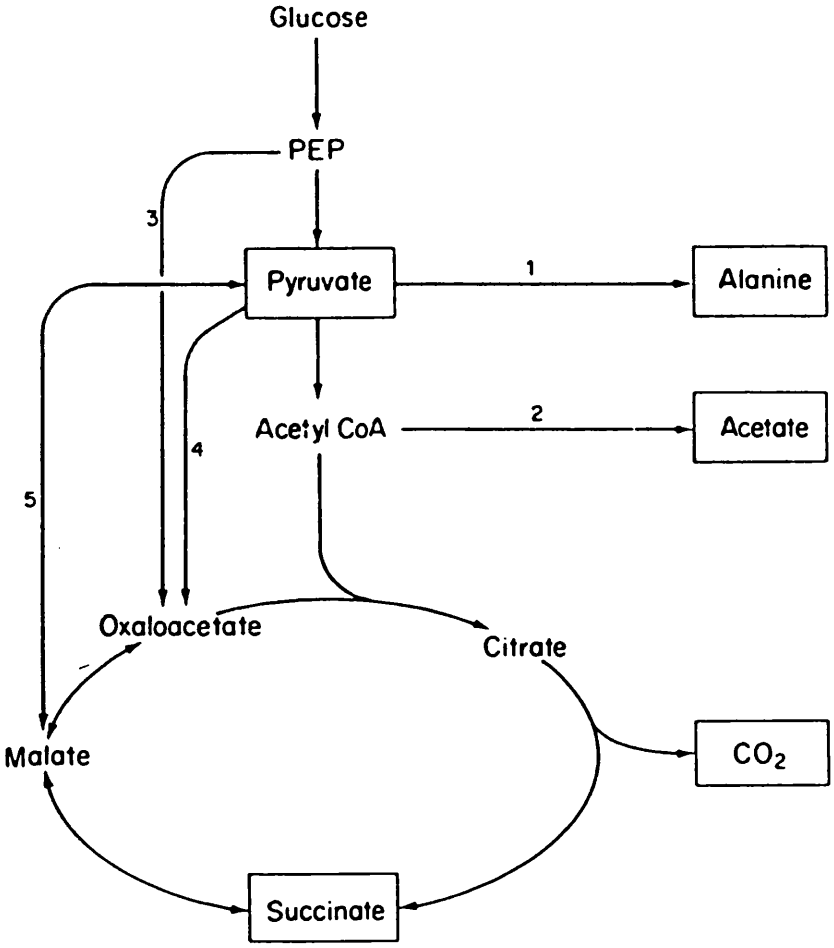



Fig. 10 Catabolism of glucose in the culture forms of T. brucei. End products are boxed. Enzymes: 1, alanine aminotransferase; 2, carnitine acyltransferase; 3, PEP carboxykinase; 4,pyruvate carboxylase; 5, malic enzyme. Adapted from (28).

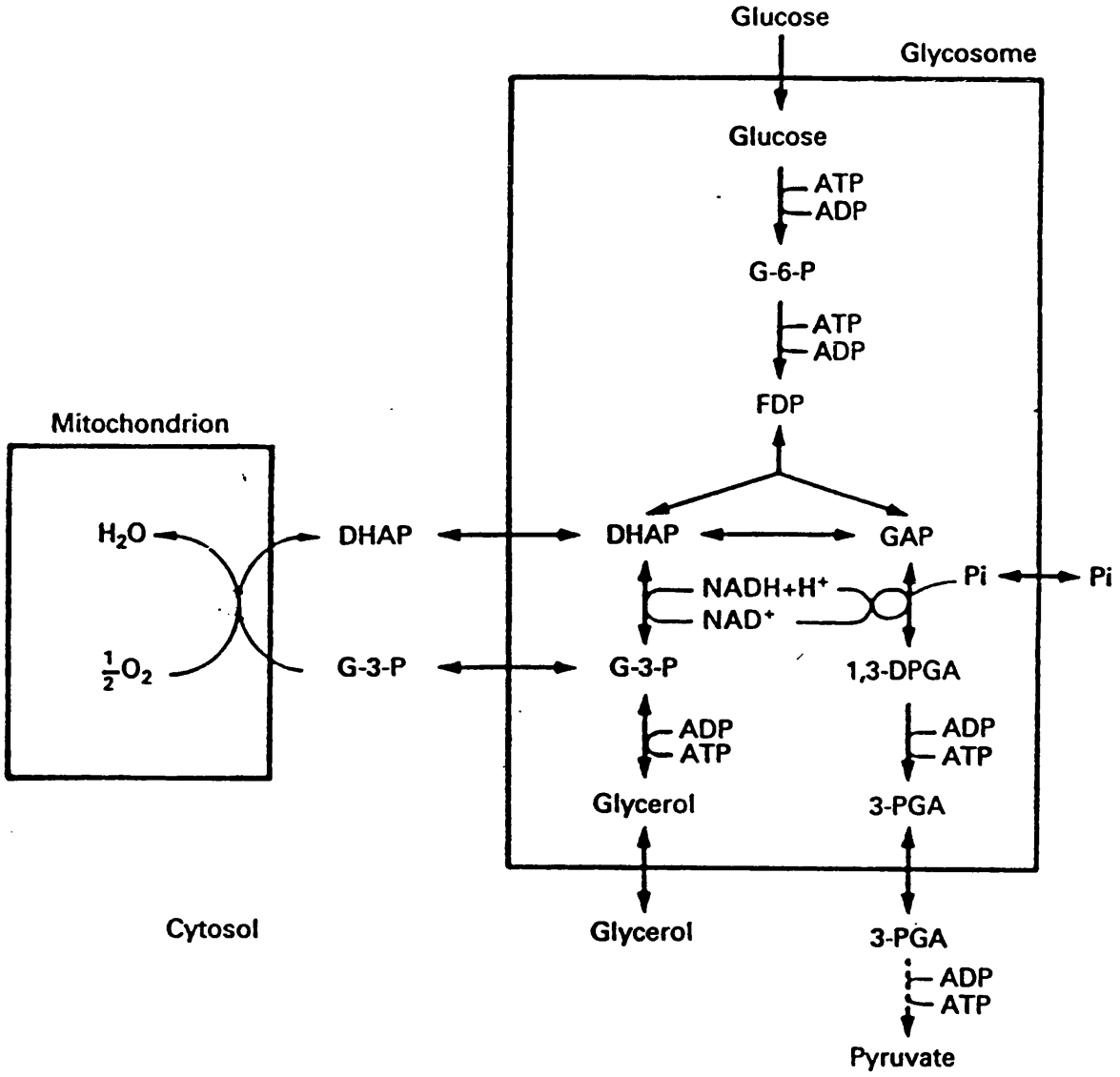



Fig. 11 Compartmentalization of glycolysis and glycolytic intermediates in the bloodstream form. Abbreviations: G-6-P, glucose-6-phosphate; FDP, fructose-1,6-bisphosphate; DHAP, dihydroxyacetone phosphate; G-3-P, glycerol-3-phosphate; 1,3-DPGA, 1,3-diphosphoglycerate; 3-PGA, 3-phosphoglycerate; Pi, inorganic phosphate. Reproduced from (16).

### 1.3.3.2 Genomic organization and housekeeping genes

Like most eukaryotic genomes, that of T. brucei comprises a significant percentage (32%) of repetitive sequences. The remaining 68% is made up of single copy gene sequences (29). It has been found that the trypanosome is diploid for all its house-keeping genes, i.e there are two copies for every house-keeping gene (30). Some of these genes are in tandem linkage or are found in multiple copies e.g the genes for aldolase (22) and glyceraldehyde phosphate dehydrogenase (31) are duplicated, whereas those of phosphoglycerate kinase (32) and calmodulin exist as three copies (33). Others such as the VSG genes which exist in about 1000 copies, are highly clustered in the genome (34) much like the a- and b-tubulin genes which are in tight tandem linkage of 15 alternating alleles for these forms (35). The occurrence of several copies of the same gene may be linked to the gene regulation machinery of the trypanosome in consonance with its life cycle. By appropriately shifting copies of a gene to an expression locus, the synthesis of gene products may be modulated at the transcriptional level.

| Gene | Number of genes per cluster | Number of clusters per cell | Indications for polycistronic transcription |
|---|---|---|---|
| RNA pol I | 1 | 2 | |
| RNA pol II | 1 | 4 | |
| RNA pol III | 1 | 2 | |
| ODC | 1 | 2 | |
| g PGI | 1 | 2 | |
| g TIM | 1 | 2 | |
| c GAPDH | 1 | 2 | |
| g GAPDH | 2 | 2 | + |
| g ALD | 2 | 2 | + |
| c PYK | 2 | 2 | |
| g+c PGK | 3×1 | 2 | + |
| PARP | 2 | 4 | + |
| Actin | 2–4 | 2 | + |
| Calmodulin | 3 | 2 | + |
| hsp 70 | 5 | 2 | + |
| $\alpha/\beta$ tubulin | 15 | 2 | + |
| Cysteine proteinase | ~20 | 2 | |

Fig. 12 Organization of genes in T. brucei. Adapted from (40).

### 1.3.3.3 Ploidy

The use of schizodeme and zymodeme analyses and RFLPs have not only been invaluable in strain identification and classification, but also provided information regarding the possible mechanisms of gene exchange in the African trypanosome. These methods have established mating in the metacyclic forms, which follows classical diploid Mendelian inheritance (36,37).

The karyotype of T. brucei as analyzed by pulse field gel electrophoresis has shown a spectrum of chromosomes ranging in size from 30-150kb for the

minichromosomes to over 2000kb making up a total of 20 bands encompassing the whole genome of the parasite (38). There are about a hundred minichromosomes in band one and these may be involved in conferring antigenic expression plasticity; 5-7 moderately sized chromosomes (200-700kb) and five others of about 2000kb (39). There are significant numerical and morphological differences of chromosomes from one deme to another within the Trypanosomatids due primarily to chromosomal rearrangements and suggestive of an underlying mechanism for gene exchange or expression.

#### 1.3.3.4 Gene structure

##### 1.3.3.4.1 Codon usage

The degeneracy of the universal genetic code coupled with the number of transfer RNA (tRNA) molecules available per codon places restrictions on codon choice and usage. Whilst it may be said that there is a canalised pattern of codon usage for some genes, there can still be found a broad spectrum of codon preferences in the trypanosome. It has been observed that most or all housekeeping genes have no preferences in terms of the overall nucleotide composition, although they have a predilection for C as the third base of their codons and a strong bias against A in this position (40). However, there is no clear bias in the codon usage of VSG genes and expression site-associated genes or ESAGs. Highly expressed genes are the most biased in their use of codons and are predominantly GC-rich. Lowly expressed ones use a greater fraction of the rare codons but still use the common ones as well (41).

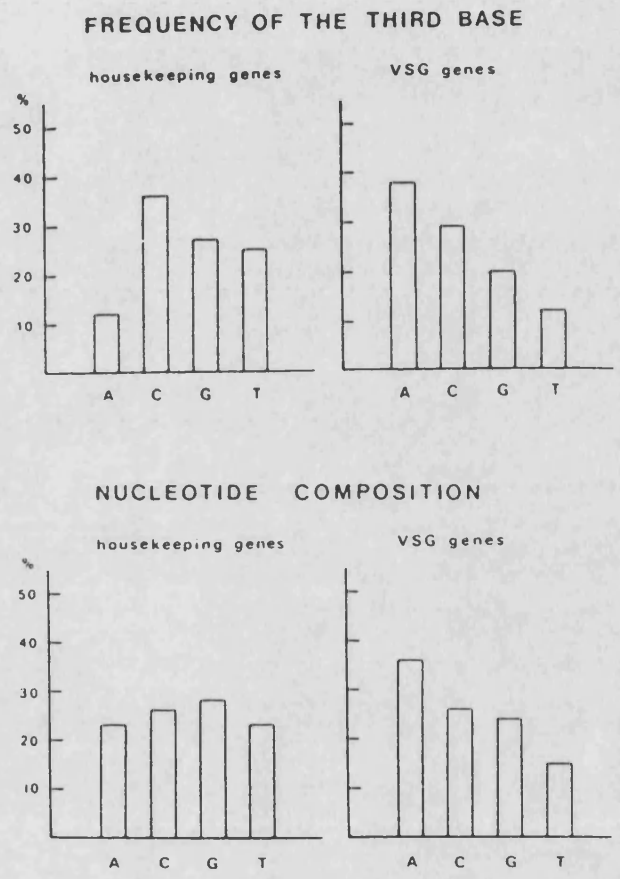



Fig. 13 The frequency of the third base of the codons used in eight housekeeping genes and three VSG genes of T. brucei, and the overall nucleotide composition of those genes. Reproduced from (40).

### 1.3.3.4.2 Nature of flanking sequences/regulatory elements

Classical canonical gene regulatory sequences viz CCAAT and TATA boxes which are associated with promoter activity in most eukarytoic genes (11,42), are absent from all trypanosome genes analyzed so far, and neither do they have any consensus polyadenylation signals (AAUAAA) (43). However, deletion mapping of a VSG gene with the aid of chloramphenicol acetyltransferase (CAT) reporter activity, identified a putative promoter within 90bp upstream from the transcription site (44). This promoter is markedly different from the promoters of T. brucei ribosomal RNA (45) and procyclin genes (46), which are transcribed in a similar fashion by an a-amanitin resistant RNA polymerase 1. Sequences 3' to this promoter appeared essential for expression in a stage-specific manner as shown by expression using chimaeric constructs of VSG and procyclin genes. The result indicated a decrease in activity in procyclics compared to bloodstream forms. The PARP promoter is presumed to be located between 200bp upstream, and the presumptive 3'-splice acceptor site (46).

For a majority of eukaryotic genes, the most important determinant(s) of translation are the cap sites, and the sequences immediately upstream and downstream of the start codon AUG (47,48). A consensus sequence was compiled from about 211 eukaryotic genes (49) as CC(A,G)CCATGG, positions -3 and +4 being the most conserved bases. This compares favorably with the consensus derived from some protozoan genes below. The G position at +4 was suggested to base pair with the fourth base in the anticodon of Met-tRNA (50,51) for efficient translation.

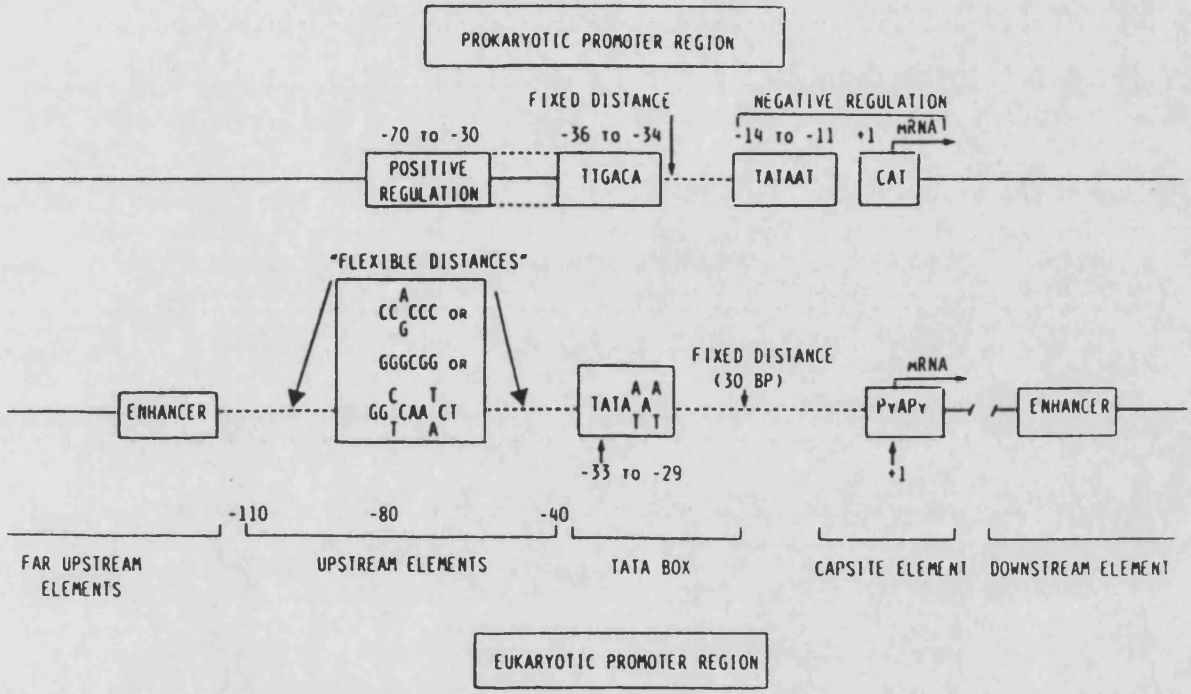



Fig. 14 Structural elements of prokaryotic and eukaryotic genes. Promoter, enhancer and cap sites are boxed.

| Trypanosomabrucei | | Reference |
|---|---|---|
| VSG | GGAGCGACTACAATGGA | (52) |
| Calmodulin | ACTTGATTTACGATGCC | (33) |
| PARP | GTAAAATTCACAATGGC | (53) |
| Aldolase | ACTGCAACGAAGATGTC | (22) |
| Heat shock protein 70 | CCTCTTTGAAGGATGAC | (54) |
| Trypanothione reductase | AATCGCTTTTCTATGTC | (55) |
| Leishmania enrietti | | |
| Membrane transport protein | TTCACTAGAATCATGAG | (56) |
| Plasmodium falciparum | | |
| Multidrug resistance protein | TTGTGTTGAAAGATGGG | (57) |
| Glycophorin binding protein | TTTTGTGTAAATATGCG | (58) |
| | -7 -3 +1 +5 | |
| CONSENSUS FOR PROTOZOAN NUCLEAR GENES | TT--A--ATG-G/C | |

Fig. 15 Flanking nucleotide sequences in some protozoan genes. The first base of the coding region is indicated by +1.

### 1.3.4 Variant surface glycoproteins (VSGs)

#### 1.3.4.1 Nature and structure of VSGs

The differential expression of VSGs or antigenic variation is the basis for the survival of the African trypanosome within the mammalian host (59,60). Each VSG is made up of 450-500 amino acids [50-60 kd] with a 7-17% carbohydrate content. This forms a monolayer of homodimers on the surface of the parasite of about 12-15 nm thickness and 10 million molecules per cell (61-63). There are two domains: the N-terminal domain of about 400 amino acids which constitutes the variable region and which is distinct for each variant antigenic type (VAT). This is the region also presented to host antibodies. The C-terminus is relatively more conserved with about 110 amino acid residues. The entire VSG layer is sequestered to the membrane via a hydrophobic glycophosphatidylinositol (GPI) anchor at the C-terminus. The fatty acid moiety of the phospholipid is myristic acid [C14:0] (64). The nascent VSG has a 20-40 amino acid signal sequence which aids in sorting and insertion into the membrane, the latter function being more the prerogative of the 17-23 residue extension at the C-terminus. Both the N- and C- terminal extensions are cleaved off post-translationally and after the incorporation of carbohydrate moieties to form the mature glycoprotein.

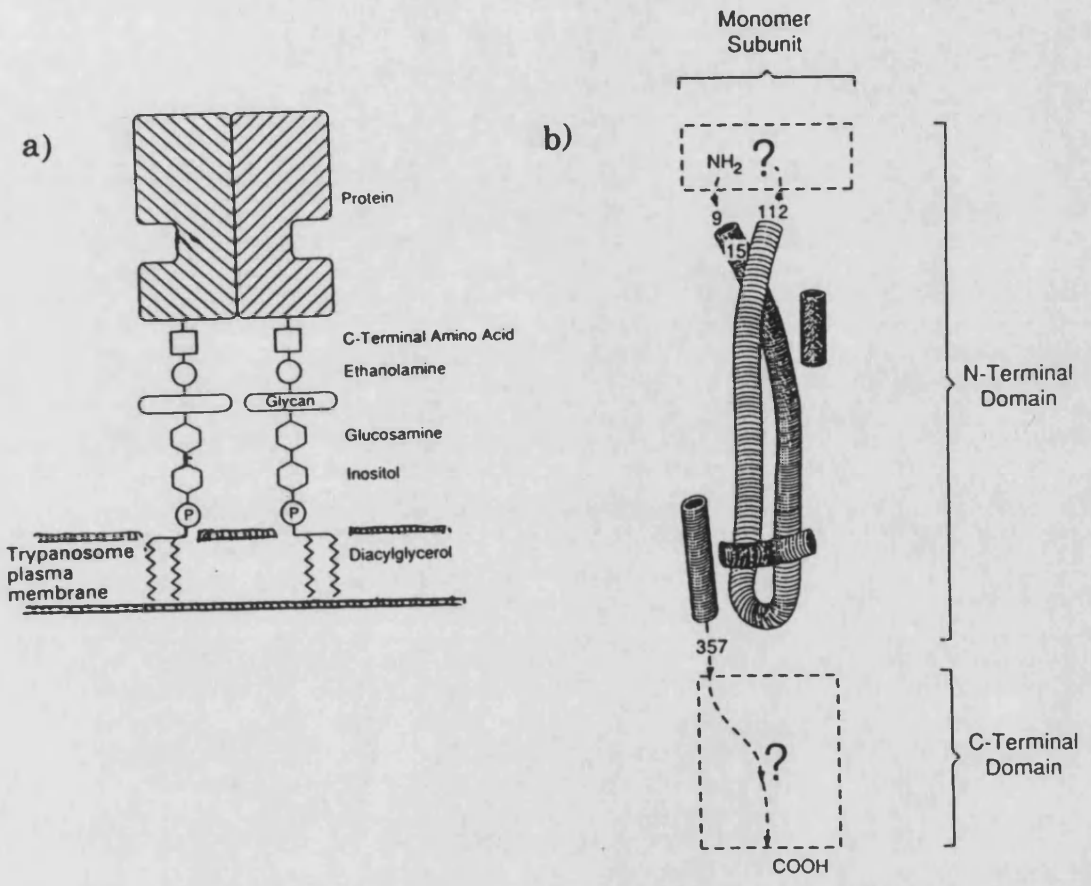



Fig. 16 Structure of a VSG in T. brucei. (a) Schematic representation of the trypanosome membrane. The shaded sections represent the protein portion of the molecule to which is attached a glycan containing a Cross Reactive Determinant, and a diacylglycerol anchor by which the whole molecule is sequestered to the plasma membrane. (b) Sketch of the subunit structure of the VSG. Alpha helices are represented by cylinders. Reproduced from (65).

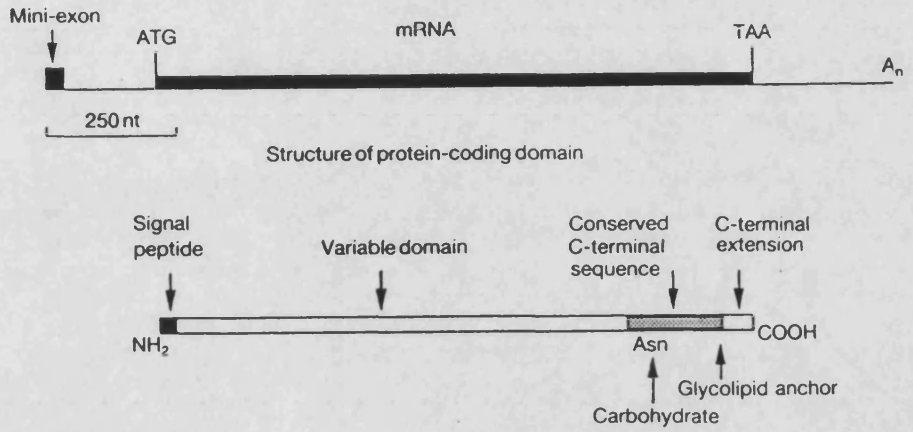



Fig. 17 VSG gene organization showing the mini-exon and structure of the protein-coding domain. Reproduced from (11).

### 1.3.4.2 Antigenic variation

In a typical clonal infection with trypanosomes, the course of the disease is presented in a series of remissions and recrudescences, which are reflected in the levels of parasitaemia. The initial rise in parasitaemia following an infection is followed by a remission, consequent upon the release of antibody by the host and specific to each VAT (66). However, because the parasite has the ability to alter its surface coat, it is able to evade host immune attack by expressing a serologically distinct antigenic coat (10,60,67). In this way, different immunologically distinct trypanosome populations appear at every

7-10 day interval. Peak parasitaemia may reach antilog 8-9 cells/ml of blood, and at a switch rate of $10^{-6}$-$10^{-7}$, spontaneous recovery almost never occurs (63,68-70). As a result of such a sinusoidal relationship between parasite load and host response, infections are protracted. This is a teleological adaptation of the parasite to ensure its survival without having to eliminate its host prematurely. This relationship can only continue for as long as the host can respond immunologically to the antigenic repertoire of the parasite. However, because of the plastic nature of this repertoire involving about 100 VSGs (71), the host humoral pressure cannot sustain it and as such the host becomes immunocompromised before the trypanosome runs out of antigenic repertoire.

This phenomenon of antigenic variation is common to the Salivaria, and is mediated by the sequential expression of the extensive series of VSGs. It has not been observed in other trypanosomatids like T. cruzi, which expresses a VSG-like glycoprotein in the trypomastigotes (72), or the Leishmania which have evolved alternative defence mechanisms against host immune pressure by dwelling within the macrophages (1). In Babesia and malaria infections, the antigenic variation which has been reported is attributed to antigens located on the surface of the erythrocyte membrane (73). This has only recently been shown to occur in P. falciparum (74). Other protozoa such as Paramecium caudatum and Tetrahymena also exhibit antigenic variation, though the functional significance of this has not been established. Antigenic variation in the latter has been thought to occur through a series of repressor and derepressor genes (75).

Inspite of the versatility of the trypanosome to change its surface coat, the pattern of antigenic variation is not completely random, given the size of the repertoire.

All VATs have an intrinsic tendency to produce variable antigens in a sequential yet unpredictable pattern, and to revert after each cyclical transmission to one and the same VAT i.e the parent antigenic type of the strain, thus resetting the repertoire (76,77). This suggests that re-expression of early VATs occurs continuously in a chronic infection, these being almost inevitably eliminated by host antibody. The reversion to a common basic antigenic type, regardless of the nature of their surface antigens at the time of infection, held hopes of a vaccine from inactivated trypanosomes carrying the basic antigens. However, it has been shown that there is a great deal of heterogeneity among metacylic trypanosomes- the infective forms (78,79).

The kinetics of appearance of new VATs suggests that population survival is at the expense of 99.9% trypanosome destruction (69,80). The maximum number of different VATs observed in a clonal infection is 101 for T. equiperdum (81).

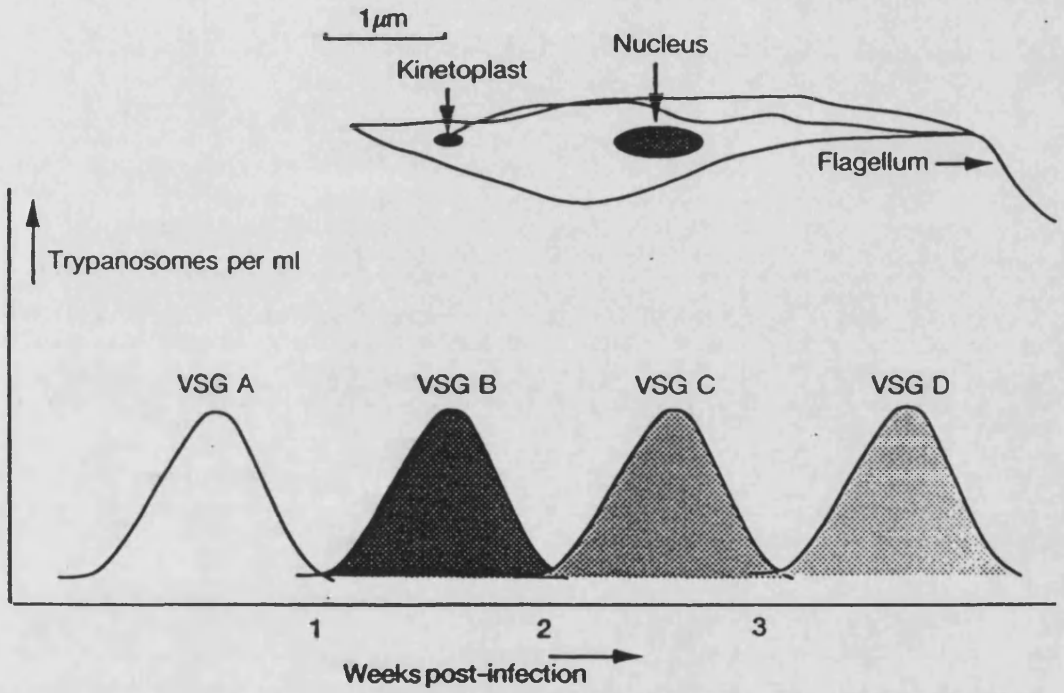



Fig. 18 Schematic representation of antigenic variation. The sequence of appearance of the VATs is represented by the VSGs expressed in a time series, VSG A-D. The peaks may represent parasitaemias of antilog 9. A simplified model of the trypanosome is inserted. Reproduced from (11).

### 1.3.4.3 Modulation of VSG expression

As earlier indicated, the ability of the African trypanosome to switch coats is the function of the repertoire of about 1000 VSG-encoding nuclear genes (34,82). The temporal expression of these genes is mutually exclusive both at the levels of transcription and antigenic presentation. Their products (VSGs) are immunologically distinct and unique even though mosaic coats have been observed (83). Switching may be facilitated by cleavage by phospholipase C thereby promoting VSG turnover during antigenic switch. Although the rate or frequency of switching is very low, it is enough to maintain a chronic infection indefinitely, during which there is a preferential predetermined expression of early and late VATS.

Expression of these antigenic variants is developmentally integrated and regulated. The insect mid-gut forms or procyclics assume a unique coat protein, procyclin, upon differentiation. The expression of this coat is regulated by TCA cycle intermediates and a temperature shift to 27°C (84,85). Procyclin genes comprise a polymorphic gene family found in two genetic loci A and B and in two tandem copies per locus, alpha and beta (53). As indicated in Fig.12, there are eight PARP genes encoded at two A loci, one B1 locus and one B2 locus distributed among four different chromosomes. The encoded protein [PARP], is not expressed in the bloodstream form but is rapidly induced upon differentiation to procyclics (86). Like the VSGs, transcription of PARP genes is insensitive to a-amanitin, an inhibitor of RNA polII (87). RFLPs analysis indicate that there are structural differences between PARP genes in both bloodstream and procylic forms of T. brucei variant 118, ruling out genomic

rearrangement as a mechanism to control expression as observed in VSGs (11). The metacyclic (salivary gland) forms possess a coat analogous to the VSG, whose expression is determined by a repertoire of 10-20 strain-specific VSG genes or the mVATs (88). The size of this repertoire varies from one isolate to another (89). These genes are expressed at the last developmental stage in the vector prior to infection. They are then employed in the very initial stages of establishment within the host to expedite switching at a very high rate of 1 per every cell division (90-92). The VSGs of mVATs are telomeric and may be expressed in situ. Each bloodstream form VSG gene is flanked by 70bp repeat units with a strong Z-DNA potential. This may be a chiral determinant of recombination or segmental gene conversions between basic copy (BC) and expression linked copies (ELC). The absence of these repeats in mVAT genes means that the metacyclic VSG repertoire is relatively more stable and therefore accounts for the absence of rearrangements in these genes (90,93).

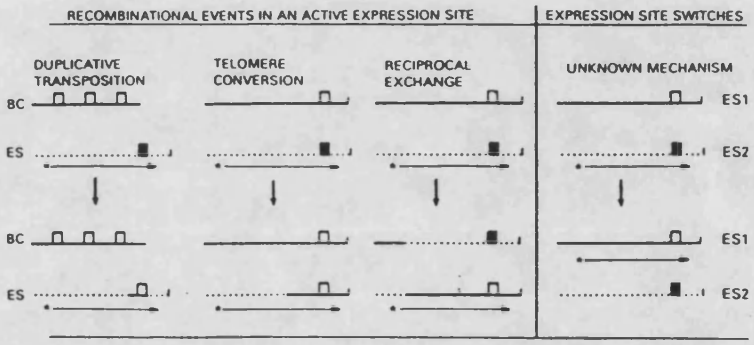



Fig. 19 Mechanisms of VSG expression. Open boxes represent basic copy (BC) genes, and solid boxes represent expression linked copies located within an expression site. The direction of transcription is indicated by a horizontal arrow. Reproduced from (9).

### 1.3.5 RNA metabolism

#### 1.3.5.1 RNA splicing

In general, splicing reactions occur at consensus splice sites or signals [AG/GT and C(T)AG/GT], and are catalyzed by small nuclear ribonucleoprotein particles [snRNPs] (94). The result is a removal of introns or intervening sequences (IVS), which are non-coding, and a splicing together of coding, translatable mRNA sequences.

In the African trypanosome, two types of splicing mechanisms may be involved, although there is no evidence for cis-splicing in these organisms. This may be due to the absence of introns in the genes that have been characterized so far.

However, all trypanosome pre-mRNAs so far identified are derived from polycistronic templates and mature by trans-splicing onto a splice leader sequence known as the mini-exon (96-100), and by polyadenylation following cleavage, to generate multiple mRNA templates. This mini-exon comprises 39 nucleotides 5' to the coding region and is encoded in arrays of 1.35kb repeats of about 200 copies that are clustered in the genome, and is transcribed as a 140-nt RNA species termed medRNA (mini-exon derived RNA) (95). It has the same cap structure (trimethylguanosine cap) as the mature RNA species. However, the 5' most nucleotide of the cap is 7-methylguanosine. Nuclear run-on transcription assays showed a 700-fold higher rate of transcription compared to other RNA species indicating that the medRNA is transcribed independently of the rest of the mini-exon repeat sequence (99). In addition, it has been found that there is no mini-exon in the intergenic region of the ab tubulin gene cluster before transcription: it is spliced in trans after transcription of the gene (98). The sensitivity of the medRNA

transcription to a-amanitin, in sharp contrast to the insensitivity of VSG gene transcription (1mg/ml), further lends credence, albeit tenuously, to the supposition that the two RNA species are not of the same precursor and that transcription is indeed discontinuous in the trypanosome.

This form of splicing is by no means peculiar to T. brucei. It has been found as a regular feature in mRNA processing in other kinetoplastids (99), and in Caenorhabditis elegans which has a 22nt splice leader (100).

| Organism | Mini-exon sequence |
|---|---|
| T. brucei | AAC GCTATTA TTAGAACAGT TTCTGTACTA TA [TTG/GTATG |
| T. vivax | AAA GCTTTTA TTAGAACAGT TTCTGTACTA TA [TTG/GTATG |
| T. cruzi | AAC GCTATTA TTGATACAGT TTCTGTACTA TA [TTG/GTACG |
| CONSENSUS | AA GCT TTA TT ACAGT TTCTGTACTA TA [TTG/GTA G |

Fig. 20 Comparison of kinetoplastid mini-exons. Splice junctions are indicated with a slash. Reproduced from (99).

Immediately downstream of the mini-exon is a putative splice donor site [5'TTG/GTAPyG 3'] which resembles the consensus (101) for splice donor site [5'CAG/GTPuAG 3']. This is assumed to be the mini-exon splice site during mRNA maturation.

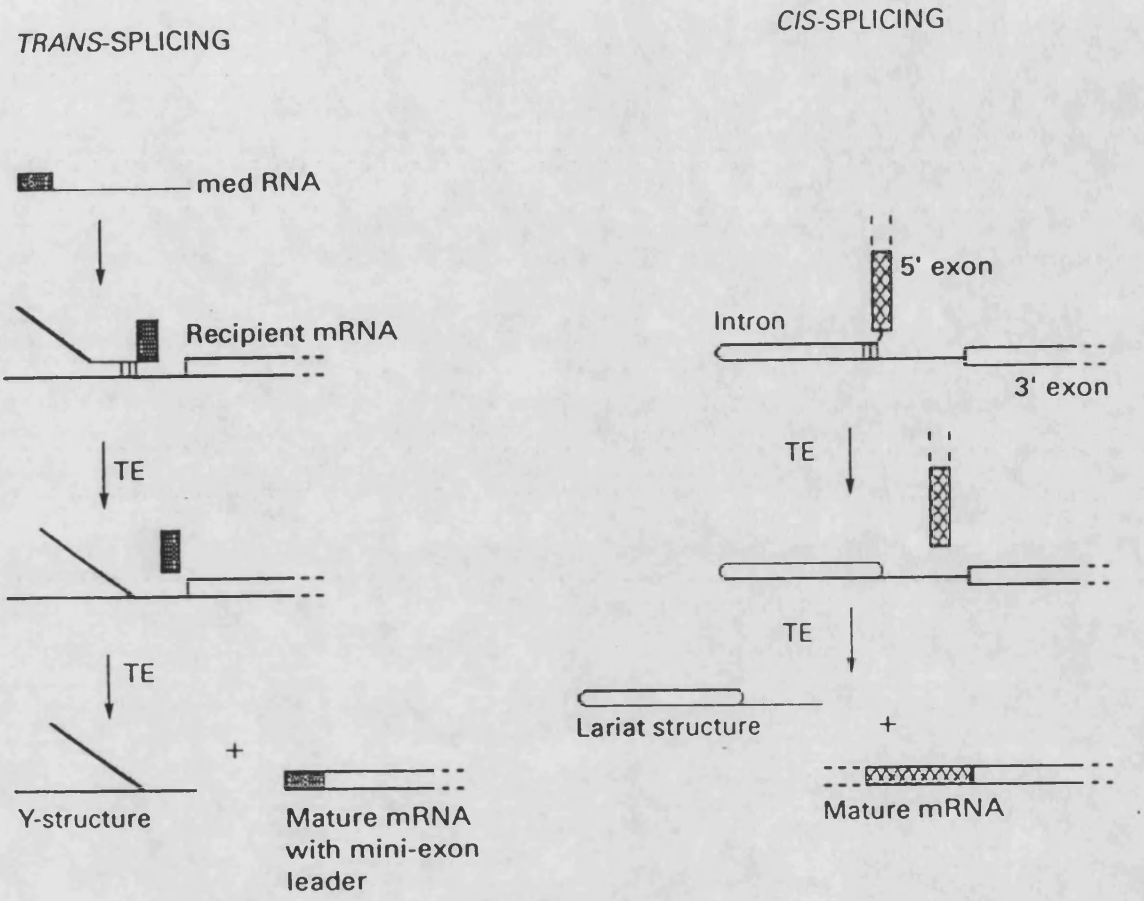



Fig. 21 Trans-splicing as elucidated in the Kinetoplastidae compared to the conventional cis-splicing mechanisms of intron removal from primary RNA transcripts. Transesterification reactions (TE) lead to the formation of a 2'-5' phosphodiester bond which constitutes the branch point for both the lariat and the Y-structure. Reproduced from (9).

### 1.3.5.2 RNA editing

As indicated above, the kinetoplast or mitochondrial DNA (mtDNA) in the trypanosome is varied, comprising the mini and maxi circles. The mini circles encode the guide RNAs that are involved in editing (278) whilst the maxi circles are the trypanosome equivalent of higher eukaryotic mtDNA with regards to energy metabolism. Among others, it is here that the cytochrome genes are transcribed and processed, notably the subunits I, II, and III of cytochrome c oxidase. Paradoxically however, mitochondrial RNA metabolism seems to defy the central dogma in molecular biology viz: DNA makes RNA makes protein. By and large, transcripts do not totally correspond to their encoding genes in nucleotide sequence. This is due to a deletion/insertion strategy involving uridine residues (102). As a result, while some transcripts are only moderately edited e.g apocytochrome b (103), others such as coIII are edited by up to 55%, with the gene significantly different from its message in nucleotide composition (104,105). The significance of this is that editing may be used to regulate the expression of mitochondrial genes in the trypanosome. Frame shifts or the creation of start codons might make this possible (106).

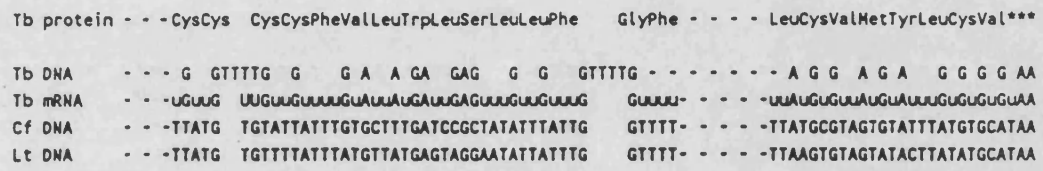



Fig. 22 RNA editing in cytochrome c oxidase III (COIII) in T. brucei. Insertions and deletions of uridines (u) are indicated in the mRNA. A termination codon, TAA has been created by editing. A comparison is made with corresponding genes of the gene in Crithidia fasciculata (Cf) and Leishmania tarentolae (Lt) where RNA editing does not occur. Reproduced with permission from (9).

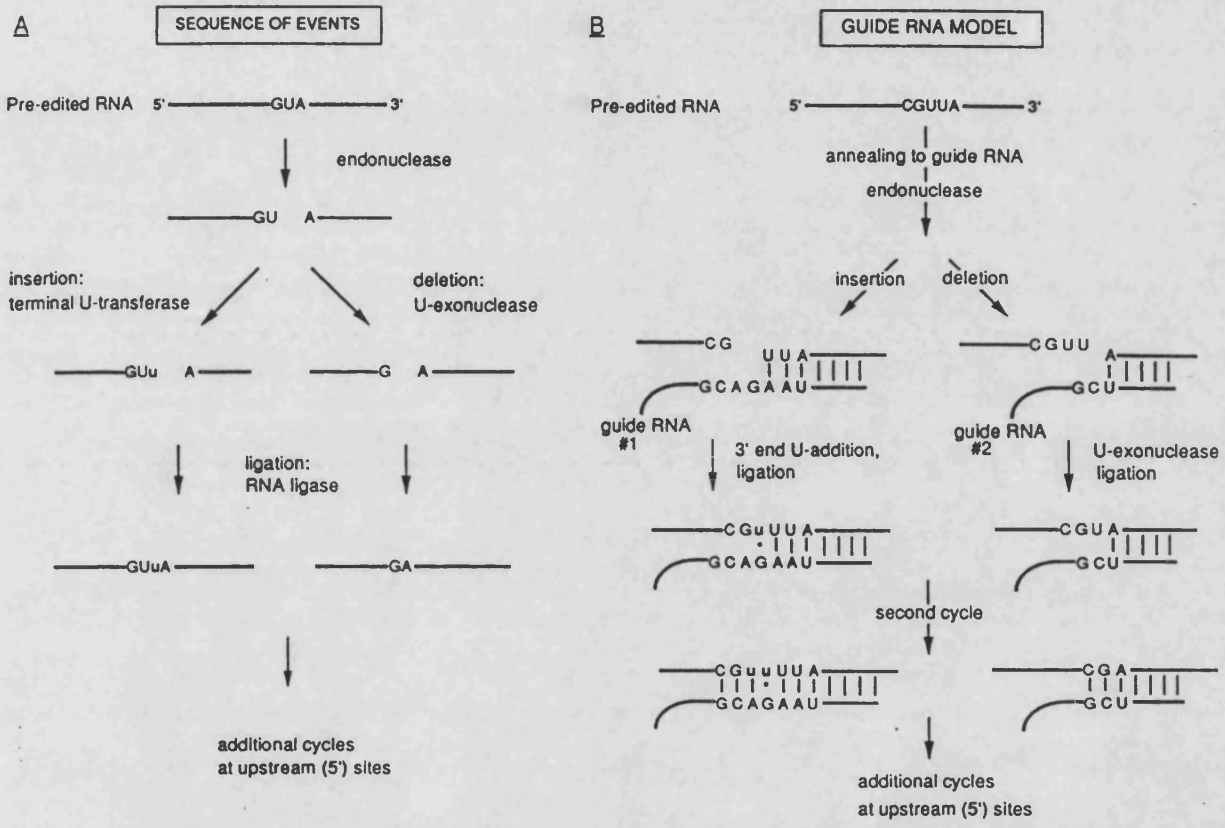



Fig. 23 Models of RNA editing. (A) Sequence of events and enzymatic reactions leading to the insertion or deletion of uridines. (B) The guide RNA model. Two hypothetical, different guide RNAs are shown that would promote insertion (#1) or deletion (#2) of U residues at the same editing site. Asterisks represent a G.U base pair. Vertical lines indicate conventional base pairing. Adapted from (106).

## 1.4 Control of African trypanosomiasis

### 1.4.1 Vector control

Targeting the tsetse fly earlier presented an opportunity for the elimination of the disease by the use of non-residual insecticides such as DDT, Aldrex T etc. Unfortunately, the tsetse fly developed resistance to these. Other tsetse control measures have included baited traps, sterile male techniques, clearing vegetation, and biological control using insectivorous predators.

### 1.4.2 Chemotherapy

The path of the chemotherapy of African sleeping sickness has been marred not so much by the refractory nature of the disease to the available drugs, as by the lack of sufficient information about the basic biology of the parasite- information that may provide leads for the rational development of safe drugs.

#### 1.4.2.1 Current treatment regimens: drugs and their limitations

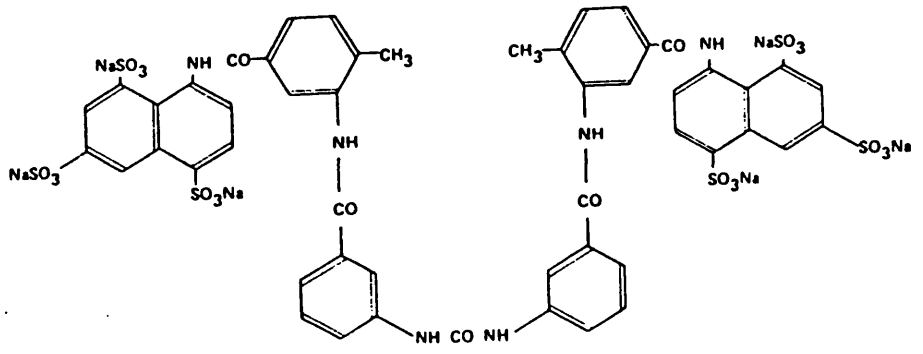
Although chemotherapy is as old as the trypanosome (in terms of discovery), there is little advance in the treatment of African sleeping sickness compared to the rapid treatment regimens available for other diseases. Current treatment regimens rely on suramin, melarsoprol and pentamidine which have largely been derived empirically (107). Their inefficacies stem from poor pharmacokinetics to death in the subjects. They are not only unduly toxic - principally because they have not been designed for any specific target molecule in the parasite-, but also have a very limited

effect on parasite elimination. As a result, there may be periods of recrudescence of parasitaemia with a vengeance, due to survivors that may be increasingly developing resistance to the drug in use. At the same time, their level of toxicity means that they should be administered under clinical supervision, which in itself is limited in the African setting as some areas are inaccessible.

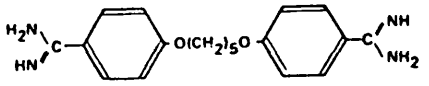
Summary of some drugs and their limitations in the treatment of African sleeping sickness.

| Drug | Chemistry | Selectivity | Majorlimitations |
|---|---|---|---|
| Pentamidine | Diamidine | Poor; may impede parasite nucleic acid metabolism | Hypoglycaemia renal damage Preclinical (PCI) intervention |
| Suramin | Sulphated naphthylamine | Very poor; parasite energy metabolism | Renal damage, optic atrophy, poor pharmacokinetics |
| Tryparsamide | Arsenical | Unknown; suitable for advanced and preclinical use | Optic atrophy (blindness) Only for T. gambiense infections |

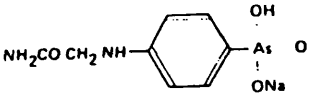
| | | | |
|---|---|---|---|
| Melarsoprol | Arsenical | Good; energy metabolism; advanced form treatment | Reactive encephalopathy |



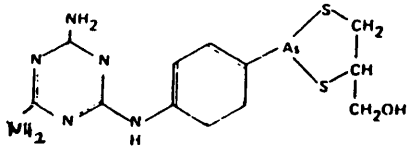
Suramin



Pentamidine



Tryparsamide



Melarsoprol

Berenil

Fig. 24 Structures of drugs in current use for African trypanosomiasis. Reproduced from (28).

### 1.4.2.2 Current rational strategies

Attempts are being made to optimize the exploitation of trypanosome biochemistry based on differences at the host-trypanosome interface (108). One of these relied on the presence of a unique glycerol phosphate oxidase (GPO comprises a glycerol phosphate dehydrogenase, ubiquinone and a terminal oxidase) which is absent from mammals (109,110). Using chelators like benzhydroxamic acid, diphenyl dithiocarbazone, salicylhydroxamic acid (SHAM), it was possible to inhibit GPO non-competitively, blocking electron transport between ubiquinone and the terminal oxidase and killing them by O2/ATP starvation. However, the parasites employed an alternative pathway involving glycerol kinase:

$$\text{G-3-P} + \text{ADP} \quad > \quad \text{glycerol} + \text{ATP},$$

in order to survive. Conceptually, increasing the [glycerol] should block the forward reaction by a mass action effect. This was observed to be the case as SHAM and glycerol synergistically killed the parasites. Infusion of this cocktail produced a temporary remission of parasitaemia only to recrudesce a week later. Aside from this limitation, excessively large amounts of it (which was unduly toxic) were required and these factors made it an impractical regimen to employ.

Another therapeutic strategy relied on the fact that trypanosomes do not synthesise purines de novo, and have to salvage them somehow (111). Therefore using purine analogues like cordycepin (3' deoxyadenosine), nucleocidin and puromycin proved trypanocidal albeit toxic to the host. Allopurinol, a hitherto favourite drug for the

treatment of gout, has been found to be particularly effective and at the same time of tolerable toxicity (112). The apparent selective toxicity of this drug resides in the fact that the step from allopurinol catabolism catalyzed by xanthine-guanine phosphoribosyl transferase, is of little significance in mammals. Furthermore the nucleotide produced is aminated to an analogue of AMP, aminopyrazaopyrimidine ribonucleoside monophosphate, a reaction which mammalian enzymes are incapable of catalyzing. The incorporation of this analogue into RNA interferes with translation and protein synthesis.

The observation that trypomastigotes do not synthesise spermine offered a fresh target. They synthesise putrescine and spermidine de novo from ornithine and methionine. The rate-limiting step in polyamine synthesis is thought to be catalyzed by ornithine decarboxylase, (ODC) which has a high turn-over rate in mammals, and is susceptible to external inducers (113). Using an analogue of ornithine, a-difluoromethylornithine (DFMO) it was observed that polyamine synthesis was blocked at the ODC step. This mammalian enzyme is 60X less sensitive, therefore providing a suitable target for selection (114,115). One reason for this selective toxicity was that the trypanosome enzyme has a relatively lower turnover rate compared to the host enzyme (116), thereby allowing for inactivation of the former over a long period. However, the drug is incapable of curing advanced infections due to its inability to cross the blood-brain barrier.

The African trypanosome is also incapable of synthesising haem and therefore lacks catalase, thus providing a channel for treatment by inducing a free radical-mediated lipid peroxidation. The accumulation of $H_2O_2$ (up to 100X more than in mammalian cells) generates HO and HOO radicals in the presence of metalloporphyrins e.g hematoporphyrin D. This may be exacerbated by sequestering free radical scavengers like riboflavin and glutathione with the naphthoquinones.

Other promising strategies involve exploiting and targeting differences between glycosomal enzymes in trypanosomes and mammals (117). Lately, it has been observed that trypanosomal glutathione reductases (trypanothione reductases) differ from their mammalian counterparts in their absolute requirement for a thiol-containing cofactor, trypanothione (118). This cofactor contains a spermidine moiety covalently attached to glutathione (119). Inhibition of polyamine synthesis by DFMO is thought to be linked to this site, precluding the provision of trypanothione for glutathione reduction which would protect it from free radical damage.

It has also been shown that the GPI-anchor of the VSG in bloodstream trypanosomes is unique in having myristic acid. In contrast, anchors in mammalian cells lack the myristoyl moiety. By introducing analogues of myristic acid in an in vitro system designed to monitor VSG biogenesis, it was found that incorporation and synthesis of the anchor was impeded, resulting in cell death (120). In contrast procyclics, which lack a VSG coat (have a PARP coat), and Saccharomyces cerevisiae with no myristate, survived. Other compounds which affected trypanosome morphogenesis such as vinblastine, disrupted microtubule formation and trypanosome replication.

No doubt, the litany of strategies could continue, but suffice to say that in a search for selectively toxic drugs, it is not enough to identify a unique pathway as a target. Rather, it is imperative to identify strong functional differences between that pathway in both parasite and host, and to demonstrate an absolute dependence of the trypanosome on the identified pathway. This is where glucose transport comes in.

## 1.5 Hexose transport: an overview

### Membranes in transport and metabolic regulation

Intermediary metabolism in all living cells is largely dependent on the availability- chemical or physical- not only of enzymes, but also of substrates. Physical compartmentation of the latter seems to be the main determinant of the rate of metabolism. Biological membranes present significant cellular barriers to the utilization of substrates. Apart from acting as barriers to substrate flow, they also elicit a selective property with regards to the nature of the transported molecule.

### 1.5.1 Mechanisms of substrate transport

There are four basic mechanisms by which cells may transport substrate across their membranes.

i. Simple diffusion

This is a typical scalar phenomenon characterized by first-order kinetics. It is essentially stochastic and involves non-specific permeation of solutes, which as it were "roll" down their concentration gradients, in consonance with Fick's law at rates directly proportional to the concentration differentials between the two compartments. There is no specific interaction between the permeant molecule and the membrane.

ii. Facilitated diffusion

This is a saturable uptake process which relies on a carrier molecule to traffic solute from one compartment to another. Unlike simple diffusion, there is specific interaction between the permeant and the carrier, much in keeping with stereochemical

requirements. Because there is a finite number of binding sites, which therefore sets a limit to the number of solute molecules that can be transported, there is competition between substrate analogues and a predilection for saturation of these binding sites. The process is also sensitive to the effect of inhibitors, and shows uphill counter-transport characteristics.

iii. Active transport.

This requires metabolic energy for the translocation of substrate. It is basically a unidirectional flux of solute from a lower to a higher concentration. Because this requires energy, the process is susceptible to metabolic inhibitors. Whilst the simplest form of this mechanism involves only a single molecular species, more complex transfers may occur which involve the coupled transport of solutes simultaneously across the membrane in the same direction or symport (121). In this case, there is an obligatory sequential attachment of the solutes to the transport site before either can be translocated. The energy for transport may also be provided in an electrochemical potential gradient, as in the mitochondrial ATPase (122), or the coupled system may involve an exchange process as in antiport systems. Like facilitated diffusion, the process shows saturation or Michaelis-Menten kinetics and is specific.

iv. Group transfer

This form of substrate transport is accompanied by a covalent modification of the transported substrate, relative to its initial state. This is aptly represented in the phosphoenoltransferase (PTS) system in bacteria (123).

Glucose transport has been observed to occur largely by facilitated

diffusion in most cells including the T. brucei transporter and is thought to be the rate limiting step in glucose metabolism in the trypanosome (17,124,125).

### 1.5.2 Phenomenology of the facilitated glucose transporter

The main features of the facilitated transporter enunciated in (ii) above, engender a quasi-cyclical process of binding and translocating substrate. This may embrace four sequential cyclic steps that describe the carrier as a mobile entity. These steps are depicted in:

1. Recognition and binding of substrate on the cis side,
2. Translocation of substrate across the membrane
3. Release of substrate on the trans side
4. Restitution of the carrier to its initial state.

Within this schema, only steps (2) and possibly (4) are vectorial, whereas (1) and (3) may be scalar. This model is consistent with enzyme-type catalysis; it is strongly compatible with both substrate structural specificities and sensitivities to inhibitors, and intuitively involves cyclical changes in affinity of the transporter for substrate on either side of the cell membrane. As such, typical Michaelis-Menten kinetics may be assigned in the equation:

$$V = \frac{Vmax\ S}{Km + S}$$

Implicit in this equation are all the determinants of the phenomenon of substrate transport, with Vmax and Km being, operationally, mutually dependent variables at a given concentration of substrate [S]. The rate of transport would be a function of the number of binding sites, of a total carrying capacity Vmax. Conformational transitions may occur between the two states of free transporter and transporter-substrate complex (upon binding substrate), which may result in the lowering of transition/activation energy enabling internalization of the complex. The net rate of internalization of substrate will therefore be dictated by the proportion of high affinity sites available cis-trans or for steps 2 and 4.

Evidence for conformational change as a prelude to facilitated transport accrues from the analysis of circular dichroic spectra of the human erythrocyte glucose transporter. This indicated that the a-helices which constitute 50-80% of the protein, have a preferential orientation perpendicular to the plane of the membrane, with an effective tilt angle below 38° (126). This preferential orientation was slightly improved in the presence of D-glucose, suggesting that the binding of substrate may be associated with changes in conformation of the transporter. In addition, a comparison of the infra-red spectra recorded under conditions of the inward or introverted form of the protein with those measured under the outward facing form, revealed small shifts in the bands assigned to a-helices and b-strands (127) supporting alternating conformation as the mechanism of glucose transport.

Cytochalasin B (cytB), a fungal metabolite and a known inhibitor of

glucose transport, is thought to bind to the inward facing conformation of the transporter as it is a competitive inhibitor of glucose efflux but a non-competitive inhibitor of influx. The binding of cytB precludes the binding of ASA-BMPA [azidosalicoyl bis(D-mannos-4-yloxy)-2-propyamine] which is an impermeant competitive inhibitor of glucose uptake (128). Consequently, ASA-BMPA is thought to bind to the outward-facing conformation. Binding of these ligands was seen to quench the intrinsic fluorescence of the transporter, indicating a conformational change. These observations were confirmed from limited cleavage with thermolysin (128). Based on altered sensitivity at defined cleavage sites determined by limited proteolysis in the presence of various substrates or inhibitors of transport, Gibbs et al. (129) were able to detect conformational changes associated with the function of the erythrocyte transporter. A similar observation has been made with the Na+-linked renal glucose transporter (130), where a ligand/glucose-induced conformational change affected the binding of eight different monoclonal antibodies. Other observations were made using 1-fluoro 2,4-dinitrobenzene FDNB (131), ethylmaleimide and chloronitrobenzoxadiazole (132) and ethylidene glucose (133) which induced a fluorescent transient and hence a conformational change in the transporter. This transient has been attributed to a reorientation of the transporter as a result of ethylidene binding to the exofacial site thus inducing a "flip" from the inward- to the outward-facing conformation.

With this information however, it is subjective to link all segments of the transporter in the conformational change. The reality was that these changes were rather localised. Labelling with the above ligands therefore implicated transmembrane region

(TM9) as the mobile region conferring conformational change, as it is presumed to form a hydrophobic cleft in the substrate-binding site. Pawagi and Deber (134) suggest however that TM10 may be the dynamic segment of the glucose transporter.

### 1.5.3 Structure and function correlations in glucose transport.

The pyranose chair configuration of glucose appears to be the preferred substrate for the erythrocyte-type transporter. Kahlenberg and Dolansky (135) provided earlier indications of the structural requirements of the transporter for its substrate. They established that substrate docking requires hydrogen-bonding at the anomeric carbon (C-1) hydroxyl of the pyranose chair configuration. This (anomeric hydroxyl) was found to be the most critical, although all of them had the propensity to H-bond to varying degrees (136). Using glucose analogues, it was observed that the C-1, C-3, and C-6 hydroxyls of glucose were the most important moieties in binding to the erythrocyte glucose transporter (137). Substitution of OH groups with bulky hydrophobic groups at C-1, C-4 and C-6 positions suggested the existence of a hydrophobic cleft adjacent to the C-4 and C-6 positions of bound glucose, and that it was the C-1 and C-6 hydroxyls of the sugar that initially interact with the inward and outward facing conformations of the transporter respectively. The general picture that emerges from these studies is the polarised binding of glucose i.e that the glucose molecule enters with the C-1 end, which is free from steric hindrance, and tails in with the C-6 where there is a larger spatial constraint. This view is consistent with the "reactive site-directed" proposal of Berlin (138), that since phosphorylation of the glucose molecule by hexokinase occurs at the C-6

position, this position would be protected from direct interaction with the transporter during translocation of the glucose/transporter complex. This specificity requirement is lacking in situations where hexokinase is absent, e.g gluconeogenesis, further confirming the positional importance of C-1 with respect to the binding site and a lack of protection of C-6 on glucose.

These observations were translated to specificity studies of the glucose transporter in T. brucei (17). It was established that H-bond interactions occurred at C-1, C-3, C-4 and C-5 between the transporter and glucose, C-2 and C-6 appearing to have hydrophobic interactions with the transporter. By systematically substituting various groups at the various carbon positions of the pyranose ring, it has been possible to map out the chemical and steric determinants of the glucose-transporter interactions. Spatial constraints were ascribed to the C-4 position in contrast to the mammalian forms where this constraint resides at the C-6 position.

The use of infra-red spectroscopy enabled the analysis of the kinetics of hydrogen-deuterium exchange in the erythrocyte GT (139). These experiments suggested that a large proportion of amide hydrogens in the membrane-spanning segments of the protein were easily accessible to water and hence support for the possible existence of an aqueous channel (139). The exchange kinetics of these channel-associated hydrogens were drastically altered in the presence of glucose and cytB. It may be surmised here that the glucose molecule is heavily solvated in transit through the lipid bilayer, from the exterior to the inside of the cell.

The structural model predicted (see below) for the GT is largely based on

multiple alignments of cloned sequences. These alignments highlight residues conserved within the superfamily and may subserve the common function, directly or indirectly, of glucose transport. These residues include the serine and glutamic acid residues in TM7 which may provide OH and amino groups in building an aqueous channel or H-bonding to glucose. The largely amphipathic nature of TM5 and TM8 makes them candidates for building this channel. The H-bonding interactions between glucose and the binding site may be occurring between helices TM7, TM8 and TM11, while the hydrophobic cleft may occur between helices TM9 and TM10 (140). Evidence of this comes from the observation of hydrogen exchange experiments that the aqueous channel collapses on binding glucose (127). This was corroborated by tagging the transporter with ASA-BMPA and cytB, followed by proteolysis. Hence, it is clear that some residues at the glucose binding site provide the relevant bonding interactions to facilitate binding and translocation. The relevance or role of more distal residues may involve conferring conformational changes to, or anchoring the entire transporter within the membrane.

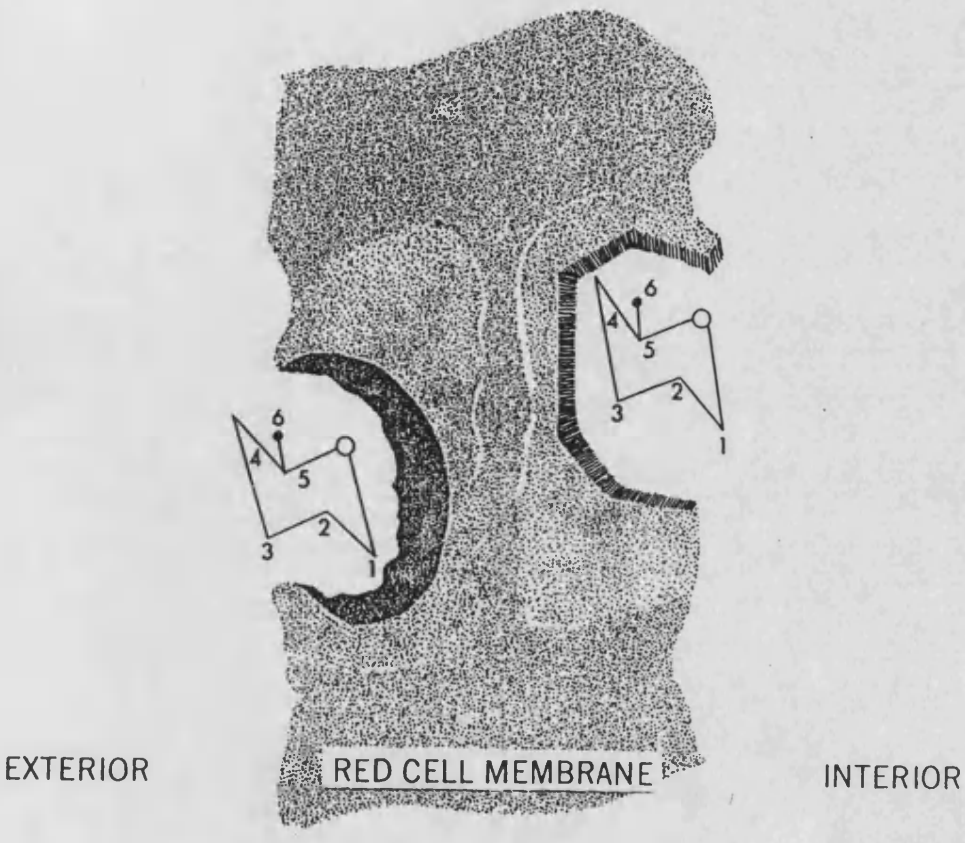



Fig. 25 Spatial disposition of GT and glucose in binding and translocation across the erythrocyte membrane. The carbons are numbered 1-6. Reproduced from (140).

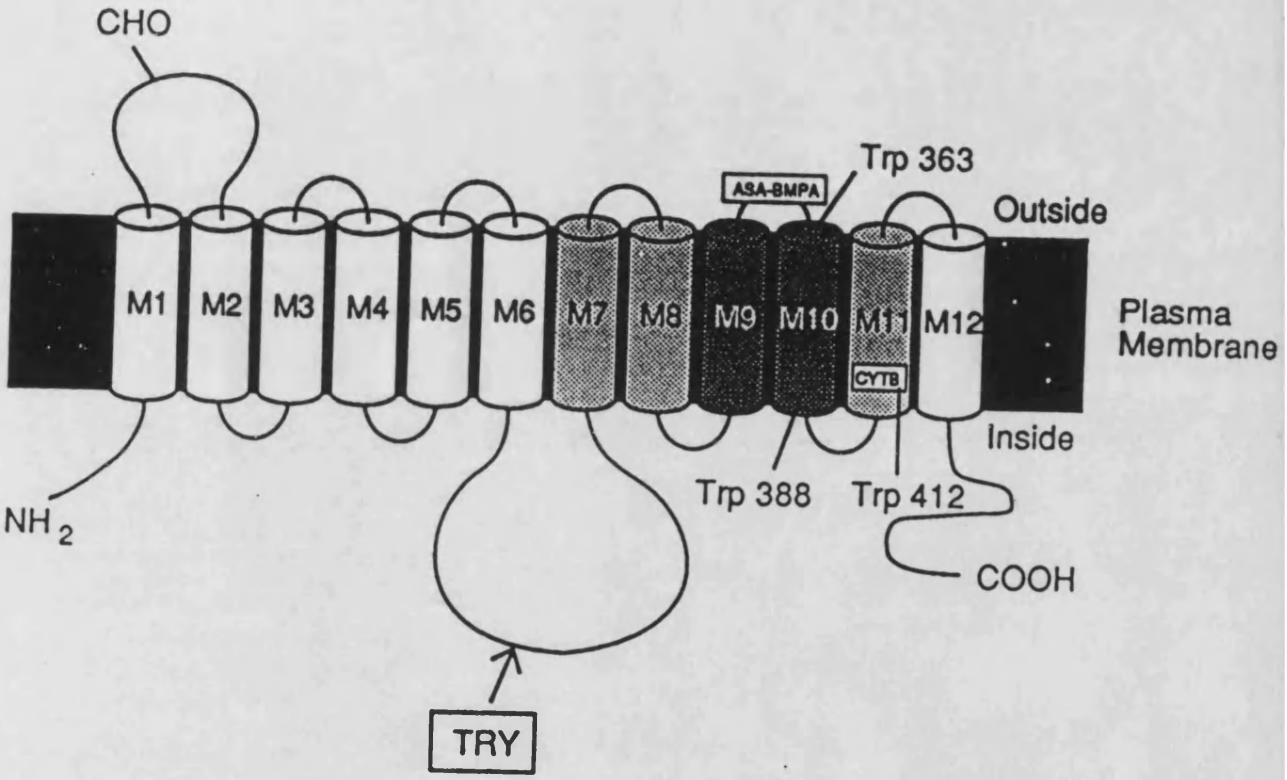



Fig. 26 Structural dissection of the erythrocyte GT. The sites for ASA-BMPA, cytochalasin B (cytB) binding and tryptic cleavage are indicated. The potential glycosylation site is indicated by CHO. Transmembrane domains are numbered from M1-M12. Reproduced from (140).

## 1.6 Molecular biology of glucose transport

### 1.6.1 Prokaryotes

As may be inferred from the phenomenology of sugar transport, transport proteins are as multiple as they are ubiquitous. Among the prokaryotes, there are the PEP transferase system, the cotransporters, and the sugar binding proteins. The PEP system has been referred to above and will not be mentioned any further.

The sugar binding proteins are a class of perisplasmic proteins which serve as the initial components of high-affinity active transport systems. They may act as initial receptors for chemotaxis, and aside from sugars, they may also carry amino acids and ions in Gram-negative bacteria (141). These proteins include the arabinose-binding protein as well as those specific for galactose and maltose.

The cotransporters are essentially active transporters involved in the sequential binding of ion and sugar and their simultaneous translocation. Invariably, the ion of choice is $H^+$, the electrochemical gradient of which provides the energy for transport. The essential pre-requisite for coupled transport is a binding site for both ion and sugar, and the binding of one should not induce a conformational change that precludes the binding of the other.

A number of coupled transporters have been cloned and sequenced especially for $H^+$/xylose (142), $H^+$/arabinose (143) and lac permease (144) of Escherichia coli. Significantly, these have been found to be homologous to each other and to the eukaryotic transporters (145-147).

### 1.6.2 Eukaryotes

Glucose transporters in eukaryotes are essentially of two kinds: cotransporters/symporters and facilitated transporters.

Since the studies using the red blood cell as a model system, there has been a burgeoning collection of information regarding not only the ubiquity and similarity of function of these molecules, but also basic identities in structure.

The cotransporters, as alluded to above, involve $H^+$ or $Na^+$. The yeast GAL2 transporter (148) is an example of this group. Among the lower eukaryotes, it has been observed that Leishmania transports sugar coupled to a $H^+$ gradient (149,150). A similar observation has been made in P. yoelii (151). Even though the bloodstream form of T. brucei has been shown to transport sugar via a facilitated transporter (17,124,125), other work indicates that the procyclics may employ a $H^+$ gradient while the bloodstream forms couple transport to a $Na^+$ gradient (152,153). Whichever way this may be considered, there is evidence both for a facilitated and a coupled carrier system in the trypanosome. This links the lower eukaryotes in an evolutionary relationship to the prokaryotes in terms of function.

In the higher eukaryotes, the $Na^+$/glucose transporter in the renal and intestinal brush-border membrane has been the most studied (154), and even though it may be functionally identical to the facilitated and cotransporters, the two types of transporter are significantly structurally different. In either coupled transport systems ( prokaryotes and eukaryotes), binding sites are accessible to both ion and sugar, and the stoichiometry approaches 1:1.

In contrast to symporters, the facilitated carriers are uniport systems that equilibrate across the membrane in response to sugar. It is becoming increasingly clear that these carriers exist in isoforms in mammals, differing more in their stage and/or tissue specificity and to some extent their kinetics, than in their structure. No isoforms have been identified in the prokaryotes or the lower eukaryotes, possibly due to their comparatively simpler energy pathways/requirements or smaller genome sizes. By and large, both transport systems have similar domain organizations as elaborated below.

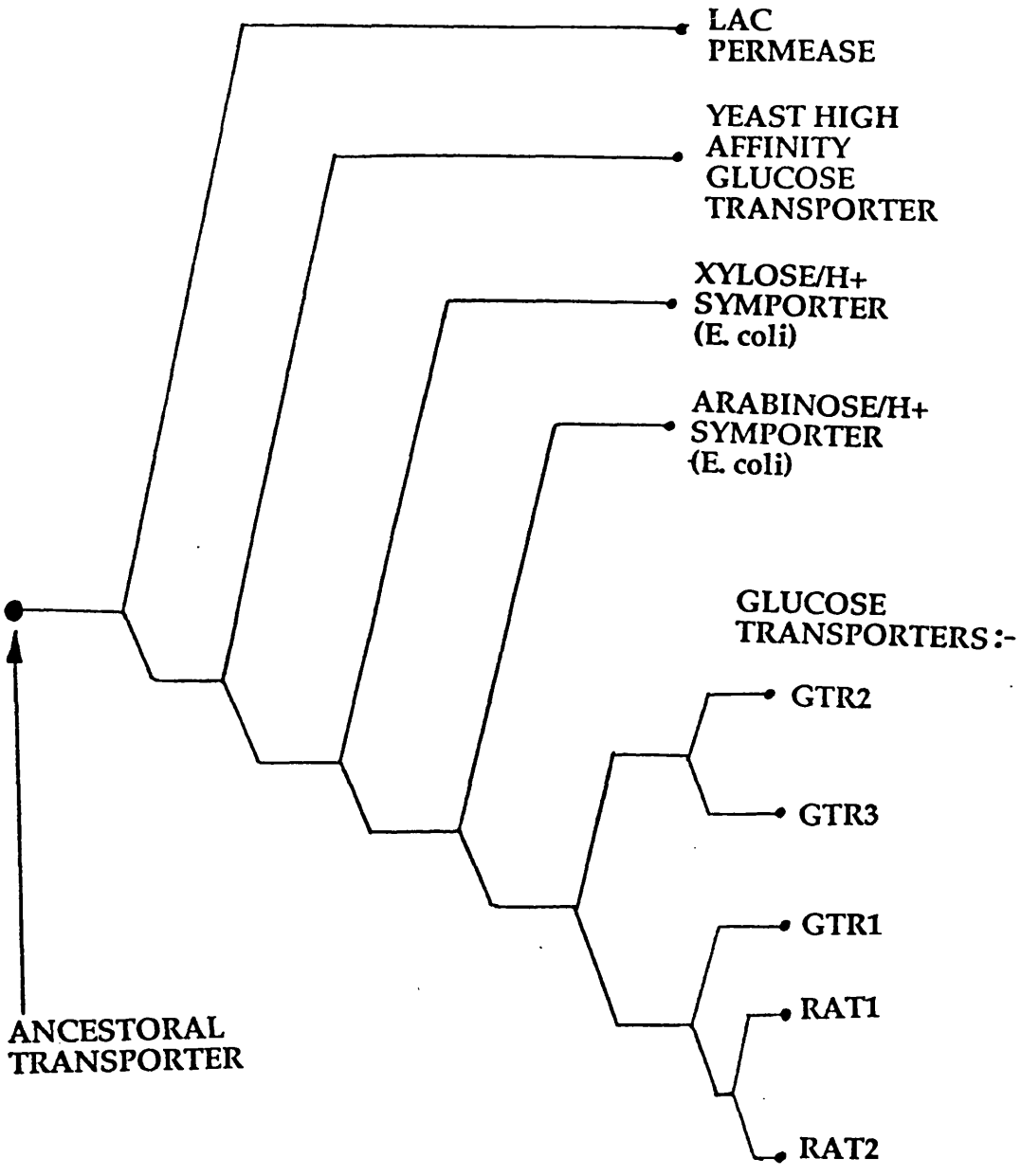



Fig. 27 The Glucose Transporter Superfamily: some evolutionary ramifications. Adapted from (155).

Fig. 28 Multiple sequence alignment of the GTs.

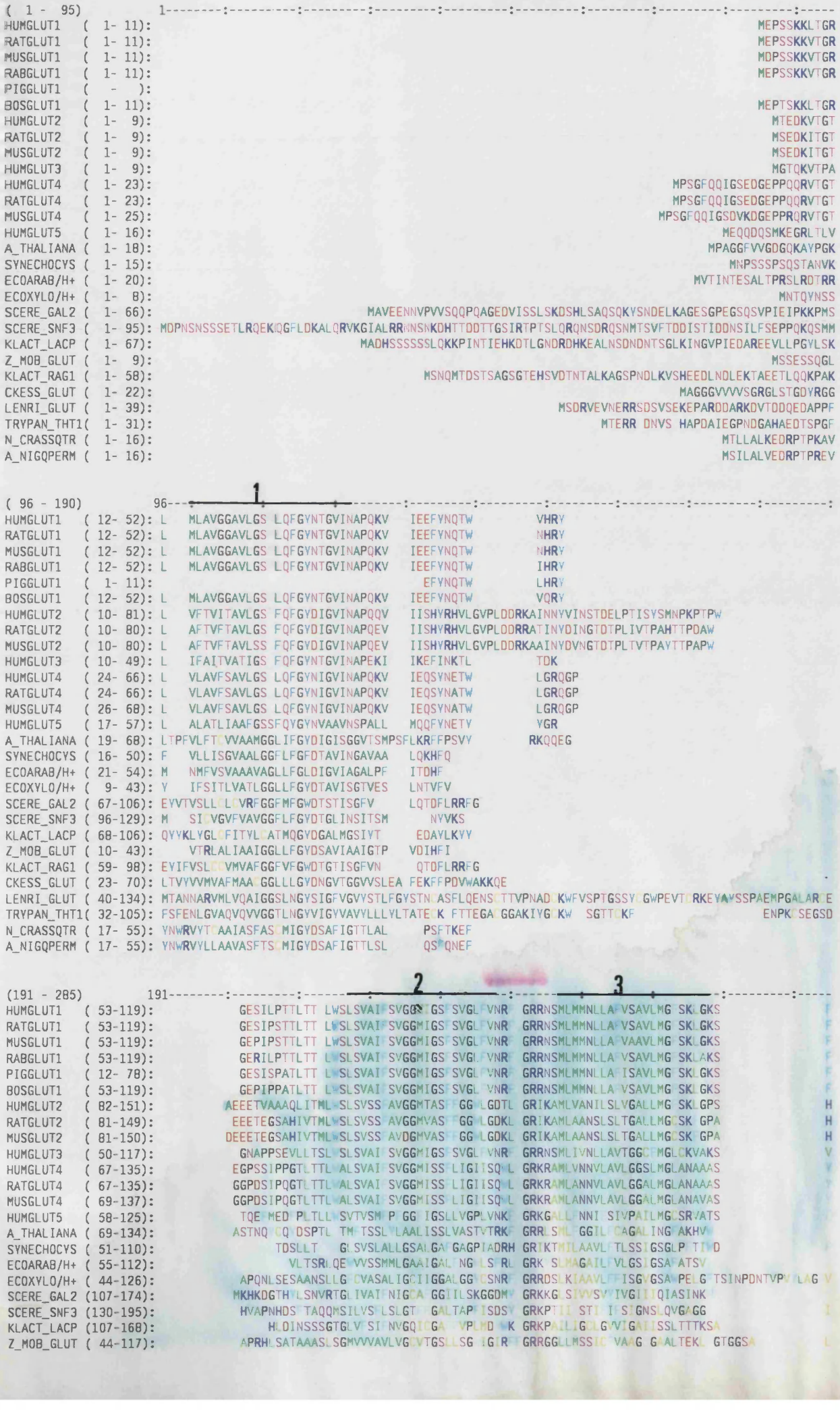
Legend:

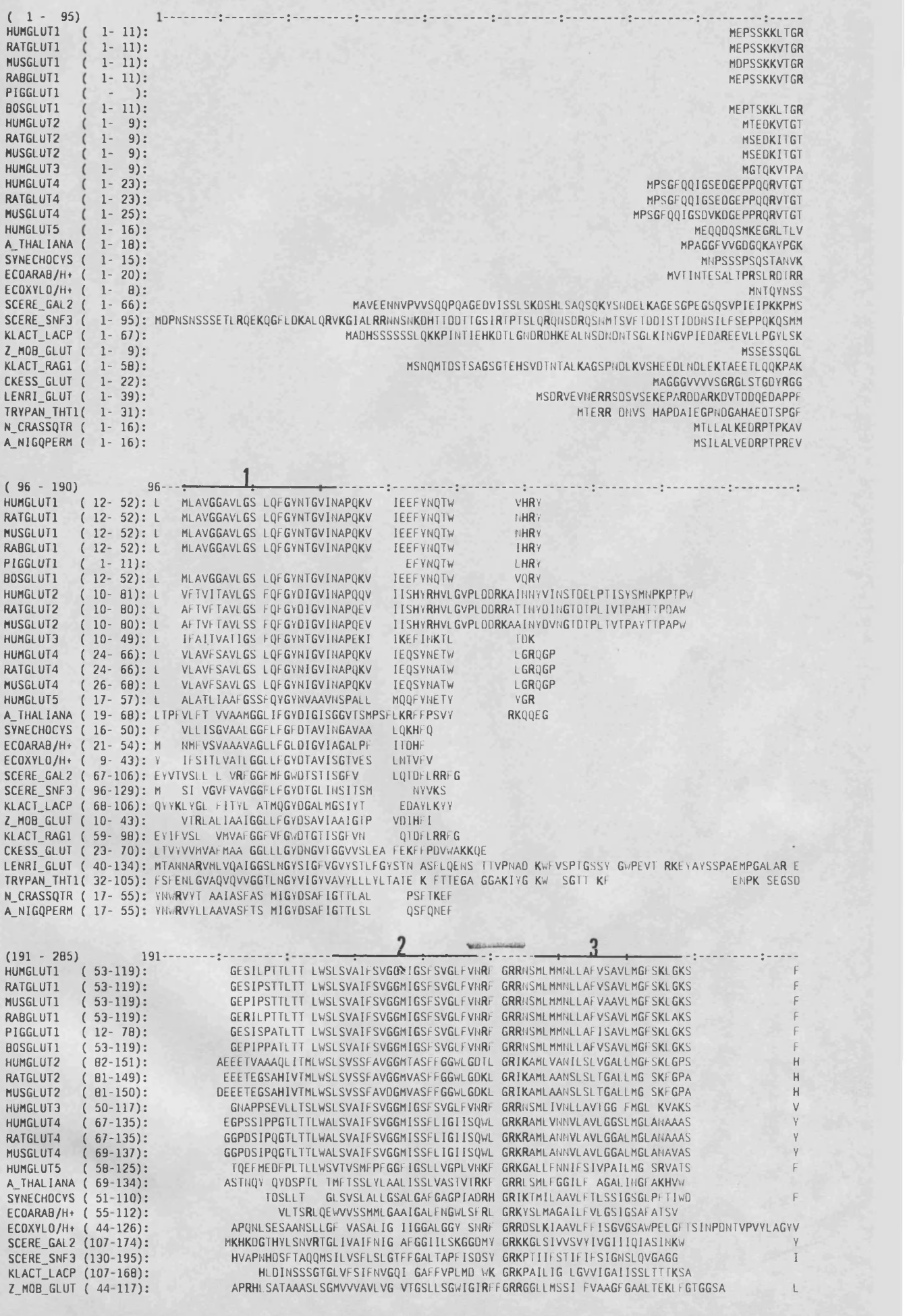
| | |
|---|---|
| GLUT1 | GT isoform 1 from Human (HUM), rat, mouse (MUS), pig and cow (BOS) |
| GLUT2 | GT isoform 2 from human, rat and mouse |
| GLUT3 | Human GT isoform 3 |
| GLUT4 | GT isoform 4 from definitions above. |
| GLUT5 | Human GT isoform 5 |
| A-thaliana | GT from the plant A. thaliana |
| ECOXYLO/H+ | E. coli xylose/H+ symporter |
| ECOARAB/H+ | E. coli arabinose/H+ symporter |
| SCERE_GAL2 | Yeast galactose transporter |
| SCERE_SNF3 | Yeast high affinity glucose transporter |
| KLACT_LACP | Lactose permease of Kluyveromyces lactis |
| Z_MOM_GLUT | GT of Zymomonas mobilis |
| LENRI_GLUT | Putative GT for Leishmania enriettii |
| TRYPAN_THT1 | Trypanosome putative hexose transporter |
| N_CRASSQTR | Quinate transporter of Neurospora crassa |
| A_NIGQPERM | Quinate permease of A. nigeriensis |

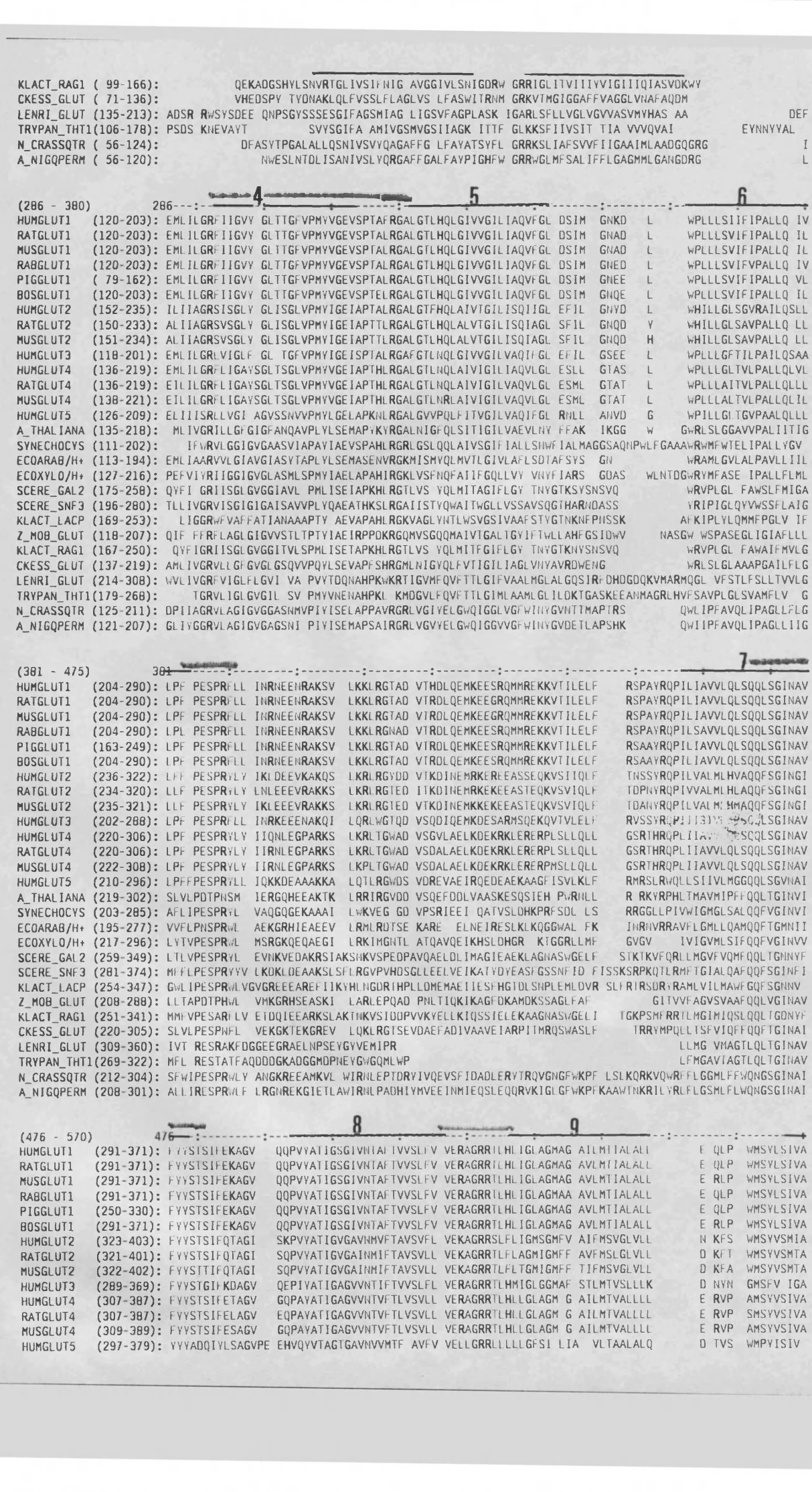
## COLOUR CODES FOR AMINO ACIDS IN SEQUENCE ALIGNMENT

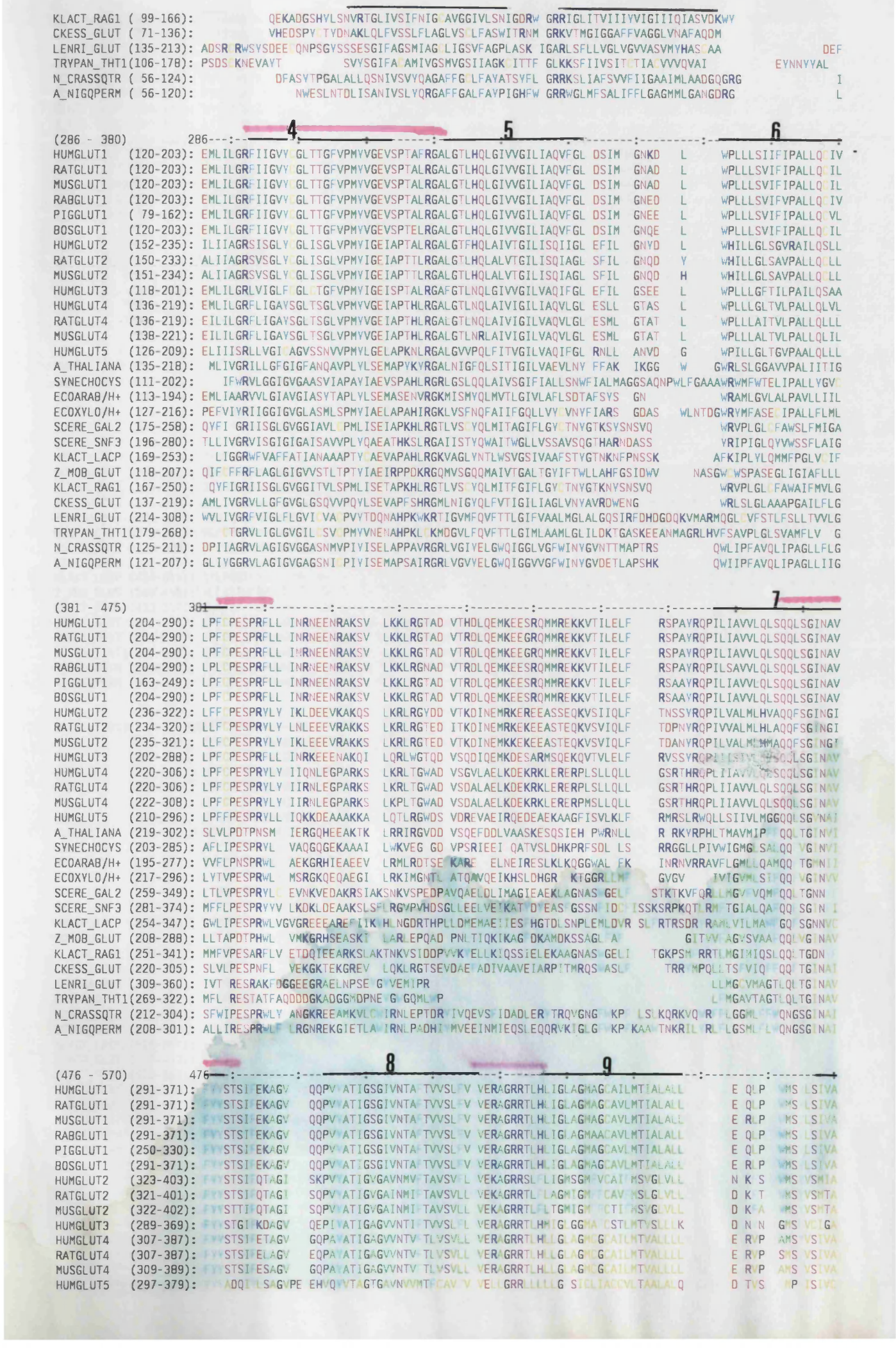
| Colour | Code |
|---|---|
| Yellow | cysteins |
| Black | prolines and glycines |
| Magenta | hydrophobic side chains |
| Green | aliphatic side chains |
| Cyan | aliphatic hydrocarbon side chains |
| Red | negatively charged/acidic side chains |
| Blue | positively charged/basic side chains |

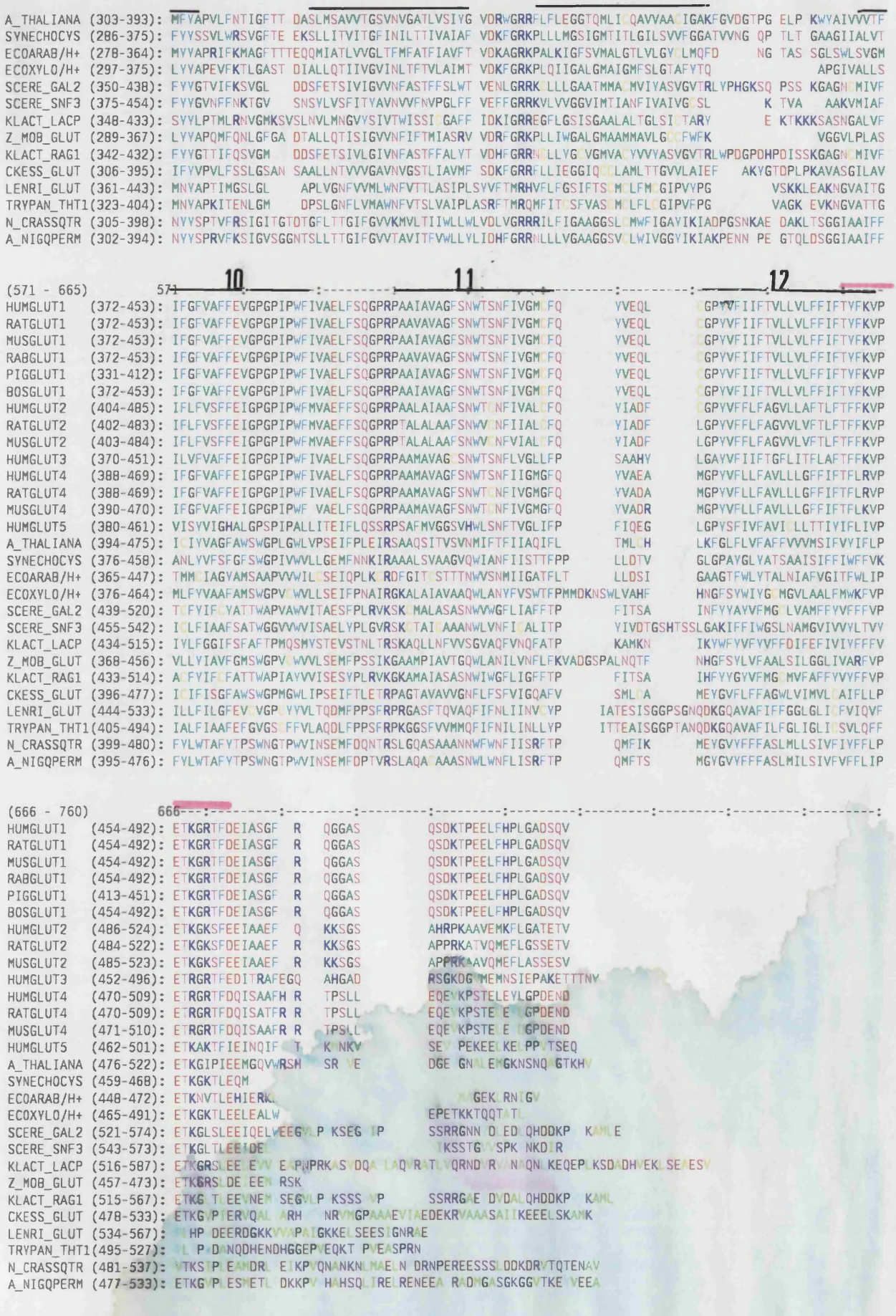
Sequence aligment was performed according to Barton and Sternberg (208) and of Stockwell (209). Conserved amino acid residues/motifs are overlined( in pink)
Transmembrane domains (1 - 12) are overlined in black.






KLACT_RAG1 ( 99-166): QEKADGSHYLSNVRTGLIVSIFNIG AVGGIVLSNIGDRW GRRIGLITVIIIYVIGIIIQIASVDKWY
CKESS_GLUT ( 71-136): VHEDSPY TYDNAKLQLFVSSLFLAGLVS LFASWITRNM GRKVTMGIGGAFFVAGGLVNAFAQDM
LENRI_GLUT (135-213): ADSR RWSYSDEE QNPSGYSSSESGIFAGSMIAG LIGSVFAGPLASK IGARLSFLLVGLVGVVASVMYHAS AA DEF
TRYPAN_THT1(106-178): PSDS KNEVAYT SVYSGIFA AMIVGSMVGSIIAGK ITTF GLKKSFIIVSIT TIA VVVQVAI EYNNYYAL
N_CRASSQTR ( 56-124): DFASYTPGALALLQSNIVSVYQAGAFFG LFAYATSYFL GRRKSLIAFSVVFIIGAAIMLAADGQGRG I
A_NIGQPERM ( 56-120): NWESLNTDLISANIVSLYQRGAFFGALFAYPIGHFW GRRWGLMFSALIFFLGAGMMLGANGDRG L
(286 - 380)
HUMGLUT1 (120-203): EMLILGRFIIGVY GLTTGFVPMYVGEVSPTAFRGALGTLHQLGIVVGILIAQVFGL DSIM GNKD L WPLLLSIIFIPALLQ IV
RATGLUT1 (120-203): EMLILGRFIIGVY GLTTGFVPMYVGEVSPTALRGALGTLHQLGIVVGILIAQVFGL DSIM GNAD L WPLLLSVIFIPALLQ IL
MUSGLUT1 (120-203): EMLILGRFIIGVY GLTTGFVPMYVGEVSPTALRGALGTLHQLGIVVGILIAQVFGL DSIM GNAD L WPLLLSVIFIPALLQ IL
RABGLUT1 (120-203): EMLILGRFIIGVY GLTTGFVPMYVGEVSPTALRGALGTLHQLGIVVGILIAQVFGL DSIM GNED L WPLLLSVIFVPALLQ IV
PIGGLUT1 ( 79-162): EMLILGRFIIGVY GLTTGFVPMYVGEVSPTALRGALGTLHQLGIVVGILIAQVFGL DSIM GNEE L WPLLLSVIFIPALLQ VL
BOSGLUT1 (120-203): EMLILGRFIIGVY GLTTGFVPMYVGEVSPTELRGALGTLHQLGIVVGILIAQVFGL DSIM GNQE L WPLLLSVIFIPALLQ IL
HUMGLUT2 (152-235): ILIIAGRSISGLY GLISGLVPMYIGEIAPTALRGALGTFHQLAIVTGILISQIIGL EFIL GNYD L WHILLGLSGVRAILQSLL
RATGLUT2 (150-233): ALIIAGRSVSGLY GLISGLVPMYIGEIAPTTLRGALGTLHQLALVTGILISQIAGL SFIL GNQD Y WHILLGLSAVPALLQ LL
MUSGLUT2 (151-234): ALIIAGRSVSGLY GLISGLVPMYIGEIAPTTLRGALGTLHQLALVTGILISQIAGL SFIL GNQD H WHILLGLSAVPALLQ LL
HUMGLUT3 (118-201): EMLILGRLVIGLF GL TGFVPMYIGEISPTALRGAFGTLNQLGIVVGILVAQIFGL EFIL GSEE L WPLLLGFTILPAILQSAA
HUMGLUT4 (136-219): EMLILGRFLIGAYSGLTSGLVPMYVGEIAPTHLRGALGTLNQLAIVIGILIAQVLGL ESLL GTAS L WPLLLGLTVLPALLQLVL
RATGLUT4 (136-219): EILILGRFLIGAYSGLTSGLVPMYVGEIAPTHLRGALGTLNQLAIVIGILVAQVLGL ESML GTAT L WPLLLAITVLPALLQLLL
MUSGLUT4 (138-221): EILILGRFLIGAYSGLTSGLVPMYVGEIAPTHLRGALGTLNRLAIVIGILVAQVLGL ESML GTAT L WPLLLALTVLPALLQLIL
HUMGLUT5 (126-209): ELIIISRLLVGI AGVSSNVVPMYLGELAPKNLRGALGVVPQLFITVGILVAQIFGL RNLL ANVD G WPILLGLTGVPAALQLLL
A_THALIANA (135-218): MLIVGRILLGFGIGFANQAVPLYLSEMAPYKYRGALNIGFQLSITIGILVAEVLNY FFAK IKGG W GWRLSLGGAVVPALIITIG
SYNECHOCYS (111-202): IFWRVLGGIGVGAASVIAPAYIAEVSPAHLRGRLGSLQQLAIVSGIFIALLSNWFIALMAGGSAQNPWLFGAAAWRWMFWTELIPALLYGV
ECOARAB/H+ (113-194): EMLIAARVVLGIAVGIASYTAPLYLSEMASENVRGKMISMVQLMVTLGIVLAFLSDTAFSYS GN WRAMLGVLALPAVLLIIL
ECOXYLO/H+ (127-216): PEFVIYRIIGGIGVGLASMLSPMYIAELAPAHIRGKLVSFNQFAIIFGQLLVY VNYFIARS GDAS WLNTDGWRYMFASE IPALLFLML
SCERE_GAL2 (175-258): QYFI GRIISGLGVGGIAVL PMLISEIAPKHLRGTLVS YQLMITAGIFLGY TNYGTKSYSNSVQ WRVPLGL FAWSLFMIGA
SCERE_SNF3 (196-280): TLLIVGRVISGIGIGAISAVVPLYQAEATHKSLRGAIISTYQWAITWGLLVSSAVSQGTHARNDASS YRIPIGLQYVWSSFLAIG
KLACT_LACP (169-253): LIGGRWFVAFFATIANAAAPTY AEVAPAHLRGKVAGLYNTLWSVGSIVAAFSTYGTNKNFPNSSK AFKIPLYLQMMFPGLV IF
Z_MOB_GLUT (118-207): QIF FFRFLAGLGIGVVSTLTPTYIAEIRPPDKRGQMVSGQQMAIVTGALIGYIFTWLLAHFGSIDWV NASGW WSPASEGLIGIAFLLL
KLACT_RAG1 (167-250): QYFIGRIISGLGVGGITVLSPMLISETAPKHLRGTLVS YQLMITFGIFLGY TNYGTKNYSNSVQ WRVPLGL FAWAIFMVLG
CKESS_GLUT (137-219): AMLIVGRVLLGFGVGLGSQVVPQYLSEVAPFSHRGMLNIGYQLFVTIGILIAGLVNYAVRDWENG WRLSLGLAAAPGAILFLG
LENRI_GLUT (214-308): WVLIVGRFVIGLFLGVI VA PVYTDQNAHPKWKRTIGVMFQVFTTLGIFVAALMGLALGQSIRFDHDGDQKVMARMQGL VFSTLFSLLTVVLG
TRYPAN_THT1(179-268): TGRVLIGLGVGIL SV PMYVNENAHPKL KMDGVLFQVFTTLGIMLAAMLGLILDKTGASKEEANMAGRLHVFSAVPLGLSVAMFLV G
N_CRASSQTR (125-211): DPIIAGRVLAGIGVGGASNMVPIYISELAPPAVRGRLVGIYELGWQIGGLVGFWINYGVNTTMAPIRS QWLIPFAVQLIPAGLLFLG
A_NIGQPERM (121-207): GLIVGGRVLAGIGVGAGSNI PIYISEMAPSAIRGRLVGVYELGWQIGGVVGFWINYGVDETLAPSHK QWIIPFAVQLIPAGLLIIG
(381 - 475)
HUMGLUT1 (204-290): LPF PESPRFLL INRNEENRAKSV LKKLRGTAD VTHDLQEMKEESRQMMREKKVTILELF RSPAYRQPILIAVVLQLSQQLSGINAV
RATGLUT1 (204-290): LPF PESPRFLL INRNEENRAKSV LKKLRGTAD VTRDLQEMKEEGRQMMREKKVTILELF RSPAYRQPILIAVVLQLSQQLSGINAV
MUSGLUT1 (204-290): LPF PESPRFLL INRNEENRAKSV LKKLRGTAD VTRDLQEMKEEGRQMMREKKVTILELF RSPAYRQPILIAVVLQLSQQLSGINAV
RABGLUT1 (204-290): LPL PESPRFLL INRNEENRAKSV LKKLRGNAD VTRDLQEMKEESRQMMREKKVTILELF RSPAYRQPILSAVVLQLSQQLSGINAV
PIGGLUT1 (163-249): LPF PESPRFLL INRNEENRAKSV LKKLRGTAD VTRDLQEMKEESRQMMREKKVTILELF RSAAYRQPILIAVVLQLSQQLSGINAV
BOSGLUT1 (204-290): LPF PESPRFLL INRNEENRAKSV LKKLRGTAD VTRDLQEMKEESRQMMREKKVTILELF RSAAYRQPILIAVVLQLSQQLSGINAV
HUMGLUT2 (236-322): LFF PESPRYLY IKLDEEVKAKQS LKRLRGYDD VTKDINEMRKERKEEASSEQKVSIIQLF TNSSYRQPILVALMLHVAQQFSGINGI
RATGLUT2 (234-320): LLF PESPRYLY LNLEEEVRAKKS LKRLRGTED ITKDINEMRKEKEEEASTEQKVSVIQLF TDPNYRQPIVVALMLHLAQQFSGINGI
MUSGLUT2 (235-321): LLF PESPRYLY IKLEEEVRAKKS LKRLRGTED VTKDINEMKKEKEEEASTEQKVSVIQLF TDANYRQPILVALMLHMAQQFSGINGI
HUMGLUT3 (202-288): LPF PESPRFLL INRKEEENAKQI LQRLWGTQD VSQDIQEMKDESARMSQEKQVTVLELF RVSSYRQPIIIAVVLQLSQQLSGINAV
HUMGLUT4 (220-306): LPF PESPRYLY IIQNLEGPARKS LKRLTGWAD VSGVLAELKDEKRKLERERPLSLLQLL GSRTHRQPLIIAVVLQLSQQLSGINAV
RATGLUT4 (220-306): LPF PESPRYLY IIRNLEGPARKS LKRLTGWAD VSDALAELKDEKRKLERERPLSLLQLL GSRTHRQPLIIAVVLQLSQQLSGINAV
MUSGLUT4 (222-308): LPF PESPRYLY IIRNLEGPARKS LKPLTGWAD VSDALAELKDEKRKLERERPMSLLQLL GSRTHRQPLIIAVVLQLSQQLSGINAV
HUMGLUT5 (210-296): LPFFPESPRYLL IQKKDEAAAKKA LQTLRGWDS VDREVAEIRQEDEAEKAAGFISVLKLF RMRSLRWQLLSIIVLMGGQQLSGVNAI
A_THALIANA (219-302): SLVLPDTPNSM IERGQHEEAKTK LRRIRGVDD VSQEFDDLVAASKESQSIEH PWRNLL R RKVRPHLTMAVMIPFFQQLTGINVI
SYNECHOCYS (203-285): AFLIPESPRYL VAQGQGEKAAAI LWKVEG GD VPSRIEEI QATVSLDHKPRFSDL LS RRGGLLPIVWIGMGLSALQQFVGINVI
ECOARAB/H+ (195-277): VVFLPNSPRWL AEKGRHIEAEEV LRMLRDTSE KARE ELNEIRESLKLKQGGWAL FK INRNVRRAVFLGMLLQAMQQFTGMNII
ECOXYLO/H+ (217-296): LYTVPESPRWL MSRGKQEQAEGI LRKIMGNTL ATQAVQEIKHSLDHGR KTGGRLLMF GVGV IVIGVMLSIFQQFVGINVV
SCERE_GAL2 (259-349): LTLVPESPRYL EVNKVEDAKRSIAKSHKVSPEDPAVQAELDLIMAGIEAEKLAGNASWGELF STKTKVFQRLLMGVFVQMFQQLTGNNYF
SCERE_SNF3 (281-374): MFFLPESPRYYV LKDKLDEAAKSLSFLRGVPVHDSGLLEELVEIKATYDYEASFGSSNFID FISSKSRPKQTLRMFTGIALQAFQQFSGINFI
KLACT_LACP (254-347): GWLIPESPRWLVGVGREEEAREFIIKYHLNGDRTHPLLDMEMAEIIESFHGTDLSNPLEMLDVR SLFRTRSDRYRAMLVILMAWFGQFSGNNV
Z_MOB_GLUT (208-288): LLTAPDTPHWL VMKGRHSEASKI LARLEPQAD PNLTIQKIKAGFDKAMDKSSAGLFAF GITVVFAGVSVAAFQQLVGINAV
KLACT_RAG1 (251-341): MMFVPESARFLV ETDQIEEARKSLAKTNKVSIDDPVVKYELLKIQSSIELEKAAGNASWGELI TGKPSMFRRTILMGIMIQSLQQLTGDNYF
CKESS_GLUT (220-305): SLVLPESPNFL VEKGKTEKGREV LQKLRGISEVDAEFADIVAAVEIARPIIMRQSWASLF IRRYMPQLLISFVIQFFQQFTGINAI
LENRI_GLUT (309-360): IVT RESRAKFDGGEEGRAELNPSEYGYVEMIPR LLMG VMAGTLQLTGINAV
TRYPAN_THT1(269-322): MFL RESTATFAQDDDGKADGGMDPNEYGWGQMLWP LFMGAVIAGTLQLTGINAV
N_CRASSQTR (212-304): SFWIPESPRWLY ANGKREEAMKVL WIRNLEPTDRYIVQEVSFIDADLERVTRQVGNGFWKPF LSLKQRKVQWRFFLGGMLFFWQNGSGINAI
A_NIGQPERM (208-301): ALLIRESPRWLF LRGNREKGIETLAWIRNLPADHIYMVEEINMIEQSLEQQRVKIGLGFWKPFKAAWTNKRILYRLFLGSMLFLWQNGSGINAI
(476 - 570)
HUMGLUT1 (291-371): FYYSTSIFEKAGV QQPVYATIGSGIVNTAFTVVSLFV VERAGRRTLHLIGLAGMAG AILMTIALALL E QLP WMSYLSIVA
RATGLUT1 (291-371): FYYSTSIFEKAGV QQPVYATIGSGIVNTAFTVVSLFV VERAGRRTLHLIGLAGMAG AVLMTIALALL E QLP WMSYLSIVA
MUSGLUT1 (291-371): FYYSTSIFEKAGV QQPVYATIGSGIVNTAFTVVSLFV VERAGRRTLHLIGLAGMAG AVLMTIALALL E RLP WMSYLSIVA
RABGLUT1 (291-371): FYYSTSIFEKAGV QQPVYATIGSGIVNTAFTVVSLFV VERAGRRTLHLIGLAGMAA AVLMTIALALL E QLP WMSYLSIVA
PIGGLUT1 (250-330): FYYSTSIFEKAGV QQPVYATIGSGIVNTAFTVVSLFV VERAGRRTLHLIGLAGMAG AVLMTIALALL E QLP WMSYLSIVA
BOSGLUT1 (291-371): FYYSTSIFEKAGV QQPVYATIGSGIVNTAFTVVSLFV VERAGRRTLHLIGLAGMAG AVLMTIALALL E RLP WMSYLSIVA
HUMGLUT2 (323-403): FYYSTSIFQTAGI SKPVYATIGVGAVNMVFTAVSVFL VEKAGRRSLFLIGMSGMFV AIFMSVGLVLL N KFS WMSYVSMIA
RATGLUT2 (321-401): FYYSTSIFQTAGI SQPVYATIGVGAINMIFTAVSVLL VEKAGRRTLFLAGMIGMFF AVFMSLGLVLL D KFT WMSYVSMTA
MUSGLUT2 (322-402): FYYSTTIFQTAGI SQPVYATIGVGAINMIFTAVSVLL VEKAGRRTLFLTGMIGMFF TIFMSVGLVLL D KFA WMSYVSMTA
HUMGLUT3 (289-369): FYYSTGIFKDAGV QEPIYATIGAGVVNTIFTVVSLFL VERAGRRTLHMIGLGGMAF STLMTVSLLLK D NYN GMSFV IGA
HUMGLUT4 (307-387): FYYSTSIFETAGV GQPAYATIGAGVVNTVFTLVSVLL VERAGRRTLHLLGLAGM G AILMTVALLLL E RVP AMSYVSIVA
RATGLUT4 (307-387): FYYSTSIFELAGV EQPAVATIGAGVVNTVFTLVSVLL VERAGRRTLHLLGLAGM G AILMTVALLLL E RVP SMSYVSIVA
MUSGLUT4 (309-389): FYYSTSIFESAGV GQPAYATIGAGVVNTVFTLVSVLL VERAGRRTLHLLGLAGM G AILMTVALLLL E RVP AMSYVSIVA
HUMGLUT5 (297-379): VYYADQIYLSAGVPE EHVQYVTAGTGAVNVVMTF AVFV VELLGRRLLLLLGFSI LIA VLTAALALQ D TVS WMPYISIV





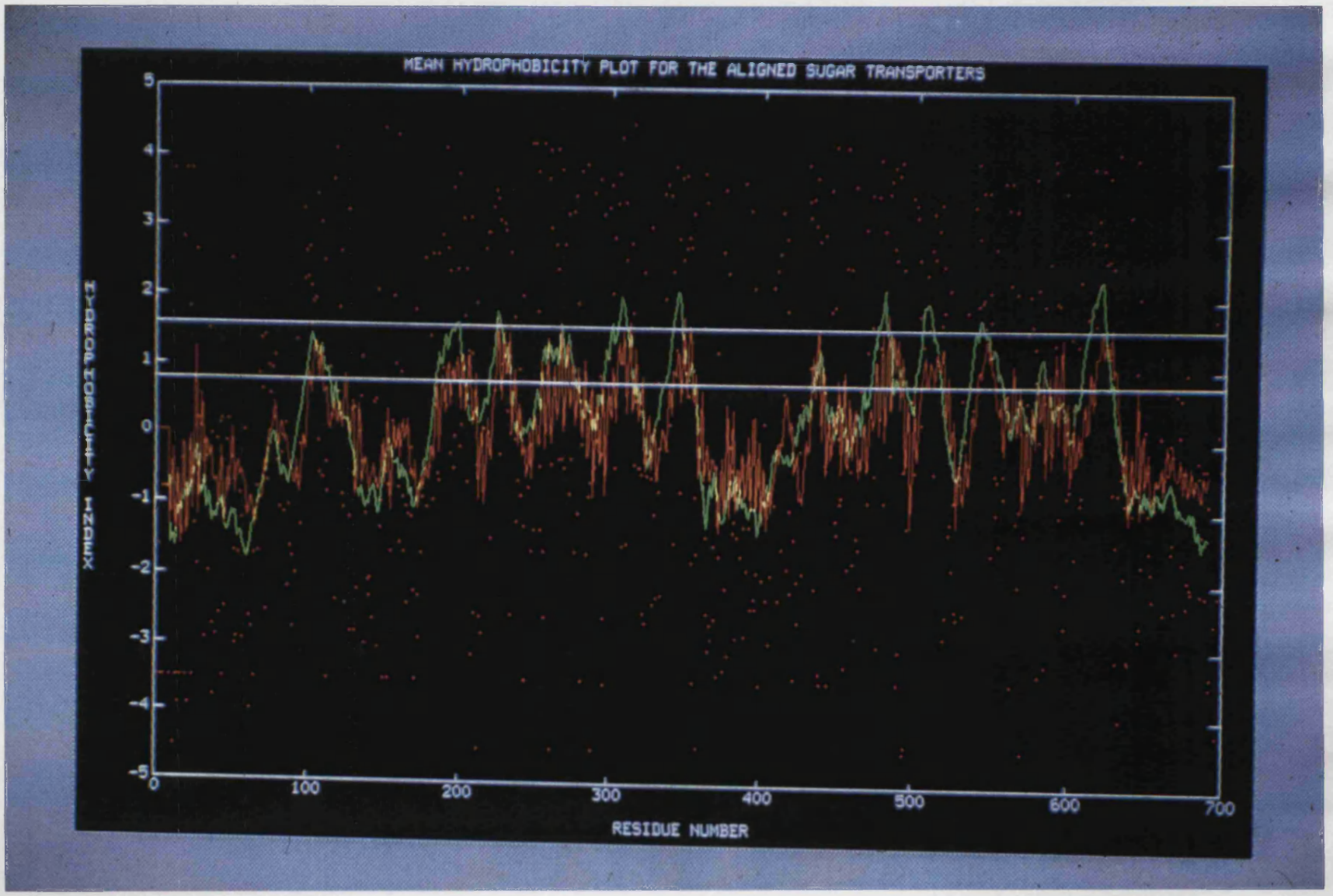
### 1.6.4 Molecular organization of glucose transporters

#### 1.6.4.1 Molecular cloning and characterization of glucose transporters

The advent of recombinant DNA technology enabled the isolation, cloning and sequence analysis of the structural genes encoding GTs. This was the first step in an attempt at rationalising the kinetics/phenomenology of transport as enunciated above, in terms of structure. The first of these was the Hepatoma G2 (HepG2) glucose transporter (156). This was cloned by using antiserum raised against the human red blood cell glucose transporter, on an expression library in lambda gt11. The predicted amino acid sequence of the encoded protein was identical to that of the red blood cell glucose transporter which was determined by N-terminal sequencing and carboxypeptidase digestion at the C-terminal. Structural analysis by fast atom bombardment (FAB), proteolysis and N-terminal sequencing of peptides generated thereof, indicated that not only was the protein hydrophobic, but it was also glycosylated. Its hydrophobic nature was confirmed by a Kyte-Doolittle hydropathy plot (157). Further characterization of the transmembrane topology using the Eisenberg algorithm (158) indicated 12 anti-parallel putative membrane spanning segments with an average hydrophobicity over 21 residues greater than 0.42. B-turns were predicted to loop between the 12 a-helices using the Chou-Fasman algorithm (159). Hence, an a-helix content of 70%, a B-strand component of 10%, and a B-turn contribution of 15% derived from Raman spectroscopy, have been assigned to the lac permease (160). Notably, TM3, TM5, TM7, TM8, and TM11 were amphipathic and contained serine, threonine, aspartic acid and asparagine residues which may be on the same face of the a-helix. These may provide hydroxyl and amide side-

chains to build an aqueous channel for glucose transport. This characterization is the prototype upon which other glucose transporters were cloned and founded, and provided impetus for the search and isolation of related genes (161-166).

Fig. 29 Hydrophobicity profile for the aligned sugar transporters. The plot is based on the algorithm of Kyte and Doolittle, using a scanning window of 19 residues (160). Transmembrane segments are represented by peaks (green line) at about 1.58. Potential amphipathic helices (red and yellow lines) span the line at about 0.8. The Hydrophilic segments have negative hydrophobicities.


MEAN HYDROPHOBICITY PLOT FOR THE ALIGNED SUGAR TRANSPORTERS
HYDROPHOBICITY INDEX
5
4
3
2
1
0
-1
-2
-3
-4
-5
0
100
200
300
400
500
600
700
RESIDUE NUMBER

All the mammalian GT genes that have been cloned contain introns, which span several kilo base pairs, but with identical exon-intron organization. This arrangement may be important as a mechanism for exon shuffling in the regulation, stage- and tissue-specific expression of the mammalian GTs.

Another feature of these transporters is a pseudo-symmetrical arrangement of the domains, which may be suggestive of the duplication of a primordial six membrane-spanning transporter. This is made all the more credible by the identities shared between TM1 and TM7, TM2 and TM8, TM3 and TM9, etc. Other workers have earlier tried to compare the N- and C-termini in an attempt at rationalizing the duplication concept (167). This, I consider, is inappropriate because the two termini are the ends of the protein and should be viewed as such.

Subsequent searches using heterologous gene probes on the gene libraries of other organisms ranging from E. coli to plants (142,168,169), identified genes whose products were homologous at the amino acid level and were predicted to assume identical secondary structures (145). Based on these structural identities with the cloned GTs, a presumptive membrane protein cloned from L. enrietti has been assigned to this family of membrane proteins although its ligand is at best conjectural (56). From the outset, there appears to be a major discrepancy between the sizes of the predicted amino acid sequence and that deciphered from affinity labelling with cytB (170), notwithstanding reducing SDS-PAGE conditions, proteolysis or biosynthetic modifications such as glycosylation.

The function of the products of these genes was variously expressed: in

E. coli (162), Xenopus laevis oocytes (163), or as complements to sugar transporter lesions (168). Others used functional expression as the tack to isolate the related transporter, Na+/glucose cotransporter, from intestinal brush border using Xenopus oocytes (171). Significantly, the predicted amino acid sequence and secondary structure of this transporter was different from the family of facilitated or proton-coupled transporters.

Having established a relationship between transport and transporters, a solid basis was provided for further structural characterization of the transporters in relation to function.

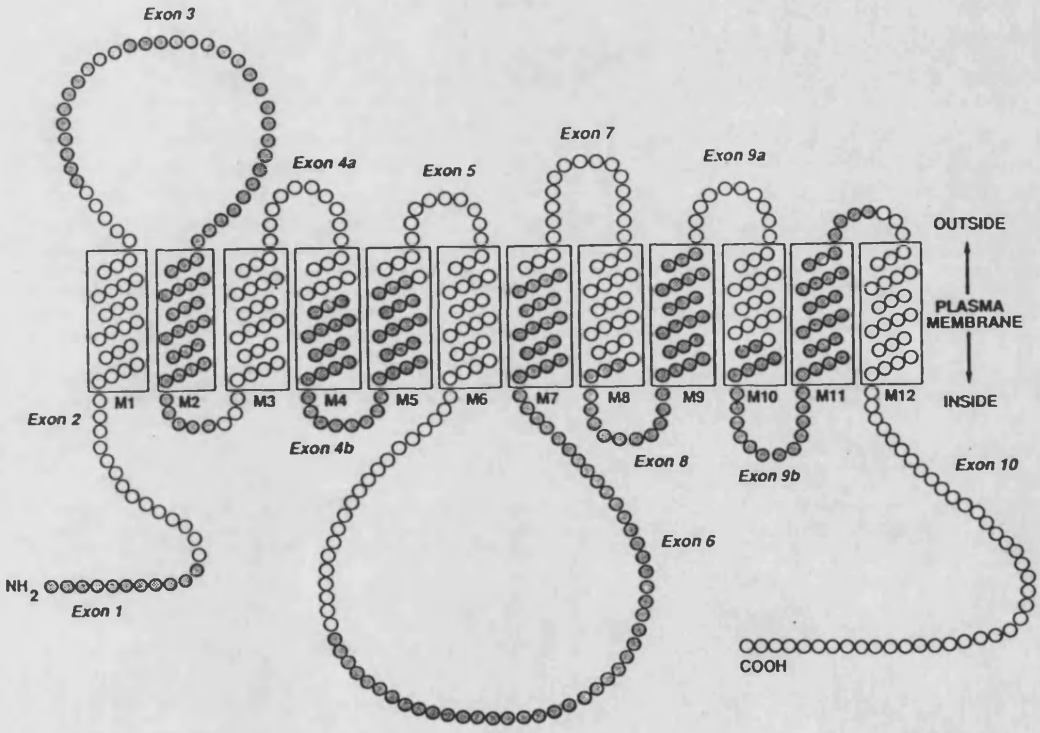



Fig. 30 Molecular organization of the (representative) facilitated GT. Exon-intron junctions are shown between light and dark circles (amino acids). a-helices are numbered from M1-M12. Reproduced from (170).

| Intron Number | Exon Size | Location | Donor | Intron Size | Acceptor |
|---|---|---|---|---|---|
| | | | SerLys | | LysVal |
| I | 226 bps | 226 | AGCAAGgtgagt... | - ~10,000 bps - | ...ttttttccctctgcagAAGGTG |
| | | | GlnLys | | ValIle |
| II | 96 bps | 322 | CAGAAGgtaagt... | - ~13,000 bps - | ...gctcttctccctgcagGTAATT |
| | | | eGlyAr | | gArgAs |
| III | 161 bps | 483 | TGGCAGgtaagc... | - 145 bps - | ...tgcgacccattcccagGCGGAA |
| | | | AlaGln | | ValPhe |
| IV | 241 bps | 724 | GCCCAGgtaagc... | - 684 bps - | ...cctagttcttgtccagGTGTTC |
| | | | ysSerV | | alLeuL |
| V | 163 bps | 887 | AGAGTGgtacgg... | - 90 bps - | ...cctcoggtcccaccagTGCTGA |
| | | | AsnAla | | ValPhe |
| VI | 188 bps | 1075 | AATGCTgtgagt... | - 278 bps - | ...tctccctttctggcagGTGTTC |
| | | | ValSer | | LeuPhe |
| VII | 105 bps | 1180 | GTGTCGgtaagt... | - 398 bps - | ...tttctcctttccacagCTGTTC |
| | | | LeuLeu | | GluGln |
| VIII | 102 bps | 1282 | CTGCTGgtgagt... | - 1228 bps - | ...cattctccctcaacagGAGCAG |
| | | | ValGlu | | GlnLeu |
| IX | 204 bps | 1486 | GTGGAGgtgagc... | - 274 bps - | ...tgtgtttttttccagCAACTG |
| | 1089 bps | 2575 | TAGAAA | | |

Fig. 31 Exon-intron boundaries of the rat glucose transporter gene. Exon sequences are indicated by capital letters, and introns by small letters. Amino acids are shown above the corresponding codons. Splice donor and acceptor sites are indicated, and fit the consensus splice junction sequences (101). Reproduced from (172).

### 1.6.4.2 Structural characterization of glucose transporters

The topology of membrane spanning proteins and for that matter glucose transporters, has been determined from a combination of biochemical and molecular approaches. The use of limited proteolysis involving tryptic cleavage for instance would not only locate accessible or exposed residues, but also those regions of relatively high mobility or fluidity (173). This would be expected of surface loop regions between helices or similarly exposed segments lacking defined secondary structures. Hence such cleavage would be invaluable in delineating structural domains. By this approach, two membrane-associated fragments were obtained from the tryptic cleavage of GLUT1, as trypsin cleaved on the cytoplasmic face of the transporter. Of these two fragments, one bore the N-linked glycosylated residue and the other contained the cytB binding site. In addition, labelling this transporter with ASA-BMPA and cytB, followed by proteolytic digestion with either chymotrypsin or subtilisin yielded a 6-7kD fragment which was labelled with either ligand (128). This enabled the localization of ASA-BMPA and cytB binding sites to between M9 and M11, after chemical cleavage at cysteine and tyrosine residues using 2-nitro-5-thiocyanobenzoic acid and N-bromosuccinimide, respectively in the erythrocyte GT (128,174,175).

Monoclonal antibodies have also been extremely helpful in this regard. These antibodies would only react with exposed epitopes which are the structural features of the protein. The specificity of this interaction makes this approach important in evaluating structure-function relationships. A combined strategy of limited proteolysis and epitope mapping would identify epitopic structural regions associated with certain

functions when combined with immunoblot analyses. It was therefore possible to establish that the C-terminal of the E. coli lactose and the red cell glucose transporters were cytoplasmic (176,177). The antigenicity of a structural feature may therefore be correlated with hydrophilicity (178), surface protrusions (179,180), and mobility (181), even though the hydrophilic character, propensity to form turns and mobility may be predicted from the amino acid sequence a priori.

However, the major drawback to these methods is that with polytopic membrane proteins (i.e proteins with multiple spanning segments) such as the GTs, some hydrophilic domains, especially small ones may not be accessible to cleavage or antibody. Therefore, a molecular genetic approach was adopted (182). Based on the fact that alkaline phosphatase (phoA) is enzymatically active only when it is exported to the periplasm and inactive in the cytoplasm, it was possible to delineate the orientations of specific segments of the gene sequence of the maltose transporter. Fusions to periplasmic domains resulted in high phoA activity. These findings were corroborated by protease-accessibility studies in relation to structure and function (183). Similar results were obtained with B-galactosidase fusions (which show increased activity only when B-galactosidase is cytoplasmic) and fusions with B-lactamase which confers ampicillin resistance only when it is exported to the periplasm (184). No fusions of this nature have been performed with the mammalian or lower eukaryote GTs even though this is conceptually possible. Hence, the coordinate use of reporter enzyme/gene fusions, proteolysis and immunoblotting, may enhance an accurate deduction of the topology of these proteins.

### 1.6.4.3 The domain organization of glucose transporters

From the earlier studies of bacteriorhodopsin (185,186), it has been found that integral membrane proteins derive their transmembrane topology from the spatial arrangement of a-helices and b-turns. This was the basis for the structural model of the first glucose transporter to be characterized (156).

All the GTs studied conform to a common domain structure based on the degree of homology that exists between them, homology of residues that may subserve similar functions and may therefore assume identical spatial dispositions within the protein.

Three major domains can be deciphered, but a fourth domain of leucine zippers is also gaining recognition.

a. The 12 a-helices which constitute the polypeptide backbone with the N- and C-terminals bathed in the cytoplasm. These helices are built of hydrophobic residues of 18-21 (the equivalent of 30 A), each of which is long enough to span the lipid bilayer. These residues are often followed by positively charged residues at the cytoplasmic side. [The nature of the factors dictating the orientation and stability of these helices is not clear, though it is thought that a dipole-dipole interaction may be at play as a result of the hairpin anti-parallel arrangement of the helices, aside from the propensities or conformational preferences of the constituent amino acids (187)]. This morphological model subsumes a channel between the helices with charged residues oriented inward and the non-polar ones outward. There are however, some regions of low hydrophobicity especially TM5 and TM8. These regions are coincident with regions of high hydrophobic

moments and with channel forming functions. This has been the observation in bacteriorhodopsin, rhodopsin and the acetylcholine receptor (188).

b. An exofacial loop of about 33 amino acids between TM1 and TM2, with a site for N-linked glycosylation at Asp 45. This position is not however universal among the GTs. It is peculiar to the mammalian GTs. The alternative arrangement is a potential glycosylation site in the loop between TM6 and TM7.

c. A hydrophilic loop of about 65 amino acids (on the average) joining TM6 and TM7 and bathed in the cytoplasm. The length of this region varies from one transporter to the other.

d. Leucine-zipper motif within TM2.

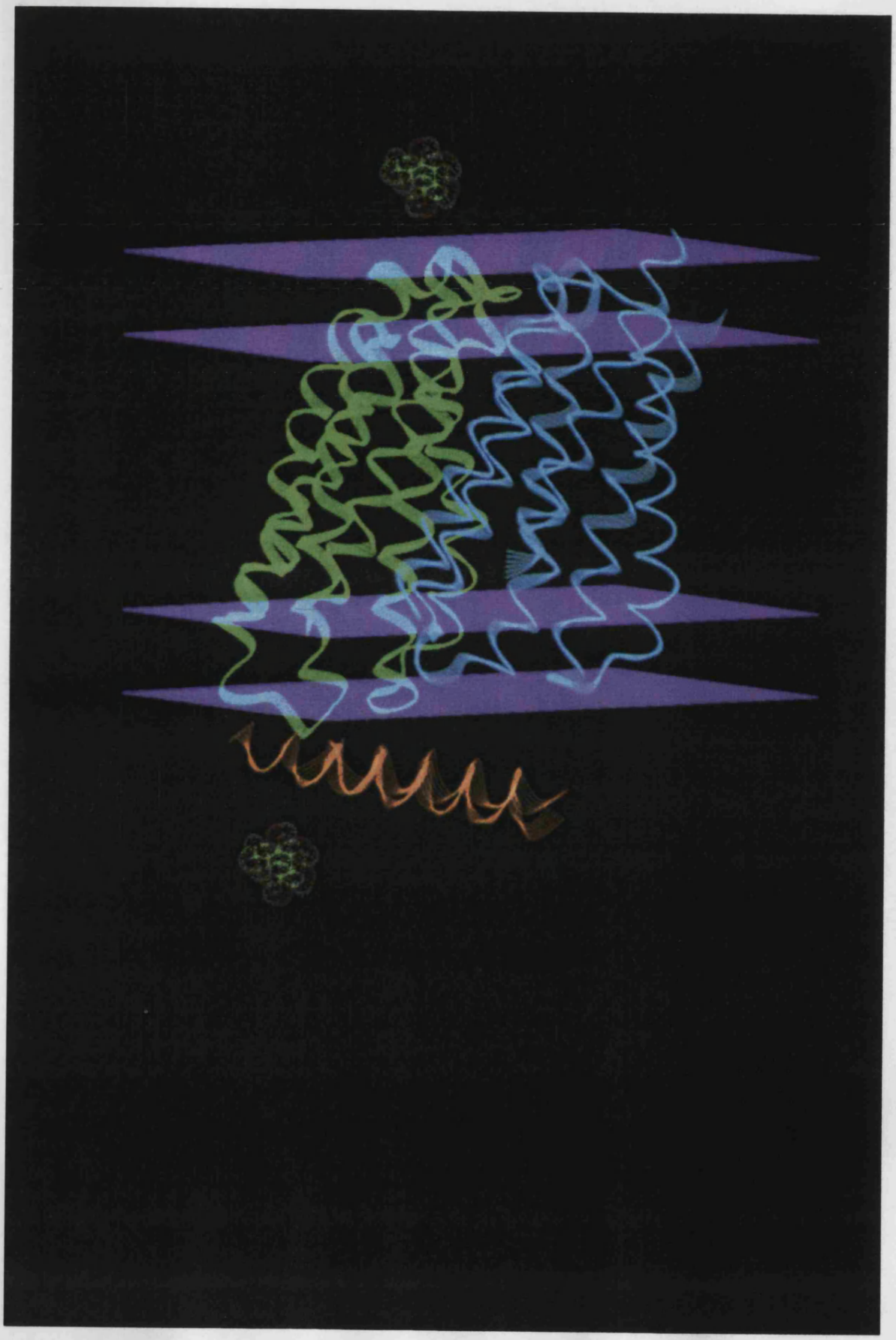
This is a recent addition to the domain structure of the GTs. Leucine-zippers have earlier been associated with the dimerization of DNA binding proteins (189-191). They have recently been found to be associated with or adjacent to membrane spanning regions of transport or receptor proteins (192). Their presence in the $K^+$-channels is thought to be linked to subunit interactions that may mediate voltage-dependent gating of the channel (192). The same motif has been found in TM2 of some of the glucose transporters where there is a leucine or isoleucine at every other seventh residue for at least eight turns of an a-helix (192). Radiation inactivation studies of intact red cell membranes suggest that the native GT may exist as a homotetramer. Indeed, phorbol ester-induced in vivo phosphorylation of these cells identified a band that ran at 80Kd (193). Cuppoletti and Jung (1980) also found that radiation inactivation of the erythrocyte transporter yielded a multimer of 185kD, indicating that this GT probably exists as a

multimer in the native state and in the absence of substrate (194). Similar conclusions were made from inactivation of the lac permease of E. coli. (195). These observations may lend credence to the possible role of the leucine-zipper in maintaining the transporter in an oligomeric form. The potential of this motif to act as a swivel point in the GT monomer-dimer equilibrium in response to cellular signals, is an attractive one in terms of transport regulation. It has been reported that galactose and maltose induce a monomerization of their cognate binding proteins transporters; these are otherwise purified as dimers in their native form (196). Furthermore, the affinity of the ligand was found to vary directly with the degree of monomerization, showing a preferential affinity for the monomer rather than the dimer. Although this feature is not universal among the GTs, the aforementioned observation may spell differences in transport regulation within the superfamily.

| Transporter | Sequence | Reference |
|---|---|---|
| CEF | tsLwslsvaIfsvggmIgsfsvsLff | 197 |
| HepG2 | ttLwslsvaIfsvggmIgsfsvgLfv | 156 |
| Ratb | ttLwslsvaIfsvggmIgsfsvgLfv | 161 |
| IRGT | ttLwalsvaIfsvggmIssfligIis | 163 |
| Arab/H+ | evLtsrlqewvvssmmLgaaigaLfn | 143 |
| GAL2 | hyLsnvrtgLivaifnIgcafggIil | 148 |
| Klact | hyLsnvrtglivsifnIgcavggIvl | 198 |

Fig. 32 Comparison of leucine-zippers from various glucose transporters: CEF, chicken embryo fibroblasts; Ratb, rat brain glucose transporter (GLUT1); IRGT, insulin-regulatable glucose transporter; HepG2, human hepatoma transporter (GLUT1); Arab/H+, E. coli arabinose/H+ symporter; GAL2, galactose transporter from Saccharomyces cerevisiae; Klact, lactose transporter from Kluyveromyces lactis. Leucine and Isoleucine identities are in capital letters as shown.

Fig. 33 Domain organization/model of the glucose transporter. a-helices are represented by ribbons. The entire transporter is viewed as two halves of six helices each, and encaving the glucose pocket. The glucose molecule is seen docking and leaving the GT in the cytosol. The yellow ribbon represents the long intracellular loop connecting TM6 and TM7. Kindly provided by Dr. P. Hodgson, Molecular Graphics Unit, Bath University.



## 1.7 AIMS OF THIS STUDY

Baseline studies on glucose transport in T. brucei indicate that there is a stereospecific, saturable transporter. Its substrate specificity is akin to those of the mammalian hosts. However, their kinetics of transport differ significantly. These differences are thought to reside in the structures of their cognate transporters. Hence the primary aim of this study is to clone the gene encoding the trypanosome glucose transporter as a prelude to understanding the molecular or structural basis for the observed differences in specificity and kinetics between that transporter and that of the mammalian host. Since the trypanosome is absolutely dependent on the glucose transporter for its survival within the mammalian bloodstream, it is hoped that the differences in their structures might be exploited to synthesise selective trypanocidal drugs in the long term.

## CHAPTER 2

## 2.0 Functional identification of the trypanosome glucose transporter

### Materials and Methods

### 2.1 Infection of rats

Previously cryopreserved stabilates of T.b.brucei clone MiTat 1.1, were resuspended in 1% glucose/3% sucrose in Krebs Ringer Phosphate Buffer (KRPB) pH8.0 to give about $10^8$ viable cells/ml in a Neubauer counting chamber. The cells were then injected intraperitoneally at doses of $1\text{-}2x10^7$ cells per 300g male Wistar rat. Parasitaemia was allowed to develop over 72hrs. Infection was verified by carrying out a tail count prior to exsanguination, which would give an approximate level of parasite load per rat.

### 2.1.1 Tail count

The tip of the rat tail was snipped off with a pair of scissors to provide about 5-50ul of blood. This was then appropriately diluted into KRPB/glucose/sucrose, and counted in a Neubauer chamber.

### 2.1.2 Harvest and purification of trypanosomes

Rats were exsanguinated under diethyl ether anaesthesia, into 1ml heparin (200u/ml) in 10-20ml syringes. Bleeding was done from the iliac artery at the point of bifurcation. They were laid dorsally, the abdomen cut open, and the viscera drawn aside to reveal the bifurcation. Care was taken to avoid coagulation by mixing with heparin

gently. All tubes and blood were kept on ice until they were centrifuged at 2800rpm and 4°C in conical propylene tubes using a bench top refrigerated centrifuge (IEC-Centra 3R) for 10mins. The upper layer (plasma) was aspirated and the trypanosomes, which appeared as a white band between plasma and red blood cells, were resuspended in KRPB/glucose/sucrose. They were recentrifuged as was necessary to remove contaminating blood.

Purification followed the method of Lanham (199). Briefly, trypanosomes were run down a DEAE-cellulose column (5x1.5cm) equilibrated in KRPB supplemented with 10mM glucose and eluted with the same. Elution was monitored by observing a decrease in opalescence of the effluent. The cells were then counted as previously and total cell harvest calculated from effluent volume, cell dilution factor and volume of the counting chamber.

### 2.1.3 Preparation of stabilates

Some of the harvested blood was resuspended in 15% glycerol (w/v) with gentle agitation to ensure aeration, and taken up by capillary action into haematocrit tubes. Both ends of the tubes were carefully sealed off first with wax and then on a Bunsen flame; all tubes were placed in a flask of liquid nitrogen to equilibrate for 2-3hrs before being transferred for long-term storage in a jar of the same.

### 2.1.4 Metabolic study: rate of glycolysis

The rate of energy metabolism was measured using the Clark (Oxygen) electrode as previously described (200).

The effect of cytB (up to 50uM concentration) was similarly investigated. Chart recorder speeds were adjusted appropriately and points of substrate or inhibitor addition carefully noted for each batch of cells.

### 2.1.5 Reconstitution of the trypanosome glucose transporter

#### 2.1.5.1 Purification of plasma membranes

The procedure for the isolation and purification of plasma membranes followed that of Voorheis (201), based on mechanical shearing of cells under osmotic swelling conditions.

#### 2.1.5.2 Protein determinations

These were done by the BioRad method. Membranes and bovine serum albumin standard were solubilized in 0.5% (final concentration) MEGA-10. Absorbances were measured at 595nm and sample protein contents determined from a calibration curve of the standard.

### 2.1.5.3 Reconstitution

The method of reconstitution by detergent dilution followed that of Kasahara and Hinkle (202) with some modifications.

Trypanosome plasma membranes were solubilized in 5% MEGA-10 to give a final concentration of 1% detergent and 1mg/ml of membrane protein. 30mg of E. coli L-a-phosphatidylethanolamine was dissolved in 1ml of 100mM sodium phosphate buffer pH7.4, containing 1% MEGA-10, and sonicated under an atmosphere of nitrogen for 1-3hrs (or until dissolution was total) in a bath-type Decon FS100 sonicator. The sample was centrifuged at 20000rpm for 30mins and 4°C. The solubilized membrane was similarly treated. 450ul of membrane solubilate was mixed with the phospholipid solution in a 4ml polyethylene tube and allowed to stand on ice for 30-60mins to allow vesicles to form. The mixture was then diluted to 4ml with assay (phosphate) buffer, and the vesicles used immediately or stored at -70°C until required.

### 2.1.5.4 Infinite-trans/counterflow vesicular transport

Frozen proteoliposomes were thawed slowly at ambient temperature and subsequently centrifuged at 20000rpm for 30mins at 4°C. Detergent was removed by cyclical spinning with 2.5ml volumes of assay buffer. Vesicles were resuspended in 500ul of 200mM D-glucose. After another centrifugation, the pellet was taken up in 50ul of 100mM D-glucose.

Transport assay was initiated by adding 5ul of preloaded vesicles to 500ul of assay buffer containing 4uCi $^{14}$C-D-glucose, and mixed rapidly, carefully but thoroughly. This produced an internal concentration of 100mM inside the vesicles and 1mM outside. Uptake of

tracer/label was monitored over a time-course of 1min. All time-points were run in duplicate. Each uptake was terminated with 1ml 10uM (0.1%) phloridzin in assay buffer. After quenching, the reaction mixture was immediately filtered (under 5secs) through 0.22um Millipore filters previously soaked in stopper, and mounted on a filtration turret linked to a suction pressure tap. Filters were washed 2X with 10ml of stopper, and allowed to dry briefly before being transferred to scintillation vials. Background or zero-time point values were obtained by adding vesicles to the stopper before adding the label.

Scintillation cocktail (5ml) was added to each filter and left to dissolve overnight. They were then counted and uptake values obtained by deducting ackground values from those at the respective time-points.

## 2.2 Molecular cloning of the trypanosome glucose transporter

### 2.2.1 Materials and Methods

Two trypanosome gene libraries- a complementary DNA (cDNA) library in the bacteriophage vector lambda gt11, and a genomic library in lambda EMBL4 were gifts from Dr. J.D. Barry, University of Glasgow. A glucose transporter cDNA from rat liver was obtained from Dr. Harvey F. Lodish (MIT/Whitehead Institute, Cambridge, Mass. USA). COS-7 and pSVL were kind gifts from Dr. Adrian Wolstenholme, Department of Biochemistry, Bath University.

Restriction endonucleases, DNA modifying enzymes, IPTG, X-gal, DNA markers (Pst and HindIII digested lambda DNA) were purchased from Northumbria Biologicals Limited (NBL), Cramlington, UK. Cloning vectors- M13 and pUC, ribonucleotides (rNTPs), deoxy NTPs and dideoxy NTPs, T7 polymerase DNA sequencing kit, Escherichia coli poly(A) polymerase, Sepharose CL6B, PD-10 columns, were obtained from Pharmacia LKB, Uppsala, Sweden; pBluescript SK(+/-), T3 and T7 polymerases, helper phages VCSM13 and R408 were from Stratagene; pGEM7 was from Promega, Madison, USA; all bacterial culture media (bacto-agar, bactotryptone, yeast extract) were from Difco, Michigan, USA; mammalian cell culture media, foetal calf serum, trypsin/EDTA and lipofectin were obtained from GIBCO BRL, Gaithersburg, USA; cell culture dishes, bottles, tubes, pipettes were from Sterilin, Nunc, and Costar; PBS tablets were purchased from Oxoid, Basingstoke, U.K; ampicillin, agarose (normal and low melting) were from Sigma, Poole, UK; radioactive isotopes 35-S ATP (1000 Ci/mmol), 35-S methionine, a- and g-32P (5000 Ci/mmol), NEN GeneScreen Plus hybridization

membranes, NENsorb columns and En$^3$Hance were obtained from NEN (Du Pont), Stevenage, UK; Hybond N and N+, random primer labelling kit from Amersham International, Bucks., UK; acrylamide Sequagel kits were from National Diagnostics, New Jersey, USA; Polyallomer Quick-seal ultracentrifuge tubes were obtained from Beckman Instruments Inc., Palo Alto, USA; Sequenase kits were purchased from United States Biochemical Corporation, Cleveland, Ohio, USA; Polaroid films were from Polaroid Corporation, Cambridge, USA; Millex-GS 0.22um filters were from Millipore S.A, France; DE-81 and chromatography paper (46x57cm), pH strips were from Whatmann Ltd, Maidstone, UK; redistilled phenol was from Rathburn Chemicals Ltd., Walkerburn, UK; all other chemicals of analytical grade (Electran grade ammonium persulphate, acrylamide, and N,N,N'N'-tetramethylene diamine, TEMED) were obtained from BDH Chemicals Ltd, Poole, UK, Fisons, Loughborough, UK and Sigma, UK.

### 2.2.2 Isolation and purification of trypanosome genomic DNA

High molecular weight genomic DNA was prepared from bloodstream form trypanosomes purified as above. The cells were centrifuged at 1000g for 5mins and the KRPB decanted. The cells were resuspended in Proteinase K buffer (PK 1mg/ml, TE pH 7.4, 150mM NaCl and 0.4% SDS) at $10^6$-$10^8$ cells/ml and incubated at 37°C overnight in a 50ml Falcon tube. The DNA solution was extracted with 1vol Tris-saturated phenol by a series of gentle inversions. The phases were separated by spinning the solution at room temperature in a bench centrifuge. The upper phase was transferred to a fresh tube using a 1ml tip with the bottom half cut off to provide a wide bore. If the solution was

sufficiently viscous, it was decanted. This phase was again extracted with chloroform:isoamyl alcohol (24:1) and this was repeated twice. The sample was centrifuged and the aqueous phase transferred to a Corex tube and mixed with 2.5vols of absolute ethanol and 1/10th volume of 3M NaAc pH 5.2. The precipitated DNA was spooled out with a glass rod, washed with 70% ethanol and dried under a vacuum. It was then dissolved in TE at 4°C overnight.

### 2.2.3 Estimation of yield and purity of DNA

The amount of DNA isolated was measured using a double beam UV spectrophotometer. The number of optical density (OD) units obtained at 260nm was multiplied by 50ug (for double strand DNA) or 37ug for single strand DNA or oligonucleotides. Purity was based on the ratio of the absorbances at 260/280 nm. A value of 1.8-2.0 for double strand templates was an indication of acceptable purity. DNA with a purity below this was re-extracted with phenol:chloroform and reprecipitated with ethanol.

### 2.2.4 Restriction analysis of genomic DNA

The DNA was digested with various restriction endonucleases either singly or in combination, at 2-5U/ug of DNA for 2-3hrs in a 37°C water bath. The samples were then electrophoresed at 100V in 1X TBE on a 0.8% agarose gel containing 1ug/ml EtBr. The integrity of the DNA was checked by running uncut trypanosome genomic DNA in parallel with uncut lambda DNA, and HindIII or Pst - cut lambda size

markers. After electrophoresis, the DNA fragments were visualised on a UV transluminator. A photograph of the gel was taken where necessary.

### 2.2.5 Southern transfer and hybridization of immobilized DNA

The gel was transferred to a sandwich box and the DNA fragments depurinated in 0.25M HCl for 30mins at room temperature with gentle rocking, to facilitate the transfer of large fragments. The acid was subsequently decanted and replaced with 5-10 vols of denaturing solution, and then with neutralizing solution for 30mins each time. This was poured off and the gel rinsed in several vols of distilled water. Nylon membrane (Hybond-N) was cut to size, soaked for 5mins in distilled water and then in 20X SSC. The nylon membrane was placed on the vacuum blotter and the gel placed upon the membrane. Blotting was done at 60-80mm Hg pressure for 2-6hrs (preferably overnight) and aided by layering 20-50mls 20X SSC on the gel. The traditional system of Southern transfer (203) was used on occasion, using a platform to support the gel and a wick of four layers of 3MM Whatman paper upon which the gel rests. Another set of 3MM paper was placed over the gel and the uncovered areas around the gel were covered with cling film to avoid the buffer being soaked up by the blotting tissue. The latter was placed about 10cm high on top of which a weight of about 500g was placed. Transfer was allowed onto the nylon membrane by upward capillary action in 20X SSC over several hours. Following transfer in either method, the set up was dismantled and the membrane rinsed in 2X SSC, blotted on 3MM paper, air-dried and baked for 2hrs at 80°C to fix the transferred fragments.

The membrane was prehybridized in 50% deionized formamide/1% SDS/5X Denhardts/5X SSC/5mM EDTA pH 7.5/50mM sodium phosphate buffer pH 7.0 and 250ug per ml of salmon sperm DNA, at 42°C overnight. Hybridization was carried out in the same buffer by adding random primer-labelled rat liver glucose transporter cDNA at the same temperature for 16-24 hrs with shaking. The membrane was then washed in steps of increasing stringency from 5X SSC/0.5% SDS at room temperature for 15mins (X2); 2X SSC/0.5% SDS at 42°C (2X 15mins); to 1X SSC/0.5% SDS at 55°C for 30mins. It was then rinsed in 2X SSC and blotted on 3MM, sealed in cling film and exposed to presensitized film for 24hrs. The film was developed and hybridizing fragments identified and sized.

### 2.2.6 Isolation and purification of specific DNA fragments

The freeze-squeeze technique was routinely used for the isolation and purification of DNA fragments following their identification on a Southern blot.

A 1% LMP agarose gel was prepared using Tris-acetate-EDTA (TAE) buffer. After electrophoresis and visualization, the band of interest was cut out in a slice of gel using a sharp razor blade. The slice was taken in a microfuge tube plugged with glass wool and frozen in liquid nitrogen for 2-5mins. The tube was then quickly centrifuged over a 4ml polypropylene tube in a microfuge. The eluted DNA was then phenol-chloroformed and ethanol-precipitated as previously described. It was further purified by any of the methods for purifying probes as described below, though at this stage, it was good enough for restriction or ligations after ethanol precipitation.

### 2.2.7 Construction of an enriched plasmid library

Following Southern transfer, hybridizing fragments of genomic DNA were located with the help of lambda markers. A restriction digest in which the fragments had asymmetrical ends was selected and a large volume of this electrophoresed in 0.8% low melting point agarose gel and 1X TAE buffer. The gel fraction containing those fragments was cut out in a slice of agarose. The fragments were purified as in 2.2.6 above.

The plasmid vector, pUC 19 was digested with the same enzymes as those yielding the hybridizing fragments (BamH1 and EcoR1). The cut vector was purified and ligated with the genomic fragments overnight in 1X ligase buffer and 16U T4 ligase in a total volume of 100ul, at 15°C. The ligation mixture was then centrifuged through a Sepharose CL6B column to remove buffer salts, and the eluate collected in a 1.5ml microfuge tube. Hence a trypanosome library enriched for fragments hybridizing to the rat glucose transporter cDNA was generated.

## 2.3 Preparation of probes

### 2.3.1 Random-primer extension/oligo labelling of cDNA

The rat liver cDNA probe was purified as a 1.5kb fragment (containing more than 95% of the reading frame) from the vector pGEM7 by restriction with EcoR1 and freeze-squeezing as earlier described. This was radiolabeled according to the method of Feinberg and Vogelstein (204). Briefly, 100-200ng of the cDNA was heat-denatured, snap-cooled on ice and labelled with the random-primer DNA labelling kit and 50-100uCi of a-32P using the Klenow fragment of DNA polymerase 1, and incubated for 2-3hrs at room temperature, as indicated in the manufacturers protocol. Unincorporated

radiolabel was removed on a spun mini- column of DE-52.

### 2.3.2 Design and synthesis of oligodeoxynucelotides

Oligonucleotides were designed from various conserved motifs of the glucose transporter superfamily and based on the codon usage of T. brucei genes. Basically, the relevant amino acid residues were reverse translated into DNA taking the codon preferences into account and allowing for some degree of redundancy, as indicated below.

Synthesis was based on the solid phase B-cyanoethyl phosphoramidite chemistry of Caruthers (205) using an Applied Biosystems DNA Synthesizer Model 381A.

The oligonucleotide was cleaved from the solid support by a 1-2hr elution from the column with 1ml of concentrated ammonium hydroxide. The DNA solution was then incubated overnight at 55°C to remove protecting groups on the exocyclic amines of the bases, and then lyophilized prior to dissolving in 1ml of water and precipitating with 1/10 volume of 3M NaAc pH5.2 and 3vols of absolute ethanol at -20°C for 1hr. The DNA was spun down for 15mins at 12,000 rpm, 4°C and washed with 70% ethanol before drying under vacuum for 15mins. It was redissolved in 1ml of water. Occasionally, the lyophilization step was omitted and the crude DNA was precipitated from the ammonia solution and contaminating deprotection products as above, and used without purification.

**Oligonucleotide probes:**

```
Oligo 1 : CCC GA(A/G) AGC CCN CGC T(T/A)(C/T) (C/T)T
Peptide :  P     E     S   P   R      Y/F       L
```

---

```
Oligo 2 : CA(G/A) CA(G/A) (C/T)TN ACN GGN AT(A/C/T) AA(C/T)
Peptide : Q       Q         L    T    G    I          N
Oligo 2 (cont'd): GCN GTN ATG TA(C/T) TA(C/T) GCN C
Peptide :          A   V   M    Y       Y      A
```

---

```
Oligo 3 : GTN GAA CGC GCN GGN CGC CGC ACN (C/T)T
Peptide : V   E   R   A   G   R   R   T    L
```

---

```
Oligo 4 : CAA ATA GGT CCA GGT CCA ATA CCA TGG TTT
Peptide : Q   I   G   P   G   P   I   P   W   F
```

---

```
Oligo 5 : GT(A/T) CC(A/T) GAA ACT AAA
Peptide : V       P       E   T   K
```

---

These oligonucleotides were synthesised from a consensus of the aligned glucose transporters. These regions are indicated in **Fig. 28**. Oligo 4 was initially designed for the Plasmodium GT, employing that codon usage. Obviously, this was inappropriate. In addition this was a wrong choice of consensus as this region is very poorly conserved in the prokaryote and lower eukaryote GTs.

### 2.3.3 5' end-labelling of oligonucleotides

Oligonucleotide (20-100pmol) were labelled with gamma 32-P in the presence of T4 polynucleotide kinase and 1X kinase buffer in a total reaction volume of 50ul. The reaction mix was briefly centrifuged in a microfuge and incubated for 45mins at 37°C. Labelling was terminated by adding 0.5M EDTA to 1mM final concentration or heating at 65°C for 10mins.

### 2.3.4 Purification of DNA probes

#### 2.3.4.1 Purification on Quickspin columns

A DEAE-Sephadex 50 mini-column of 0.5ml bed volume was prepared in a 1ml syringe. The column was washed twice with distilled water by centrifugation at 2000rpm for 3mins. The probe was added to the column and recentrifuged. Unincorporated label remained on the column and the labelled DNA was collected in a propylene tube.

#### 2.3.4.2 Ethanol precipitation

0.5vol of 5M ammonium acetate was added to the DNA sample and precipitated with absolute ethanol on dry ice for 15mins. The probe was collected by spinning at 15000rpm for 15mins, 4°C and was washed with 70% ethanol, briefly dried and resuspended in 20ul TE pH 8.0.

### 2.3.4.3 Denaturing PAGE/"crush and soak" method

A 20% gel was prepared from a 40% stock of acrylamide:bisacrylamide (39:1) and allowed to polymerise over 1-2hrs at room temperature. The DNA sample was mixed with 0.5vol of sample buffer (60% deionised formamide, 0.6% bromophenol blue and 6% TBE) and heated for 3mins at 95°C. Samples were then loaded into pre-flushed wells and the gel was electrophoresed at 100V until the dye front was 75% down the gel. The latter was removed, covered in Saran wrap and exposed to X-ray film for 1hr. The film was developed and the location of the labelled DNA identified and cut out with a razor blade. The gel slice was crushed in a 1.5ml microfuge tube and 1vol of elution buffer (TE pH 7.4/50mM NaCl/0.1% SDS) was added and left to rock overnight on a table shaker at room temperature. The gel debris was spun down and the DNA ethanol-precipitated from the supernatant. The pellet obtained from spinning at 15000rpm and 4°C, was washed with 70% ethanol and dried. It was then resuspended in TE.

### 2.3.4.4 Hydrophobic chromatography on Nensorb columns

Nensorb 20 cartridges were prewashed with 1ml methanol to activate the resin. It was again washed with 2ml Reagent A (0.1M Tris-HCl pH 8.0/10mM triethylamine, TEA) to equilibrate the resin. The DNA was applied to the column in 1ml TEA, eluted and repassaged two more times. The column was then washed with 2ml TEA and again with 2ml water to remove the former. 1ml of Reagent B (50% methanol) was then added to the cartridge and the DNA was eluted from the column. It was evaporated to dryness and resuspended in an appropriate volume of TE.

### 2.3.4.5 Estimation of specific activity of probes

Aliquots of the probe (1-2ul) were spotted onto DE-81 paper. One of them was washed extensively with several volume changes of 0.5M sodium phosphate buffer pH 7.0. The washed filter was then dried and the bound radioactivity on both filters was estimated using a Geiger counter (Cerenkov counting) or counted in a scintillation counter using a toluene-based scintillant. The percentage incorporation of label was estimated by dividing the counts on the washed filter by those of the unwashed filter. The specific activity of the probe was calculated as the cpm of the washed filter divided by the number of micrograms of starting template.

## 2.4 Cloning in bacteriophage lambda

### 2.4.1 Plaque blotting and hybridization

The libraries were plated out on 24x24cm LB plates at $10^6$-$10^7$ plaque forming units (pfu) and incubated overnight at 37°C using Q359 as the host E. coli strain. Plaques were allowed to grow to confluence, and then blotted onto asymmetrically marked nylon membranes (Hybond-N) for 1min and denatured 5min on 3MM paper soaked with denaturing solution. They were neutralized on 3MM paper soaked in neutralizing solution for 7mins. The membranes were rinsed thoroughly in 2X SSC and any agar sticking to them was gently teased off. They were subsequently air-dried, baked for 2hrs at 80°C, and prehybridized and then hybridized using the same regime described earlier [see **2.2.5**].

Positive clones were identified and carried through two more cycles of purification. This was done by orienting the membranes with the agar plates using the asymmetrical marks to locate hybridizing plaques. These plaques were picked with the broad end

of a Pasteur pipette into storage medium (SM) in a universal bottle. The phage were eluted overnight by gentle rocking on a table-top shaker at room temperature, and then centrifuged to remove cell debris. Dilutions of the phage were made such that isolated plaques would be obtained for easy identification. These were then used to reinfect fresh host cells and replated. The cycle of plating, plaque blotting, hybridization, and clone isolation were repeated at decreasing plaque densities until single positive plaques could be identified.

### 2.4.2 Preparation of lambda lysates and isolation of lambda DNA

Phage were eluted from positive clones and replated on fresh 10cm diameter LB plates. Plate lysates were prepared from these by adding 5ml of SM directly to the plates and the phage was eluted. Bacterial cell debris was removed by centrifugation, and the lysate stored in a drop of chloroform. This was subsequently used to grow a lytic culture of phage at a multiplicity of infection of $10^9$, in 21 flasks. The phage was isolated as before, and precipitated in 35%NaCl/40% polyethylene glycol (PEG) at 4°C overnight. The PEG-phage precipitate was DNAse-treated (100ug/ml), extracted with an equal volume of chloroform and centrifuged on a CsCl gradient (0.77g/ml lysate) in polyallomer tubes and in a Beckman ultracentrifuge, at 35K for 40hrs at room temperature. The phage band was removed with a 18mm gauge needle and dialyzed overnight with two changes of dialysis buffer (10mM MgSO4)/10mM Tris-HCl pH 7.5). To the dialyzed phage was added a tenth volume of 500mM Tris-HCl/100mM NaCl/100mM EDTA, and 100-200ug/ml of Proteinase K. The mixture was incubated at 37°C for 1hr. One volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added and mixed gently for 5mins and then centrifuged for 5mins at 1000g at room temperature. To the aqueous

phase was added one volume of chloroform:isoamyl alcohol and treated as before. To the final aqueous phase was added 3M NaAc and three volumes of cold absolute ethanol. The precipitated DNA was spooled out, dried under vacuum and allowed to dissolve slowly at 4°C.

### 2.4.3 Restriction analysis/mapping of lambda clones

Lambda DNA isolated above was digested to completion with various restriction enzymes (2-5units/ug DNA) as indicated in the legend on the figure below. The restriction digests were electrophoresed on an 0.8% agarose gel containing 1ug/ml of EtBr (1mg/ml), at 50 volts overnight in 1X TBE buffer. Fragments were visualized over an ultraviolet transilluminator and sized based on the HindIII restriction profile of lambda DNA.

### 2.4.4 Southern transfer, identification and isolation of inserts

Southern transfer and hybridizations were essentially as described above. Hybridizing fragments were identified with various clones and sized. Hybridizing clones were digested in a volume of 500ul with the restriction enzyme whose digest showed a positive fragment. About half of this digest was run on a 1% low melting point agarose gel and the fragment of interest was cut out and freeze-squeezed as earlier described.

## 2.5 Subcloning strategy

### 2.5.1 Vectors and vector preparation

0.5 ug of EcoR1-cut and phosphatased pUC or M13 was digested with HindIII at 37°C for 1hr. The digest was extracted with an equal volume of TE-saturated phenol:chloroform:isoamyl alcohol (25:24:1) by vortexing for 1min and spinning in a microfuge for another minute. The aqueous phase was transferred to a clean tube and the DNA precipitated as before.

### 2.5.2 Ligation/construction of recombinant plasmids

100ng of the insert DNA purified in **2.4.4** above, was ligated with 200ng of pUC or M13 in a reaction mixture containing 2 units of T4 ligase, 1X ligase buffer in a total volume of 20ul. Control ligations included vectors without inserts and vice versa. Ligation mixes were incubated at 15°C overnight. They were subsequently spin-dialyzed through a small Sepharose CL6B column to remove buffer salts.

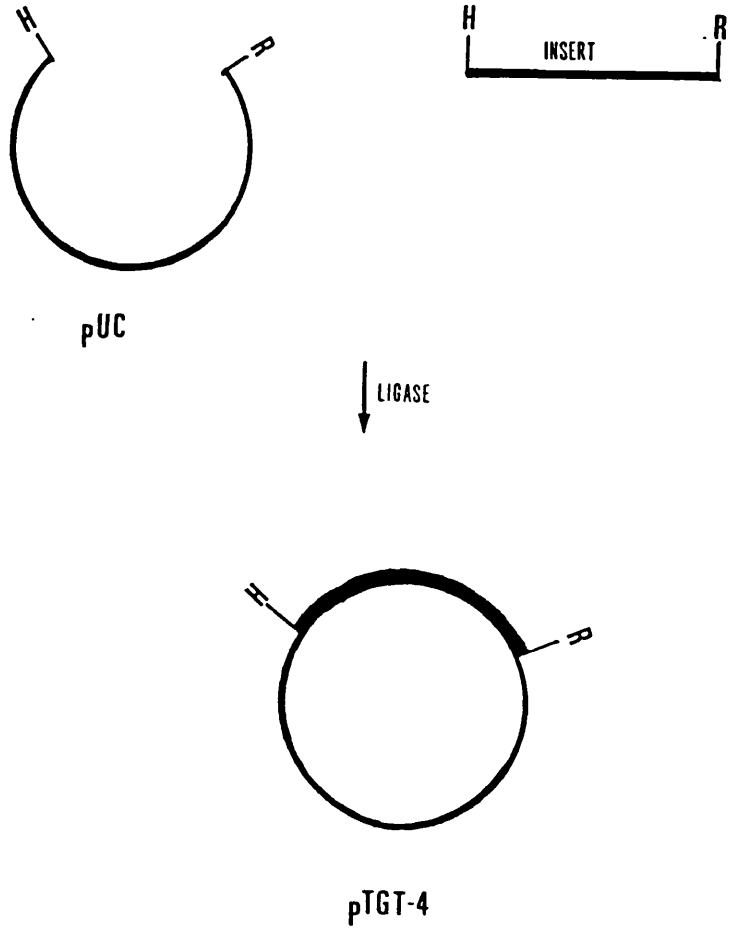



Fig. 34 Scheme for construction of pTGT-4 (subcloning strategy).

R : EcoR1 H : Hind III

### 2.5.3 Preparation of competent cells

An overnight culture of TG1 (see Appendix for genotype) cells was grown in 50ml DYT broth. A 1 in 50 dilution of the culture was made in 100ml fresh broth and cell growth was monitored every 20-30mins until the exponential phase of growth i.e the optical density reached 0.4-0.6 corresponding to $10^8$ cells/ml. The cells were transferred to sterile Falcon tubes and chilled on ice for 10mins. They were centrifuged at 7000rpm for 10mins at 4°C in a Sorvall. The supernatant was decanted and traces of media drained by standing tubes inverted. Each pellet was resuspended in 50ml of ice-cold 50mM $CaCl_2$ and placed on ice for 30mins. They were then centrifuged as above, drained and the pellet resuspended in a total volume of 5ml of 50mM $CaCl_2$ for each 50ml of original culture. The cells were left on ice in the cold room for 24-48hrs to maximise their competence.

### 2.5.4 Transformation with recombinant constructs

300ul of competent cells was added to the ligation mixes (1-5ul) and kept on ice for 40mins. They were then heat shocked at 42°C for 2mins. 1ml of fresh DYT broth was added and incubated for 1hr at 37°C. This incubation step was omitted in M13 transformations. They were centrifuged briefly, the supernatant decanted, and 10ul of IPTG (100mM), 50ul X-gal (2%) was added. They were spread with a sterile glass rod using a potters wheel, onto agar plates containing 100ug/ml ampicillin (sodium salt). The latter was eliminated in transformation into M13. For the latter, the transformed competent cells were added to 3mls of DYT broth at 48°C and quickly poured onto agar plates. Plates were allowed to dry and then incubated inverted at 37°C overnight.

Recombinants (white colonies or clear plaques) were identified and picked for regrowth in 3-5ml of fresh broth containing 100ug/ml ampicillin (for pUC recombinants) or without ampicillin (for M13 recombinants).

### 2.5.5 Cloning in plasmid

The enriched plasmid library constructed in **2.2.7** above was transformed into competent TG1 or HB101 cells and plated onto 20cm-diameter agar plates containing ampicillin as described above. Colonies were lifted onto nylon membranes as described for plaque lifts. After lifting, the plates were re-incubated for 1hr to allow the colonies to regrow to ease identification after hybridization. The membranes were treated as for plaque lifts and hybridized with oligo 2 and the cDNA probes. Hybridizing colonies were picked and processed as below.

### 2.5.6 Plasmid isolation, purification and restriction mapping

#### 2.5.6.1 Small-scale plasmid isolation (mini-preps)

Recombinant plasmids were grown in 5mls of fresh broth overnight. The cells (1.5ml) were pelleted at 5000rpm (12,000g) for 5mins in an Ole Dich refridgerated microfuge (Model 154, Camlab). The supernatant was discarded and 100ul of lysis buffer (50mM glucose/25mM Tris-HCl pH 8.0/10mM EDTA/2mg per ml fresh lysozyme) added to resuspend the pellet. The mixture was incubated on ice for 10mins. 200ul of freshly made 0.2M NaOH/1% SDS was added and incubation continued for a further 10mins. 150ul of 3M KAc pH4.8 was added and left on ice for 10mins. The contents were centrifuged at 12,000g for 15mins at 4°C. The supernatant was treated with 2.5 vols of cold EtOH and incubated on ice for 10mins. The

DNA was pelleted by spinning for 10mins at 15,000g in a microfuge and washed with cold 70% EtOH. It was briefly dried, redissolved in 50ul of water and purified by spinning through a Sepharose CL6B column. An aliquot of each clone was cut with HindIII for 1hr and then with EcoR1 and electrophoresed in parallel with uncut plasmid and lambda markers on a 1% agarose gel to check and verify the presence and size of the inserts. Plasmids isolated from the plasmid library were cut with BamH1 and EcoR1 as these were the sites at which genomic inserts were cloned into the vector.

### 2.5.6.2 Large-scale plasmid isolation (maxi-preps)

Clones with inserts were grown as 1l cultures in 2l flasks overnight. The cells were harvested by spinning in a Mistral 6L centrifuge in 1l bottles at 4000rpm, for 15mins at 4°C. The supernatant was decanted and the pellet washed with 100ml of STE. The cells were respun and resuspended in 10ml of lysis buffer, and left at room temperature for 5mins. 20ml of freshly prepared 0.2M NaOH/1% SDS was added and the tube inverted several times and then allowed to sit at room temperature for 10mins. 15ml of ice-cold KAc was added, the tube capped and its contents mixed and placed on ice for 10mins. It was spun at 4000rpm for 15mins at 4°C in a Sorvall. The latter was allowed to stop with the break off. The supernatant was filtered through four layers of muslin or gauze. Isopropanol was added to 0.6vol, mixed and stood at room temperature for 10mins. The DNA was recovered by centrifuging at 5000rpm for 15mins at room temperature. The supernatant was decanted and the pellet washed with 70% EtOH and evaporated to dryness.

### 2.5.6.3 Equilibrium gradient purification of plasmids

The DNA was dissolved in 3ml TE pH 8.0. To each ml of DNA was added 1g of CsCl and 100ul of EtBr (10mg/ml). The mixture was centrifuged in a Sorvall at 8000rpm for 5mins at room temperature to remove the pellicle. The clear purple solution of DNA/CsCl/EtBr was transferred by means of a Pasteur pipette into Beckman quickseal polyalloma ultracentrifuge vials. The remainder of the vials was filled with 1mg/ml CsCl in TE. The vials were sealed and centrifuged at 60000rpm in a Ti70 rotor for 24hrs. The plasmid bands were removed as for the phage. The EtBr was removed from the plasmid by a series of vortexing, spinning and removal, with equal volumes of isoamyl alcohol. This was continued until both aqueous and organic phases were clear. The plasmid/CsCl mixture was diluted out with 3vols of water, and 2vols of cold EtOH added. The DNA was precipitated at 4°C for 15mins, and spun out at 10000rpm for 15mins at 4°C. The pellet was vacuum dried and resuspended in 1ml TE pH 8.0. The yield of DNA was estimated as previously described. Aliquots were digested as for the mini-prep with the relevant restriction enzymes to verify the presence and size of insert.

### 2.5.7 Isolation of single-strand DNA from M13 clones

Clear plaques of M13 recombinants were picked and grown in 2-5ml exponential cultures of TG1 cells and cultured overnight (6-10hrs) at 37°C with shaking. 1.5ml of the culture was centrifuged in an eppendorf tube for 10mins. The phage-containing supernatant was carefully transferred to a clean tube containing 1/10vol 20% PEG/2.5M NaCl. The phage was precipitated for 15mins on ice and was spun down for 10mins. The supernatant was decanted and the tube respun and all traces of PEG carefully removed. The pellet was resuspended in

200ul of TE and extracted with an equal volume of phenol:chloroform (3:1) by vortexing and centrifuging. The DNA in the final aqueous phase was isolated as previously described, dried and resuspended in 20ul of TE.

### 2.5.8 Single-strand plasmid rescue

A fresh plasmid colony was picked into 5ml of DYT containing 100ug/ml of ampicillin. The helper phage VCSM13 was added at $10^7$-$10^8$ pfu/ml or a multiplicity of infection of 10. The culture was grown for 2hrs at 37°C after which kanamycin was added to 70ug/ml final concentration, and growth continued overnight. Single-strand DNA was then prepared as earlier described.

## 2.6 DNA sequencing

Sequencing was performed according to Sanger's dideoxy chain termination method (206). All reactions were primer-directed.

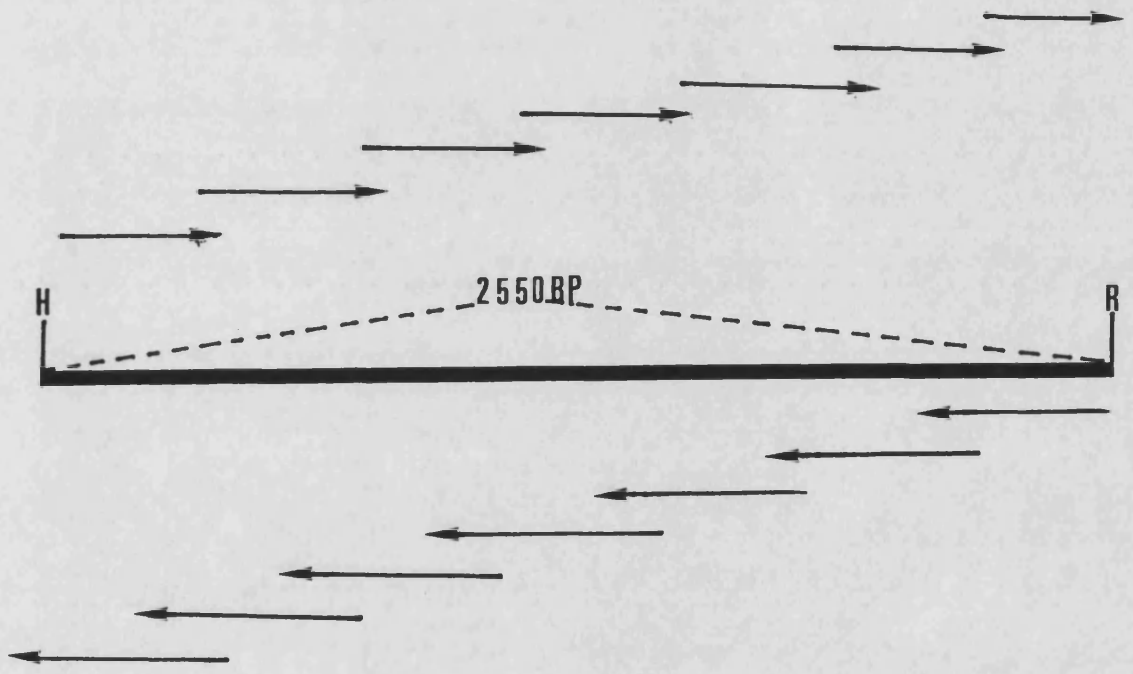



Fig. 35 DNA sequencing strategy. Horizontal bar represents double-strand template. Arrows above and below the template indicate the strand and direction of sequencing. Regions of overlapping arrows represent regions for primer-initiation of sequencing. Cloning sites are indicated. R= EcoR1 and H= HindIII

### 2.6.1 Template preparation and sequencing

Double-strand pUC templates (2-5ug per sequencing reaction) were denatured in a total volume of 30ul with 0.2M NaOH/2mM EDTA for 10mins. This was then spin-dialyzed through a Sepharose CL6B mini-column after repeatedly (2-4 times) washing out the resin storage-buffer with distilled water. The denatured template was then used immediately. Single-strand M13 templates prepared above were used without this step. DNA templates were sequenced using the sequenase kit as instructed in the manufacturer's manual. Sequencing reactions were subsequently electrophoresed on 6-8% polyacrylamide gels of 0.2mm-0.35mm thickness, made using the recipe of the supplier of Sequagel. Gels were processed as described in the sequenase manual. Sequencing was done on both strands (pUC) and oligonucleotide primers complementary to the template sequence were synthesised as and when necessary.

## 2.7 Sequence analysis

### 2.7.1 Analysis of gene structure

The sequence was analyzed using the programme of Staden (207). As sequencing was primer-directed, new sequences were added onto those of the preceding segments. This avoided using the contig assembly menu, DBAUTIL as would be the case with a shotgun sequencing strategy. Sequence analysis was then performed using the ANALYSEQ menu to study the structure (searching for Z-DNA, palindromes, restriction enzyme sites etc); for statistical analysis of the sequence with respect to base composition, dinucleotide frequencies and codons; to translate the nucleotide sequence into amino

acids, and to search for open reading frames and possible promoter regions. The translated sequence was then analyzed further on the ANALYSEP and Sequence Comparison menus. The former enabled the identification of possible hydrophobic segments, and the latter was helpful in a direct segmental sequence comparison to identify the binding/ hybridization sites of the probes used in cloning.

### 2.7.2 Multiple sequence alignment and homology search

This was performed according to Barton and Sternberg (208) and the HOMED programme of Stockwell (209). Briefly, all the available glucose transporter sequences were aligned pair-wise i.e the third sequence was aligned with the best alignment of the first two sequences, the fourth with the best of three etc. The trypanosome sequence was added to the optimal/consensus template alignment of the preceding sequences. Comparisons of the trypanosome sequence were also made with a selection of mammalian glucose transporters on one hand, and with prokaryote and lower eukaryote transporters on the other. Gaps were manually included to optimize the alignment.

### 2.7.3 Secondary structure prediction

This was performed using the Chou-Fasman algorithm (159), to predict potential a-helices, b-sheets or regions with a propensity to form turns.

## 2.8 Database search

The trypanosome protein sequence was used to search several databases (EMBL, GeneBank, Swissprot etc) using the FASTP programme of Lipman and Pearson (210).

## CHAPTER 3

### 3.0 Molecular characterization of the putative glucose transporter clone

### 3.1.1 Construction of transcription plasmids

The insert from pTGT-4 was isolated from the latter by digesting with EcoR1 and Apa1, and purified after low gelling agarose gel electrophoresis. This was then ligated directionally into pBluescript cut with the same restriction enzymes to produce the plasmid pBS.TGT-4. Insertion of the fragment into the vector was such as to enable transcription from the T7 promoter of pBluescript.

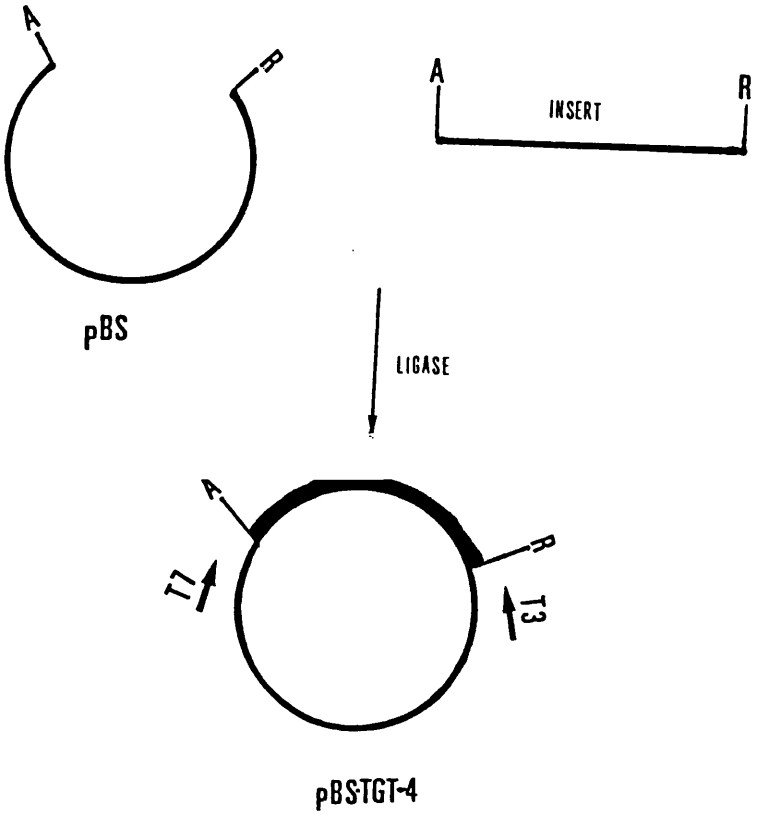



Fig. 36 Scheme for construction of transcription vector.

Enzymes: A= Apa1 and R= EcoR1. The direction of transcription is specified by the orientation of the promoters T3 and T7.

### 3.1.2 Cloning and purification of plasmid constructs

The plasmid was subsequently transformed into competent TG1 cells as previously described. Transformants were picked and grown in 5ml of fresh broth. Plasmid DNA was prepared on the small scale as detailed above. The presence of inserts was verified by restriction mapping of the plasmids with the above restriction enzymes and running on a 1% agarose gel. Plasmids with the expected size of insert were grown in 500ml DYT broth and the DNA purified as previously on a CsCl gradient.

### 3.1.3 Transcription and capping of transcripts

Sterile procedures were stringently adhered to in this step. RNAse-free solutions were prepared in DEPC-treated water (0.1%) and oven-baked or autoclaved bottles and tubes. Pipette tips were autoclaved for 70mins on the dry cycle. Pipettes were cleaned in DEPC-water and then rinsed with 3% hydrogen peroxide.

The plasmid was linearized downstream of the insert (using EcoR1) and downstream of the T7 polymerase promoter to enable transcription of the coding strand. Proteinase K was added to the digest to 50ug/ml at 37°C for 30mins. The template was then extracted two times with 1:1 phenol/chloroform, and precipitated with ethanol. The pellet was vacuum-dried and resuspended at 1mg/ml in DEPC-treated water.

Run-off transcripts were prepared according to Melton et al(211) with some modifications as follows:

| | |
|---|---|
| 5ul | 5X transcription buffer (200mM Tris-HCl pH8.0/40mM $MgCl_2$/10mM spermidine/250mM NaCl) |
| 1ul | 0.5M DTT |
| 1ul | RNAguard |
| 4ul | 10mM rNTP cocktail |
| 2.5ul | 5mM m7G(5')ppp(5')G capping agent |
| 1ul | DNA template |
| 20U | T7 polymerase |

Water was added to a final volume of 25ul.

The transcription mix was incubated for 1hr at 37°C. After this, 10U of RNAse-free DNAse was added and the incubation continued for 15mins. The RNA was then purified by phenol:chloroform extraction and precipitated with 1/10th volume 3M NaAc pH 5.2, and 2.5 volumes of absolute ethanol. The pellet was briefly dried and resuspended in 50ul DEPC-treated water. The yield of RNA was estimated by running 5ul on a normal/non-denaturing agarose gel, and the concentration adjusted to 1mg/ml.

### 3.1.4 In vitro translation

This was performed using rabbit reticulocyte lysate in the following set up:

| | |
|---|---|
| 2ul | translation (amino acid) mix |
| 1ul | RNAguard |
| 1ul | KAc |
| 1.5 ul | $Mg(Ac)_2$ |
| 4ul | $^{35}$[S]-methionine (1200Ci/mmol) |
| 2ul | RNA |
| 10ul | reticulocyte lysate |

The mix was made up to 25ul with DEPC-treated water and incubated for 1hr at 30°C. A positive control was set up to run in parallel using Xenopus laevis oocyte B-globin RNA.

### 3.1.5 SDS-PAGE

The translation products were separated by electrophoresis on a 10% SDS-polyacrylamide gel prepared from a 40% stock. Samples were mixed with 50ul sample buffer (6M urea/10%SDS/35% glycerol/0.5% bromophenol blue supplemented with 10% mercaptoethanol), and boiled for 2min. They were then loaded onto the gel and run together with molecular weight markers at 100mA constant current in 1X electrophoresis buffer. After electrophoresis, the gel was stained in Coomassie brilliant blue (30% methanol, 10% glacial acetic acid, 0.1% Coomassie stain) with gentle rocking overnight. It was then destained in 30% methanol/10% acetic acid.

### 3.1.6 Fluorography

The gel was rinsed with distilled water and blotted dry on 3MM paper. En$^3$Hance was added to the gel and rocked for 1hr at room temperature. It was subsequently rinsed 2X for 10mins each time, with distilled water, fixed onto 3MM paper and dried on a vacuum slab gel dryer. The dried gel was then exposed to X-ray film as before.

## 3.2 Developmental expression of the glucose transporter-related gene

### 3.2.1 Isolation of trypanosome RNAs

Bloodstream trypanosomes were purified as above, and culture forms (procyclics) were provided by Ms M. Kasraeian, Department of Biochemistry, Bath University. All cells were frozen in liquid nitrogen until used. Precautions pertaining to RNA handling were adhered to as above.

RNA isolation and purification was carried out essentially by the method of Chomczynski and Sacchi (212). Briefly, trypanosomes were resuspended to $10^6$-$10^8$ cells/ml of guanidinium thiocyanate (GIT) in a sterile falcon tube whereupon they lysed immediately. In sequence, 0.1vol of NaAc pH 4.0, 1 vol of water-saturated phenol (Tris-treated) and 0.2vol of chloroform:isoamyl alcohol (49:1) were added to the cell lysate and thoroughly mixed after each reagent was added. The mixture was briefly vortexed and then chilled on ice for 15mins. It was spun for 20min at 10,000g and 4°C. After centrifugation the supernatant containing RNA was transferred to a fresh tube and the RNA precipitated with 1vol of isopropanol for 1hr at -20°C. The RNA was spun down as above and redissolved in 0.3vol (original lysate) of GIT in a 1.5ml Eppendorf tube. The RNA was reprecipitated and centrifuged for 10mins. The pellet was rinsed

with 70% ethanol, vacuum-dried and stored in ethanol or as a dry pellet at -70°C until required..

### 3.2.2 Formaldehyde gel electrophoresis

A 1% formaldehyde gel was prepared by replacing 1X TBE buffer for a normal agarose gel, with 1X MOPS buffer prepared in DEPC-water. The gel was allowed to cool to about 50°C when it was transferred to a fume hood and 37% formaldehyde added to a final concentration of 0.66M. The gel was left to cool to about 35°C and then poured in the hood and allowed to set. 10-20ug total RNAs (1mg/ml) were mixed with 2vols of sample buffer (50% formamide/2.2M formaldehyde/1X MOPS) and denatured by heating at 65°C for 10mins. They were snap-chilled on ice and then mixed with sample loading buffer (0.1% SDS/5% glycerol/50ug per ml bromophenol blue). The samples were loaded onto the gel together with RNA markers and electrophoresed at a constant voltage of 100V in 1X MOPS. After electrophoresis, the gel was stained for 30mins in 1mg/ml EtBr to visualise RNA.

### 3.2.3 Northern blotting and hybridization

Following visualization, the gel was rinsed in several volumes of 10X SSC to remove formaldehyde. All other procedures for Northern blotting were as for those described in Southern blotting above. RNA was immobilised onto Hybond-N+ membranes by baking at 80°C for 2hrs or by UV irradiation for 5mins.

The insert from pTGT-4 was purified as described and radiolabeled by random-primer extension as above. The rat liver GT cDNA was similarly purified and treated. Oligo (2) was end-labelled as above. All three probes (previously used to isolate the gene) were used to

hybridize the Northern blots to detect glucose transporter mRNA transcripts and transcripts relating to the trypanosome gene, in both bloodstream and procyclic forms. Hybridization and subsequent processing were as for Southern blotting.

## 3.3 Functional expression of the gene

### 3.3.1 Construction of expression vector

The eukaryotic expression vector pSVL was digested with Sma1 within the polylinker. pTGT-4 was also digested with HindIII and EcoR1, electrophoresed on a 0.8% low melting point agarose gel and the insert cut out from the gel. The fragment produced from this digest contained all the putative promoter and enhancer elements of the gene. The DNA was isolated from the digest by freeze-squeezing and the insert and vector were purified as before. The insert was end-filled using Klenow enzyme. Briefly, the fragment was dissolved in 20ul TE, 5ul of 10X Klenow buffer, 1ul of 15mM dNTP mix, 1ul (4U) Klenow, and water to a final volume of 50ul. The mixture was incubated for 15mins at room temperature. The reaction was stopped by phenol:chloroform extraction and the DNA ethanol-precipitated.

Vector and inserts were ligated in a ratio of 0.5 (vector:insert) in a total volume of 20ul as detailed above. Hence pSVL/2.5 was constructed.

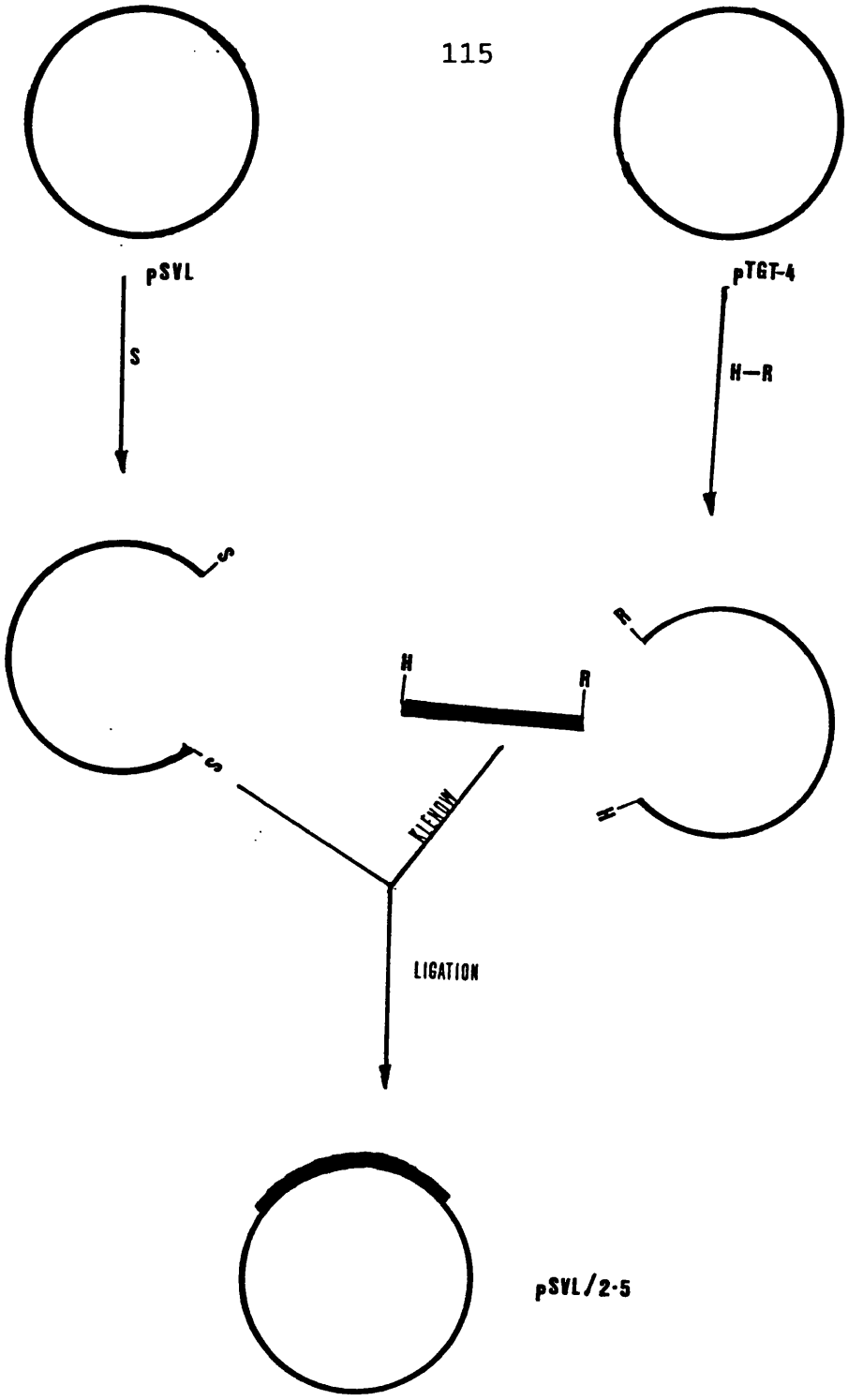



Fig. 37 Construction of eukaryotic expression vector, pSVL/2.5

Enzymes: H= HindIII; R= EcoR1; S= Sma1;

### 3.3.2 Cloning and purification of vector

The expression construct was transformed into competent HB101 cells and plated on DYT plates containing 100ug/ml ampicillin. After an overnight incubation at 37°C, transformants were picked and regrown in 5ml of DYT broth. Plasmid mini-preps were prepared and cut with restriction enzymes with recognition sites in the polylinker of the cloning vector viz. Xhol and Sacl. (Blunt-end ligations result in the loss of original restriction sites, in this case, EcoR1 and HindIII. pSVL/2.5. Restriction was performed with the knowledge of the restriction maps of both insert and vector). Restriction digests were run on a 1% agarose gel and the sizes of the inserts noted with respect to the beginning of the coding strand and the regulatory regions of pSVL. Plasmids with the expected sizes and native pSVL were grown on a large scale as previously described and purified on a CsCl gradient. The maxi preps were again cut with the above enzymes (Xhol and Sacl) to confirm the presence of inserts of the expected sizes.

### 3.3.3 Determination of insert orientation

Even though the orientation was determined by restriction mapping of the constructs, an unequivocal result was obtained by cutting out the insert as before and subcloning into Xhol and BamH1-cut M13mp19. Single-strand DNA was prepared and sequenced with the universal forward primer. The sequences indicated the orientation of the coding strand with respect to the vector promoter and polyadenylation signals.

### 3.3.4 Cell culture

COS-7 cells were removed from storage in liquid nitrogen and quickly thawed in a water bath at 37°C. Just before the lump completely thawed, the cells were diluted with 1ml DMEM supplemented with 10% foetal calf serum (FCS), hereafter referred to as "medium" i.e DMEM+FCS. They were then seeded on a 10cm diameter petri dish and 9ml of more medium was added. The cells were then incubated at 37°C and 10% $CO_2$. The medium was changed after every 48-72hrs. At about 90% confluence, the cells were washed 2X with PBS equilibrated at 37°C, and the PBS aspirated. They were subsequently trypsinized with 1ml of 1X trypsin/EDTA (0.05% trypsin/0.02% EDTA). As the cells detached (3-5mins) from the plate, 10ml of medium was added and the cells collected in polystyrene bottles with a rubber policeman. They were centrifuged at 1500 rpm for 3mins and the medium carefully decanted. The cell pellet was resuspended in 10ml of fresh medium.

### 3.3.5 Trypan Blue exclusion

To 50ul of the cell suspension was added an equal volume of 0.1% trypan blue for 15mins. The number of live cells was counted in a haemocytometer at x40 magnification. The total harvest of cells per ml was estimated from the number of cells counted in the five wells of the haemocytometer, divided by 5 and multiplied by the dilution factor of 2 and the volume of the well ($10^4$). An appropriate volume of medium was added to give a cell count of $10^4$-$10^6$cells/ml.

### 3.3.6 Transfection of COS-7 cells

3.5cm-diameter dishes were seeded with about $10^5$ cells from the above and made up to 3ml with fresh medium. They were then cultured as before.

When cell growth attained about 80-90% confluence, they were transfected with expression constructs using lipofectin as instructed by the manufacturer. Briefly, for each plate, 10ug (10ul) of lipofectin was diluted to 25-50ul with sterile water. 1-10ug of DNA containing gene construct or native plasmid (control) was diluted to the same volume. All dilutions of lipofectin and DNA templates were done separately in polystyrene tubes. They were then gently mixed and left to coprecipitate for 15mins at room temperature. In the interval, the cells were washed twice with prewarmed PBS and then seeded with 3ml of serum-free DMEM. The lipofectin-DNA complex was added to the cells at 50-100ul per plate. The plates were gently swirled to ensure even distribution of the complex, and then incubated for 5hrs. After this time the serum-free medium was replaced with 3ml of fresh medium containing serum at 10%. Incubation was continued for 72hrs.

### 3.3.7 Transport assays

Expression of glucose transport activity was measured under zero-trans conditions. The cells were washed 2X with Krebs-Ringer-Hepes buffer (KRH) equilibrated at 37°C. 0.5ml of this buffer was added after washing and the cells incubated for 30mins at 37°C. This was subsequently aspirated and 0.5ml of 100uM/2uCi of $^3$H-2-deoxyglucose (2-DOG, 6.6Ci/mmol) or 1.2 uCi $^{14}$C-3-O-methylglucose containing 50-

100uM cold D-glucose was added to each plate. After an appropriate incubation time, glucose uptake was terminated by adding 2ml of ice-cold KRH/phloridzin, 10mg/100ml. This was quickly aspirated and the cells washed two more times. For background/nonspecific uptake, the stopping solution was added to the cells before adding substrate. All time points were run in duplicate. After the assay, the cells were solubilized with 0.5ml 1% SDS at room temperature for 30mins. Aliquots of 10ul of each sample were taken for protein determination by the BCA method (213). The remainder of the lysates were transferred to scintillation vials and counted in 10ml scintillant as previously.

# CHAPTER 4

## RESULTS

### 4.1 Functional identification and characterization of the trypanosome glucose transporter

The rate of glycolysis in T. brucei was studied in an attempt to correlate this with the rate of glucose transport. Preliminary results from the measurement of glycolysis indicated that the rate of this process was approximately twenty-fold that of mammalian systems, at 86.7nmol $O_2$/min/$10^8$ cells. This is consistent with the results of Flynn and Bowman (214).

Functional activity of the glucose transporter was assayed in both whole cells and reconstituted vesicles. Results from assays in whole cells were consistent with those reported earlier (215) in this laboratory.

An attempt to achieve transport in reconstituted vesicles was replete with several difficulties, not least the integrity of the plasma membranes prepared, and the possible proteolytic degradation of membranes in storage. The main methods employed to achieve physiological reconstitution included cycles of freeze-thawing, sonication and detergent dilution. A combination of sonication and detergent dilution achieved functionally reconstituted vesicles which transported glucose under trans-stimulation. It must be mentioned that even though this was achieved with the expected result of physiological reconstitution, it is not a very reproducible technique. This is due to a large number of reasons including the nature of detergent, membrane integrity, the sidedness of the vesicles, integrity and homogeneity of size of the vesicles, nature and integrity of phospholipid (oxidized phospholipid was no use) etc. The amount of incorporation of protein into liposomes varied from 3.8-5% of initial crude membrane

protein. Even though this level is low, it was not a critical factor for achieving reconstitution and transport. The large number of variables therefore makes this technique unavailable for routine application.

The biphasic nature of transport in the vesicles in response to trans-stimulation intrinsically may hold information regarding the binding, rate of translocation and restitution of the transporter molecules. The graph below [Fig. 38] may therefore be the embodiment of the phenomenon of facilitated transport, and of substrates which share the same carrier and whose uptake and translocation are mutually dependent. There is an initial counterflow/trans-stimulation transient as labelled substrate is transported up its concentration gradient. The movement of label is simultaneously coupled to the exit of cold intracellular glucose. Uptake peaks within 3secs, and tapers at a point where the concentration of label on both sides of the membrane are identical. This is the only criterion for delineating glucose transport in the trypanosome from speculations of endocytosis or diffusion as the mechanism of substrate uptake. Another point from this study was the observation that the stoichiometry of transport was (tentatively) about 0.8moles of glucose per mole of transporter. This was estimated from the gradient of the graph [Fig. 39] and is consistent with that observed for the arabinose binding protein/transporter (216).

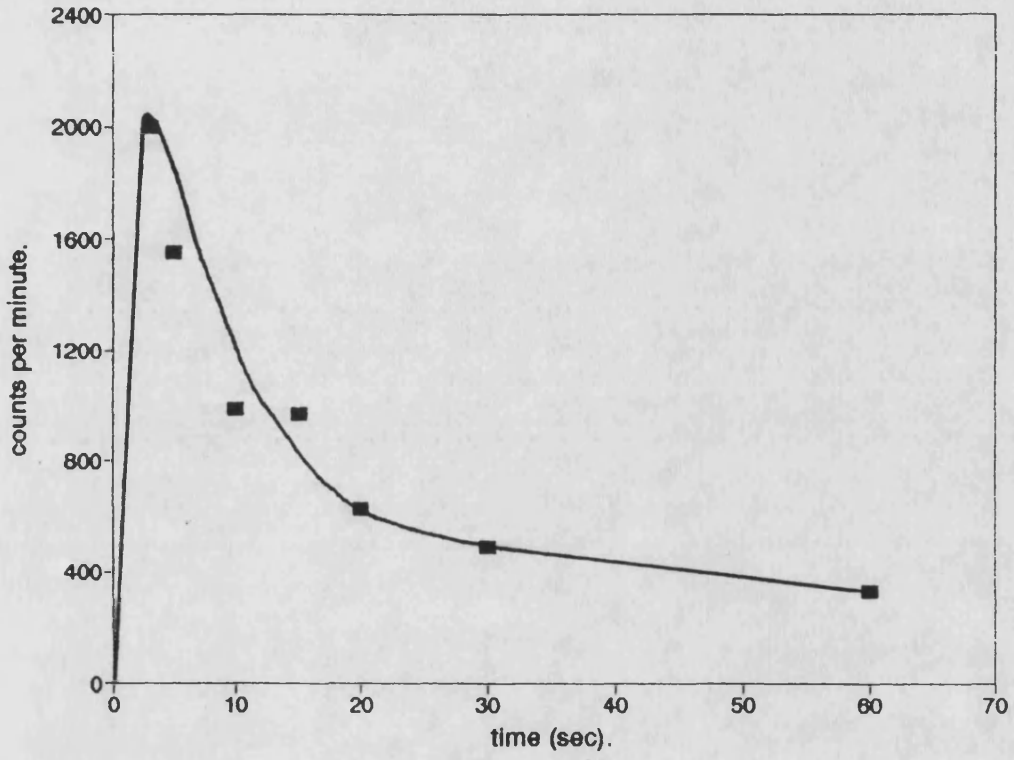



Fig. 38 Binding and translocation of radiolabeled glucose in the reconstituted trypanosome GT. Results are expressed as the amount of radiolabeled glucose transported (cpm) as a function of time (secs).

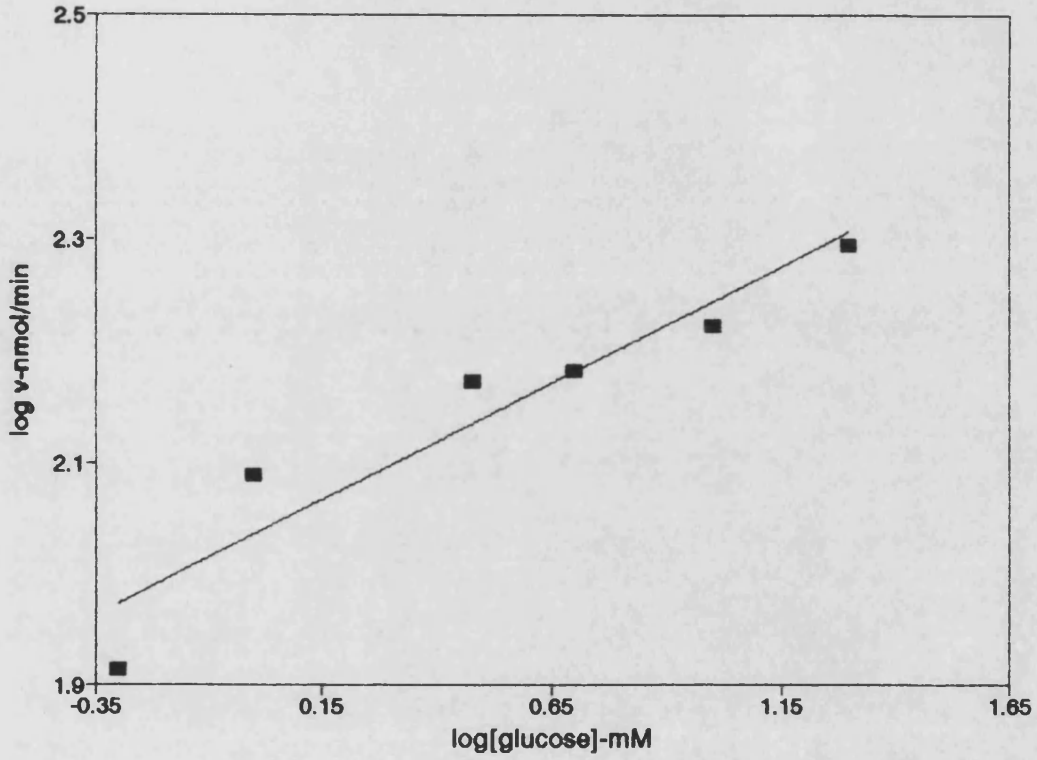



Fig. 39 Estimation of the stoichiometry of glucose transport. The gradient of the slope represents the number of moles of glucose per mole of transporter.

## 4.2 Molecular cloning of the GT-related gene

The trypanosome genomic library was screened with the rat liver glucose transporter cDNA. Hybridizing plaques were washed at room temperature with 5X SSC, and at 55°C with 2X SSC. In the first round of screening, 17 clones were identified and further selected by cross-hybridization with oligo 2 followed by oligo 3. Twenty four plaques were picked from the hybridization with oligo 3. These were then cross-hybridized with oligo 4, and 12 positive plaques were picked. Six of these clones were selected on the strength of their hybridization signals, and a large scale preparation of lambda DNA was carried out on them. These were cut with various combinations of restriction endonucleases to map and identify the inserts by Southern blotting. Hybridization with the rat cDNA identified four types of clones: 1-LT2, 1-LT5, and 1-LT7; 2-LT3; 3-LT10; 4-LT19. (The number preceding and after each clone denote its category and cloning identification number respectively. LT is an abbreviation for each trypanosome genomic clone in lambda).

Of the four categories, 1 and 3 gave hybridizing fragments of about 1200bp and 2550bp respectively, when digested with HindIII and EcoR1. Clones 2-LT3 and 4-LT19 contained feebly hybridizing inserts. Upon the strengths of their hybridization, the inserts of the positive groups represented by 1-LT2 and 3-LT10 were purified and ligated into EcoR1 and HindIII-cut pUC. The recombinants were denoted as pTGT-1 and pTGT-4, containing the 1200bp and 2550bp respectively.

DNA was isolated from maxi preps of these plasmids and cut with the above enzymes to verify the presence and size of the inserts. They were mapped by

cutting with other restriction enzymes, and subsequently hybridized with the rat cDNA again. pTGT-4 was selected for further study because it contained a more strongly hybridizing insert. Also a partial sequence of pTGT-1 showed long tracts of repetitive DNA, and a long open reading frame was not revealed. It was also impossible to perceive any homology between it and the rat cDNA or any of the oligos which could account for the strong signal obtained with this clone.

An enriched genomic library was also prepared from fragments of trypanosome genomic DNA hybridizing to the rat cDNA. This library was prepared by digesting trypanosome genomic DNA with various restriction enzymes and electrophoresis. The restriction fragments of DNA were transferred to nylon membranes by Southern blotting. They were subsequently hybridized with the rat cDNA probe. This probe hybridized to two fragments on BamH1/EcoR1-cut genomic DNA, one of which was about 2800bp and the other about 5000bp. This was possibly due to the presence of two copies of the gene or that the two fragments were parts of the same gene. On the other hand, either one could bear the gene of interest. There was another fragment from a Xho1/EcoR1 digest of about 2800bp. Upon screening the library with the cDNA and oligo 2, a number of clones were identified one of which, from a partial sequence, was identical to pTGT-4.

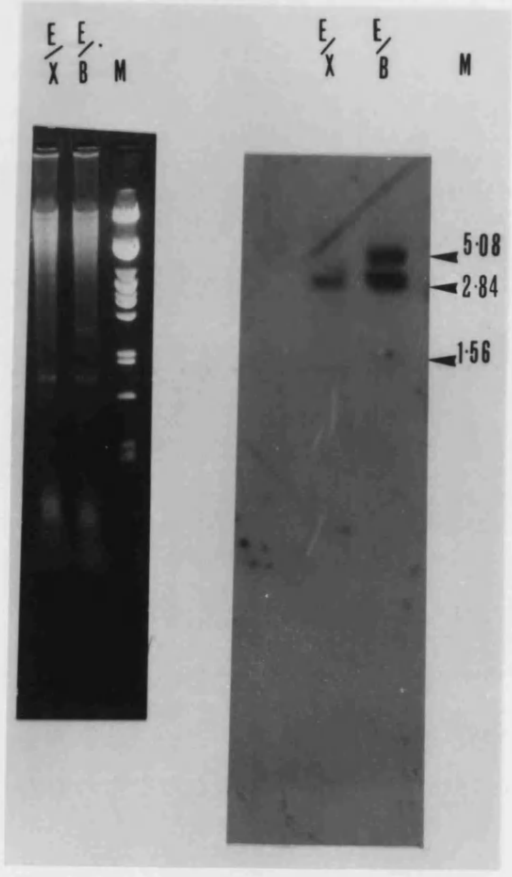
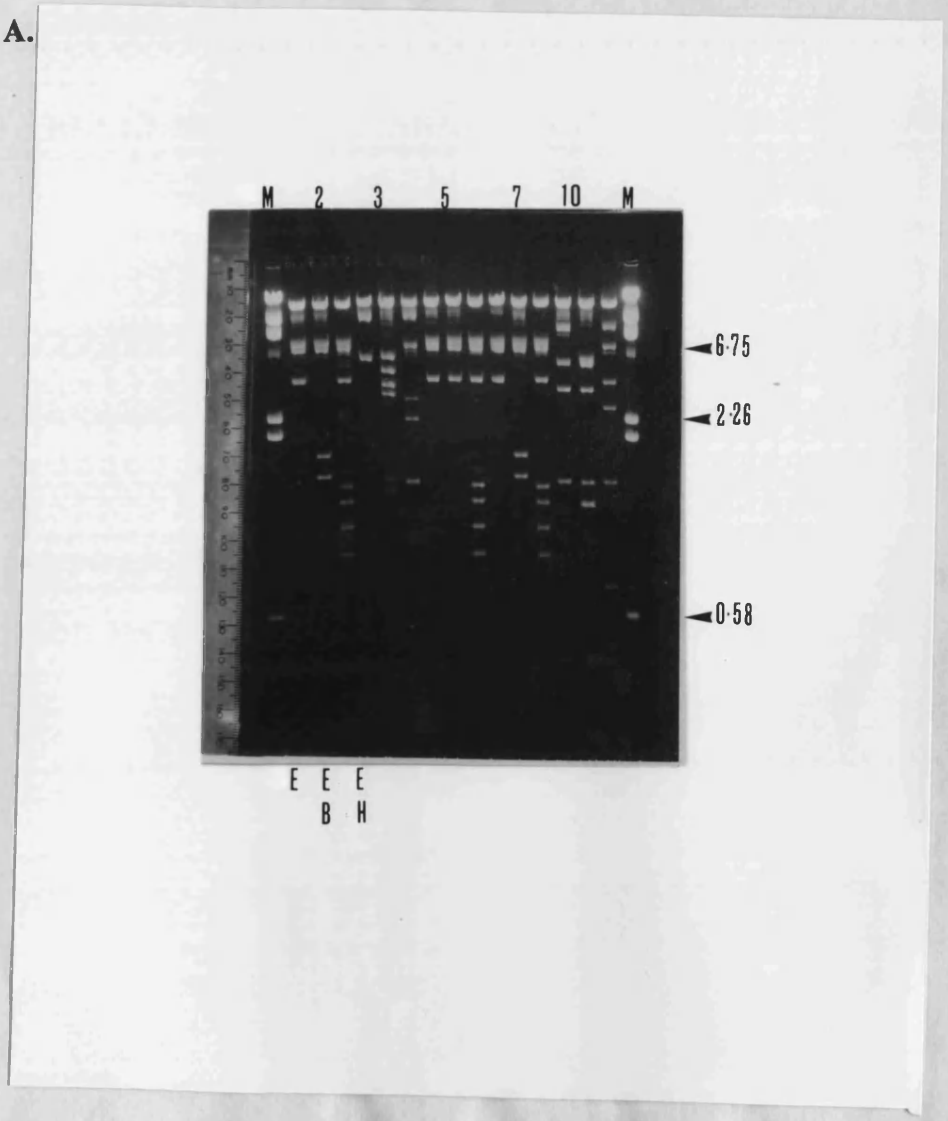


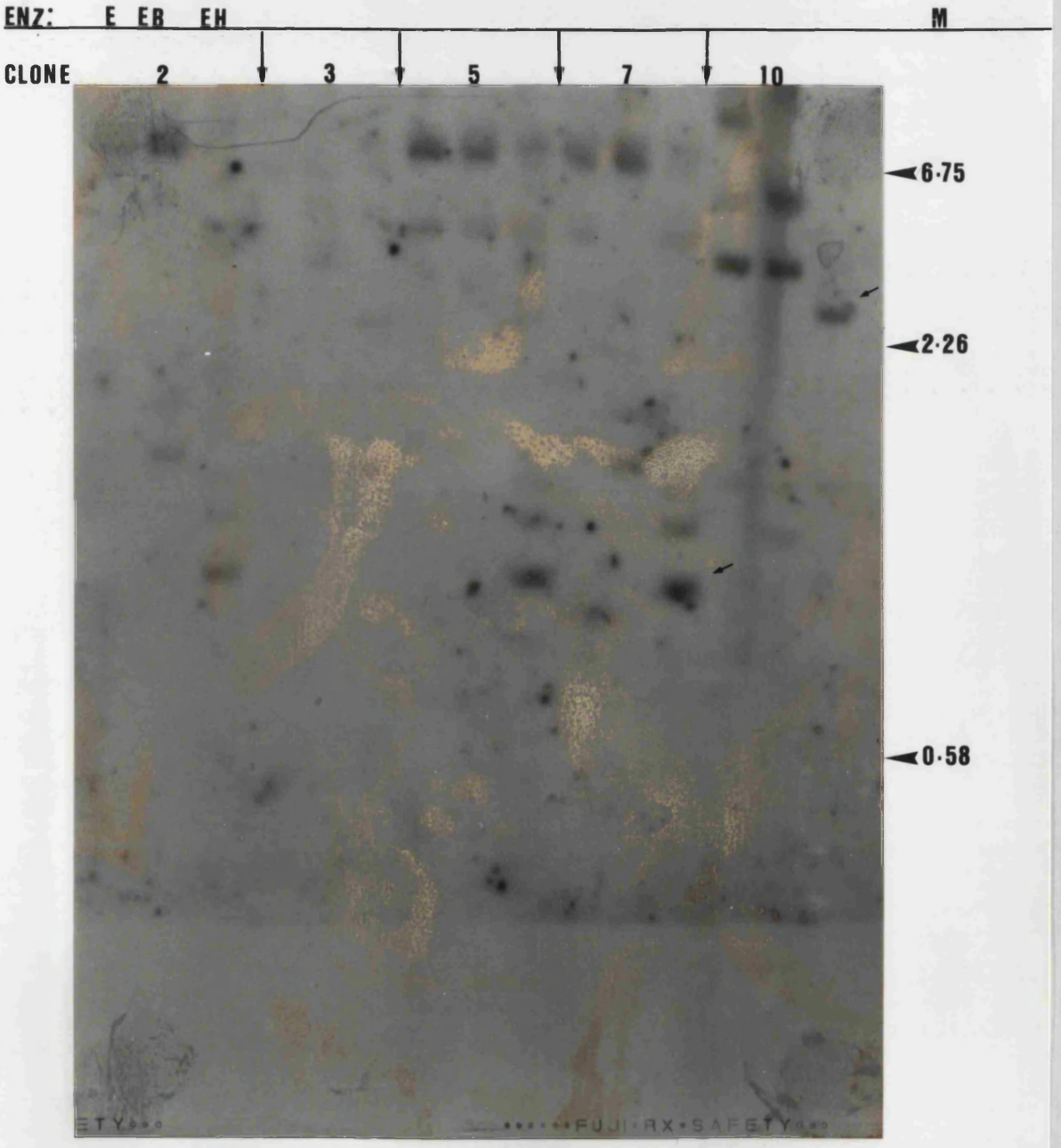

Fig. 40 Southern blot and hybridization of trypanosome genomic DNA with the rat GT cDNA probe. (A) EtBr-stained gel (B) Autoradiograph of (A). M= lambda markers; Enzymes:B= BamH1; E= EcoR1; X= Xho1.

Fig. 41 Southern blot and hybridization of rat GT cDNA to lambda clones isolated from EMBL4 library. (A) Profile of restriction digests as seen in EtBr-staining. (B) Autoradiograph of (A). M= lambda markers; Enzymes: B= BamH1; H= HindIII; R= EcoR1. 2, 3, 5, 7, 10 are clone identification numbers. All clones were digested with the same enzyme(s) or combinations thereof. Cf the restriction/hybridizing fragments from EcoR1/BamH1 digests of clone 3-LT10 and the genomic DNA.

A.



B.



### 4.2.1 DNA sequencing and sequence analysis

pTGT-4 was sequenced on both strands using the dideoxy chain termination method. The 2550bp insert was also cloned and sequenced in M13 in both orientations i.e by sequencing the template in M13mp18 and M13mp19. Oligonucleotide primers (15-20 bases) were synthesized complementary to regions and along the length of the template, to prime all sequencing reactions. The sequences of these primers were based on the sequences further into the gene to enable sequencing of the entire clone and thus avoiding subcloning smaller fragments. Sequencing of both strands in pUC templates and in M13mp18 and 19 allowed the resolution of ambiguities involving GC-rich regions of the sequence.

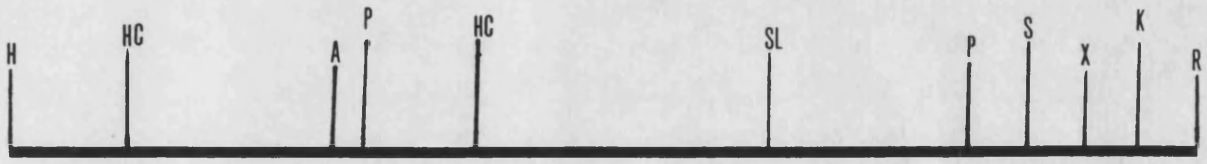



Fig. 42 Restriction map of pTGT-4. Key cloning sites are indicated. Restriction enzyme sites: A= Apa1; H= HindIII; HC= HincII; K= Kpn1; P= PvuII; R= EcoR1; Sl= Sal1; S= Sac1; X= Xho1.

### 4.2.2 Analysis of gene structure

Fig. 43 Nucleotide and predicted amino acid sequence of the trypanosome gene. Promoter structures are boxed. The two translational start sites are indicated,and the stop codon(s) is indicated with a star. Regions of oligo probe hybridization are underlined. Potential N-linked glycosylation sites are overlined, and potential cyclic-AMP dependent protein kinase C phosphorylation sites are indicated by broken lines. Two polypyrimidine tracts, which are potential tran-splicing sites are underlined with broken bars. Cloning sites are indicated by H (HindIII) and R (EcoR1).

```
                                                                      H
AAGCTTTACCACATCTATCACCACCACCACCCTCCCCGGTGGCCTTTTTTTTCCCCCCTT
        10        20        30        40        50        60

TTCTGGTGCCCCTCCCCCCCCCTCCACACACACACACACACACACACACACACACACA
        70        80        90       100       110       120

CATTAGGTCTTTTGAAAATAAGGGAACGGCATGTGCGGTTGACCCCTTTGGAGCAGCGGT
       130       140       150       160       170       180

GAATGGAGCAAACGAGCAACCCCACGCCATGCGGAGCAACAAAGAGAGACACGTATCGCC
       190       200       210       220       230       240

TACACATGAAGCGGCTAGCGGGGATGAGGGTTGGAAAACGAATCGGGAGGAAACGGCCTT
       250       260       270       280       290       300

TTTGCCGGCGAAATAAAAATGGTGTCGGCGAGCGAGTAGGTGCGGCAACGCGGGCCTTTA
       310       320       330       340       350       360

AGGCGGACACTTTCTTGATGAAGCGGATCGGCAACTTCCTTTCGGGGATGTAAGGCGTAA
       370       380       390       400       410       420

CGCGGTTGTGCGAAACCCCGCGGGGCCAATGTCAAGCACCATTTTTTTTTTTTTTTTGC
       430       440       450       460       470       480

AACGACGGTAAACGGGTAGAGAAAGAAAAATAGTAACAGACAGGGCCTCAAGCACTATTT
       490       500       510       520       530       540

                  H  A  S  L  F  I  Y  D  G  G  R  N  V  T  V
TCCCCCCTTTTACATATGGCAAGTCTATTTATATATGATGGGGGCCGCAACGTAACCGTG
       550       560       570       580       590       600

 T  T  G  R  N  H  R  E  A  E  H  P  T  A  T  G  A  G  S  C
ACAACGGGAAGAAATCATCGTGAAGCTGAGCACCCCACTGCAACGGGTGCCGGATCGTGC

       610       620       630       640       650       660

                                                  J
 N  E  V  V  K  E  T  F  Y  G  N  L  N  K  K  K  W  S  A  T
AACGAAGTGGTTAAAGAGACATTTTATGGAAATCTGAACAAAAAAAAAGTGGAGCGCCACA
       670       680       690       700       710       720

 L  L  F  S  D  C  I  L  G  P  L  H  P  A  A  A  G  R  E  H
TTGTTGTTTAGTGATTGCATTTTGGGCCCGCTGCACCCAGCTGCTGCTGGAAGGGAAATG
       730       740       750       760       770       780

 G  S  G  H  G  P  S  T  A  Q  T  A  H  C  L  D  G  K  D  V
GGTTCGGGACACGGCCCAAGCACCGCGCAGACAGCTCATTGTTTGGATGGAAAGGATGTT
       790       800       810       820       830       840

 F  M  K  E  I  G  V  R  H  W  N  L  R  E  V  V  G  E  R  E
TTCATGAAGGAAATAGGGGTTAGGCACTGGAATTTGAGAGAGGTTGTAGGAGAAAGAGAG
       850       860       870       880       890       900

 T  G  W  V  K  A  A  I  R  R  G  E  A  V  Y  L  L  R  R  A
ACTGGTTGGGTGAAAGCGGCTATAAGACGGGGAGAGGCGGTATACTTGTTGAGACGTGCG
       910       920       930       940       950       960

 D  V  E  G  E  K  K  E  G  D  I  A  S  G  N  L  C  G  L  R
GATGTTGAAGGGGAAAAAAAAGAAGGGGATATAGCTTCAGGAAATCTATGTGGTTTACGC
       970       980       990      1000      1010      1020

 V  S  L  Y  T  N  V  W  E  M  T  G  V  R  M  A  R  R  H  G
GTCAGTCTCTACACGAACGTGTGGGAGATGACGGGGGTCCGTATGGCTAGACGCCACGGA
      1030      1040      1050      1060      1070      1080

 K  G  D  R  P  S  Q  K  P  K  S  S  S  Q  K  R  C  R  H  L
AAGGGGGACCGACCCTCACAGAAACCGAAATCCAGTAGTCAAAAGCGCTGCCGACATCTG
      1090      1100      1110      1120      1130      1140
```

```
 S  G  S  S  W  R  E  P  R  R  T  I  A  P  A  R  T  A  F  E
AGCGGCAGTAGTTGGAGGGAACCGCGTCGCACAATAGCACCGGCTCGCACCGCATTTGAA
      1150       1160       1170       1180       1190       1200

 S  T  S  S  P  L  H  S  L  P  M  L  S  R  N  R  G  Y  V  V
TCCACAAGCTCCCCCCTCCACTCCCTCCCCATGCTTTCCAGGAACCGTGGGTACGTGGTA
      1210       1220       1230       1240       1250       1260

                  2
 I  P  E  K  P  N  R  S  T  A  P  Y  Y  E  V  G  P  A  K  A
ATTCCCGAAAAACCTAATCGGTCAACCGCCCCATATTATGAAGTAGGCCCGGCTAAGGCA
      1270       1280       1290       1300       1310       1320

 W  G  F  P  M  H  A  Q  R  Q  K  V  F  F  S  F  F  M  K  H
TGGGGATTCCCCATGCACGCACAGAGGCAGAAGGTTTTTTTTTCCTTTTTTATGAAGCAT
      1330       1340       1350       1360       1370       1380

 I  R  S  N  S  P  P  S  L  A  A  V  D  M  T  T  L  T  G  A

ATTAGGTCAAACTCTCCACCGTCGCTTGCCGCCGTCGACATGACGACTTTAACAGGTGCA
      1390       1400       1410       1420       1430       1440

 V  A  G  D  G  K  N  A  N  L  A  P  I  P  K  F  Q  K  T  I
GTAGCAGGTGACGGCAAAAATGCGAACCTCGCACCTATACCAAAGTTCCAGAAGACGATT
      1450       1460       1470       1480       1490       1500

          3
 K  K  R  K  T  I  C  E  T  L  F  P  D  Q  H  P  K  Q  G  P
AAAAAAAGAAAAACAATATGCGAAACCCTTTTCCCGGATCAACACCCCAAACAGGGACCA
      1510       1520       1530       1540       1550       1560

                  4                  3
 S  A  Y  I  S  K  R  I  S  N  P  S  L  F  L  E  M  R  L  P
TCCGCCTACATCAGCAAAAGAATAAGTAATCCCTCTCTCTTTTTGGAAATGCGCCTTCCC
      1570       1580       1590       1600       1610       1620

 R  V  S  N  P  V  V  L  P  R  H  L  E  G  V  S  S  T  D  W
CGAGTGTCGAACCCTGTGGTGTTGCCGCGTCACTTGGAGGGGGTTTCAAGTACAGACTGG
      1630       1640       1650       1660       1670       1680

 F  D  T  S  D  R  N  G  I  I  D  V  V  P  L  S  F  C  E  M
TTCGATACGTCGGATCGCAATGGAATAATAGATGTTGTTCCTCTTAGTTTTTGTGAGATG
      1690       1700       1710       1720       1730       1740

                                          4                  5
 Q  C  D  P  Q  P  N  S  G  I  S  N  G  N  K  T  N  K  T  N
CAATGCGACCCGCAACCCAATAGCGGCATAAGCAATGGAAACAAAACAAATAAAACAAAT
      1750       1760       1770       1780       1790       1800

                                    6
 A  N  T  N  K  E  S  S  F  H  N  V  S  R  S  G  K  G  G  S
GCCAATACTAACAAAGAGTCTTCTTTCCACAATGTTTCACGGAGTGGTAAAGGAGGCTCT
      1810       1820       1830       1840       1850       1860

 P  P  R  A  D  G  R  W  N  G  R  T  K  G  V  R  S  N  S  P
CCTCCGCGTGCCGATGGACGATGGAATGGGCGGACAAAGGGTGTAAGAAGTAATAGCCCC
      1870       1880       1890       1900       1910       1920

 R  M  S  N  Q  R  P  E  P  L  E  A  H  T  S  A  T  C  S  P
CGTATGTCTAACCAACGCCCTGAACCGTTAGAGGCCCACACTTCTGCAACCTGTTCCCCA
      1930       1940       1950       1960       1970       1980

 A  F  L  Q  A  P  A  V  F  V  L  E  M  A  G  I  R  C  T  S
GCCTTTCTTCAGGCGCCTGCCGTGTTCGTACTTGAAATGGCGGGCATTCGGTGTACTTCT
      1990       2000       2010       2020       2030       2040

 P  E  V  K  F  T  A  H  S  T  G  I  S  E  R  K  E  R  L  Q
CCTGAGGTCAAATTCACGGCCCATTCCACGGGAATATCGGAAAGGAAAGAGAGGCTTCAG
      2050       2060       2070       2080       2090       2100
```

```
 C  P  N  L  R  Q  K  T  P  T  T  C  G  K  I  D  A  A  G  V
TGTCCTAACCTCCGCCAGAAAACGCCAACTACGTGCGGTAAAATCGACGCGGCGGGGGTA
      2110       2120       2130       2140       2150       2160

 S  R  Y  P  V  I  T  D  R  R  V  I  H  P  E  T  T  I  S  G
TCACGGTATCCAGTGATTACGGATCGTCGCGTCATCCATCCAGAGACCACCATATCGGGG
      2170       2180       2190       2200       2210       2220

 P  E  T  S  K  Q  V  A  D  Q  Q  L  P  F  I  P  S  S  A  A
CCAGAGACATCCAAACAGGTTGCCGACCAGCAGCTGCCCTTTATTCCTTCGTCTGCTGCT
      2230       2240       2250       2260       2270       2280

 G  Q  *  *
GGCCAGTAGTAGCAGCAGCAAATGAAATCGAGCTCAAGAAGCCAACGCGGACCCTAAAAG
      2290       2300       2310       2320       2330       2340

AAGAAGGGTGAGGTGCCTTAGCGGTTTCCATCCTTCAGGGGCCGCAACAAATACGCAGCC
      2350       2360       2370       2380       2390       2400

TCTTGGCGGAGCTGTTGAACGAAGCAGCGGAGGCGGTGCCCGAAATAGACAAATTGGGAA
      2410       2420       2430       2440       2450       2460

CACCCGGTGTGTGCAGCCCAATAAGTGCGTTTATCGGTGCCATAGCGGGTACCTCAGCTA
      2470       2480       2490       2500       2510       2520

                                                  R
CGACTAACTCCGCATACAAGATCCGGGGAATTC
      2530       2540       2550
```

### 4.2.2.1 Structure of the promoter

The sequences upstream of the transcriptional start site in the trypanosome gene are consistent with those in a gene that is transcribed by RNA polII. This involves a putative promoter core that consists of a classical TATA box which satisfies the consensus TATA(T/A)AT (217). Twelve base pairs upstream from this element is what may the first translation initiation site, ATG. Further downstream, at about 200bp can be found a second possible translation start site. The significance of this arrangement is not clear, although it may be said that the first methionine is the more likely candidate for translation to start from.

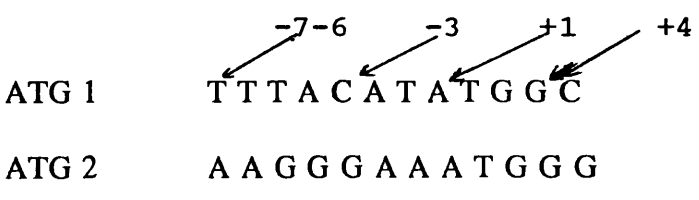
In addition to this motif, there is a classical CCAAT box (GGCCAAT) at about 330bp from the start codon. This box is usually found about 80bp from the initiation site in eukaryotic genes. The CCAAT and TATA boxes flank two stretches of polypyrimidines. These stretches have been underlined with broken lines in **Fig. 43**. Sequences of that nature have been observed in trypanosome genes and have been implicated as the "flag posts" for trans-splicing (99).

### 4.2.2.2 Translational control signals

From an analysis of a number of protozoan genes (see Fig.15), it has been observed that the consensus sequence flanking the initiation codon has the arrangement:

T T - - A - - A T G (G/C)

The sequences flanking the translational start site in the trypanosome gene comprise:



The consensus derived by Kozak comprises the sequence

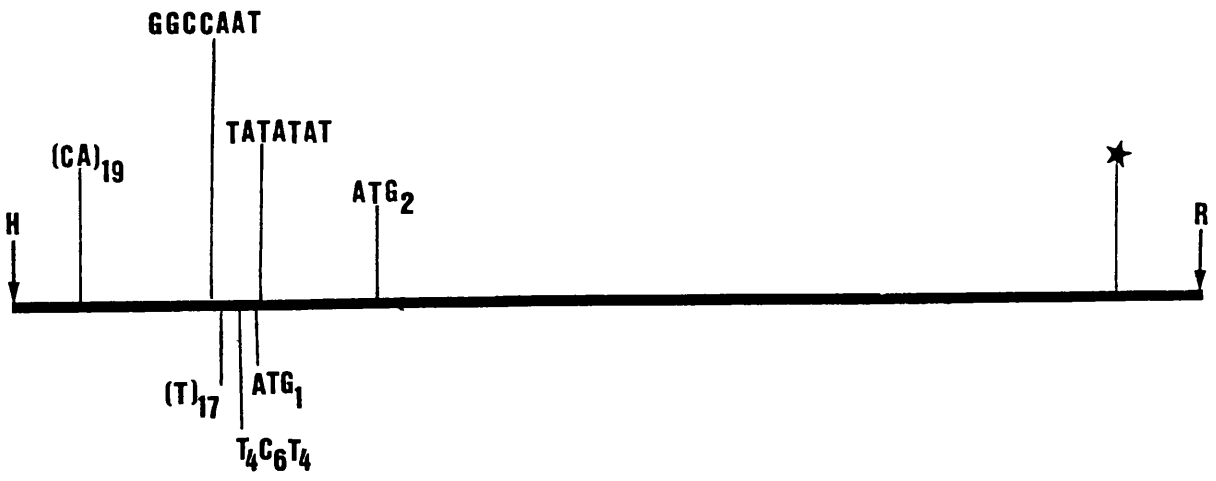
A/G - -A T G G

as the favoured sequence for eukaryote translation initiation sites. According to this rule, most authentic initiator codons are preceded by a purine (usually A, but may also be G) in position -3 (47). It may be said that the flanking sequences of the two possible initiation sites observed in the trypanosome gene sequence is consistent with that consensus or rule and with the consensus derived for protozoan genes.

There is an unusually strong stop signal represented by two (TAGTAG) amber codons. The sequence between the start codon and this stop signal constitutes an open reading frame of 1731bp or 577 amino acids. The encoded protein will therefore have a molecular weight of about 63,346 daltons.

There is no identifiable polyadenylation signal (AAUAAA). The closest identity to that signal is a stretch of nucleotides comprising AAATGAAA located ~10bp downstream of the stop codon.

Fig. 44 Summary structure of the putative trypanosome gene. The stop codon is represented by a star. The putative initiation sites and promoter and enhancer elements [4.2.2.3] are indicated, in addition to the polypyrimidine tracts. Cloning sites are HindIII (H) and EcoR1 (R).



#### 4.2.2.3 The [d(CA).d(GT)] element

There is a run of polypurine/polypyrimidine residues comprising 38bp of alternating C and A, approximately 100bp from the HindIII site. This region adds to the overall repetitive nature of the residues 5' to the coding sequence of the gene.

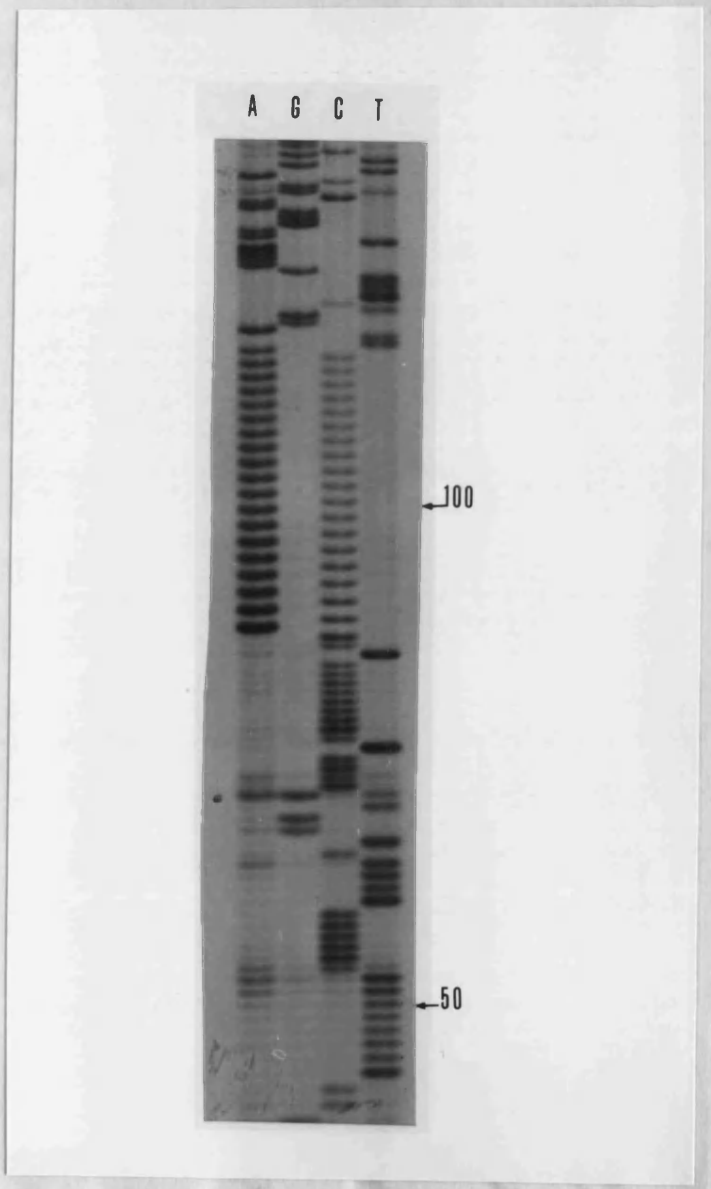



Fig. 45 Autoradiograph of the partial sequence of pTGT-4 showing $(CA)_{19}$ stretch. Distance from HindIII site is indicated.

#### 4.2.2.4 Codon usage and dinucleotide frequencies

There is little evidence of a biased codon usage in this gene. This may be due to a paucity of iso-accepting tRNAs for the various codons. What this implies is that it may not have a house-keeping function such as actin or tubulin which are continually expressed; genes in that category have a more biased codon choice, and are abundant within the cell (41).

However, there is close agreement with previous observations with respect to dinucleotide frequencies. According to Nussinov's definition of dinucleotide patterns in eukaryotes, the following hierarchical preference is adhered to: GC > CG, TG > GT, CA > AC, AT > TA, CT > TC, and AG > GA (51). In the trypanosome sequence, the following pattern has been observed: GC > CG, GT > TG, CA > AC, AT > TA, TC > CT, and GA > AG. Hence, there is a significant agreement with that prediction. In addition, the sequence agrees with the codon preferences of other trypanosome genes (41); it prefers a pyrimidine in the wobble positions of the codons.

Fig. 46 Codon usage in the trypanosome gene

| Amino acid | Codons | Number of amino acids | Subtotal | Percent of total amino acids | |
|---|---|---|---|---|---|
| Phe | TTT | 12 | | 2.08 | |
| | TTC | 8 | 20 | 1.39 | 3.47 |
| Leu | TTA | 3 | | 0.52 | |
| | TTG | 10 | | 1.73 | |
| | CTT | 8 | | 1.39 | |
| | CTC | 6 | | 1.04 | |
| | CTA | 2 | | 0.35 | |
| | CTG | 4 | 33 | 0.69 | 5.72 |
| Isoleu | ATT | 7 | | 1.21 | |
| | ATC | 3 | | 0.52 | |
| | ATA | 13 | 23 | 2.25 | 3.98 |
| Methionine | ATG | 13 | 13 | 2.25 | 2.25 |
| Val | GTT | 11 | | 1.91 | |
| | GTC | 5 | | 0.87 | |
| | GTA | 9 | | 1.56 | |
| | GTG | 10 | 35 | 1.73 | 6.06 |
| Serine | TCT | 9 | | 1.56 | |
| | TCC | 10 | | 1.73 | |
| | TCA | 7 | | 1.21 | |
| | TCG | 8 | | 1.39 | |
| | AGT | 12 | | 2.08 | |
| | AGC | 8 | 54 | 1.39 | 9.36 |

| | | | | | |
|---|---|---|---|---|---|
| Proline | CCT | 10 | | 1.73 | |
| | CCC | 12 | | 2.08 | |
| | CCA | 11 | | 1.91 | |
| | CCG | 11 | 44 | 1.91 | 7.63 |
| Threonine | ACT | 7 | | 1.21 | |
| | ACC | 8 | | 1.39 | |
| | ACA | 13 | | 2.25 | |
| | ACG | 12 | 40 | 2.08 | 6.93 |
| Alanine | GCT | 12 | | 2.08 | |
| | GCC | 13 | | 2.25 | |
| | GCA | 10 | | 1.73 | |
| | GCG | 9 | 44 | 1.56 | 7.63 |
| Tyrosine | TAT | 5 | | 0.87 | |
| | TAC | 4 | 9 | 0.69 | 1.56 |
| Histidine | CAT | 6 | | 1.04 | |
| | CAC | 11 | 17 | 1.91 | 2.95 |
| Glutamine | CAA | 5 | | 0.87 | |
| | CAG | 13 | 18 | 2.25 | 3.12 |
| Asparagine | AAT | 16 | | 2.77 | |
| | AAC | 12 | 28 | 2.08 | 4.85 |
| Lysine | AAA | 24 | | 4.16 | |
| | AAG | 11 | 35 | 1.91 | 6.07 |
| Aspartic | GAT | 12 | | 2.08 | |
| acid | GAC | 7 | 19 | 1.21 | 3.29 |
| Glutamic | GAA | 17 | | 2.95 | |
| acid | GAG | 14 | 31 | 2.43 | 5.38 |

| | | | | | |
|---|---|---|---|---|---|
| Cysteine | TGT | 6 | | 1.04 | |
| | TGC | 6 | 12 | 1.04 | 2.08 |
| Tryptophan | TGG | 8 | 8 | 1.39 | 1.39 |
| Arginine | CGT | 9 | | 1.56 | |
| | CGC | 11 | | 1.91 | |
| | CGA | 4 | | 0.69 | |
| | CGG | 6 | | 1.04 | |
| | AGA | 9 | | 1.56 | |
| | AGG | 8 | 47 | 1.39 | 8.15 |
| Glycine | GGT | 9 | | 1.56 | |
| | GGC | 10 | | 1.73 | |
| | GGA | 17 | | 2.95 | |
| | GGG | 11 | 47 | 1.91 | 8.15 |
| Stop | TGA | 0 | | | |
| | TAG | 1 | | | |
| | TAA | 0 | | | |
| Total codons | | | 577 | | 100 |

## 4.3 Database search

When the entire putative trypanosome GT sequence was used to search the databases, a number of protein sequences mainly membrane proteins unrelated to glucose transport, were identified.

An alternative approach to searching with the entire sequence was to use a part of the sequence which bears a subtle identity to the putative glucose binding site in the GT superfamily. Previous attempts at searching with the various conserved motifs of the GTs as

strings (PESPR, VERAGRRTL, VPETK) had yielded many proteins with no relationship to glucose transport whatsoever. Searching with QQLSGINAVFYY, which is implicated in glucose binding, however identified glucose transporters exclusively (21 out of a total of 22). Hence this is the only diagnostic feature of the GTs. On this premise, QXXSGI from the trypanosome sequence was used as the search string. As a result, several of the GTs were identified from the database. This is by no means suggestive of its identity.

## 4.4 Sequence alignment and homology search

The alignments below underscore the low level of homology between the trypanosome sequence and the GTs. There is however a better consensus between that sequence and the mammalian GTs than with the GTs of other eukaryotes and prokaryotes. A number of deletions and insertions made alignment difficult. Even though the hybridizing regions can be seen, the overall identity is very poor. This low identity may be largely ascribed to the evolutionary distance between the trypanosome and other organisms, and to possible selection pressures that resulted in replacement substitutions. This is assuming that this sequence is in any way related to the GTs. Nonetheless, there are significant pockets of homology corresponding to some of the key conserved modules within the superfamily. There is no sequence information regarding intrinsic membrane proteins in trypanosomes to enable an objective comparison within the context of hydrophobicity profiles, deletions or insertions. This makes it extremely difficult to say one way or the other what the nature of the trypanosome sequence is, from which a functional inference may be made. At the same time, the spatial or sequential arrangement of some motifs (if taken in isolation) is coincident with that in the GTs; this is an arrangement that

may not occur by chance.

```
Protein sequence     : Q        Q        L       T    G    I
Oligo probe sequence: CA(G/A) CA(G/A) C(T)TN ACN GGN AT(A/C/T)
Trypanosome sequence: CAA      CCC      AAT     AGC GGC ATA [AGC]
Peptide sequence     : Q        P        N       S    G    I    S*

(continued)
Protein sequence     : N        A    V    M    Y        Y        A
Oligo probe sequence: AA(C/T) GCN GTN ATG TA(C/T) TA(C/T) GCN
Trypanosome sequence: AAT       GGA AAC AAA ACA       AAT       AAA
Peptide sequence     : N        G    N    K    T        N        K
```

Fig. 47A (a) Partial sequence alignment and comparison of the gene with Oligo 2. (*) is a serine insertion.

Fig. 47A (b) Alignment with rat liver GT cDNA. Colour connotations are as previously defined.

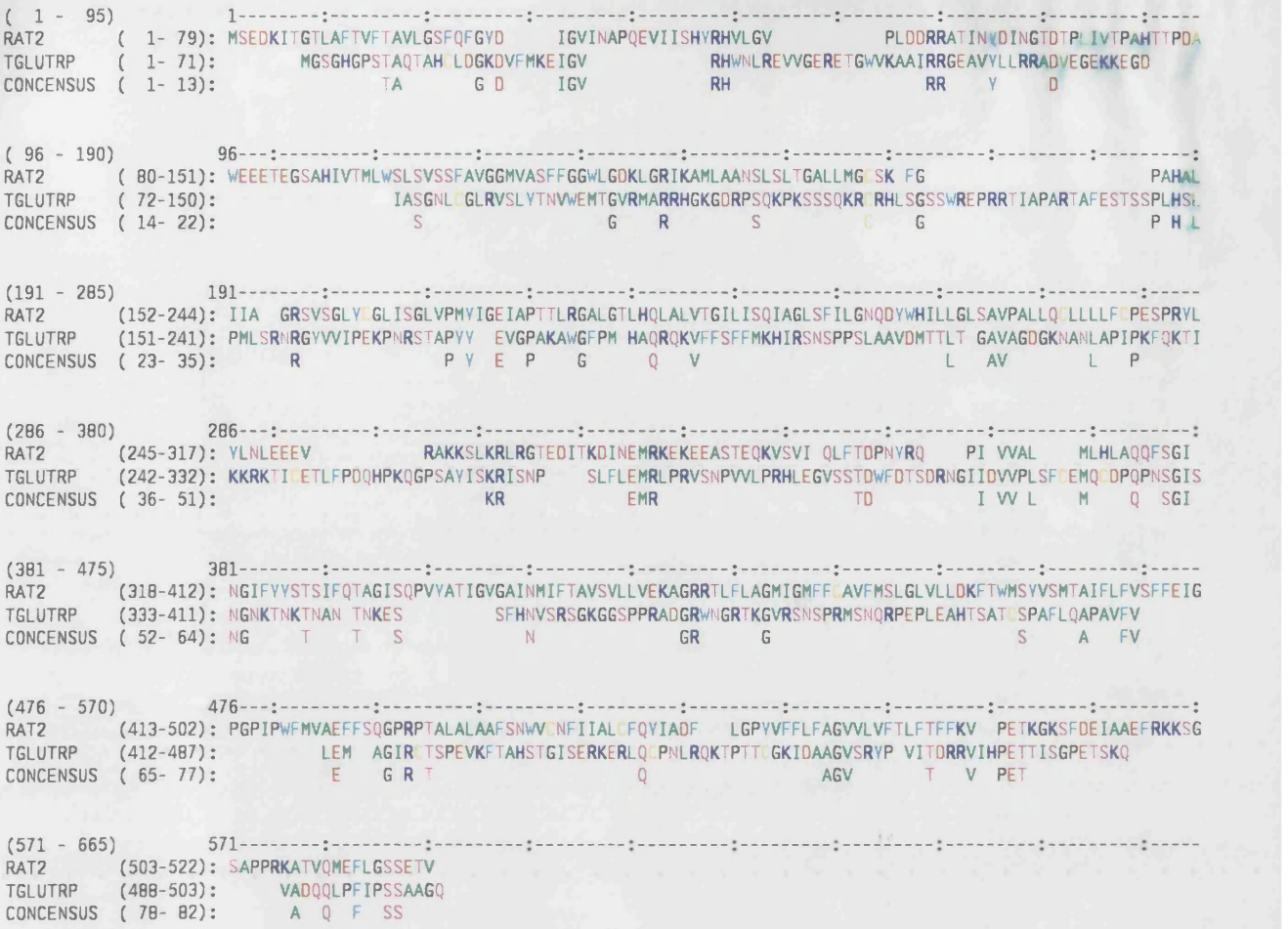


Fig. 47B Multiple sequence alignment with mammalian GTs, representing GLUTs 1-5. Definitions and colour connotations are as earlier indicated in **Fig. 28.** GT definitions are as before.

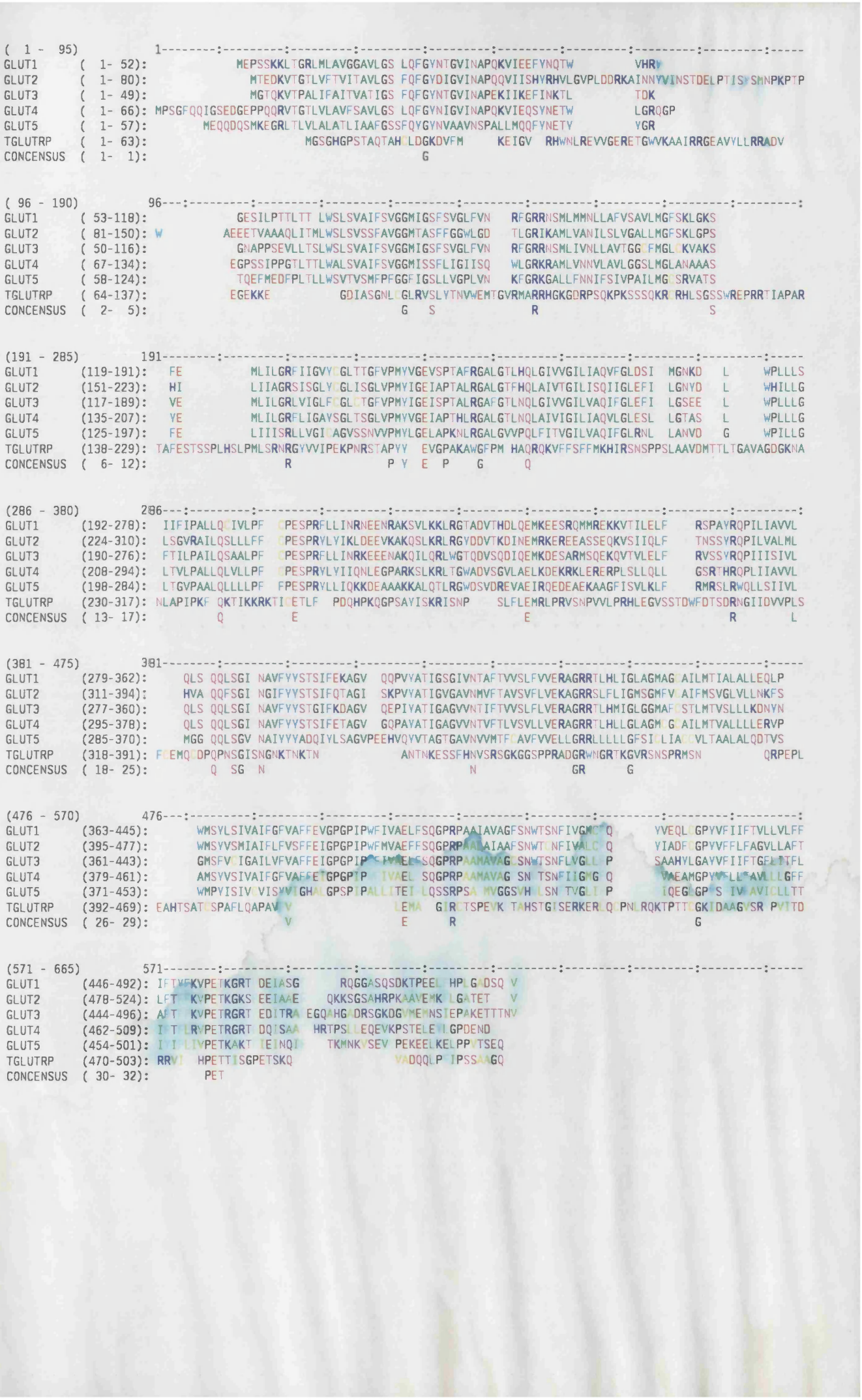
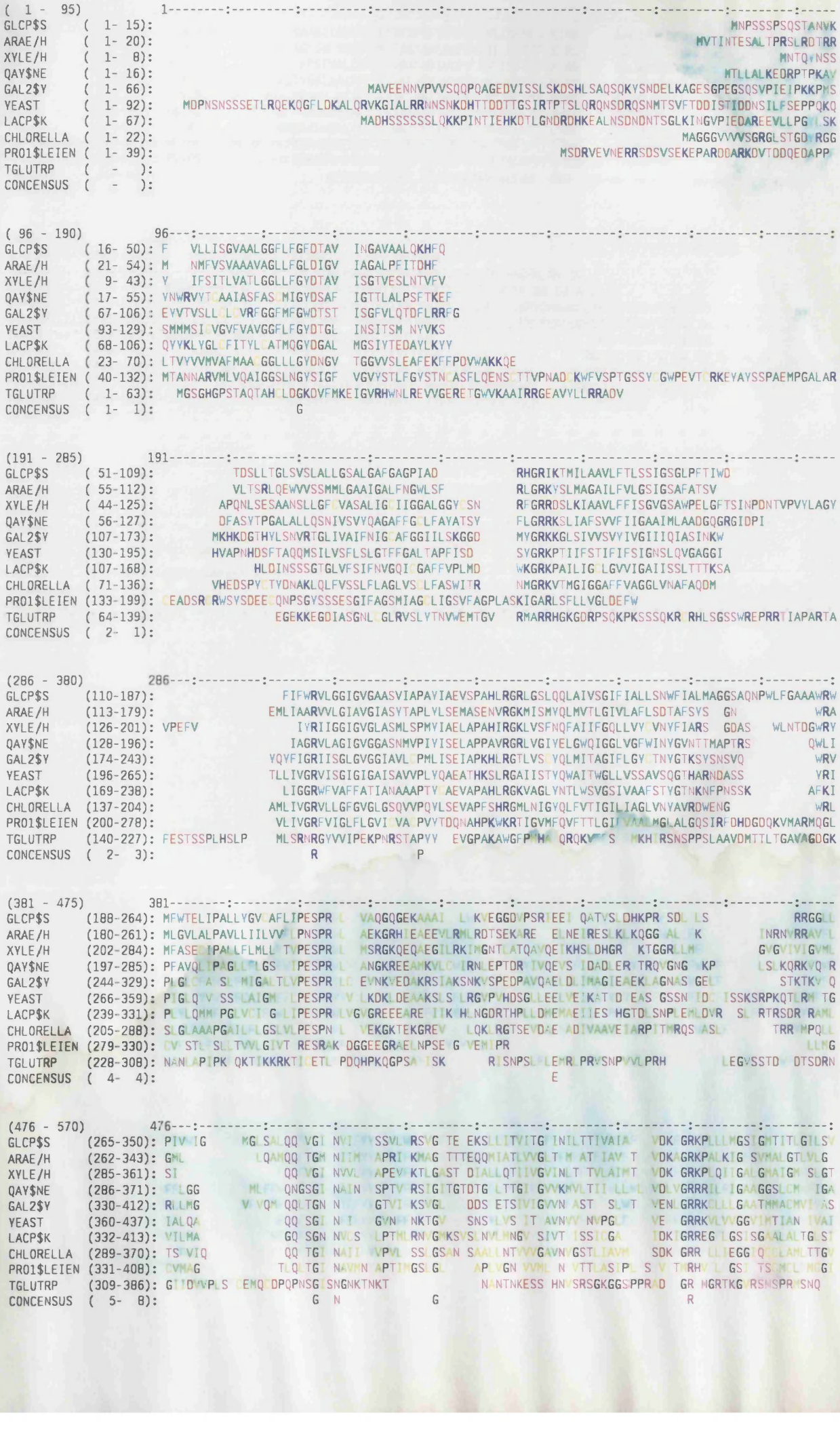


Fig. 47C Multiple sequence alignment of select GTs of prokaryotic and eukaryotic origins. Colour and transporter definitions from previous figures apply. Previous GT definitions apply.



```
(571 - 665)        571--------:---------:---------:---------:---------:---------:---------:---------:---------:-----
GLCP$S     (351-428): VFGGATVVNG QP TLT GAAGIIALVTANLYVFSFGFSWGPIVWVLLGEMFNNKIRA      AALSVAAGVQWIAN IISTT     PP
ARAE/H     (344-417): CLMQFD     NG TAS SGLSWLSVGMTMMCIAGYAMSAAPVVWILCSEIQPLKGRD      FGITCSTTTNWVSNMIIGAT     LT
XYLE/H     (362-430): AFYTQ              APGIVALLSMLFYVAAFAMSWGPVCWVLLSEIFPNAIRG      KALAIAVAAQWLANWVSMT      PMM   D
QAY$NE     (372-450): YIKIADPGSNKAE DAKLTSGGIAAIFFFYLWTAFYTPSWNGTPWVINSEMFDQNTRS      LGQASAAANNWFWNFIISR      TP
GAL2$Y     (413-490): VGVTRLYPHGKSQ PSS KGAGNCMIVFTCFYIFCYATTWAPVAWVITAESFPLRVKS      KCMALASASNWVWGFLIAFF      TP
YEAST      (438-508): VGCSL       K TVA   AAKVMIAFICLFIAAFSATWGGVVWVISAELYPLGVRS      KCTAICAAANWLVNFICALI      TP    I
LACP$K     (414-485): CTARY        E KTKKKSASNGALVFIYLFGGIFSFAFTPMQSMVSTEVSTNLTRS      KAQLLNFVVSGVAQFVNQIA      TP
CHLORELLA  (371-447): VLAIEF    AKYGTDPLPKAVASGILAVICIFISGFAWSWGPMGWLIPSEIFTLETRP      AGTAVAVVGNFLFSFVIGQA      LV
PRO1$LEIEN (409-490): PVYPG         VSKKLEAKNGVAITGILLFILGFEVCVGPCYYVLTQDMFPPSFRPRGASFTQVAQFIFNLIINVCVPIATESISGGP
TGLUTRP    (387-452):                      RPEPLEAHTSATCSPAFLQAPAVFVLEMA  GIRCTSPEVKFTAHSTGISERKERLQCPNLRQKTPTT
CONCENSUS  (  9-  8):

(666 - 760)        666---:---------:---------:---------:---------:---------:---------:---------:---------:---------:
GLCP$S     (429-468):     LLDTVGLGPAVGLVATSAAISIFFIWFFVKETKGKTLEQM
ARAE/H     (418-472):     LLDSIGAAGTFWLYTALNIAFVGITFWLIPETKNVTLEHIERKLMAGEKLRNIGV
XYLE/H     (431-491): KNSWLVAHFHNGFSYWIYGCMGVLAALFMWKFVPETKGKTLEELEALWEPETKKTQQTATL
QAY$NE     (451-537):     QMFIKMEYGVYFFFASLMLLSIVFIYFFLPVTKSIPLEAMDRLFEIKPVQNANKNLMAELNFDRNPEREESSSLDDKDRVTQTENAV
GAL2$Y     (491-574):     FITSAINFYYAYVFMGCLVAMFFYVFFFVPETKGLSLEEIQELWEEGVLPWKSEGWIPSSRRGNNYDLEDLQHDDKPWYKAMLE
YEAST      (509-573): VDTGSHTSSLGAKIFFIWGSLNAMGVIVVYLTVYETKGLTLEEIDELY            IKSSTGVVSPKFNKDIR
LACP$K     (486-570):     KAMKNIKVWFYVFYVFFDIFEFIVIYFFFVETKGRSLEELEVVFEA      PNPRKASVDQAFLAQVRATLVQRNDVRVANAQNLKEQEP
CHLORELLA  (448-533):     SMLCAMEYGVFLFFAGWLVIMVLCAIFLLPETKGVPIERVQALYARHW    FWNRVMGPAAAEVIAEDEKRVAAASAIIKEEELSKAMK
PRO1$LEIEN (491-552):        SGNQDKGQAVAFIFFGGLGLICFVIQVFFLHPWDEERDGKKVVAP    AIGKKELSEESIGNRAE
TGLUTRP    (453-503):           CGKIDAAGVSRYPVITDRRVIHPETTISGPETSKQVADQQLP      FIPSSAAGQ
CONCENSUS  (  9-  9):                                        E

(761 - 855)        761--------:---------:---------:---------:---------:---------:---------:---------:---------:-----
GLCP$S     (   -   ):
ARAE/H     (   -   ):
XYLE/H     (   -   ):
QAY$NE     (   -   ):
GAL2$Y     (   -   ):
YEAST      (   -   ):
LACP$K     (571-587): LKSDADHVEKLSEAESV
CHLORELLA  (   -   ):
PRO1$LEIEN (   -   ):
TGLUTRP    (   -   ):
CONCENSUS  ( 10-  9):
```

## 4.5 Hydrophobicity profile

The overall charge distribution in Tglutrp is inconsistent with that either of the GTs or membrane spanning proteins in general. As a result this protein is largely amphipathic. Even though charged groups have been found in some membrane spanning domains of bacteriorhodopsin (186) and lac permease (144), the preponderance of these groups in Tglutrp makes it difficult for its inclusion in that category. However, it may be said that the encumbrance of the VSGs and the fact that the overall surface charge of the trypanosome is negative (which facilitates their purification from red blood cells on a DEAE-cellulose column), may justify the presence of an usually large number of positive charges in this protein. These charges would therefore enable it to anchor itself by classical ionic interactions to the VSG coat.

The apparent paucity in hydrophobic segments in the protein should however not preclude it from being considered as a membrane protein. It has been noted that several other membrane proteins contain charged residues within their presumptive transmembrane segments especially the Na+ channel protein of the eel (218). The latter is the strongest case in point with about 4-7 arginine residues within the S4 domain, which are presumed to line the walls of the channel. The other domains contain several charges as well.

The poor segmental homology observed between Tglutrp and the GTs implies that any attempt at delineating hydrophobic segments may meet with only short stretches (8-14) of hydrophobic residues. By the dictates of the various algorithms in current use, none of these can or should span the membrane. However, there is a possibility that relatively short a-helices may be formed by these segments as has been proposed for the eel Na+ channel and depicted and discussed in the model below **[5.4.2]**. Hence, by comparison to classical GTs, Tglutrp has a

relatively diminished hydrophobicity profile with no discernible 12 putatively membrane spanning helices. The only discernible hydrophobic and potentially membrane spanning segments are indicated below. The relevance of this arrangement if far from clear.

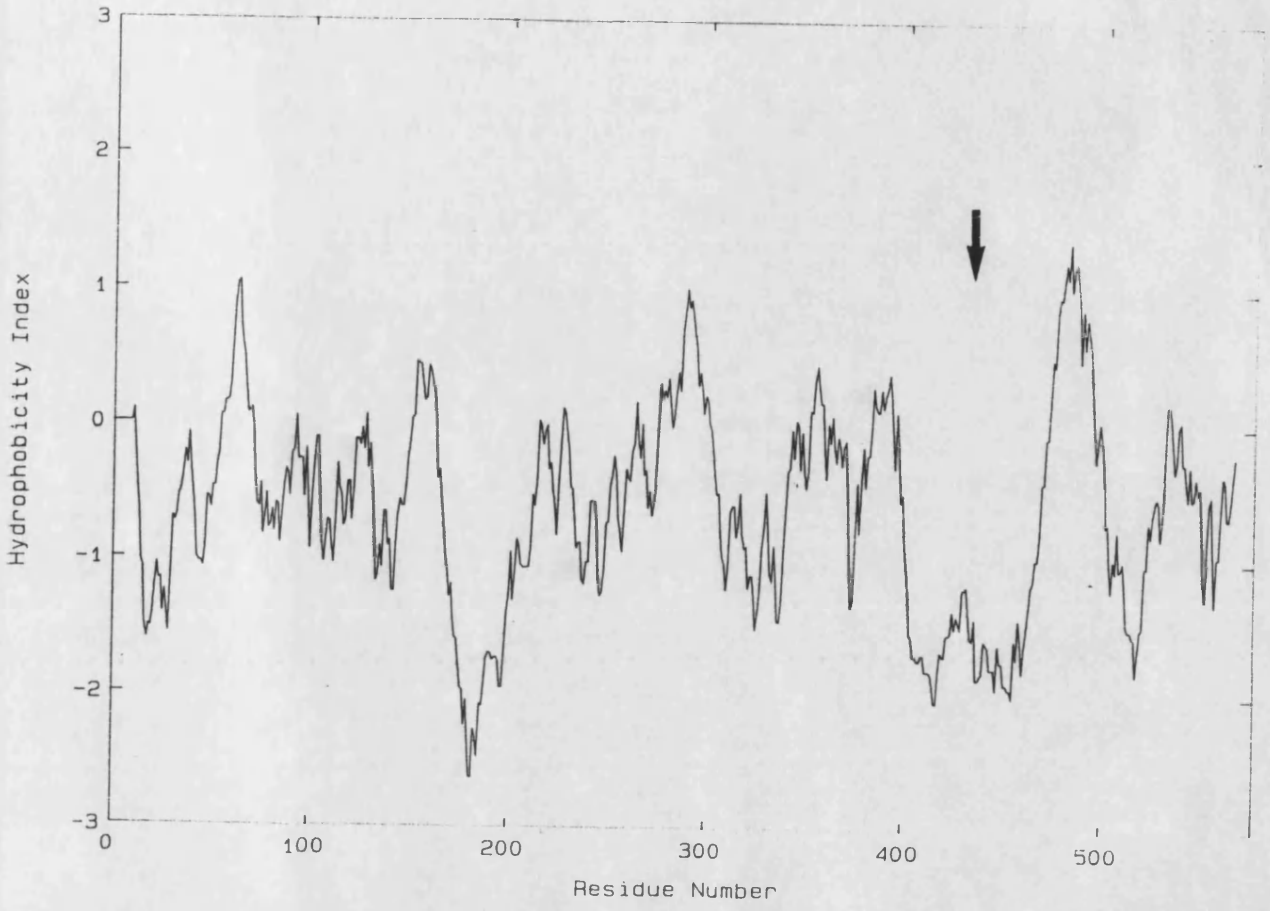



Fig. 48 Hydrophobicity profile of the trypanosome sequence based on the Kyte-Doolittle algorithm. The arrow indicates a region with a channel forming potential. Horizontal bar represents the threshold hydropathy for membrane-spanning domains.

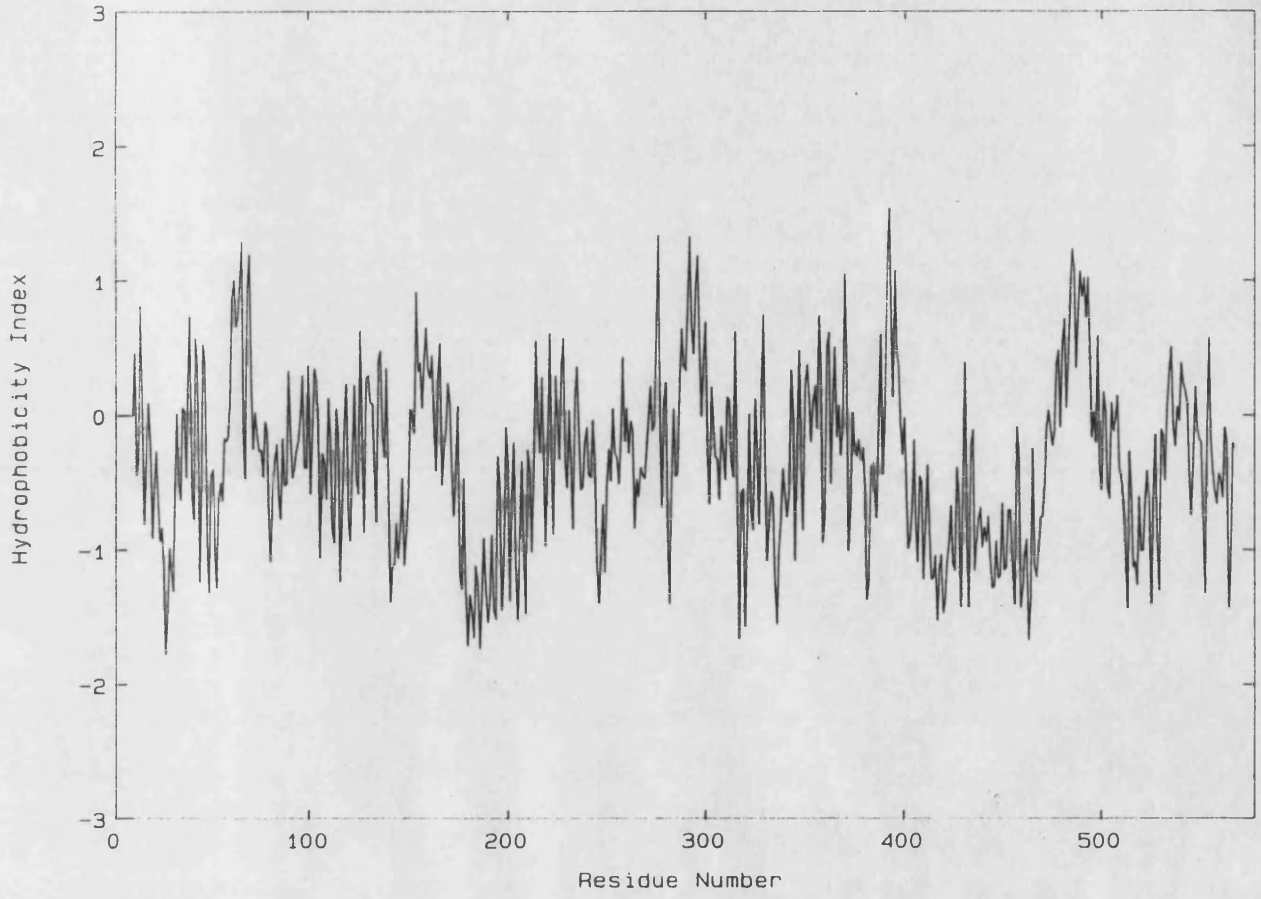



Fig. 49 Amphipathy profile of the trypanosome sequence based on the Eisenberg algorithm (158).

## 4.6 Secondary structure prediction

The a-helix content of the entire protein is low although it is fair to say that it contains equal numbers of residues with preference for a-helices as it has for b-sheets. This means that the contributions of a-helices and b-sheets are about the same, which again is not compatible with previous observations on the erythrocyte GT. Only 4 and possibly six potential a-helix nucleation sites can be deciphered.

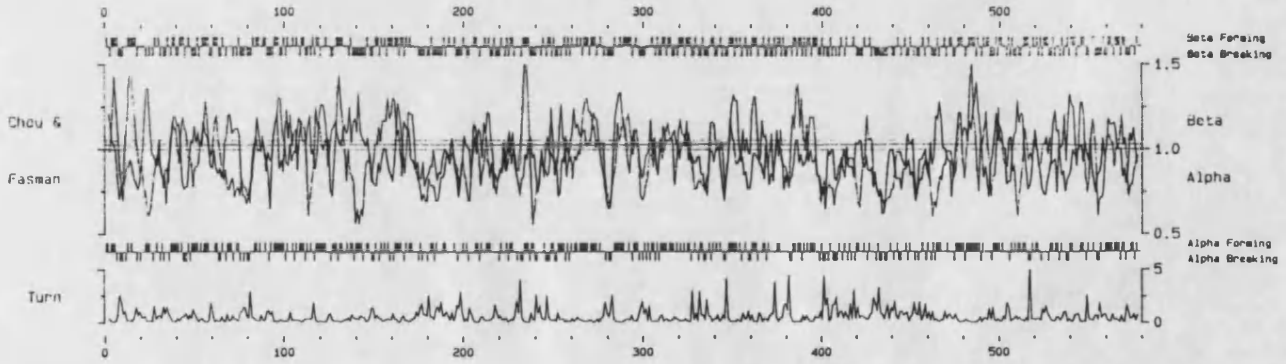



Fig. 50 Secondary structure prediction for the trypanosome protein.

## 4.7 In vitro translation

Run-off RNA transcripts of the gene (cloned in pBluescript) were used for translation in a rabbit reticulocyte system. Translation products were run on a 10% SDS-PAGE gel under denaturing conditions. The gel was impregnated with a fluor (En$^3$Hance), dried and exposed to X-ray film.

A band running at 34-36kD was observed. This is not consistent with the molecular weight of ~55kD from the predicted amino acid sequence. This result may be due to an incomplete translation product. However, the observed molecular weight compares well with the sizes of glucose transporters that have been similarly treated. For example, the xylose transporter runs at 37-41kD, the H+/arabinose transporter at about 36kD, and the H+/galactose transporter (GalP) at about 37kD (219).

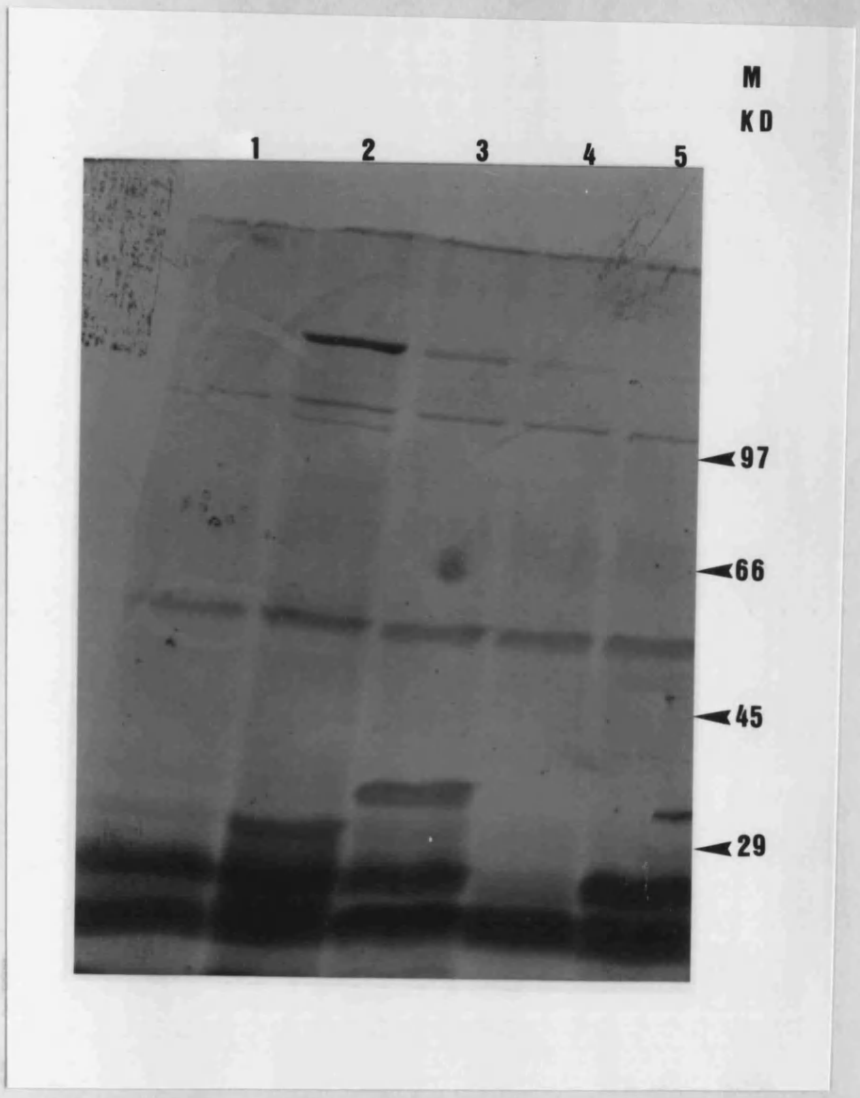



Fig. 44 SDS-PAGE fluorogram of <u>in vitro</u> translation products of pBS.TG-4. [1] and [2] are translations with uncapped transcripts from the sense and antisense strands respectively. [3] and [4] are capped transcripts of the sense and antisense strands respectively. [5] is a translation with water as a control. [1], [2], [4] were internal controls. Molecular weight markers are indicated by M.

## 4.8 Developmental expression of the gene

Northern blot analysis, using this clone as probe on RNA of the developmental forms (bloodstream and culture/procyclic) of T. brucei, showed a band that ran at about 5.2kb in bloodstream form RNA. This is inconsistent with the size of the reading frame observed in the gene. This may be due to the presence of large stretches of trailer sequences, or the message is still in the pre-maturation state.

Significantly, when the oligonucleotide probe encompassing the putative glucose binding site (Oligo 2) was used to probe the blots, a band of a similar size was identified. In addition, the rat cDNA probe seemed to (faintly) identify the same or an identical transcript. This band may be one and the same species in all three cases. Although the poor signalling from the hybridizations may be attributed to probe insensitivity, it is possible that this is due to a low cellular content of the transcript in question i.e it is a rare or lowly expressed message. This probably suggests that it may be an inducible gene. Constitutive transcripts would normally be present at much higher levels than that observed here.

However, the probes did not hybridize to RNA of the culture (procyclic) forms of T. brucei. This is probably (preliminary) evidence that the trypanosome gene is developmentally regulated at the transcriptional level, as it shuttles between the mammalian bloodstream and the insect host. The fact that the identified message is recognized by both the gene and the oligonucleotide probe in particular, makes room for optimism in further characterizing this gene in terms of glucose transport.

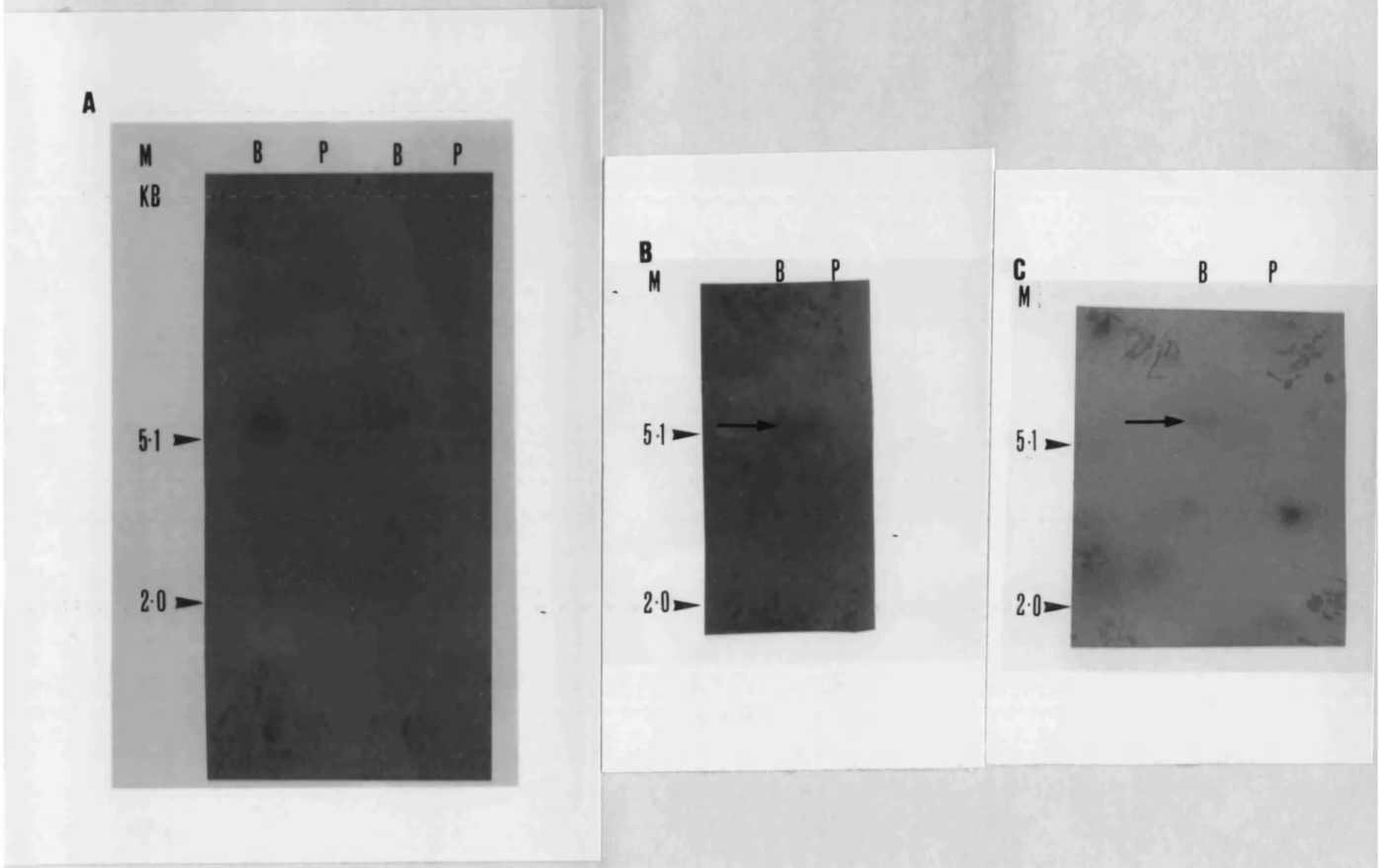



Fig. 52 Northern blot analysis of the RNA of the developmental forms of T. brucei. Hybridizations with (A) trypanosome gene; (B) oligo 2; (C) rat GT cDNA. M= molecular weight markers; B= bloodstream form; P= procyclics.

## 4.9 Functional expression and characterization

An attempt was made at expressing the gene in a mammalian cell line (COS-7) after engineering it into the eukaryotic expression vector pSVL.

Transport data were taken by measuring the uptake of radioactive glucose analogues ($^3$H 2-deoxyglucose) in COS-7 cells transfected with the gene, and cells transfected with only the vector or salmon sperm DNA. There was no difference in uptake between cells with the gene and cells without (controls). A number of explanations may be advanced to account for this observation. The gene may not encode glucose transport; if it does, it is not expressed because it was probably not transfected into the cell. On the other hand, the intrinsic glucose transporter activity of the cultured cells may be masking the activity of the transfected gene.

# CHAPTER 5

# DISCUSSION

## 5.1 Relationship between trypanosome glucose transport and metabolism

There are two basic factors that may dictate the course and rate of metabolism in the bloodstream trypanosome:

a. there are no energy stores or reserves

b. there are no alternative substrates for producing energy apart from glucose.

These factors compound the already inefficient energy producing step in glycolysis. Both factors are equally important and together, may constitute the determinant(s) of the rate of glucose uptake.

The bloodstream trypanosome has an extremely high rate of respiration/oxidation, and this is justified on the basis of (a) above. In this study, it was found that the rate of glycolysis is 86.7nmol $O_2$/min/$10^8$ cells or 173.4nmol $O_2$/mg of protein. This is consistent with previous observations of 60-90nmol $O_2$/min/$10^8$ cells or 120-180nmol $O_2$/min/mg of protein (214). This value is about 20-50X higher than that observed in mammalian cells. This rapid oxidation of glucose can only be matched with a consonant rapid flux of glucose into the cell. This suggests that the rate of glucose transport may be correspondingly 20-50 fold higher than in mammalian cells.

The results of this study identified a saturable carrier consistent with Michaelis-Menten kinetics and in line with previous observations (17,124,125). A Km of 1.13mM for zero trans entry of 2-deoxyglucose was deduced, in tight agreement with 1.54mM for 6-deoxyglucose (17) and 1.98mM for D-glucose (125), and satisfying some

criteria for carrier-mediated transport. These observations are in contrast to others which suggest that the glucose transporter in bloodstream forms is coupled to a Na+ gradient (153). It is certainly possible to overlook this phenomenon in view of the inclusion of 98mM NaCl in the assay buffer used in previous assays, in which case Na+-dependence would have been severely masked. An assay system devoid of this ion may be required to confirm or deny either observation. Kinetic characterization of the trypanosome GT by Eisenthal et al, indicated that the Km for efflux was higher than for influx (17), but this was discounted in view of the inherent difficulties in the assay procedure. They therefore suggested that glucose transport across the membrane was kinetically symmetric. However, Ter Kuile and Opperdoes [1991,(125)] apart from confirming the Km differentials, established an asymmetry factor of 1.79 in close agreement with 1.8 of Eisenthal et al Hence, it is possible that their previous observation was not artifactual afterall, but that the trypanosome GT is indeed kinetically asymmetric in disposition, which is a totally plausible concept.

However, the observation of counterflow from this study in the reconstituted transporter provides proof of a facilitated carrier although it could be argued that such a trans-stimulation profile may be observed in the active transport of two substrates sharing a common carrier. Incidentally, H+-dependent transport has been observed in the procyclics as well (152). The differences in cation specificity has not been explained. Arguably, the observed differences/specificities in glucose transport may be attributable to kinetic differences by way of affinities of the transporters. Whilst the procyclics have a higher affinity (Km 23uM), the corresponding Km for transport in the bloodstream forms is 10X higher at 237uM. This affinity is consistent with that in Na+-dependent glucose uptake (220). The observed affinity of the

procyclic transporter is similar to that observed in Leishmania donovani promastigotes which have a Km of 24uM (149). This is understandable on the basis of the fact that they live in similar environments at that stage of their life-cycle in their respective insect vectors. What may defy reason is the need for a coupled glucose transporter in the bloodstream forms since they are bathed in 5mM glucose.

The sequestration of glycolytic enzymes within the glycosomes may be a teleological adaption to maximise energy flux, but at the same time presents a problem for the localization of the glucose transporter within the trypanosome. Pulse-labelling with $^{14}$-C glucose indicated that the glycosomal components were rapidly labelled compared to their cytosolic counterparts (26). This suggests that the glycosomal membrane presents a permeation barrier to glucose which can only be taken up via a transport protein within the membrane. Unfortunately, there is no information on glycosomal glucose transport. As transport has been observed both in whole cells and in reconstituted vesicles, there must be a plasma membrane glucose transporter. If indeed that is the case as it might, one may be looking at two or three isoforms of the glucose transporter differing more in their stage and tissue specificity than in their structure.

There are other speculations with respect to the subcellular localization of receptor/transport proteins. It has earlier been thought that trypanosomes take up nutrients via the flagellar pocket by endocytosis (221). It is possible that transport proteins are within the flagellar pocket rather than delocalized or "smeared" over the entire plasma membrane. Such an arrangement presupposes that the protein would radiate from this site to expand the repertoire of transporter molecules, and would be anchored to the VSG-plasma membrane interface via a

GPI linker. The VSG coating would also impose some "accessibility constraints" with respect to glucose molecules which further amplifies the need for a localised but accessible transporter shielded from VSG encumbrance.

Whilst it may be fair to say that the two forms of transporter identified in the trypanosome differ kinetically, it is too early to dichotomise them in the absence of more kinetic data and molecular evidence based on the sequences of their cognate structural genes. Although purely conjectural, one could conceive of two forms within the bloodstream: the facilitated transporter being required at the prodromal phase of infection, or probably being glycosomal, whereas another transporter (probably Na+-dependent) becomes operative in a chronic infection (increased parasitaemia), where increased competition for glucose among the parasites may require some form of concentrative mechanism. The latter form of transporter may then be modified for use in the late stumpy (infective) forms prior to its deployment in the insect vector. Undoubtedly, these proteins will have very high turnover rates. The picture that emerges is a glucose transporter(s) under tight molecular controls through and at every stage of the parasite's development. Having said that, one is also looking at a transporter that is potentially structurally malleable.

## 5.2 Molecular cloning of the putative glucose transporter-related gene

The preferred strategy for cloning genes is to use homologous nucleic acid probes derived from the partial amino acid sequence of the encoded protein. Alternatively, the peptide may be synthesized and used to raise antibodies with which to screen an expression library. In many cases, proteins can be purified easily and in sufficient quantities for an N-terminal peptide sequence to be obtained. This is especially possible if the purification is of an enzyme which relies on a simple assay system. However, it is very difficult to obtain pure intrinsic membrane proteins because their purification often relies on using detergents on account of their lipid environment, there is no easy assay system to follow, and they are often present in very small amounts within the membrane. The only exception is the human red blood cell glucose transporter which constitutes about 5% of the total cell protein which eased its purification. Antibodies raised against a peptide from a partial sequence of the protein enabled the isolation of the HepG2 transporter from an expression library (156).

On the other hand, the binding specificities of the protein may give clues as to what tags may be used to identify it to ease purification. Under those circumstances, the protein of interest is easily identifiable on an SDS-PAGE gel if some of it can be labeled with radioactively-labeled ligands, and electrophoresed in parallel with a cocktail of unlabeled proteins from the organism. It has been possible to attach photolabile moieties to some sugar analogues to enable a structural-functional dissection of mammalian transporters as mentioned above. But this has been possible because of the wealth of information regarding the kinetics of those transporters and their presumed

similarity in structure. This could hardly the case with the trypanosome glucose transporter.

In the absence of a protein sequence, the only choice is to employ probes based on the sequences of genes within the family of glucose transporters. It was on this basis that the rat liver glucose transporter was used, although with hindsight a yeast transporter might have been a better candidate in view of the positions of yeasts and trypanosomes on the scale of evolution. A multiple sequence alignment of the GTs enabled the identification of structural motifs that are presumed conserved in the superfamily. These motifs and their translated nucleotide sequences are presented in **[2.3.2]**. In principle such motifs should exist in the trypanosome transporter as well, especially oligo 2, whose sequence is thought to constitute the glucose binding site of the superfamily of GTs. As such, probes were synthesised from those regions by reverse translation taking into account the codon usage of trypanosome genes and the degeneracy of the genetic code.

The cDNA library was screened with the rat liver glucose transporter cDNA, and with oligo 1. However, the positives that were picked were not hybridizing consistently with these probes or the other oligos. An alternative to screening this library was to make primers complementary to the vector lambda gt11 reverse and forward sequencing primers, and to use these to prime the amplification of the GT gene using pairwise combinations of GT-specific primers and the vector primers in a polymerase chain reaction (PCR). This technique was also used on trypanosome genomic DNA. The amplified DNA fragments that were obtained from this approach were hybridized with

the primers to ensure that they were not artifacts and then blunt-end cloned into pGEM-7 and sequenced. It was found that although the fragments contained the primer binding sites, the remainder of the sequences did not code for extensive open reading frames.

The genomic library was then chosen for screening in view of the possibility that the gene might be differentially processed and expressed, and therefore may not be represented in the cDNA library. This was the only way to ensure that the gene was present since such a library would contain all the genes of the trypanosome. In addition, the apparent lack of introns in trypanosome genes means that if the glucose transporter gene was isolated, there would be no problems relating to that factor.

Due to the relative redundancy of the probes, the genomic library was screened at room temperature and 5X SSC to 55°C and 2X SSC. As a result several clones were isolated and carried through three more cycles of hybridization with oligo probes 2 and 4. Of the 17 primary isolates, 6 were selected on the strength of their hybridization to the probes. After Southern blot analysis, only one of these carried a strongly hybridizing insert. This insert was therefore selected and subcloned directionally into pUC to produce pTGT-4 for sequencing.

The predicted amino acid sequence of pTGT-4 was compared with the GTs to identify diagnostic features and to locate oligonucleotide probe binding sites. The poor homology between the sequence of pTGT-4 and the GTs masked the identification of any sequences that could place it among that family or at least to account for the observed hybridization pattern.

As a second approach, a size-selected plasmid library was constructed from

genomic fragments of trypanosome DNA that hybridized strongly to the rat GT cDNA. These fragments were identified from a Southern blot, purified from a BamH1/EcoR1-digest that was electrophoresed on low melting point agarose, and ligated into BamH1/EcoR1-cut pBluescript. The library was then screened with oligo 2. The hybridizing colonies were further selected by probing again with the rat GT cDNA. As a result, a number of clones were isolated from this library and sequenced with oligo 2. Although the degeneracy of this oligo made sequencing imperfect, the legible portion of one of the clones, pUP-1, was identical to pTGT-4. (This was the only plasmid that sequenced at all with oligo 2). Although there is a possibility that pUP-1 could be a contaminant of the enriched plasmid library, this was unlikely to be the case because the two plasmids are of different sizes and they have different cloning ends: pUP-1 has EcoR1/BamH1 ends whilst pTGT-4 has EcoR1/HindIII ends. This is all the more credible because the Southern blot of the lambda clone, 3-LT10 from which pTGT-4 was derived, has an identical insert to pUP-1 when restricted with EcoR1/BamH1. In, addition there is no BamH1 site in pTGT-4. This was evidence that the probes were adequately specific and that the techniques of clone isolation were satisfactory. On this basis pTGT-4 was characterized further. A partial peptide sequence comparison between the two clones is as follows:

```
pTGT-4    NTNKESSFHNVSRSGKGGSPPRADGRWNGRTKGV
pUP-1          KESSFHNVSRSGKGGSPPRADGRWNGRTKGV
```

In the course of screening the plasmid library, a long open reading frame was identified which sequenced with oligo 1, but which did not contain the peptide sequence of that oligo, and had no structural resemblance to either of the clones above or with the GTs.

## 5.3 Gene structure and implications for developmental regulation and expression

The presence of the canonical regulatory sequences, viz CCAAT and TATA boxes in the trypanosome gene is consistent with an RNA pol11 gene. Classical promoter sequences in Kinetoplastid genes have not been easy to identify. This gene may therefore be one of the first of such findings in these genes in general.

The TATA box is probably the most prevalent cis-acting element in eukaryotic genes, and it is usually located about 25-32bp from the cap or initiation site. However, it has been found 109bp and 795bp from the transcription sites in the MAL6S and MAL6T genes respectively (223). According to the "TATA box-trap" model, the transcription of a gene is enhanced when the TATA box is closely flanked by GC-rich sequences as observed in the trypanosome gene (224).

The CCAAT box is normally located about 80bp from the initiation site, but it has also been found to occur up to 225bp from that site in the mouse immunoglobulin heavy chain gene promoter (225). Although the CCAAT box is not as ubiquitous as the TATA box, it is thought to be linked to structural genes whose expression is stringently controlled, as for inducible genes, and whose products may be required in large quantities at specific stages of the cell's development. This is especially so for the human a- and b-globin and chicken ovalbumin genes (226,227). The promoters of these genes are transcriptionally active at their appropriate developmental stages, and all have the CCAAT and TATA boxes which may act in concert in the regulation of these genes. In fact, deletion analysis of the b-globin gene showed that its transcription was dependent on the TATA box, decreasing about 20-fold, and that the CCAAT

box played the strongest role in determining the efficiency of the promoter (228). Genes that are efficiently transcribed, such as the enzymes of the glycolytic pathway, are known to contain the CCAAT box (229). Overall, the modular nature of the trypanosome gene promoter as seen in the two boxes, identifies it with class II genes i.e genes that are transcribed by RNA polII. The putative trypanosome gene is therefore a likely class two gene.

The other significant feature in the arrangement of this gene is the unusually strong stop signal, TAGTAG. Taken together with the CCAAT and TATA boxes, this may be indicative of the stringent molecular controls that may regulate its expression.

Structural microheterogeneity in the DNA sequence of the trypanosome gene features a run of $(CA)_{19}$, which has a propensity for Z-DNA formation (230). This middle-repetitive element is ubiquitous in nature and has been found to punctuate the genomes of eukaryotes through man to yeast (231). It has been found associated with introns and has also been mapped to the 5' flanking regions of actively transcribed genes (232). Such sequences have therefore been associated with the regulation of those genes of which they constitute an integral part, including the globins (233), prolactin (234) etc. They have also been implicated as enhancers of replication, transcription and translation (235), and in viral recombination and integration (236). Due to the repetitive nature of these elements, they are potentially structurally labile. (Hence there is a dynamic equilibrium between Z-DNA and B-DNA as they interconvert between the two forms of DNA). These elements (Z-DNA) have been associated with gene conversions and rearrangements as observed in the developmental regulation of the globin variants (237). They have therefore been regarded as the "hot-spots" for gene conversions or recombinations. The identity of such sequences to the consensus DNA topoisomerase II

recognition site suggests that cleavage by this enzyme may precede gene rearrangement for regulation (238). This consensus is derived as follows:

5′ R N Y N C N G Y ^ N G K T N Y N Y 3′

where R is any purine, Y is pyrimidine, K is G or T and N is any base; cleavage site is denoted by a caret (^).

The presence and significance of this element within the gene under study, is not clear at the moment although a plausible conjecture may be made regarding its regulation. If Z-DNA stretches of this kind have all the functions detailed above, then this element may represent a "cross-over" point between the bloodstream and insect/culture forms of the trypanosome regarding the expression of this gene. By appropriately shifting to different expression loci, or rearranging itself, expression may be regulated presumably at the level of transcription. One way to ascertain this is to perform a deletion mapping coupled with reporter enzyme transfections, in order to determine the direction or levels of expression of the gene.

a) Tissue specific genes

| | |
|---|---|
| Immune system:– | Immunoglobulin,complement C4 |
| Red blood cells:– | Alpha, beta and epsilon globin |
| Liver:– | Alpha foetoprotein, serum albumin |
| Nervous system:– | Acetylcholine receptor |
| Pancreas:– | Preproinsulin, elastase |
| Connective tissue:– | Collagen |
| Pituitary gland:– | Prolactin |
| Salivary gland:– | *Drosophila* glue proteins |
| Silk gland:– | Silk moth fibroin |

b )Inducible genes

| | |
|---|---|
| Steroid hormones:– | Ovalbumin, vitellogenin, tyrosine amino-transferase |
| Stress:– | Heat shock proteins |
| Viral infection:– | Beta interferon |
| Amino acid starvation:– | Yeast HIS 3 gene |
| Carbon source:– | Yeast GAL genes, yeast ADH II gene |

Fig. 53 Examples of genes with DNase 1 hypersensitive sites, and with a propensity to have Z-DNA. Adapted from (239).

There is a direct correlation between the relative abundance of a protein and its pattern of codon usage. Highly expressed genes have a biased codon usage i.e they contain mostly the "preferred codons". In contrast, poorly expressed genes do not show this bias, and they tend to use many "rare" codons. Analysis of the codon preferences of the gene under study, shows that it has no particularly strong preferences. This is illustrated below.

Fig. 54 Comparison of Tglutrp codon usage to the codon bias in other trypanosome genes as defined in (40).

| Amino acid | Codons | Genes 1 | 2 | 3 | 4 | 5 | Amino acid | Codons | Genes 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Phe | TTT | 12 | 2 | 4 | 3 | 7 | Pro | CCT | 10 | 6 | 0 | 2 | 3 |
| | TTC | 8 | 6 | 5 | 8 | 3 | | CCC | 12 | 5 | 1 | 1 | 2 |
| Leu | TTA | 3 | 1 | 0 | 3 | 3 | | CCA | 11 | 3 | 1 | 5 | 1 |
| | TTG | 10 | 0 | 1 | 6 | 11 | | CCG | 11 | 2 | 0 | 5 | 1 |
| | CTT | 8 | 4 | 1 | 7 | 8 | Thr | ACT | 7 | 5 | 2 | 6 | 8 |
| | CTC | 6 | 19 | 5 | 6 | 6 | | ACC | 8 | 10 | 2 | 10 | 6 |
| | CTA | 2 | 1 | 2 | 16 | 1 | | ACA | 13 | 3 | 1 | 24 | 4 |
| | CTG | 4 | 9 | 1 | 10 | 4 | | ACG | 12 | 2 | 4 | 12 | 4 |
| Ile | ATT | 7 | 4 | 3 | 2 | 6 | Ala | GCT | 12 | 13 | 3 | 5 | 9 |
| | ATC | 3 | 8 | 6 | 6 | 2 | | GCC | 13 | 12 | 1 | 16 | 5 |
| | ATA | 13 | 1 | 0 | 9 | 8 | | GCA | 10 | 4 | 0 | 27 | 9 |
| Met | ATG | 13 | 11 | 10 | 9 | 6 | | GCG | 9 | 6 | 4 | 16 | 2 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Val | GTT | 11 | 11 | 2 | 5 | 7 | Tyr | TAT | 5 | 2 | 0 | 3 | 6 |
| | GTC | 5 | 6 | 1 | 7 | 5 | | TAC | 4 | 15 | 1 | 8 | 7 |
| | GTA | 9 | 1 | 0 | 5 | 11 | His | CAT | 6 | 2 | 0 | 2 | 3 |
| | GTG | 10 | 11 | 3 | 5 | 7 | | CAC | 11 | 6 | 1 | 8 | 4 |
| Ser | TCT | 9 | 5 | 1 | 7 | 6 | Gln | CAA | 5 | 4 | 2 | 14 | 5 |
| | TCC | 10 | 6 | 4 | 2 | 1 | | CAG | 13 | 10 | 5 | 9 | 1 |
| | TCA | 7 | 3 | 1 | 7 | 3 | Asn | AAT | 16 | 0 | 1 | 6 | 10 |
| | TCG | 8 | 4 | 1 | 1 | 2 | | AAC | 12 | 12 | 5 | 12 | 6 |
| | AGT | 12 | 1 | 0 | 2 | 5 | Lys | AAA | 24 | 7 | 4 | 47 | 14 |
| | AGC | 8 | 3 | 2 | 14 | 4 | | AAG | 11 | 17 | 5 | 13 | 7 |
| Asp | GAT | 12 | 5 | 11 | 5 | 12 | Gly | GGT | 9 | 16 | 4 | 2 | 7 |
| | GAC | 7 | 5 | 6 | 19 | 3 | | GGC | 10 | 15 | 5 | 17 | 1 |
| Glu | GAA | 17 | 11 | 9 | 26 | 18 | | GGA | 17 | 2 | 2 | 18 | 11 |
| | GAG | 14 | 17 | 11 | 15 | 13 | | GGG | 11 | 0 | 0 | 6 | 4 |
| Cys | TGT | 6 | 0 | 0 | 0 | 5 | Stop | TAA | 0 | 1 | 0 | 1 | 1 |
| | TGC | 6 | 9 | 0 | 14 | 4 | | TAG | 1 | 0 | 1 | 0 | 0 |
| Trp | TGG | 8 | 5 | 0 | 6 | 5 | | TGA | 0 | 0 | 0 | 0 | 0 |
| Arg | CGT | 9 | 10 | 2 | 0 | 4 | | | | | | | |
| | CGC | 11 | 12 | 2 | 0 | 0 | | | | | | | |
| | CGA | 4 | 0 | 0 | 3 | 0 | | | | | | | |
| | CGG | 6 | 2 | 1 | 2 | 1 | | | | | | | |
| | AGA | 9 | 0 | 0 | 1 | 3 | | | | | | | |
| | AGG | 8 | 0 | 0 | 1 | 5 | | | | | | | |

Genes:

1. Tglutrp
2. Glycolytic enzyme, aldolase
3. Structural protein, calmodulin
4. Variant surface glycoprotein (VSG MiTat 1.4)
5. Expression site-associated gene (221a ESAG 1 prot.).

From the foregoing, it is clear that there is no strong bias in codon usage in Tglutrp. It uses all the 61 possible codons unlike, for example, calmodulin which is highly expressed both in terms of mRNA and protein, is highly selective in its choice of codons and employs only 45 of the 61 codons (40). Its pattern of codon usage is very similar to those of the differentially expressed genes- VSG genes and ESAGs, although even these are slightly more biased than Tglutrp. It may therefore be assumed that the cloned gene is lowly or poorly expressed in comparison to the housekeeping gene for calmodulin for instance, and therefore will be poorly represented in cellular mRNA species. This confirms the finding and the allusions made on the possible developmental regulation of this gene, on the basis of the fact that a gene that may be rearranged, repressed and derepressed may not have any stringent requirements in its codon selection. The speed of assembly of the codons/amino acids may take priority over the choice of codons under the circumstances.

An analysis of the nucleotide preference for the third or wobble position of the codons showed that a pyrimidine (C or T) is the candidate for this position. This is consistent with that previously observed (40). What this implies is that there is a higher frequency of

purine dinucleotides as suggested in Nussinov's hierarchy (50,51), which is indicative of strand asymmetry i.e there is a purine excess of about 7.33%. This was calculated using the total base composition in the coding strand, and based on the equation:

$$Strand\ asymmetry = \frac{A + G - T - C \times 100\%}{A + G + T + C}$$

Expressed another way, the ratio of the total purine dinucleotides to total pyrimidine dinucleotides is 1.4 i.e 168/118. This is in consonance with most eukaryotic genes, even for those with unusual nucleotide composition such as the genes of P. falciparum which are predominantly A/T rich (240). The significance of this observation of strand asymmetry is that the gene sequence under study is a coding sequence, adding to the evidence of differential expression observed above.

A number of conclusions can therefore be made on the structure of this gene. The poor bias in codon choice implies that it is lowly expressed, i.e it is not constitutive. Its expression may be inextricably linked to precursor controls, namely mRNA abundance. The turnover rate of the mRNA may be very high. Such a molecular control strategy is peculiar to genes with strong promoters. Therefore, the TATA and CCAAT boxes, and the (CA)-enhancer element in the trypanosome gene may in concert, constitute the pivot of a highly mobile gene or one that may be switched on and off as and when its product is required. Within the context of trypanosome survival in the bloodstream of the host, it is envisaged here a gene that may be subject to developmental regulation. This may be supported by the finding that the gene is differentially expressed in the bloodstream forms only. Although this may be tenuous, it is fair

to suggest that this gene is vital to its survival under the exigencies of the vector and host environments otherwise it would not control its expression. It may well be premature to attempt to justify the working assumption that the digenetic life-cycle of the trypanosome requires a mobile developmentally regulated GT gene that enables it to shuttle between amino acid and glucose dependence. There is kinetic evidence for this possibility (241). If Tglutrp is the de facto gene for glucose transport in the trypanosome, then the present molecular information (regarding differential expression) lends credence to the kinetic observations, and this would also fall into line with the developmental regulation of the enzymes of glycolysis (25).

## 5.4 Structural profile: a functional and evolutionary comparison with "classical" glucose transporters.

### 5.4.1 Sequence comparison and secondary structure prediction

The secondary structure of a protein may be predicted from a combination of algorithms or by relating conserved residues or sequences in a multiple alignment approach. The latter is currently the most widely used technique as a result of a burgeoning of sequence information from molecular cloning. Since structural homology and hence functional identity may be better conserved than sequence homology per se, it is assumed that the conformations of conserved residues or groups of residues which may subserve a common function, will remain conserved within a family inspite of some form of mutational pressure.

An averaged combinatorial prediction approach involving the use of the Chou-Fasman or CHF (159), Garnier-Osguthorpe-Robson or GOR (242) and the Kyte-Doolittle (157) methods have been employed to ascribe hydrophobicity, turns or loops, and flexibility to some regions of homologous sequences. The larger the number of these sequences, the greater the level or accuracy of prediction. However, none of these algorithms or combinations thereof, provide tailored solutions to structural predictions. Few predictions have stood the test of an in vivo structural dissection, and there are a myriad conflicting views on this point.

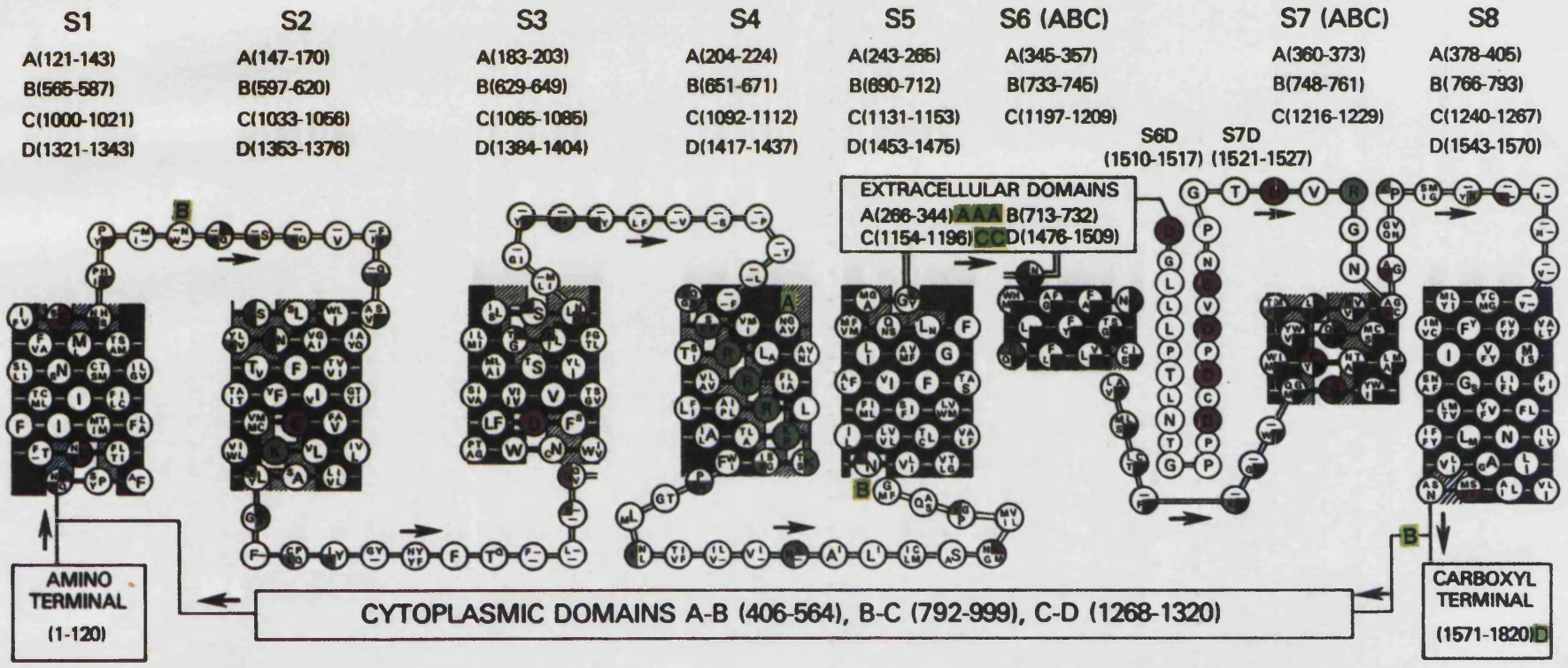
A priori predictions have also been used based largely on the knowledge of amino acid propensities.

From this study it is difficult to ascribe any secondary structural designation or feature to the trypanosome protein with any level of certainty. As already mentioned above, what may be said for now is that a large part of the structure is attributed to equal contributions of

a-helices and b-sheets. The turn component might also be in a significant proportion.

If Tglutrp is the authentic glucose transporter in the trypanosome, it has a very poor homology to the rest of the family. This is more so for its level of hydrophobicity, which should be the over-riding criterion for its inclusion among that group. What may be said about its low hydrophobicity is that it is probably an amphipathic protein, with a-helices that are shorter than usual. These would just about touch the lipid bilayer. In fact, recent evidence from experiments with model peptides and phospholipids, suggests that an effective lipid-interacting amphipathic a-helix need only be about 8-10 residues in length (243). This has been the model suggested by Lodish [1988, (244)] and for the Na+ channel [see fig. 47 (218)], since the available algorithms cannot predict or identify short helices.

Fig. 47 Model of the sodium channel protein of the eel, Electrophorus electricus. The two short helices are indicated. S1-S8 are the building subunits. Charged residues (basic and acidic) are in colour. Amino acid designations apply as before. Reproduced from (218).


S1
A(121-143)
B(565-587)
C(1000-1021)
D(1321-1343)
S2
A(147-170)
B(597-620)
C(1033-1056)
D(1353-1376)
S3
A(183-203)
B(629-649)
C(1065-1085)
D(1384-1404)
S4
A(204-224)
B(651-671)
C(1092-1112)
D(1417-1437)
S5
A(243-265)
B(690-712)
C(1131-1153)
D(1453-1475)
S6 (ABC)
A(345-357)
B(733-745)
C(1197-1209)
S6D
(1510-1517)
S7D
(1521-1527)
S7 (ABC)
A(360-373)
B(748-761)
C(1216-1229)
S8
A(378-405)
B(766-793)
C(1240-1267)
D(1543-1570)
EXTRACELLULAR DOMAINS
A(266-344) AAA B(713-732)
C(1154-1196) CC D(1476-1509)
AMINO TERMINAL
(1-120)
CYTOPLASMIC DOMAINS A-B (406-564), B-C (792-999), C-D (1268-1320)
CARBOXYL TERMINAL
(1571-1820) D

All the (possible) conserved motifs within the superfamily were used as markers to search the trypanosome sequence. There is a diffused motif spanning TM4 and TM5 represented by:

R---G---G------P-Y--E------RG

Within the trypanosome sequence, a very similar run exists in an analogous position of the sequence except for a two-residue deletion between Y and E. A comparison is made between the consensus sequence, Tglutrp, and lac permease of K. lactis as follows:

```
Consensus    R---G---G------P-Y--E------RG
Tglutrp      R--------------P-Y  E------WG
Klact        R--------------P-Y--E------RG
```

Both Tglutrp and Klact lack glycines in the consensus positions.

The sequence has no PESPR motif, but neither does the putative Leishmania or citrate transporters (56,245). The lac permease of E. coli is probably the most salient example of a hexose transporter with none of the modules typical of members of that family (144).

The location of the motif QPNSGI(S)NGNKT within Tglutrp would be identical to that in the GTs except for a serine insertion. Insertions and deletions of 2-10 amino acids are regular features in trypanosome proteins as has been found in the glycosomal isoenzymes of glycolysis (27).

```
Tglutrp                  QPNSGISNGNKTN
GLUT2 (probe)            QQFSGI-NGIFYY
```

As already mentioned above, this region is the only diagnostic feature among the GTs. That this may be the case is borne out by the fact that although the phosphate transporter has all the other

conserved motifs seen within the GT superfamily, it partly lacks this module, the putative glucose binding site (246). It may be surmised that these regions compared above are in fact homologous. The substitution of the Y residues in a majority of GTs for T and N in the trypanosome sequence is not significant because all these residues are functionally identical i.e they are all polar residues. From a structural point of view, T and Y are both ambivalent i.e neither of them has any special preference for the interior or exterior of the membrane.

In comparison to the conserved motif VERAGRRT beyond TM8 and VPETK beyond TM12 in the mammalian GTs, the closest identities are as follows:

| | | |
|---|---|---|
| Tglutrp | AD--GRWN | HPETT |
| Klact | IDKIGRRE | FVETK |
| GLUT2 | VEKAGRRT | VPETK |

Neither of these motifs is present in the putative Leishmania protein. The only criteria for its inclusion in this family are its relatively higher hydropathy index, and the presence of a motif (LQLTGINAVMNY) that is identical to the putative glucose binding site.

The conserved tryptophan residue (W) in TM11 of the superfamily has been implicated in cytochalasin B binding (247). If Tglutrp is the trypanosome GT, the absence of this residue in that position of the sequence may account for the failure of cytochalasin B to inhibit glucose transport in the trypanosomes (Bayele, unpublished observations).

Generally, trypanosome proteins have a relatively low homology compared to their counterparts in other organisms, with identities between 38-57% (40). This is a significant pointer to an evolutionarily distant relationship between them and other eukaryotes. The several peculiarities alluded to above lend credence to this view, and to the conclusions based on rRNA

analysis (8), cytochrome c (248) and mitochondrial genes (249), that the trypanosomes diverged very early from other eukaryotes. As such, structural comparisons between trypanosome proteins or genes and those of other organisms should be done within the context of that tenuous relationship.

Fig. 56

Proximate homologies between trypanosome gene sequence and conserved modules of the Glucose Tranporter Superfamily; A priori comparisons.

| Transporter | Subdomain 1 | | | Subdomain 2 | | |
|---|---|---|---|---|---|---|
| Humglut1 | R F G R R | (89) | | P M Y V G E V S P T A F R G | | (142) |
| Ratglut2 | K L G R I | (92) | | P M Y I G E I A P T T L R G | | (144) |
| Ecoaraea | R L G R K | (84) | | P L Y L S E M A S E N V R G | | (134) |
| Ecoxyle | R F G R R | (81) | | P M Y I A E L A P A H I R G | | (151) |
| Yeastgal2 | M Y G R K | (165) | | P L Y Q A E A T H K L S R G | | (215) |
| Chlorella | N M G R K | (110) | | P Q Y L S E V A P F S H R G | | (155) |
| T. brucei | R M A R R | (168) | | P Y Y - - E V G P A K A W G | | (245) |
| Leishman | K I G A R | (179) | | P V Y T D Q N A H P K W K R | | (211) |
| Ecocitr | R I G R R | (78) | | S V Y L S E I A T P G M K G | | (130) |

| | 3 | | | 4 | | |
|---|---|---|---|---|---|---|
| Humglut1 | P E S P R F L | (210) | | Q Q L S G I | N A V F Y Y | (286) |
| Ratglut2 | P E S P R Y L | (213) | | Q Q F S G I | N G I F Y Y | (289) |
| Ecoaraea | P N S P R W L | (199) | | Q Q F T G M | N I I M Y Y | (260) |
| Ecoxyle | P E S P R W L | (224) | | Q Q F V G I | N V V L Y Y | (302) |
| Yeastgal2 | P E S P R Y - | (284) | | Q Q L T G N | N Y F F Y Y | (367) |
| Chlorella | P E S P N F L | (221) | | Q Q F T G I | N A I I F Y | (281) |
| T. brucei | - - E T L F P | (321) | | Q P N S G I S | N G N K T N | (399) |
| Leishman | - - S T L F S | (275) | | L Q L T G I | N A V M N Y | (331) |
| Ecocitr | Q E T E A F S | (202) | | T T - T T F | Y F I T V Y | (240) |

| | 5 | | 6 | |
|---|---|---|---|---|
| Humglut1 | V E R A G R R | (332) | V P E T K G R T F D E I | (459) |
| Ratglut2 | V E K A G R R | (335) | V P E T K G K S F D E I | (461) |
| Ecoaraea | V D K A G R K | (310) | I P E T K N V T L E H I | (438) |
| Ecoxyle | V D K F G R K | (349) | V P E T K G K T L E E L | (476) |
| Yeastgal2 | V E N L G R R | (414) | V P E T K G L S L E E I | (554) |
| Chlorella | S D K F G R R | (330) | L P E T K G V P I E R V | (461) |
| T. brucei | A D - - G R W | (438) | H P E T T I S G P E T S | (548) |
| Leishman | V T F M - - R | (377) | I Q V F F L H P W D E E | (508) |
| Ecocitr | S D R I G R R | (390) | R L S S G Y Q T W E N K | (418) |

Comparison of residue numbers and calculated molecular weights

| | Number of residues | Calc. M.wt |
|---|---|---|
| Humglut1 | 492 | 54,117 |
| Ratglut2 | 498 | 55,333 |
| Ecoaraea | 464 | 51,556 |
| Ecoxyle | 491 | 53,607 |
| Yeastgal2 | 574 | 63,789 |
| Chlorella | 519 | 57,667 |
| T. brucei | 577 | 63,346 |
| Leishman | 567 | 61,449 |
| Ecocitr | 431 | 46,979 |

### 5.4.2 Structural motifs: a functional appraisal

A number of motifs may be found in the primary structure of the trypanosome sequence. In comparison with the glucose transporters, the N-linked glycosylation site at Asp45 is not found, but neither is it present in most other GTs of the lower organisms. From the consensus N-glycosylation sequence, N-X-S/T (where N is asparagine, X is any amino acid, S is serine and T is threonine), six potential N-glycosylation sites as indicated in Fig. 39 above, can be identified. It is unlikely however, that all these sites are glycosylated in vivo. Site 3 has the same glycosylation sequence to that in the putative Leishmania transport protein. Sites 4 and 5 (in tandem) are identical both in position and sequence to that predicted for SNF3 at Asp383 of yeast (168) and almost the same position of the presumed helix. If the GTs are duplications of a primordial 6 membrane-spanning transporter as has been suggested (145), this position would correspond to the identical position in the second half of the duplicate. It may be surmised that the relatively larger number of potential glycosylation sites compared to the GTs, may be associated with the heavy glycoprotein environment of the protein involving VSGs. Upon this finding it may be said that the trypanosome protein is a (membrane) glycoprotein.

Other features of functional significance include cyclic AMP-dependent protein kinase C sites. These have the consensi:

R-R-S/T

R-R-X-S/T

R-R-X-X-S/T,

where any R (arginine) may be replaced by a lysine (K); S and T are as defined above and X is any residue. These may give an indication of the regulatory cascades involved in its operation. One of these sites, site 4, is found in an analogous position in GAL2 and the maltose

transporter. Indeed the lower forms have their phosphorylation sites within their large cytoplasmic loops corresponding to that position in Tglutrp. By their structure, all the GTs have several potential phosphorylation sites including the R-R-T motif beyond TM8 in the mammalian GTs. This position is absent, except the R-R repeat, in the lower GTs. However, the yeast high affinity glucose and galactose transporters are known to be kinase-dependent transport systems (250). A number of membrane proteins have been reported to act as substrates for protein kinase C including the glucose transporter (193). Hence, the regulation of glucose transport may involve the coordination of signal transduction mediated by the phosphorylation of the transporter by protein kinase C. The location of these potential phosphorylation sites in the various GTs sets a dichotomy with respect to their mechanisms of regulation. The presence and relevance of such sites in Tglutrp are open to question.

A general feature in all the GTs examined is the location of acidic [glutamic acid (E) and aspartic acid (D)], proline (P), serine (S), and threonine (T) residues at their C-termini. This region, the so-called "PEST" cluster, is invariably flanked by the basic residues arginine, lysine or histidine. These sequences have been thought to have strong implications for protein turnover. It has been suggested that proteins with local concentrations of acidic residues have very short half-lives and are therefore rapidly turned-over (251). This finding has been confirmed with ornithine decarboxylase in mammals, and constitutes the basis for the selective toxicity of difluoromethylornithine [DFMO] in trypanosomiasis treatment (252).

The occurrence of these residues at an identical position in Tglutrp puts it in this category of PEST-cluster-containing proteins.

Fig. 57 PEST clusters in proteins with high turn-over rates

| Protein | Sequence | Half-life/hrs | Ref |
|---|---|---|---|
| ODC | HGFPPEVEEQDDGTLPMSCAQESGMDR | 0.5 | (252) |
| HMG-CoA reductase | KLSSVEEEPGVSQDR | 1.5-3 | (253) |
| c-myc | HEETPPTTSSDSEEEQEDEEEIDVVSVEK | 0.5 | (254) |
| B-casein | RELEELNVPGEIVESLSSSEESITR | 2-5 | (255) |
| P730 | HLQYMENMNSIASLVMAVVVNENEEDDEAESEQPAQQQK | 1.0 | (256) |

Abbreviations are as follows:ODC, ornithine decarboxylase; HMG-CoA, hydroxymethylglutaryl-CoA; P730, phytochrome 730.

The "PEST" cluster in the trypanosome sequence is as follows:

```
          +          +
          HPETTISGPETSK
```

By comparison to the proteins above, the trypanosome protein would have a PEST score that would put it among those proteins that have half-lives of 0.5-1.0hr, on the basis of the relative mole percentages of P, E, S, and T residues. Such a high turnover rate would imply a rapidly turned-over message/mRNA which would be consistent with an inducible tightly controlled structural gene.

Another structural element observed is an almost perfect repeat of six R residues after every 3 and then 4 residues in Tglutrp. This region corresponds to TM9 in the GTs, and may hold some intrinsic function particularly in channel building. The sequence runs as follows:

```
R---R---R----R----R----R
<          24 residues   >
```

where (-) is any residue. Within the same run are 3 glycines separated from each other by three residues (G---G---G). The entire stretch of 24 residues is enough to cross the membrane. This repeat unit may act as voltage sensor or a charge relay grid. Similar units have been found in the S4 region of K+-channels, the Na+-channel of the eel (218,257,258) and the acetylcholine receptor which has 4-6 R (188). In addition, dipole shift measurements in the squid Na+ channel indicate that there are about six charges that cross the membrane. Where they occur, they have been suggested to form amphipathic helices with the Rs lining the walls of the ion channel, and are responsible for sensing membrane electrical potential [(218) see Fig. 45]. Changes in membrane potential accordingly induces a translocation of the S4 helix with a resulting gating of the channel in a helical screw manner (257). This arrangement of positive charges is proposed to be stabilized by ionic interactions/salt bridging with proximal (or distal, depending on the manner of folding) negatively charged residues. Such a view has been noted for the acetylcholine receptor (259,260). It is too early to extrapolate this result to the trypanosome sequence, but the identity of the repeat units makes this possible. In the context of glucose transport, it may be considered that if Tglutrp is the authentic GT in the trypanosome, then it is probably coupled to a cation gradient i.e it is a cotransporter.

The implication is that a gating mechanism in that region affords a conformational change that may facilitate ligand binding. There is no structural equivalent in the GTs that may perform this function though it is fair to say that the corresponding transmembrane region is only weakly hydrophobic.

From the point of view of the spatial organization of these residues to form a channel in the trypanosome, this region may be assumed to span the membrane with the R residues located at periodic intervals. If an a-helix is thus assigned, the R residues would be distributed along a spiral peptide backbone making one complete turn. They may alternatively lie on one side of the helix so that this region would be strongly amphipathic. If they are assumed to form a b-sheet structure, they would alternate with each other on both sides of the peptide backbone. Negative charges from adjacent or distal segments may be proposed to form ion pairs with the R residues of the helix. The positioning of glycine residues at periodic intervals among the Rs also signifies a tunnel/channel formation. The glycines would confer turns or bumps as they have the propensity to create turns and are therefore found in hinge regions of proteins. Proline residues (not shown) at both ends of the helix would break the helix at the endoface and exoface, inducing sharp turns that connect to the neighbouring segments. Such a model would be water-filled in order to traffic the hydrated cation the gradient of which enables ligand binding consequent upon an induced conformational change.

This arrangement has been depicted in the models below. Both models (the helical wheel and the helical net) confirm the above prediction. One turn of an a-helix requires approximately 3.6 residues. Therefore, in any peptide that constitutes an

amphipathic a-helix, there will be a periodic variation in the hydrophobicity values with a periodicity of 3.6 residues per cycle. The helical wheel particularly holds this point well, as seen in the clustering of almost all the R residues on one side of the helix, and the glycines interspersed with them. This half of the helix is absolutely polar. The other half has one potential salt-bridge which would nullify the charge contributions of R and D. The arrangement of the Rs and the Gs in the tube model also supports the prediction.

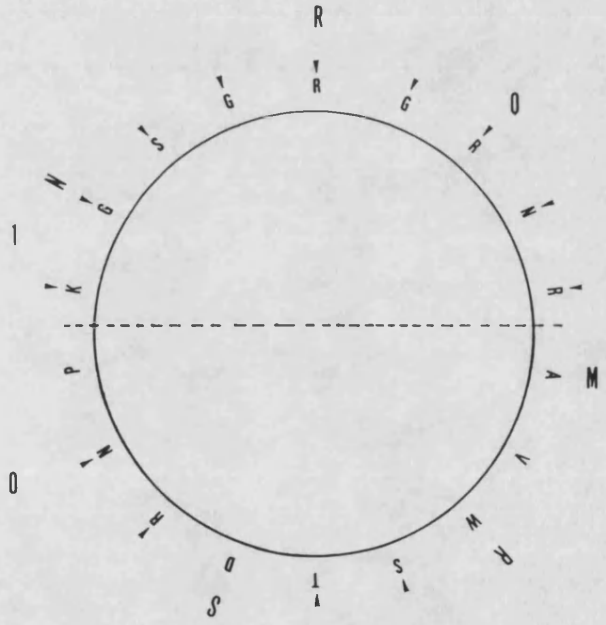



Fig. 58 Helical wheel of residues with a channel-forming/amphipathic helix potential. Residues are arranged 100° or 3.6 residues (360°/100°) apart. Hemisphere [1] is entirely/strongly polar.

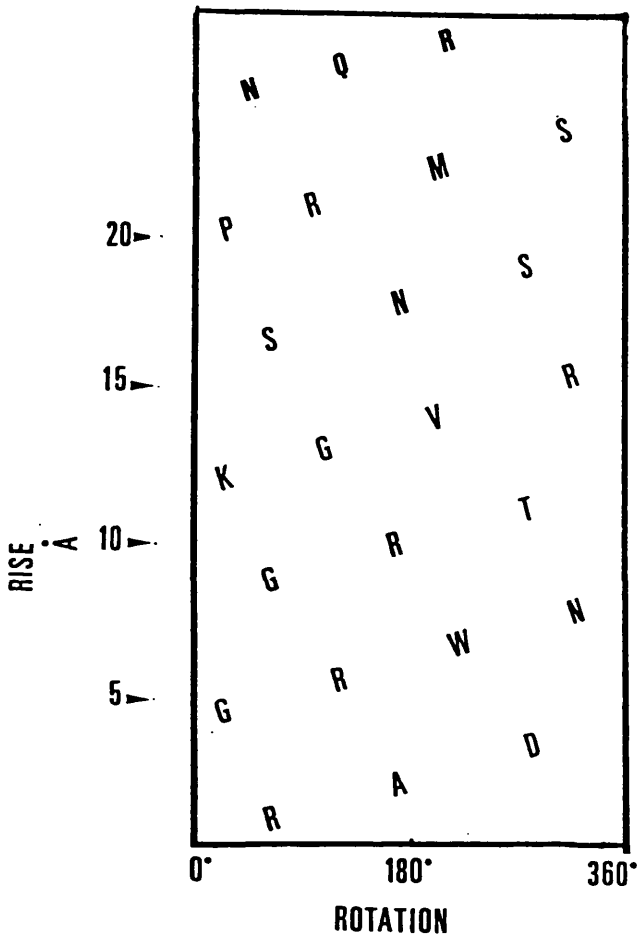



Fig. 59 Helical net presentation of Fig.48. Rotation along the axis of the helix is indicated in degrees. Distance between residues is in angstroms

Na+ ion binding to the intestinal glucose transporter has been observed to produce a conformational change at the glucose binding site which results in an increase in affinity for glucose binding (261). By using fluorescence energy transfer, it was found that the Na+ and glucose binding sites in the intestinal brush border symporter were separated by about 30-35 A (262). This corresponds to about 20-25 amino acids or one transmembrane segment; this is also a perfect fit to the channel length proposed above. Combined with earlier indications that ASA-BMPA binds to TM10, it may be said that these observations satisfy the arrangement for a Na+/glucose transporter. If this is the case then the finding that the trypanosome transporter is coupled to a Na+ gradient is credible (153), depending on whether Tglutrp is **the** trypanosome glucose transporter.

**Are there any topogenic or targeting signals that might give a clue to the subcellular location of the cloned trypanosome protein?**

Until recently, there have been speculations as to the nature of the targeting signals involved in the post-translational exporting and sorting of microbody (peroxisomal, glyoxysomal and glycosomal) proteins (27). Since these proteins lack a cleavable signal peptide, their topogenic signals may be internal. In some peroxisomal proteins, targeting signal sequences have been identified at the C-terminus with the aid of site-directed mutagenesis. These sequences comprise the consensus tripeptide S-K-L

to within 30 residues of the carboxy terminus. Targeting was abolished with the replacement of K with N, probably suggesting a key role for this residue (263,264).

A comparison of the C-terminus of Tglutrp with those of peroxisomal proteins identified a very similar and possibly homologous sequence located 19 residues from the carboxy terminus. Incidentally, this region of Tglutrp also constitutes the putative hot-spot for its turnover. The implications of this coincidence could be very strong indeed. An alignment of these sequences is presented as follows:

| Protein | C-terminal sequence | Ref |
|---|---|---|
| Acyl coA oxidase | K H L K P L Q S K L | 265 |
| Bifunctional enzyme | S L A G P H G S K L | 266 |
| Malate synthase | I H H P R E L S K L | 267 |
| Firefly luciferase | K A K K G G K S K L | 268 |
| Tglutrp | T I S G P E T S K Q | |

Although the identity between these sequences and Tglutrp is significant, it is not certain whether there is any functional relationship between them in terms of targeting. What is certain is that those sequences have been confirmed for the referenced proteins. The functional and morphogenetic relationship between the glycosomes of trypanosomatids on the one hand, and peroxisomes and glyoxysomes on the other, may justify a qualified speculation. If a topogenic function can be ascribed to that region of Tglutrp, then the latter is probably a glycosomal membrane protein. The charge distribution in Tglutrp, and the presence of insertions and deletions, are consistent with those reported for glycosomal proteins (27). IF it is a glucose transporter, then it probably is the glycosomal isoform.

## 5.5 CONCLUSIONS AND SUGGESTIONS FOR FURTHER STUDY

There are a number of specific differences between the trypanosome and mammalian GTs.

a. The trypanosome GT is not inhibitable by phloretin which is routinely employed in quenching glucose transport in mammalian systems; the glycone analogue, phloridzin, is a more potent inhibitor of transport in the trypanosome. The inhibition kinetics of phloridzin in that transporter is identical to those of the Na+/glucose cotransporter. Is the trypanosome GT a Na+/glucose cotransporter?

b. Cytochalasin B has no inhibitor effect even though this is the classical tag for studying the transporter in mammalian systems.

c. It has a 20-50 fold higher affinity for glucose compared to mammalian GTs.

d. Photolabeling of the transporter with ATB-BMPA is not displaceable by phloridzin. This implies that the two ligands probably have different binding sites on the transporter (Fry A, unpublished observations).

e. The trypanosome GT is capable of transporting fructose (17) whereas the mammalian erythrocyte GT is not. Besides, the furanose form of this sugar is the preferred configuration for transport by the former.

These, largely kinetic, differences spell very basic differences in the structures of their cognate glucose transporters. The one motif that is unlikely to differ is the glucose binding site in both forms of transporter. It may therefore be envisioned here that the trypanosome transporter may contain residues proximal or distal to this site, that directly or indirectly modulate the kinetics of transport, and to give the observations

mentioned here.

It may be said that the added specificity requirements of hexokinase, with inference from Berlin's work (138), may place further constraints on the specificity of the glucose transporter in the trypanosome. This is especially so since hexokinase is ensconced within the glycosome. This may account for a number of positional preferences of the transporter for the various carbons or their substituents. It is therefore possible that a form of synergistic specificity is imposed stringently on every carbon by both the transporter and hexokinase to result in the differences in substrate specificity observed between trypanosomes and mammalian systems. Having said that, the question is where is the glucose transporter in the trypanosome with respect to hexokinase? Is it glycosomal or is it on the plasma membrane?

In this study, a gene has been cloned that bears only segmental homology to the glucose transporters. Some key conserved domains or residues are identifiable with the cloned sequence. However, its hydrophobicity profile is inconsistent with the GTs or with the probe that was used to clone it. This is the one factor that disqualifies it. If it were sufficiently hydrophobic, it would pass as a GT. However, if the above differences in function or kinetics may be extrapolated to a structural elucidation, it is probably only fair to have an open mind in that respect. Therefore, the interpretation of any molecular characteristics of the two forms of transporter should be done within the context of those observations, i.e the trypanosome GT may be significantly (structurally) different from its homologues in other organisms.

However, if it is not a glucose transporter, it is difficult to make any

functional inferences especially as it shows specific hybridization (from two gene libraries) to a well characterized glucose transporter, the rat GT cDNA.

Obviously, the only way to ascribe a function to this gene is to express it and measure glucose transport. It would then be possible to raise antibodies to specific segments of the protein to assess the transmembrane topology and to provide proof of or deny the predictions based on the algorithms used to define its hydrophobicity in this study. It may also be possible to create chimaeras with mammalian GTs to characterize transport kinetics with the aim of locating regions of "kinetic sensitivity", especially if the trypanosome GT may be used as a chemotherapeutic target. The contributions of various residues to transport may also be studied by site-directed mutagenesis etc.

**ADDENDUM**

At the close of this thesis, two isoforms of a putative hexose transporter cDNA (THT) were reportedly cloned using a VSG-specific probe (269). The two are highly homologous to the putative Leishmania transporter (LTP). Like the LTP, neither isoform has any of the conserved motifs (except the putative glucose binding site) discussed in the text, and which have earlier been used as probes to clone the trypanosome GT. In terms of motif representation, Tglutrp has a higher score than LTP or THT. If the cloned genes are the authentic GTs, the absence of these motifs explains the failure to isolate those genes from the gene libraries, using those motifs as probes.

A number of pertinent questions remain unanswered:

a. Is LTP the Leishmania glucose transporter? No evidence has been advanced within the last three years to support the speculation on this issue. The cloned putative GTs from the trypanosome will have to satisfy the functional test of glucose transport. The purely conjectural nature (as to function) of those genes, underscores the need for expression studies or at least to demonstrate hybridization to a better characterized and de facto glucose transporter. Besides, the homology of the citrate transporter (245), the phosphate transporter (246), and the a-ketoglutarate transporter (270) to the GT superfamily, makes no room for complacency in respect of ascribing glucose transport function to the cloned Kinetoplastid genes.

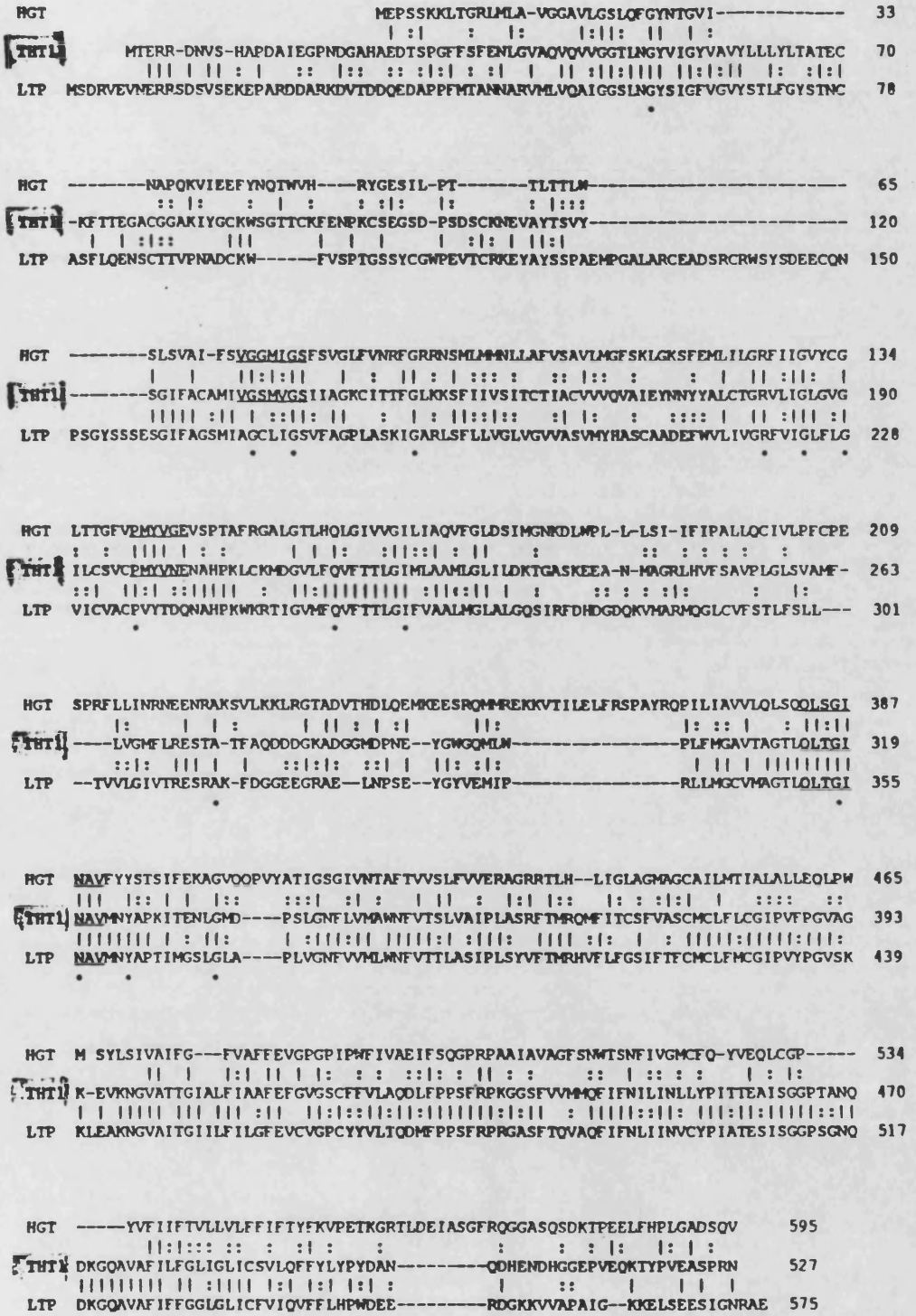
b. If the THTs are true to type, what is the function of a gene that shows specific hybridization to a well characterized GT viz. the rat GT cDNA. [Hybridization is

deemed specific on the basis of having cloned the same gene from two different libraries using cDNA and glucose-binding site probes].

c. The homology of THT to LTP suggests a functional identity. There is evidence for H+/glucose symport in both Leishmania and T. brucei, with similarities in Kms. If the two proteins are GTs, can it be inferred that THT is a coupled transporter? What then comes of the evidence of facilitated transport accruing from this study and others before?

d. The hydrophobicity profile of THT illustrates a typical 12 membrane-spanner. Which membrane does it span? The plasma membrane or the glycosomal membrane? [The compartmentation within the trypanosome, teleologically, would require a glycosomal GT as well]. Could there be plasma membrane-specific and glycosomal membrane-specific isoforms of the trypanosome GT? How structurally identical or different are they?

Needless to say, there are many more questions than answers with the new information, and unfortunately there will not be answers for some time yet.



Addendum: Comparison of THT with the erythrocyte GT and LTP

## APPENDIX A: REAGENTS

### General Buffers

**Krebs Ringer Phosphate, pH8.0**

| | |
|---|---|
| $KH_2PO_4$ | 22mM |
| NaCl | 98mM |
| KCl | 2 mM |
| $MgSO_4$ | 1 mM |

**SSC, 20X**

| | |
|---|---|
| NaCl | 3 M |
| $Na_2$ citrate | 0.3M |

**SSPE, 20X pH7.4**

| | |
|---|---|
| NaCl | 3M |
| $NaH_2PO_4$ | 0.2M |
| EDTA | 20mM |

**TE, pH7.4, 7.6, 8.0**

| | |
|---|---|
| Tris-HCl | 10mM |
| EDTA | 1 mM |

**MOPs, 10X**

| | |
|---|---|
| MOPS | 200mM |
| NaAc | 50 mM |
| EDTA | 10 mM |

**Krebs Ringer Hepes, pH7.6**

| | |
|---|---|
| Hepes | 10mM |
| NaCl | 140mM |
| KCl | 2mM |
| $MgCl_2$ | 1mM |
| $KH_2PO_4$ | 2mM |
| $CaCl_2$ | 1mM |

**Tris-acetate, TAE, 20X**

| | |
|---|---|
| Tris-acetate | 0.8M |
| NaAc | 1.0M |
| EDTA | 0.02M |

**Tris-Borate, 20X**

| | |
|---|---|
| Tris base | 1M |
| Boric acid | 1M |
| EDTA | 20mM |

**Tris-glyine, 10X**

| | |
|---|---|
| Tris-base | 250mM |
| Glycine | 2.5M |
| SDS | 1% |

**Gel loading buffer for DNA**

| | |
|---|---|
| bromophenol blue | 0.25% |
| xylene cyanol FF | 0.25% |

**Klenow, 10X**

| | |
|---|---|
| Tris-HCl, pH7.4 | 500mM |
| $MgCl_2$ | 100mM |
| DTT | 10mM |
| BSA | 500ug/ml |

**T4 DNA Ligase, 10X**

| | |
|---|---|
| Tris-HCl, pH7.5 | 500mM |
| $MgCl_2$ | 100mM |
| DTT | 10mM |
| ATP | 10mM |
| BSA | 1mg/ml |

**Denhardt's solution, 100X**

| | |
|---|---|
| Ficoll | 2g/100ml |
| polyvinylpyrrol) | 2g/100 ml |
| BSA (Fraction V) | 2g/100 ml |

**T4 polynucleotide kinase, 10X**

| | |
|---|---|
| Tris-HCl, pH7.5 | 100mM |
| $MgCl_2$ | 100mM |
| DTT | 50mM |
| spermidine | 1mM |

**One-Phor-All Buffer PLUS, 10X**

| | |
|---|---|
| Tris-acetate, pH7.5 | 100mM |
| MgAc | 100mM |
| KAc | 500mM |

**T3/T7 RNA polymerase, 10X**

| | |
|---|---|
| Tris-HCl, pH8.4 | 1000mM |
| $MgCl_2$ | 80mM |
| NaCl | 500mM |
| DTT | 300mM |
| spermidine | 20mM |

**DNA transfer solutions**

**Denaturing solution**

| | |
|---|---|
| NaOH | 0.4M |
| NaCl | 0.6M |

**Neutralizing solution**

| | |
|---|---|
| NaCl | 1.5M |
| Tris-HCl, pH7.5 | 0.5M |

## APPENDIX B

**Nutrient/growth media**

**Luria-Bertani (LB) medium**

| | |
|---|---|
| Bactotryptone | 10g/l |
| Yeast extract | 5g/l |
| NaCl | 10g/l |

**Luria-Bertani (LB) Agar**

LB medium plus: for

| | |
|---|---|
| Top agar | 7g/l |
| Plate agar | 15g/l |

**Lambda storage medium (SM)**

| | |
|---|---|
| NaCl | 5g/l |
| $MgSO_4.7H_2O$ | 2g/l |
| 1M Tris-HCl, pH 7.5 | 50ml/l |
| 2% gelatin | 5ml/l |

**2X yeast extract-tryptone (DYT)**

| | |
|---|---|
| Bactotryptone | 16g/l |
| yeast extract | 10g/l |
| NaCl | 5g/l |

**DYT plates as for LB plates.**

## APPENDIX C

**Recognition sequences for enzymes**

| **Enzyme** | **Site** |
|---|---|
| Apa1 | GGGCC^C |
| BamH1 | G^GATCC |
| EcoR1 | G^AATTC |
| HindIII | A^AGCTT |
| Kpn1 | GGTAC^C |
| Pst1 | CTGCA^G |
| PvuII | CAG^CTG |
| Sac1 | GAGCT^C |
| Sal1 | G^TCGAC |
| Sau3A | ^GATC |
| Sma1 | CCC^GGG |
| Xho1 | C^TCGAG. |

(^) Endonuclease cleavage site.

# APPENDIX D

## a. DNA Data

**Physical Constants of the Nucleoside Triphosphates.**

| Compound | M.W. | $\lambda_{max}$ (pH 7.0) | Absorbance at $\lambda_{max}$ for 1M Solution (E) (pH 7.0) |
|---|---|---|---|
| ATP | 507.2 | 259 | 15,400 |
| CTP | 483.2 | 271 | 9,000 |
| GTP | 523.2 | 253 | 13,700 |
| UTP | 484.2 | 262 | 10,000 |
| dATP | 491.2 | 259 | 15,200 |
| dCTP | 467.2 | 271 | 9,300 |
| dGTP | 507.2 | 253 | 13,700 |
| dTTP | 482.2 | 267 | 9,600 |

**Conversion formula**

$$\frac{\text{(observed absorbance at } \lambda_{max})}{\text{absorbance at } \lambda_{max} \text{ for 1M solution}} = \text{molar concentration of nucleic acid}$$

**Reference:**

Freifelder, D. (1982) in *Physical Biochemistry*, W.H. Freeman and Company, 494-536.

**Table 2. Lengths and Molecular Weights of Common Nucleic Acids.**

| Nucleic Acid | Number of Nucleotides | Molecular Weight |
|---|---|---|
| Lambda DNA | 48,502 (circular, dsDNA) | $3.0 \times 10^7$ |
| pBR322 DNA | 4,363 (dsDNA) | $2.8 \times 10^6$ |
| 28S rRNA | 4,800 | $1.6 \times 10^6$ |
| 23S rRNA | 3,700 | $1.2 \times 10^6$ |
| 18S rRNA | 1,900 | $6.1 \times 10^5$ |
| 16S rRNA | 1,700 | $5.5 \times 10^5$ |
| 5S rRNA | 120 | $3.6 \times 10^4$ |
| tRNA (*E. coli*) | 75 | $2.5 \times 10^4$ |

**Standards**

1. 1kb of double-stranded DNA (sodium salt) = $6.6 \times 10^5$ Daltons.
2. 1kb of single-stranded DNA (sodium salt) = $3.3 \times 10^5$ Daltons.
3. 1kb of single-stranded RNA (sodium salt) = $3.4 \times 10^5$ Daltons.
4. The average MW of a deoxynucleotide base = 324.5 Daltons.

## b. Codon Dictionary

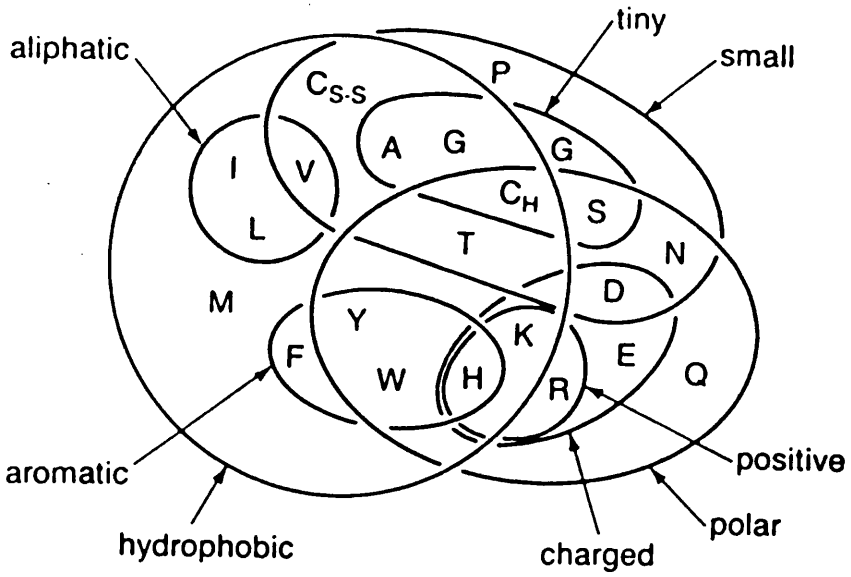
| 1st Position | 2nd Position U | 2nd Position C | 2nd Position A | 2nd Position G | 3rd Position |
|---|---|---|---|---|---|
| U | UUU Phe | UCU Ser | UAU Tyr | UGU Cys | U |
| | UUC Phe | UCC Ser | UAC Tyr | UGC Cys | C |
| | UUA Leu | UCA Ser | **UAA** End | **UGA** End | A |
| | UUG Leu | UCG Ser | **UAG** End | UGG Trp | G |
| C | CUU Leu | CUU Pro | CAU His | CGU Arg | U |
| | CUC Leu | CCC Pro | CAC His | CGC Arg | C |
| | CUA Leu | CCA Pro | CAA Gln | CGA Arg | A |
| | CUG Leu | CCG Pro | CAG Gln | CGG Arg | G |
| A | AUU Ile | ACU Thr | AAU Asn | AGU Ser | U |
| | AUC Ile | ACC Thr | AAC Asn | AGC Ser | C |
| | AUA Ile | ACA Thr | AAA Lys | AGA Arg | A |
| | AUG Met | ACG Thr | AAG Lys | AGG Arg | G |
| G | GUU Val | GCU Ala | GAU Asp | GGU Gly | U |
| | GUC Val | GCC Ala | GAC Asp | GGC Gly | C |
| | GUA Val | GCA Ala | GAA Glu | GGA Gly | A |
| | GUG Val | GCG Ala | GAG Glu | GGG Gly | G |

The codons read in the 5′→ 3′ direction.
Termination codons are in bold.

## c. Single letter annotation and molecular weights of amino acids

| Amino Acid | Three-letter Abbreviation | One-letter Symbol | Molecular Weight |
|---|---|---|---|
| Alanine | Ala | A | 89 |
| Arginine | Arg | R | 174 |
| Asparagine | Asn | N | 132 |
| Aspartic acid | Asp | D | 133 |
| Asparagine or aspartic acid | Asx | B | — |
| Cysteine | Cys | C | 121 |
| Glutamine | Gln | Q | 146 |
| Glutamic Acid | Glu | E | 147 |
| Glutamine or glutamic acid | Glx | Z | — |
| Glycine | Gly | G | 75 |
| Histidine | His | H | 155 |
| Isoleucine | Ile | I | 131 |
| Leucine | Leu | L | 131 |
| Lysine | Lys | K | 146 |
| Methionine | Met | M | 149 |
| Phenylalanine | Phe | F | 165 |
| Proline | Pro | P | 115 |
| Serine | Ser | S | 105 |
| Threonine | Thr | T | 119 |
| Tryptophan | Trp | W | 204 |
| Tyrosine | Tyr | Y | 181 |
| Valine | Val | V | 117 |

**d. Venn diagram of amino acid propensities**



## E. coli strains used and their genotypes

| | |
|---|---|
| Q359 | *supE hsdR* $\phi 80^r$ P2 |
| TG1 | *supE hsd*Δ5 *thi* Δ(*lac-proAB*)<br>F' [*traD*36 *proAB*$^+$ *lacI*$^q$ *lacZ*ΔM15] |
| XL1-Blue | *supE*44 *hsdR*17 *recA*1 *endA*1 *gyrA*46 *thi*<br>*relA*1 *lac*$^-$<br>F' [*proAB*$^+$ *lacI*$^q$ *lacZ*ΔM15 Tn*10*(*tet*$^r$)] |
| HB101 | *supE*44 *hsdS*20($r_B^- m_B^-$) *recA*13 *ara*-14<br>*proA*2 *lacY*1 *galK*2 *rpsL*20 *xyl*-5 *mtl*-1 |
| JM101[a] | *supE thi* Δ(*lac-proAB*)<br>F' [*traD*36 *proAB*$^+$ *lacI*$^q$ *lacZ*ΔM15] |

## REFERENCES

1. Molyneux, D.H. and Ashford, R.W. (1983) The biology of Trypanosoma and Leishmania, parasites of man and domestic animals. Taylor and Francis Ltd, London.

2. Lumsden, W.H.R. (1977) Problems in characterization and nomenclature of trypanosome populations. Ann. Soc. Belge. Med. Trop., **57**:361-368.

3. Gibson, W.C., Marshall, T.F.de C. and Godfrey, D.G. (1980) Numerical analysis of enzyme polymorphisms: new approach to the epidemiology and taxonomy of the subgenus Trypanozoon. Adv. Parasitol., **18**:175-246

4. Morel, C., Chiari, E., Mattei, D.M., Romanha, A.J. and Simpson, L. (1980) Strains and clones of Trypanosoma cruzi can be characterized by pattern of restriction endonuclease products of kinetoplast DNA minicircles. Proc. Natl. Acad. Sci. USA, **77**:6810-6814.

5. Hasegawa, M., Iida, Y., Yano, T., Takaiwa, F. and Iwabuchi, M. (1985) Phylogenetic relationships among eukaryotic kingdoms inferred from ribosomal RNA sequences. J. Mol. Evol., **22**:32-38.

6. Pace, N.R. Olsen, G.J. and Woese, C.R. (1986) Ribosomal RNA phylogeny and the primary lines of evolutionary descent. Cell, **45**:325-326.

7. Sogin, M.L., Elwood, H.J. and Gunderson, J.H. (1986) Evolutionary diversity in eukaryotic small subunit rRNA genes. Proc. Natl. Acad. Sci. USA, **83**:1383-1387.

8. Golvan, S.C., Castro, C., Segura, E., Casas, L. and Castaneda, M. (1991) Nucleotide sequences of the six very small molecules of Trypanosoma cruzi ribosomal RNA. Nucleic Acids Res., **19**:2496.

9. Hyde, J.E. (1990) Molecular Parasitology, Open University Press, Milton Keynes.

10. Vickerman, K. (1985) Developmental cycles and biology of pathogenic trypanosomes, Brit. Med. Bull., **41**:105-114.

11. Van der Ploeg, L.H.T. (1990) Antigenic variation in African trypanosomes: genetic recombination and transcriptional control of VSG genes. pp 51-97. In "Gene Rearrangement" eds Hames, B.D. and Glover, D.M., IRL Press, Oxford.

12. Grant, P.T. and Sargent, J.R (1960) Properties of L-a-glycerophosphate oxidase and its role in the respiration of Trypanosoma rhodesiense. Biochem. J., **76**:229-237

13. Opperdoes, F.R. and Borst, P. (1977). Localization of nine glycolytic enzymes in a microbody-like organelle in Trypanosoma brucei: the glycosome. FEBS Lett., **80**:360-364.

14. Bowman, I.B.R. and Flynn, I.W. (1976) Oxidative metabolism of trypanosomes. In: Lumsden, W.H.R., Evans, D.A. eds. Biology of the Kinetoplastida, vol.1 London: Academic Press, pp 435-476.

15. Opperdoes, F.R. (1985) Biochemical peculiarities of trypanosomes: African and South American. Brit. Med. Bull., **41**:130-136.

16. Opperdoes, F.R. (1987) Compartmentation of carbohydrate metabolism in trypanosomes. Annu. Rev. Microbiol., **41**:127-151

17. Eisenthal, R., Game, S., and Holman, G.D. (1989). Specificity and kinetics of hexose transport in Trypanosoma brucei. Biochim. Biophys. Acta, **985**:81-89.

18. Durieux, P.O., Schutz, P., Brun, P., and Kohler, P. (1991) Alterations in Krebs cycle enzyme activities and carbohydrate catabolism in two strains of Trypanosoma brucei during in vitro differentiation of their bloodstream to procyclic stages. Mol. Biochem. Parasitol., **45**:19-28

19. Bienen, E.J., Saric, M., Pollakis, G., Grady, R.W., and Clarkson, J. (1991). Mitochondrial development in Trypanosoma brucei brucei transitional bloodstream forms. Mol. Biochem. Parasitol., **45**:185-192.

20. Williams, N., Choi, S.Y.-W., Ruyechan, W.T. and Frank, P.H. (1991). The mitochondrial ATP synthase of Trypanosoma brucei: Developmental regulation through the life cycle. Arch. Biochem. Biophys., **288**:509-515.

21. Michelotti, E.F. and Hajduk, S.L. (1987) Developmental regulation of trypanosome mitochondrial gene expression. J. Biol. Chem., **262**:927-932.

22. Clayton, C. (1985) Structure and regulated expression of genes encoding fructose bisphosphate aldolase in Trypanosoma brucei. EMBO J., **4**:2997-3003.

23. Albert, S., Ernest, I., Poliszczak, A., Opperdoes, F.R. and Michels, P.A.M. (1991) Molecular cloning and analysis of two tandemly linked genes for pyruvate kinase of Trypanosoma brucei. Eur. J. Biochem., **200**:19-27.

24. Parsons, M. and Hill, T. (1989) Elevated phoshoglycerate kinase mRNA but not protein in monomorphic Trypanosoma brucei: implications for stage-regulation and post-transcriptional control. Mol. Biochem Parasitol., **33**:215-228.

25. Hart, D.T., Misset, O., Edwards, S.W., and Opperdoes, F.R. (1984). Comparison of the glycosomes (microbodies) isolated from Trypanosoma brucei bloodstream and cultured procyclic trypomastigotes. Mol. Biochem. Parasitol., **12**:25-35.

26. Visser, N., Opperdoes, F.R., and Borst, P. (1981) Subcellular compartmentation of glycolytic intermediates in Trypanosoma brucei. Eur. J. Biochem., **118**:521-526.

27. Opperdoes, F.R. (1988) Glycosomes may provide clues to the import of peroxisomal proteins. Trends Biochem. Sci., **13**:255-260.

28. Gutteridge, W.E. and Coombs, G.H. (1977) Biochemistry of Parasitic protozoa, The Macmillan Press Ltd, London.

29. Borst, P., Fase-Fowler, F., Frasch, A.C.C., Hoeijmakers, J.H.J., and Weijers, P.J. (1980) Characterization of DNA from Trypanosoma brucei and related trypanosomes by endonuclease digestion. Mol. Biochem. Parasitol., **1**:221-246.

30. Gibson, W.C., Osinga, K.A., Michels, P.A.M. and Borst, P. (1985) Trypanosomes of the subgenus Trypanozoon are diploid for housekeeping genes. Mol. Biochem. Parasitol., **16**:231-242.

31. Michels, P.A.M., Poliszczak A., Osinga, K.A., Misset, O., Van Beeumen, J., Wirenga, R.K., Borst, P., and Opperdoes, F.R. (1986) Two tandemly-linked identical genes code for glycosomal glyceraldehyde phosphate dehydrogenase in Trypanosoma brucei. EMBO J., **5**:1049-1056.

32. Osinga, K.A., Swinkels, B.W., Gibson, W.C., Borst, P., Veeneman, G.H., Van Boom, J.H., Michels, P.A.M., and Opperdoes, F.R. (1985) Topogenesis of microbody enzymes: a sequence comparison of the genes for the glycosomal (microbody) and cytosolic phosphoglycerate kinases of Trypanosoma brucei. EMBO J., **4**:3811-3817.

33. Tsudi, C., Young, A.S., Ruben, L., Patton, C.L., and Richards, F.F, (1985) Calmodulin genes in trypanosomes are tandemly repeated and produce multiple mRNAs with a common 5' leader sequence. Proc. Natl. Acad. Sci. USA, **82**:3998-4002.

34. Van der Ploeg, L.H.T., Valerio, D., De Lange, T., Bernards, A., Borst, P., and Grosveld, I.G. (1982c) An analysis of cosmid clones of nuclear DNA from Trypanosoma brucei shows that the genes for variant surface glycoproteins are clustered in the genome. Nucleic Acids Res., **10**:5905-5923.

35. Thomashow, S. L., Milhausen, M., Rutter, W.J., and Agabian, N. (1983) Tubulin genes are tandemly linked and clustered in the genome of Trypanosoma brucei. Cell, **32**:35-42.

36. Tait, A. (1980) Evidence for diploidy and mating in trypanosomes. Nature, **287**:536-538.

37. Tait, A., C.M.R, Turner., Le Page, R.W.F., Wells, J.M. (1989) Genetic evidence

that metacyclic forms of Trypanosoma brucei are diploid. Mol. Biochem. Parasitol., **37**:247-256.

38. Borst, P., Van der Ploeg, M., Van Hoek, J.F.M., Tas, J., and James, J. (1982) On the DNA content and ploidy of trypanosomes. Mol. Biochem. Parasitol., **6**:13-23.

39. Van der Ploeg, L.H.T., Schwartz, D.C., Cantor, C.R., Borst, P. (1984c) Antigenic variation in Trypanosoma brucei analyzed by electrophoretic separation of chromosome-sized DNA molecules. Cell, **37**:77-84.

40. Michels, P.A.M (1986) Evolutionary aspects of trypanosomes: analysis of genes. J. Mol. Evol., **24**:45-52.

41. Bennetzen, J.L. and Hall, B.D. (1982) Codon selection in yeast. J. Biol. Chem., **257**:3026-3031.

42. Dynan, W.S. (1986) Promoters for housekeeping genes. Trends Genet., **2**:196-197.

43. Michels, P.A.M. (1987) Genomic organization and gene structure in African trypanosomes. In "Gene Structure in Eukaryotic Microbes" ed. Kingston, J.R. vol. 22, pp 243-262, IRL Press, Oxford.

44. Jeffries, D., Tebabi, P. and Pays, E. (1991) Transient activity assays of the Trypanosoma brucei variant surface glycoprotein gene promoter: Control of gene expression at the transcriptional level. Mol. Cell. Biol., **11**:338-343.

45. White, T.C., Rudenko, G. and Borst, P. (1986) Three small RNAs within the 10kb trypanosome ribosomal RNA transcription unit are analogous to domain VII of other eukaryotic 28S ribosomal RNAs. Nucleic Acids Res., **14**:9471-9489.

46. Pays, E., Coquelet, H., Tebabi, P., Pays, A., Jeffries, D., Steinert, M., Koenig, E., Williams, R.O. and Roditi, I. (1990) Trypanosoma brucei: constitutive activity of the VSG and procyclin gene promoters. EMBO J., **9**:3145-3151.

47. Kozak, M. (1981) Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res., **9**:5233-5252.

48. Kozak, M. (1984) Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin in vivo. Nature, **308**:241-246.

49. Kozak, M. (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res., **12**:857-872.

50. Nussinov, R. (1984) Doublet frequencies in evolutionarily distinct groups. Nucleic

Acids Res., **12**:1749-1763.

51. Nussinov, R. (1981b). Eukaryotic dinucleotide preference rules and their implications for degenerate codon usage. J. Mol. Biol., **149**:125-131.

52. Boothroyd, J.C., Paynter, C.A., Coleman, S.L. and Cross, G.A.M. (1982) Complete nucleotide sequence of complementary DNA coding for a variant surface glycoprotein from Trypanosoma brucei. J. Mol. Biol., **157**:547-556.

53. Mowatt, M.R. and Clayton, C.E. (1987) Developmental regulation of a novel repetitive protein of Trypanosoma brucei. Mol. Cell. Biol., **7**:2838-2844.

54. Glass, D.J., Ramona, I.P. and Van der Ploeg, L.H.T. (1986) Conserved sequences and transcription of the hsp70 gene family in Trypanosoma brucei. Mol. Cell. Biol., **6**:4657-4666.

55. Shames, S.L., Kimmel, B.E., Peoples, O.P., Agabian, N. and Walsh, C.T. (1988) Trypanothione reductase of Trypanosoma cruzi: gene isolation, primary sequence determination, and comparison to glutathione reductase. Biochem., **27**:5014-5019.

56. Cairns, B.R., Collard, M.W. and Landfear, S.M. (1989) Developmentally regulated gene from Leishmania encodes a putative membrane transport protein. Proc. Natl. Acad.

Sci. USA, **86**:7682-7686.

57. Foote, S.J., Thompson, J.K., Cowman, A.F. and Kemp, D.J. (1989) Amplification of the multidrug resistance gene in some chloroquine-resistant isolates of P. falciparum. Cell, **57**:921-930.

58. Kochan, J., Perkins, M. and Ravetech, J.V. (1986) A tandemly repeated sequence determines the binding domain for the erythrocyte receptor binding protein of P. falciparum. Cell, **44**:689-696.

59. Vickerman, K. (1978). Antigenic variation in trypanosomes. Nature, **273**:613-617.

60. Steinert, M. and Pays, E. (1985) Genetic control of antigenic variation. Brit. Med. Bull., **41**:149-155.

61. Turner, M. J. (1985) Biochemistry of surface antigens. Brit. Med. Bull., **41**:137-143.

62. Vickerman, K. (1969) On the surface coat and flagellar adhesion in trypanosomes. J. Cell Sci., **5**:163-193.

63. Cross, G.A.M. (1975) Identification, purification and properties of clone-specific

glycoprotein antigens constituting the surface coat of Trypanosoma brucei. Parasitology, **71**:393-417.

64. Ferguson, M.A.J. and Cross, G.A.M (1984) Myristylation of the membrane form of Trypanosoma brucei variant surface glycoprotein. J. Biol. Chem., **259**:3011-3015.

65. Maizels, R.M. (1990) Parasite antigens. In "Parasites: immunity and pathology. The consequences of parasite infection in mammals" ed. Behnke, J.M. pp23-41. Taylor and Francis Ltd., London.

66. Borst, P. and Cross, G.A.M. (1982). Molecular basis for trypanosome antigenic variation. Cell, **29**:291-303.

67. Parsons, M., Nelson, R.G. and Agabian N. (1984a). Antigenic variation in African trypanosomes: DNA rearrangements program immune evasion. Immunol. Today, **5**:43-50.

68. Lamont, G.S., Tucker, R.S. and Cross, G.A.M. (1986) Analysis of antigen switching rates in Trypanosoma brucei. Parasitology, **92**:355-367.

69. Hudson, K.M., Taylor, A.E.R. and Elce, B.J. (1980) Antigenic changes in Trypanosoma brucei on transmission by tsetste fly. Parasite Immunol., **2**:57-69.

70. Turner, C.M.R. and Barry, J.D. (1989) High frequency of antigenic variation in Trypanosoma brucei rhodesiense infections. Parasitology, **99**:67-75.

71. Cross, G.A.M. (1978) Antigenic variation in trypanosomes. Proc. R. Soc. London B, **202**:55-72.

72. Snary, D. (1985) Biochemistry of surface antigens of Trypanosoma cruzi. Brit. Med. Bull., **41**:144-148.

73. Turner, M.J. (1980) Antigenic variation. In "The molecular basis of microbial pathogenicity" eds Smith, H., Skehel, J.J. and Turner, M.J. pp133-158, Verlag Chemie, Weinham.

74. Biggs, B.-A., Gooze, L., Wycherley, K., Woolish, W., Southwell, B., Leech, J.H. and Brown, G.V. (1991) Antigenic variation in Plasmodium falciparum. Proc. Natl. Acad. Sci. USA, **88**:9171-9174.

75. Capdeville, Y. (1979a) Regulation of surface antigen expression in Paramecium primaurelia. II. Role of the surface antigen itself. J. Cell Physiol., **99**:383-393.

76. Van Meirvenne, N. and Vervoort, T. (1977) Comparison of variable antigenic types produced by trypanosome strains of the subgenus Trypanozzon. Ann. Soc. Belge.

Med. Trop., **57**:409

77. Miller, E.N. and Turner, M.J. (1981) Analysis of antigenic types appearing in first relapse populations of clones of Trypanosoma brucei. Parasitology, **82**:63-80.

78. Le Ray, D., Barry, J.D. and Vickerman, K. (1978) Antigenic heterogeneity of metacyclic forms of Trypanosoma brucei. Nature, **273**:300-302.

79. Barry, J.D., Hadjuk, S.L., Vickerman, K. and Le Ray, D. (1979a) Detection of multiple variable antigen types in metacyclic populations of Trypanosoma brucei. Trans. R. Soc. Trop. Med. Hyg., **73**:205

80. Doyle, J.J. (1977) Antigenic variation in the salivarian trypanosomes. In Miller L.H., Pino, J.A and McKelvery J. Jnr (eds). Immunity to blood parasites of Animals and Man, p27, Plenum Press, New York.

81. Capbern, A., Giroud, C., Baltz, T. and Mattern, P. (1977a) Trypanosoma equiperdum: etude des variations antigeniques au cours de la trypanosomose experimentale du lapin. Exp. Parasitol., **42**:6

82. Borst, P. and Greaves, D.R. (1987) Programmed gene rearrangements altering gene expression. Science, **235**:658-667.

83. Esser, K.M. and Schoenbechler, M.J. (1985) Expression of two variant surface glycoproteins on individual African trypanosomes during antigen switching. Science, **228**:190-193.

84. Czichos, C.E., Nonnengaesser, C. and Overath, P. (1986) Trypanosoma brucei: cis-aconitate and temperature reduction as triggers of synchronous transformation of bloodstream to procyclic trypomastigotes in vitro. Exp. Parasitol., **62**:283-291

85. Dorn, P.L., Aman, R.A and Boothroyd, J.C. (1991) Inhibition of protein synthesis results in superinduction of procyclin (PARP) RNA levels. Mol. Biochem. Parasitol., **44**:133-140.

86. Roditi, I., Carrington, M. and Turner, M. (1987) Expression of a polypeptide containing a dipeptide repeat is confined to the insect stage of Trypanosoma brucei. Nature, **325**: 272-274.

87. Rudenko, G., Le Blancq, S., Smith, J., Gwo Shu Lee, M., Rattray, A. and Van der Ploeg, L.H.T (1990) Procyclic acidic repetitive protein (PARP) gene located in an unusually small a-amanitin resistant transcription unit: PARP promoter analysis assayed by transient DNA transfection of Trypanosoma brucei. Mol. Cell. Biol., **10**: 3492-3504.

88. Hadjuk, S.L, Cameron, C.R., Barry, J.D. and Vickerman, K. (1981) Antigenic

variation in cyclically transmitted Trypanosoma brucei: variable antigenic type composition of metacylcic populations from the salivary glands of Glossina morsitans. Parasitology, **83**:595-607.

89. Tetley, L. and Vickerman, K. (1985) Differentiation in Trypanosoma brucei: host-parasite cell junctions and their persistence during acquisition of the variable antigen coat. J. Cell Sci., **74**:1-19.

90. Cornellisen, A.W.C.A., Bakkern, G.A.M., Barry, J.D., Michels, P.A.M. and Borst, P. (1985). Characterization of trypanosome variant antigen genes active in the tsetse fly. Nucleic Acids Res., **13**:4661-4676.

91. Esser, K.M., Schoenbechler, M.J. and Gingrich, J.B (1982) Trypanosoma rhodesiense blood forms express all antigen specificities to protect against metacyclic (insect-form) challenge. J. Immunol., **129**:1715-1718.

92. Leonardo, M.J., Esser, K.M., Moon, A.M., Van der Ploeg, L.H.T. and Donelson, J.E. (1986) Metacyclic variant surface glycoprotein genes of Trypanosoma brucei subsp. rhodesiense are activated in situ, and their expression is transcriptionally regulated. Mol. Cell. Biol., **6**:1991-1997.

93. Borst, P. (1986b) Discontinuous transcription and antigenic variation in

trypanosomes. Annu. Rev. Biochem., **55**:701-732.

94. Kooter, J.M., De Lange, T. and Borst, P. (1984) Discontinuous synthesis of mRNA in trypanosomes. EMBO J., **3**:2387-2392.

95. Campbell, D.A., Thornton, D.A. and Boothroyd, J.C. (1984a) Apparent discontinuous transcription of Trypanosoma brucei variant surface antigen genes. Nature, **311**:350-355.

96. Sutton, R.E. and Boothroyd, J.C. (1986) Evidence for trans-splicing in trypanosomes. Cell, **47**:527

97. Nelson, R.G., Parsons, M., Selkirk, M., Newport, G., Barr, P. and Agabian, N. (1984) Sequences homologous to variant antigen mRNA spliced leader in trypanosomatidae which do not undergo antigenic variation. Nature, **308**:665-667.

98. Sather, S. and Agabian, N. (1985) A 5'-spliced leader is added in trans to both alpha- and beta-tubulin transcripts in Trypanosoma brucei. Proc. Natl. Acad. Sci. USA, **82**:5695-569.

99. De Lange, T., Berkvens, T.M., Veerman, H.J.G., Frasch, A.C.C., Barry, J.D. and Borst, P. (1984a) Comparison of the genes coding for the common 5' terminal

sequence of messenger RNAs in three trypanosome species. Nucleic Acids Res., **12**:4431-4443.

100. Krause, M. and Hirsh, D. (1987) A trans-spliced leader sequence on actin mRNA in Caenorhabditis elegans. Cell, **49**:753-761.

101. Mount, S.M. (1982) A catalogue of splice junction sequences. Nucleic Acids Res., **10**:459-472.

102. Eisen, H. (1988) RNA editing: who's on first? Cell, **53**:331-332.

103. Feagin, J.E., Jasmer, D.P. and Stuart, K. (1987) Developmentally regulated addition of nucleotides within apocytochrome b transcripts in Trypanosoma brucei. Cell, **49**:337-345.

104. Benne, R., Van den Burg, J., Brakenhoff, J.P.J., Sloof, P., Van Boom, J.H. and tromp, M.C. (1986) Major transcript of the frameshifted COXII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. Cell, **46**:819-826.

105. Feagin, J.E., Abraham, J.M. and Stuart, K. (1988). Extensive editing of the cytochrome c oxidase III transcripts in Trypanosoma brucei. Cell, **53**:413-422.

106. Benne, R. (1990) RNA editing in trypanosomes: Is there a message? Trends in Genets., **6**:177-181.

107. Gutteridge, W.E. (1985) Existing chemotherapy and its limitations. Brit. Med. Bull., **41**:162-168.

108. Clarkson, A.B. Jr. and Brohn, F.H. (1976) Trypanosomiasis: an approach to chemotherapy by the inhibition of carbohydrate metabolism. Science, **194**:204-206.

109. Grant, P.T. and Sargent, J.R. (1961) L-a-glycerophosphate dehydrogenase, a component of an oxidase system in Trypanosoma rhodesiense. Biochem. J., **81**:206-214.

110. Opperdoes, F.R., Borst, P. and Fonck, K. (1976) The potential use of inhibitors of glycerol-3-phosphate oxidase for chemotherapy of African trypanosomiasis. FEBS Lett., **62**:169-172.

111. Hammond, D.J. and Gutteridge, W.E. (1984) Purine and pyrimidine metabolism in the trypanosomatidae. Mol. Biochem. Parasitol., **13**:243-261.

112. Marr, J.J., Berens, R.L. and Nelson, D.J. (1978) Antitrypanosomal effect of allopurinol: conversion in vivo to aminopyrazolopyrimidine nucleotides by Trypanosoma cruzi. Science, **201**:1018-1020.

113. Jennings, F.W. (1988) Chemotherapy of trypanosomiasis: the potentiation of melarsoprol by concurrent difluoromethylornithine (DFMO) treatment. Trans. R. Soc. Trop. Med. Hyg., **82**:572-573.

114. Bacchi, C.J., Nathan, H.C., Hutner, S.H., McCann, P.P. and Sjoerdsma, A. (1980) Polyamine metabolism: A potential therapeutic target in trypanosomes. Science, **210**:332-334.

115. Fairlamb, A.H., Henderson, G.B., Bacchi, C.J. and Cerami, A. (1987) In vivo effects of difluoromethylornithine (DFMO) on trypanothione and polyamine levels in bloodstream forms of Trypanosoma brucei. Mol. Biochem. Parasitol., **24**:185-191.

116. Phillips, M.A., Coffino, P. and Wang, C.C. (1988) Trypanosoma brucei: enzyme purification, characterization, and expression in Eschericia coli. J. Biol. Chem., **263**:17933-17939.

117. Fairlamb, A.J., Oduro, K.K. and Bowman, I.B.R. (1979) Action of the trypanosomal drug suramin on the enzymes of aerobic glycolysis of Trypanosoma brucei in vivo. FEBS Special Meeting on Enzymes, Dubrovnik, Abstract S4-10.

118. Fairlamb, A.H., Henderson, G.B. and Cerami, A. (1989) Trypanothione is the primary target for arsenical drugs against African trypanosomes. Proc. Natl. Acad. Sci.

USA, **86**:2607-2611.

119. Fairlamb, A.H., Blackburn, P., Ulrich, P, Chait, B.T. and Cerami, A. (1985) Trypanothione: a novel bis(glutathionyl)spermidine cofactor for glutathione reductase in trypanosomatids. Science, **227**:1485-1487.

120. Doering, T.L., Raper, J., Buxbaum, L.U., Adams, S.P., Gordon, J.I., Hart, G.W. and Englund, P.T. (1991) An analogue of myristic acid with selective toxicity for African trypanosomes. Science, **252**:1851-1854.

121. Semenza, G., Kessler, M., Hosang, M., Weber, J. and Schmidt, U. (1979) Biochemistry of the Na+-D-glucose cotransporter of the small intestine brush-border membrane. Biochim. Biophys. Acta, **779**:343-379.

122. Skou, J.C. and Norby, J.G. (eds) (1979) Na+-K+-ATPase: Structure and kinetics. Academic Press, New York.

123. Postma, P.W. and Roseman, S. (1976) The bacterial PEP:sugar phosphotransferase system. Biochim. Biophys. Acta, **457**:213-257.

124. Gruenberg, J., Sharma, P.R. and Deshusses, J. (1978) D-Glucose transport in Trypanosoma brucei: D-glucose transport is the rate-limiting step

of its metabolism. Eur. J. Biochem., **89**:461-469.

125. Ter Kuile, B.H. and Opperdoes, F.R. (1991) Rate-limiting steps in glycolysis and regulation of the glycolytic flux. J. Biol. Chem., **266**:857-862.

126. Chin, J.J., Jung, E.K.J. and Jung, C.J. (1986) Structural basis of human erythrocyte glucose transporter function in reconstituted vesicles a-helix orientation. J. Biol. Chem., **261**:7101-7104.

127. Alvarez, J., Lee, D.C., Baldwin, S.A. and Chapman, D. (1987) Fourier transform infrared spectroscopic study of the structure and conformational changes of the human erythrocyte glucose transporter. J. Biol. Chem., **262**:3502-3509.

128. Holman, G.D. and Rees, W.D. (1987) Photolabeling of the hexose transporter at external and internal sites: fragmentation patterns and evidence for a conformational change. Biochim. Biophys. Acta, **897**:395-405.

129. Gibbs, A.F., Chapman, D. and Baldwin, S.A. (1988) Proteolytic dissection as a probe of conformational changes in the human erythrocyte glucose transporter protein. Biochem. J., **256**:421-427.

130. Koepsell, H., Korn, K., Raszeja-Specht, A, Bernotat-Danielowski S. and Ollig, D.

(1988) Monoclonal antibodies against the renal Na+-D-glucose cotransporter: Identification of antigenic polypeptides and demonstration of functional coupling of different Na+-cotransport systems. J. Biol. Chem., **263**:18419-18429.

131. Krupka, R.M. (1971) Evidence for a carrier conformational change associated with sugar transport in erythrocytes. Biochem., **10**:1143-1148.

132. Rampal, A.L. and Jung, C.Y. (1987) Substrate-induced conformational change of human erythrocyte glucose transporter: inactivation by alkylating agents. Biochim. Biophys. Acta, **896**:287-294.

133. Appleman, J.R. and Lienhard, G.E. (1985) Rapid kinetics of the glucose transporter from human erythrocytes: Detection and measurement of a half-turnover of the purified transporter. J. Biol. Chem., **260**:4575-4578.

134. Pawagi, A.B. and Deber, C.M. (1990) Ligand-dependent quenching of tryptophan fluorescence in human erythrocyte hexose transport protein. Biochem., **29**:950-955.

135. Kahlenberg, A. and Dolansky, D. (1972) Structural requirements of D-glucose for its binding to isolated human erythrocyte membranes. Can. J. Biochem., **50**:638-643.

136. Barnett, J.E.G., Holman, G.D., Chalkey, R.A. and Munday, K.A. (1975) Evidence

for two asymmetric conformational states in the human erythrocyte sugar transport system. Biochem. J., **145**:417-429.

137. Barnet, J.E.G., Holman, G.D. and Munday, K.A. (1973) Structural requirement for binding to the sugar transport system of the human erythrocyte. Biochem., **131**:211-221.

138. Berlin, R.D. (1970) Specificities of Transport systems and enzymes. Science, **168**:1539-1545.

139. Jung, E.K.Y., Chin, J.J. and Jung, C.Y. (1986) Structural basis of human erythrocyte glucose transporter function: Hydrogen exchange. J. Biol. Chem., **261**:9155-9160.

140. Silverman, M. (1991) Structure and function of hexose transporters. Annu. Rev. Biochem., **60**:757-794.

141. Ames, G. F.-L. (1986) Bacterial perisplasmic transport systems: structure, mechanism and evolution. Annu. Rev. Biochem., **55**:397-425.

142. Davis, E.O. and Henderson, P.J.F. (1987) The cloning and DNA sequence of the gene XylE for xylose-proton symport in Escherichia coli K12. J. Biol. Chem., **262**:13928-13932.

143. Maiden, M.C.J., Jones-Mortimer, M.C. and Henderson, P.J.F. (1988) Cloning, DNA sequence and amplified expression of the araE gene encoding H+-arabinose symport in Escherichia coli. J. Biol. Chem., **263**:8003-8010.

144. Buchel, D.E., Gronenborr, B. and Muller-Hill, B. (1980) Sequence of the lactose permease gene. Nature, **283**:541-545.

145. Maiden, M.C.J., Davis, E.O, Baldwin, S.A., Moore D.C.M., Henderson, P.J.F. (1987) Mammalian and bacterial sugar transport
proteins are homologous. Nature, **325**:641-643.

146. Baldwin, S.A. and Henderson, P.J.F. (1989) Homologies between sugar transporters from eukaryotes and prokaryotes. Annu. Rev. Physiol., **51**:459-471.

147. Henderson, P.J.F. and Maiden, M.C.J. (1990) Homologous sugar transport proteins in Escherichia coli and their relatives in both prokaryotes and eukaryotes. Phil.Trans. R. Soc. Lond., **326**: 391-410.

148. Szkutnicka, K., Tschopp, J.F., Andrews, L. and Cirillo, V.P. (1989) Sequence and structure of the yeast galactose transporter.
J. Bacteriol., **171**:4486-4493.

149. Zilberstein, D. and Dwyer, D.M. (1984) Glucose transport in Leishmania donovani promastigotes. Mol. Biochem. Parasitol., **12**:327-336.

150. Zilberstein, D. and Dwyer, D.M. (1985) Protonmotive force-driven active transport of D-glucose and L-proline in the protozoan parasite Leishmania donovani. Proc. Natl. Acad. Sci. USA, **82**:1716-1720.

151. Izumo, A., Tanabe, K., Kato, M., Doi, S., Maekawa, K. and Takada, S. (1989) Transport processes of 2-deoxy-D-glucose in erythrocytes infected with Plasmodium yoelii, a rodent malaria parasite. Parasitology, **98**:371-379.

152. Parsons, M. and Nielsen, B. (1990). Active transport of 2-deoxy-D-glucose in Trypanosoma brucei procyclic forms. Mol. Biochem. Parasitol., **42**:197-204.

153. Munoz-Antonia, T., Richards, F.F. and Ullu, E. (1991) Differences in glucose transport between bloodstream and procyclic forms of Trypanosoma brucei rhodesiense. Mol. Biochem. Parasitol., **47**:73-82.

154. Hediger, M.A., Turk, E. and Wright, E.M. (1989) Homology of the human intestinal Na+/glucose and Escherichia coli Na+/proline cotransporters. Proc. Natl. Acad. Sci. USA, **86**:5748-5752.

155. Mesbah, S.P. (1990) Computer-aided analysis of the extended transporter superfamily. B.Sc, Bath University.

156. Mueckler, M.M., Caruso, C., Baldwin, S.A., Panico, M., Blench, I., Morris, H.R., Allard, J.W., Lienhard, G.E. and Lodish, H.F. (1985) Sequence and structure of a human glucose transporter. Science, **229**:941-945.

157. Kyte, J. and Doolittle, R.F (1982) A simple method for displaying the hydropathic character of a protein. J. Mol. Biol., **157**:105-132.

158. Eisenberg, D., Weiss, R.M. and Terwilliger, T.C. (1984) The hydrophobic moment detects periodicity in protein hydrophobicity. Proc. Natl. Acad. Sci. USA, **81**:140-144.

159. Chou, P.Y. and Fasman, G.D. (1978) Empirical predictions of protein conformation. Annu. Rev. Biochem., **47**:251-276.

160. Vogel, H. and Jahnig, F. (1986) Models for the structure of outer-membrane proteins of Escherichia coli derived from Raman spectroscopy and prediction methods. J. Mol. Biol., **190**:191-199.

161. Birnbaum, M.J., Haspel, H.C. and Rosen, O.M. (1986). Cloning and characterization of cDNA encoding the rat brain glucose transporter protein. Proc. Natl.

Acad. Sci. USA, **83**:5784-5788.

162. Thorens, B., Sarkar, H.K., Kaback, H.R. and Lodish, H.F. (1988). Cloning and functional expression in bacteria of a novel glucose transporter present in liver, kidney, and B-pancreatic islet cells. Cell, **55**:281-290.

163. Birnbaum, M.J. (1989) Identification of a novel gene encoding an insulin-responsive glucose transporter protein. Cell, **57**:305-315.

164. James, D.E., Strube, M. and Mueckler, M. (1989) Molecular cloning and characterization of an insulin regulatable glucose transporter. Nature, **338**:83-87.

165. Charron, M.J., Brosius, F.C. III, Alper, S.L. and Lodish, H.F. (1989) A novel glucose transporter protein expressed predominantly in insulin-responsive tissues. Proc. Natl. Acad. Sci. USA, **86**:2535-2539.

166. Permutt, M.A., Koranyi, L., Keller, K., Lacey, P.E., Sharp, D.W. and Mueckler, M. (1989) Cloning and functional expression of a human pancreatic islet glucose transporter cDNA. Proc. Natl. Acad. Sci. USA, **86**:8688-8692.

167. Bell, G.I., Kayano, T., Buse, J.B., Burant, C.F., Takeda, J., Lin, D., Fukumoto, H. and Seino, S. (1990) Molecular biology of mammalian glucose transporters. Diabetes

Care, **13**:198-208.

168. Celenza, J.L., Marshall-Carlson, L. and Carlson, M. (1988) The yeast SNF3 gene encodes a glucose transporter homologous to the mammalian protein. Proc. Natl. Acad. Sci. USA, **85**:2130-2134.

169. Sauer, N., Friedlander, K. and Graml-Wicke, U. (1990) Primary structure, genomic organization and heterologous expression of a glucose transporter from Arabidopsis thaliana. EMBO J., **9**:3045-3050.

170. Zilberstein, D., Dwyer, D.M., Matthaei, S. and Horuk, R. (1986) Identification and biochemical characterization of the plasma membrane glucose transporter of Leishmania donovani. J. Biol. Chem., **261**:15053-15057.

171. Hediger, M.A., Coady, M.J., Ikeda, T.S. and Wright, E.M. (1987) Expression cloning and cDNA sequencing of the Na+/glucose transporter. Nature, **330**:379-381.

172. Williams, S.A. and Birnbaum, M.J. (1988) The rat facilitated glucose transporter gene: transformation and serum-stimulated transcription initiate from identical sites. J. Biol. Chem., **263**:19513-19518.

173. Fontana, A., Fassina, G., Vita, C., Dalzoppo, D., Zamai, M. and Zambonin, M.

(1986) Correlation between sites of limited proteolysis and segmental mobility in thermolysin. Biochem., **25**: 1847-1851.

174. Cairns, M.T., Alvarez, J., Panico, M., Gibbs, A.F., Morris, H.R., Chapman, D. and Baldwin, S.A. (1987) Investigation of the structure and function of the human erythrocyte glucose transporter by proteolytic dissection. Biochim. Biophys. Acta, **905**:295-310.

175. Shanahan, M.F. and D'Artel-Ellis, J. (1984) Orientation of the glucose transporter in the human erythrocyte membrane. Investigation by in situ proteolytic cleavage. J. Biol. Chem., **259**:13878-13884.

176. Seckler, R., Moroy, T., Wright, J.K. and Overath, P. (1986) Anti-peptide antibodies and proteases as structural probes for the lactose/H+ transporter of Escherichia coli: A loop around amino acid residue 130 faces the cytoplasmic side of the membrane. Biochem., **25**:2403-2409.

177. Andersson, L. and Lundahl, P. (1988) C-terminal-specific monoclonal antibodies against the human red cell glucose transporter. Epitope localization with synthetic peptides. J. Biol. Chem., **263**:11414-11420.

178. Hopp, T.P. and Woods, K.R. (1981) Prediction of antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA, **78**:3824-3828.

179. Novotny, J., Handschumacher, M., Haber, E., Bruccoleri, R.E., Carlson, W.B., Fanning, D.W., Smith, J.A. and Rose, G.D. (1986) Antigenic determinants in proteins coincide with surface regions accessible to large probes (antibody domains). Proc. Natl. Acad. Sci. USA, **83**:226-230.

180. Krchnak, V., Mach, O. and Maly, A. (1987) Computer prediction of potential immunogenic determinants from protein amino acid sequence. Anal. Biochem., **165**:200-207.

181. Westof, C., Altschuh, D., Moras, D., Bloomer, A.C., Mondragon, A., Klug, A. and Van Regenmortel, M.H.V. (1984) Correlation between sequential mobility and the location of antigenic determinants in proteins. Nature, **311**:123-126.

182. Manoil, C. and Beckwith, J. (1986) A genetic approach to analyzing membrane protein topology. Science, **233**:1403-1408.

183. Boyd, D., Manoil, C. and Beckwith, J. (1987) Determinants of membrane protein topology. Proc. Natl. Acad. Sci. USA, **84**:8525-8529.

184. Manoil, C., Boyd, D., Beckwith, J. (1988) Molecular genetic analysis of membrane protein topology. Trends Genet., **4**:223-226.

185. Henderson, R. and Unwin, P.N.T. (1975) Three dimensional model of purple membrane obtained by electron microscopy. Nature, **257**:28-32.

186. Engelman, D.M., Henderson, R., McLachlan, A.D. and Wallace, B.A. (1980) Path of the polypeptide in bacteriorhodopsin, Proc. Natl. Acad. Sci. USA, **77**:2023-2027.

187. Hartmann, E., Rapoport, T.A. and Lodish, H.F. (1989) Predicting the orientation of eukaryotic membrane spanning proteins. Proc. Natl. Acad. Sci. USA, **86**:5786-5790.

188. Finer-Moore, J. and Stroud, R.M. (1984) Amphipathic analysis and possible formation of the ion channel in an acetylcholine receptor. Proc. Natl. Acad. Sci. USA, **81**:155-159.

189. Landshulz, W.H., Johnson, P.F. and McKnight, S.L. (1989) The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. Science, **243**:1681-1685.

190. Buckland, R., Wild, F., Turner, R. and Tjian, R. (1989) Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFOS-cJUN heterodimers. Science, **243**:1689-1694.

191. Iverson, L.E., Tanouye, M.A., Lester, H.A., Davidson, N. and Rudy, B. (1988) A-type potassium channels expressed from shaker locus cDNA. Proc. Natl. Acad. Sci. USA,

**85**:5723-5727.

192. White, M.K. and Weber, M.J. (1989) Leucine-zipper motif update. Nature, **340**:103-104.

193. Witters, L.A., Vater, C.A. and Lienhard, G.E. (1985) Phosphorylation of the glucose transporter in vitro and in vivo by protein kinase C. Nature, **315**:777-778.

194. Cuppoletti, J. and Jung, C.Y. (1980) Glucose transport carrier of erythrocytes. Radiation target size measurement based on flux inactivation. J. Biol. Chem., **256**:1305-1306.

195. Goldkorn, T., Rimon, G., Kempner, E.S. and Kaback, H.R. (1984) Functional molecular weight of the lac carrier protein from Escherichia coli as studied by radiation inactivation analysis. Proc. Natl. Acad. Sci. USA, **81**:1021-1025.

196. Richarme, G. (1982) Associative properties of the Escherichia coli galactose binding protein and maltose binding protein. Biochim. Biophys. Res. Comm., **105**:476-481.

197. White, M.K., Rall, T.B. and Weber, M.J. (1991) Differential regulation of glucose transporter isoforms by the src oncogene in Chicken Embryo Fibroblasts. Mol. Cell. Biol.,

**11**:4448-4454.

198. Goffrini, P., Weslowski-Louvel, M., Ferrero, I. and Fukuhara, H. (1990) RAG1 gene of the yeast Kluyveromyces lactis codes for a sugar transporter. Nucleic Acids Res., **18**:5294.

199. Lanham, S.M. (1968) Separation of trypanosomes from the blood of infected rats and mice by anion exchangers. Nature, **218**:1273-1274.

200. Eisenthal, R. and Panes, A. (1985) The aerobic/anaerobic transitions of glucose metabolism in Trypanosoma brucei. FEBS Letts., **181**:23-27.

201. Voorheis, H.P., Gale, J.S., Owen, M.J. and Edwards, W. (1979) The isolation and partial characterization of the plasma membrane from Trypanosoma brucei. Biochem. J., **180**:11-24.

202. Kasahara, M. and Hinkle, P.C. (1977) Reconstitution and purification of the D-glucose transporter from human erythrocytes. J. Biol. Chem., **252**:7383-7390.

203. Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol., **98**:503-517.

204. Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem., **132**:6-13.

205. Caruthers, M.H. (1983) New methods for chemically synthesizing deoxyoligonucleotides. In Weismann, S.M. (ed) "Methods of DNA and RNA sequencing" New York, Praeger Publishers, pp1-22.

206. Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA, **74**:5463-5467.

207. Staden, R. (1982) An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences. Nucleic Acids Res., **10**:2951-2961.

208. Barton, G.J. and Sternberg, M.J.E. (1987) A strategy for the rapid multiple alignment of protein sequences. Confidence levels from tertiary structure comparisons. J. Mol. Biol., **198**:327-337.

209. Stockwell, P.A. (1988) HOMED: a homologous sequence editor. Trends Biochem. Sci., **13**:332-334.

210. Lipman, D.J. and Pearson, W.R. (1985) Rapid and sensitive protein similarity searches. Science, **227**:1435-1441.

211. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res., **18**:7035-7056.

212. Chromczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem., **162**:156-159.

213. Sorenson, K. and Brodbeck, U. (1986) A sensitive protein assay method using microtire plates. Experentia, **42**:161-162.

214. Flynn, I.W. and Bowman, I.B.R. (1973) The metabolism of carbohydrate by polymorphic African trypanosomes. Comp. Biochem. Physiol., **45B**:25-42.

215. Game, S., Holman, G. and Eisenthal, R. (1986) Sugar transport in Trypanosoma brucei: a suitable kinetic probe. FEBS Letts., **194**:126-130.

216. Schleif, R. (1969) An L-arabinose binding protein and arabinose permeation in Escherichia coli. J. Mol. Biol., **46**:185-196.

217. Breathnach, R. and Chambon, P. (1981) Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem., **50**:349-383.

218. Guy, H.R. and Seetharamulu, P. (1986) Molecular model of the action potential sodium channel. Proc. Natl. Acad. Sci. USA, **83**:508-512.

219. Henderson, P.J.F. and McPherson, A.J.S. (1986) Assay, genetics, proteins and reconstitution of proton-linked galactose, arabinose, xylose transport systems of Escherichia coli. Methods Enzymol., **125**:387-429.

220. Malo, C. (1988) Kinetic evidence for heterogeneity in Na+D-glucose cotransport systems in the human fetal small intestine. Biochim. Biophys. Acta, **938**:181-188.

221. Coppens, I., Opperdoes, F.R., Courtroy, P.J. and Baudhuin, P. (1987) Receptor-mediated endocytosis in the bloodstream form of Trypanosoma brucei. J. Protozool., **34**:465.

222. Sherman, D.R., Janz, L., Hug, M. and Clayton, C. (1991) Anatomy of the **parp** gene promoter of Trypanosoma brucei. EMBO J., **10**:3379-3386.

223. Hong, S.H. and Marmur, J. (1986) Primary structure of the maltase gene of the MAL6 locus of Saccharomyces carlbergensis. Gene, **41**:75-84.

224. Bensimhon, M., Gabarro-Arpa, J., Ehrlich, R., and Reiss, C. (1983) Physical characteristics in eukaryotic promoters. Nucleic Acids Res., **11**:4521-4540.

225. Singh, H., Sen, R., Baltimore, D. and Sharp, P.A. (1986) A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature, **319**:154-158.

226. Charney, P., Treisman, R., Mellon, P., Chao, M.V., Axel, R. and Maniatis, T. (1984) Differences in human a- and b-globin gene expression in mouse erythroleukaemia cells. Cell, **38**:251-263.

227. Montague, P. (1987) Comparisons of the molecular structure between mammalian and lower eukaryotic genes. In " Gene structure in eukaryotic microbes" ed Kinghorn, J.R., IRL press, Oxford.

228. Myers, R.M., Tilly, K. and Maniatis, T. (1986) Fine structure genetic analysis of a b-globin promoter. Science, **232**:613-618.

229. Dobson, M.J., Tuite, M.F., Roberts, N.A., Kingsman, A.J., Kingsman, S.M., Perkins, R.E. and Conroy, S.C. (1982) Conservation of high efficiency promoter sequences in Saccharomyces cerevisiae. Nucleic Acids Res., **10**:2625-2637.

230. Rich, A., Nordheim, A. and Wang, A. H.-J (1984) The chemistry and biology of left-handed Z-DNA. Annu. Rev. Biochem., **53**:791-846.

231. Hamada, H., Petrino, M.G. and Kakunaga, T. (1982) A novel repeated element with Z-DNA forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc. Natl. Acad. Sci. USA, **79**:6465-6469.

232. Evans, T., Schon, E., Gora-Mastak, G., Patterson, J. and Efstratiadis, A. (1984) S1-hypersensitive sites in eukaryotic promoter regions. Nucleic Acids Res., **12**:8043-8058.

233. Larsen, A., and Weintraub, H. (1982) An altered DNA conformation detected by S1 nuclease occurs at specific regions in active chick globin chromatin. Cell, **29**:609-622.

234. Naylor, L.H. and Clark, E.M. (1990) $d(TG)_n.d(CA)_n$ sequences upstream of the rat prolactin gene form Z-DNA and inhibit gene transcription. Nucleic Acids Res., **18**:1595-1601.

235. Wittig, B. and Berlin, R.D. (1991) Transcription is associated with Z-DNA formation in metabolically active permeabilized mammalian cell nuclei. Proc. Natl. Acad. Sci. USA, **88**:2259-2263.

236. Stringer, J. (1985) Recombination between poly [d(GT).d(CA)] sequences in simian virus 40-infected cultured cells. Mol. Cell. Biol., **5**:1247-1259.

237. Slightom, J.L., Blechl, A.E. and Smithies, O. (1980) Human $^{G}$g- and $^{A}$g-globin

genes: complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. Cell, **21**:627-638.

238. Spitzner, J.R., Chung, I.K. and Muller, M.J. (1990) Eukaryotic topoisomerase II preferentially cleaves alternating purine-pyrimidine repeats. Nucleic Acids Res., **18**:1-11.

239. Latchman, D. (1990) Gene Regulation, a eukaryotic perspective. Unwin Hyman, London.

240. Hyde, J.E. and Sims, P.F.G. (1987) Anomalous dinucleotide frequencies in both coding and non-coding regions from the genome of the human malaria parasite Plasmodium falciparum. Gene, **61**:177-187.

241. Parsons, M, Alexander, K., Hill, T., Neilsen, B., Dovey, H.F. and Wang, C.C. (1990) Glycosomal biogenesis and glucose transport in Trypanosoma brucei. In "Parasites: Molecular biology, drug and vaccine design, pp247-261. Wiley-Liss, Inc.

242. Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol., **120**:97-120.

243. McLean, L.R., Hagaman, K.A., Owen, T.J. and Krstenansky, J.L. (1991) Minimal

peptide length for interaction of amphipathic a-helical peptides with phosphatidylcholine liposomes. Biochem., **30**:31-37.

244. Lodish, H.F. (1988) Multi-spanning membrane proteins: How accurate are the models. Trends Biochem. Sci., **13**:332-334.

245. Sasatsu, M., Misra, T.K., Chu, L., Ladagu, R. and Silver, S. (1985) Cloning and DNA sequence of a plasmid-determined citrate utilization system in Escherichia coli. J. Bact., **164**983-993.

246. Bun-ya, M., Nishimura, M., Harashima, S. and Oshima, Y. (1991) The **PHO84** gene of Saccharomyces cerevisiae encodes an inorganic phosphate transporter. Mol. Cell. Biol., **11**:3229-3238.

247. Walmsley, A.R. (1988) The dynamics of the glucose transporter. Trends Biochem. Sci., **13**:226-231.

248. Pettigrew, G.W. (1972) The amino acid sequence of cytochrome c from a protozoan Crithidia oncopleti. FEBS Letts., **22**:64-66.

249. Benne, R. (1985) Mitochondrial genes in trypanosomes. Trends Genets., **1**:117-121.

250. Van Steveninck, J. (1972) Transport and transport-associated phosphorylation of galactose in Saccharomyces cerevisiae. Biochim. Biophys. Acta, **274**:575-583.

251. Rogers, S., Wells, R. and Rechsteiner, M. (1986) Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science, **234**:364-368.

252. Gupta, M. and Coffino, P. (1985) Mouse ornithine decarboxylase: complete amino acid sequence deduced from cDNA. J. Biol. Chem., **260**:2941-2944.

253. Faust, J.R., Luskey, K.L., Chin, D.J., Goldstein, J.L. and Brown, M.S. (1982) Regulation of synthesis and degradation of 3-hydroxy-3-methylglutamyl-coenzyme A-reductase by low density lipoprotein and 25-hydroxycholesterol in UT-1 cells. Proc. Natl. Acad. Sci. USA, **79**:5205-5209.

254. Ramsay, G., Evan, G.I. and Bishop, J.M. (1984) The protein encoded by the human oncogene c-myc. Proc. Natl. Acad. Sci. USA, **81**:7742-7746.

255. Mercier, J-C., Grosclaude, F. and Ribadeau-Dumas, B. (1971) Structure primaire de la caseine $a_{s1}$-bovine: sequence complete. Eur. J. Biochem., **23**:41-51.

256. Hershey, H.P., Barker, R.F., Idler, K.B., Lissemore, J.L. and Quail, P.H. (1985) Analysis of cloned cDNA and genomic sequences for phytochrome: complete amino acid

sequences for two gene products expressed in etiolated Avena. Nucleic Acids Res., **13**:8543-8559.

257. Catterall, W.A. (1988) Structure and function of voltage-sensitive ion channels. Science, **242**:50-61.

258. Noda, M., Shimizu, S., Tanabe, T., Takai, T., Kayano, T, Ikeda, T., Takahashi, H., Nakayama, H., Kanaoka, Y., Minamino, N., Kangawa, K., Matsuo, H., Raftery, M.A., Hirose, T., Inayama, S., Hayashida, H., Miyata, T. and Numa, S. (1984) Primary structure of Electrophorus electricus sodium channel deduced from cDNA sequence. Nature, **312**:121-127.

259. Lindstrom, J., Criado, M., Hochschwender, S., Fox, J.L. and Sarin, V. (1984) Immunochemical tests of acetylcholine receptor subunit models. Nature, **311**:573-575.

260. Young, E.F., Ralston, E., Blake, J., Ramachandran, J., Hall, Z.W. and Stroud, R.M. (1985) Topological mapping of acetylcholine receptor: Evidence for a model with five transmembrane segments and a cytoplasmic COOH-terminal peptide. Proc. Natl. Acad. Sci. USA, **82**:626-630.

261. Peerce, B.E. and Wright, E.M. (1984) Conformational changes in the intestinal brush-border sodium-glucose cotransporter labeled with fluorescein isothiocynate. Proc.

Natl. Acad. Sci. USA, **81**:2223-2227.

262. Peerce, B.E. and Wright, E.M. (1987) Examination of the Na+-induced conformational change of the intestinal brush-border sodium/glucose symporter using fluorescent probes. Biochem., **26**:4272-4279.

263. Swinkels, B.B., Evans, R. and Borst, P. (1988) The topogenic signal of the glycosomal (microbody) phosphoglycerate kinase of Crithidia fasciculata resides in a carboxy-terminal extension. EMBO J., **7**:1159-1165.

264. Gould, S.J., Keller, G.-A., Subramani, S. (1988) Identification of peroxisomal targeting signals located at the carboxy terminus of four peroxisomal proteins. J. Cell Biol., **107**:897-905.

265. Miyazawa, S., Osumi, T., Hashimoto, T., Ohno, K., Miura, S. and Fujiki, Y. (1989) Peroxisome targeting signal of rat liver acyl-coA oxidase resides at the carboxy terminus. Mol. Cell. Biol., **9**:83-91.

266. Osumi, T., Ishii, N., Hijikata, M., Kamigo, K., Ozasa, H., Furuta, S., Miyazawa, S., Kondo, K., Inoue, K., Kagamiyama, H. and Hashimoto, T. (1985) Molecular cloning and nucleotide sequence of the cDNA for rat peroxisomal enoyl-coA hydratase:3-hydroxyacyl-coA dehydrogenase bifunctional enzyme. J. Biol. Chem., **260**:8905-8910.

267. Smith, S.M. and Leaver, C.J. (1986) Glyoxysomal malate synthase of cucumber: molecular cloning of a cDNA and regulation of enzyme synthesis during germination. Plant Physiol., **81**:762-767.

268. Gould, S.J., Keller, G.-A. and Subramani, S. (1987) Identification of a peroxisomal targeting signal at the carboxy terminus of firefly luciferase.

269. Baltz et al. (1991) Unpublished.

270. Seol, W. and Shatkin, A.J. (1991) Escherichia coli **kgtp** encodes an a-ketoglutarate transporter. Proc. Natl. Acad. Sci., USA, **88**:3802-3806.

271. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning: a laboratory manual, 2nd edition, **vols 1-3**. Cold Spring Harbour Laboratory Press.

272. Berger, S.L. and Kimmel, A.R. (1987) Guide to molecular cloning techniques. In Methods in Enzymology, **152**, Academic Press, Inc. London.

273. Davis, L.G., Dibner, M.D. and Battey, J.F. (1986) Basic Methods in Molecular Biology, Elsevier Science Publishing C., Inc., New York.

274. Kriegler, M. (1990) Gene transfer and expression: A laboratory manual, Stockton

Press, New York.

275. Bothwell, A., Yancopoulos, G.D. and Alt, F.W. (1990) Methods for cloning and analysis of eukaryotic genes. Jones and Bartlett Publishers, Boston.

276. von Heijne, G. (1987) Sequence analysis in molecular biology, Treasure trove or trivial pursuit. Academic Press Inc. London.

277. Morel, C.M. (1984) Genes and Antigens of Parasites: a laboratory manual, 2nd ed. The Proceedings of a course sponsored by the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases, Fundacao Oswaldo Cruz, Brazil.

278. Koslowsky, D.J., Riley, G.R., Feagin, J.E. and Stuart, K. (1992) Guide RNAs for transcripts with developmentally regulated RNA editing are present in both life cycle stages of Trypanosoma brucei. Mol. Cell. Biol., **12**:2043-2049.