Biogenesis of peroxisomes and glycosomes: trypanosomatid glycosome assembly is a promising new drug target

Juliette Moyersoen a,1, Jungwoo Choe b,2, Erkang Fan b, Wim G.J. Hol b,c, Paul A.M. Michels a,*

a Research Unit for Tropical Diseases, Christian de Duve Institute of Cellular Pathology and Laboratory of Biochemistry, Université Catholique de Louvain, ICP-TROP 74.39, Avenue Hippocrate 74, B-1200 Brussels, Belgium
b Department of Biochemistry and Department of Biological Structure, Biomolecular Structure Center, University of Washington, P.O. Box 357742, 98195-7742 Seattle, USA
c Howard Hughes Medical Institute, University of Washington, Seattle, USA

Received 9 March 2004; received in revised form 14 June 2004; accepted 15 June 2004
First published online 21 July 2004

Abstract

In trypanosomatids (Trypanosoma and Leishmania), protozoa responsible for serious diseases of mankind in tropical and subtropical countries, core carbohydrate metabolism including glycolysis is compartmentalized in peculiar peroxisomes called glycosomes. Proper biogenesis of these organelles and the correct sequestering of glycolytic enzymes are essential to these parasites. Biogenesis of glycosomes in trypanosomatids and that of peroxisomes in other eukaryotes, including the human host, occur via homologous processes involving proteins called peroxins, which exert their function through multiple, transient interactions with each other. Decreased expression of peroxins leads to death of trypanosomes. Peroxins show only a low level of sequence conservation. Therefore, it seems feasible to design compounds that will prevent interactions of proteins involved in biogenesis of trypanosomatid glycosomes without interfering with peroxisome formation in the human host cells. Such compounds would be suitable as lead drugs against trypanosomatid-borne diseases.

Keywords: Trypanosomatids; Glycosome and peroxisome biogenesis; Peroxin; Protein–protein interaction; Drug design

Contents

1. Introduction .......................................................................................................................... 604
2. Glycolysis in trypanosomatids. ............................................................................................ 605
3. Glycosomes, structure and function. .................................................................................... 606
4. Proper compartmentation of glycosomal enzymes and intact glycosomes are vital to trypanosomes. ........................................................................................................... 607
5. Overview of peroxisome biogenesis .................................................................................... 610

* Corresponding author. Tel.: +32 2 7647473; fax: +32 2 7626853.
E-mail address: michels@bchm.ucl.ac.be (P.A.M. Michels).

© 2004 Federation of European Microbiological Societies. Published by Elsevier B.V. All rights reserved.

0168-6445/$22.00 © 2004 Federation of European Microbiological Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.femsre.2004.06.004
Trypanosomatids are parasitic organisms belonging to a higher taxonomic group of protists – usually assigned the rank of an order – called Kinetoplastida that probably diverged early in evolution from the other eukaryotic lineages. It is one of the several major lineages that separated from each other in the early, as yet unresolved, diversification of eukaryotes [1,2]. The Kinetoplastida are characterized by a number of peculiarities such as the presence of a special part of the single mitochondrion, where DNA with a highly unusual structure is found, called the kinetoplast, hence the name of this group of organisms; an unusual, often highly complex form of editing of the mitochondrial RNA; the organization of nuclear genes in very long multicistronic transcription units; precursor RNA molecules are processed to mRNAs by trans-splicing, probably while transcription continues downstream; the apparent absence of specific RNA polymerase II promoters; and a unique way of metabolic compartmentation [3,4]. The most peculiar aspect of the metabolic compartmentation in Kinetoplastida is the sequestering of part of the glycolytic pathway and some other enzyme systems within peroxisome-like organelles called glycosomes [5–9].

Several species within the kinetoplastid family of the parasitic Trypanosomatidae are responsible for serious, but largely neglected diseases of humans and domestic animals, particularly in tropical and subtropical areas of the world. In sub-Saharan Africa, two subspecies of *Trypanosoma brucei* (*Trypanosoma brucei gambiense* and *Trypanosoma brucei rhodesiense*) cause sleeping sickness or human African trypanosomiasis. These extracellularly multiplying parasites are transmitted from animal reservoirs to human beings by the tsetse fly. When introduced into the host by the bite of the insect taking its blood meal, the trypanosomes proliferate initially in the blood and lymphatic systems, before invading the central nervous system. *T. b. gambiense* causes a chronic disease that takes several years to progress to the meningoencephalitic stage. In contrast, infection by *T. b. rhodesiense* leads to an acute form that takes only a few weeks to reach the second stage. Without appropriate treatment, both forms of the disease are fatal. The World Health Organization estimates that nearly 60 million people in 36 of Africa’s 52 countries are at risk. An estimated half a million people are probably infected. Unfortunately, current treatments are largely unsatisfactory and the situation is even worsening [10,11]. Chemotherapy is the mainstay for control because effective vaccines are not available and antigenic variation makes the prospects for vaccination gloomy. Nevertheless, most currently used drugs are inefficient and toxic, and the development and spreading of drug-resistant parasites are becoming a major problem [12]. Furthermore, several *Trypanosoma* species (*Trypanosoma brucei brucei*, *Trypanosoma vivax* and *Trypanosoma congolense*) are responsible for cattle trypanosomiasis also called ‘nagana’. This disease has an important impact on the development of rural Africa by reducing cattle production and the availability of animal traction for farming.

In Central and South America, *Trypanosoma cruzi* causes Chagas’ disease. About 100 million people are at risk, an estimated 11–18 million people are infected and annually 13,000 deaths are reported. A large range of wild and domestic mammals act as a reservoir for the parasites, which are transmitted by a blood-sucking
sandfly, which deposits its infective faeces on the skin. After entry in human beings, *T. cruzi* invade host cells. After an initial, acute stage, usually a balance between host and parasite is established in which patients are asymptomatic. A chronic form of the disease often develops after 10–25 years, causing irreversible damage to the heart, digestive or nervous systems. Patients become progressively more ill and ultimately die, often by sudden heart failure. Also the available drugs for Chagas’ disease may induce serious side effects and lack high efficacy. Although success has been achieved in reducing transmission due to vector control programs, the large number of people still infected, and the limitations of the available drugs make the development of new medicines urgent [10,13].

At least 17 different species of the trypanosomatid genus *Leishmania* are pathogenic for humans. These parasites cause a spectrum of diseases, called leishmaniases, ranging from cutaneous and mucocutaneous lesions to life-threatening visceral diseases. About 350 million people in 88 countries worldwide, but mainly in the tropics and subtropics are at risk, with 12 million people affected and approximately two million new cases each year. In recent years, co-infection of HIV patients with *Leishmania* has also become a serious problem. The parasites are transmitted between human beings and animal reservoirs by the bite of a phlebotomine sandfly. Inside the mammalian host, *Leishmania* cells live within phagolysosomes of macrophages. Although new drugs against leishmaniasis have become available in recent years, the chemotherapy currently available is far from satisfactory [14]. Older drugs are toxic and less efficient, and resistance is spreading. Recently introduced drugs overcome these problems but are highly expensive or difficult to administer. A continued effort to develop efficient, safe and affordable drugs remains essential.

The deplorable situation with regard to treatment of these diseases should be attributed to several causes. First, vaccine development is difficult because of the intricate mechanisms used by the parasites to evade the host’s immune defense mechanisms; African trypanosomes employ a highly elaborate system of antigenic variation (reviewed in [15]). American trypanosomes and *Leishmania* species hide intracellularly (reviewed in [16]). But also very importantly, these diseases affect mainly the poorest people in the world. There is little financial incentive for pharmaceutical companies to invest in research and trials for new drugs against these tropical diseases. Fortunately, the many peculiarities in the cell biology and metabolism of kinetoplastids have been subject of intense research. This has led to the identification of potential targets — enzymes or subcellular structures — for new trypanocidal drugs. The potential targets identified are unique to the parasites, or at least sufficiently different from corresponding host molecules or structures, rendering feasible the development of compounds that may selectively interfere with their function in trypanosomatids without affecting a corresponding host function (reviewed in [17–19]). Moreover, several of these peculiar traits have been validated as drug targets; they proved to be essential for the survival of the parasites.

In this review, we will discuss the structure, function and biogenesis of glycosomes in the trypanosomatids, and compare the biogenesis of glycosomes with that of the related peroxisomes in other organisms. We provide arguments that interference with biogenesis offers good prospects for the development of selective trypanocidal drugs.

## 2. Glycolysis in trypanosomatids

Glycolysis has been perceived as a potentially excellent target for the design of new anti-trypanosomatid drugs [6,19–23]. This is particularly the case for African trypanosomes which are entirely dependent on glycolysis for their ATP supply when living in the mammalian bloodstream. The mitochondrion is poorly developed during this stage of the life cycle; the organelle has no enzymes of the tricarboxylic acid cycle and no cytochrome-containing respiratory chain. Glucose, abundantly present in the blood, is taken up by the parasites and metabolized into pyruvate that is excreted. Indeed, bloodstream-form trypanosomes starved at 37 °C without glucose, or incubated with an inhibitor of the plasma-membrane glucose transporter such as phloretin, die within several minutes [24]. The trypanosomes are also killed when incubated with pentenalac-tone, a specific inhibitor of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) [25], or α-cyanocinnamic acid derivatives, which specifically inhibit transporters of monocarboxylic acids such as pyruvate [26]. However, these inhibitors are not selective for the enzymes of the parasite. Therefore, structure-based drug design is being employed to develop compounds that only inhibit trypanosomatid glycolytic enzymes without having any effect on corresponding human enzymes. Indeed, adenosine derivatives were designed that compete with NAD + binding by the trypanosomatid GAPDH without affecting the human enzyme, even at 1000-fold higher concentration. Within minutes after adding these compounds to cultured bloodstream-form *T. brucei*, production of glucose-derived pyruvate ceased, parasite motility was lost, and lysis of parasites was observed [27,28]. Similarly, the designed compounds stunted the growth of intracellular *T. cruzi* without having any effect on growth of cultured human fibroblasts. Structure- and catalytic-mechanism-based approaches are also applied to design selective inhibitors of other glycolytic enzymes.
of trypanosomes. The progress in this research has been summarized in a recent review [19]. Whereas glycolysis is the only ATP-generating process in the bloodstream-form of African trypanosomes, this is not the case for the mammalian-intracellular stage (amastigote) of *Leishmania* spp. *Leishmania* amastigotes have a functional mitochondrion with enzymes of the tricarboxylic acid cycle and a respiratory chain [29–32]. Fatty acids and amino acids may play a more important role in the free energy supply of these cells [33]. However, a complete set of enzymes for carbohydrate metabolism is present [30–32,34] and indeed glucose catabolism has been measured albeit at a lower rate than in bloodstream-form *T. brucei* [35] or cultured insect-stage (promastigote) *Leishmania* cells. Moreover, both the presence of glucose and the possession of a functional glucose transporter have been shown to be essential for cultured *Leishmania mexicana* amastigotes [36].

Not much information is as yet available about the energy and carbohydrate metabolism of *T. cruzi* amastigotes. However, it has been shown that cultured amastigote-like forms readily ferment glucose to succinate, have a high activity of some glycolytic enzymes, a low activity of tricarboxylic acid cycle enzymes and a low cytochrome content [37]. The insect stages of all human-pathogenic trypanosomatids (*T. brucei*, *T. cruzi* and *Leishmania* spp.) seem to be capable of metabolizing both amino acids and sugars for their free energy supply; they have a large repertoire of enzymes for carbohydrate metabolism, including the glycolytic pathway, and a well-developed mitochondrion with a respiratory-chain-linked system for oxidative phosphorylation (reviewed in [9,38,39]).

3. Glycosomes, structure and function

Whereas in almost all organisms studied glycolysis is a process that occurs in the cytosol, in Kinetoplastida the major part of the pathway is localized in organelles called glycosomes. The glycosomes of bloodstream-form *T. brucei* contain the enzymes converting glucose into 3-phosphoglycerate; only the last three enzymes of the glycolytic pathway are present in the cytosol (Fig. 1) [5–9]. A consequence of this organization is that, inside the organelle, the consumption of ATP by hexokinase (HXK) and phosphofructokinase (PFK) equals the organelle, the consumption of ATP by hexokinase (HXK) and phosphofructokinase (PFK) equals the

energy and carbohydrate metabolism of *T. cruzi* is present in the cytosol (Fig. 1) [5–9]. A consequence of this organization is that, inside the organelle, the consumption of ATP by hexokinase (HXK) and phosphofructokinase (PFK) equals the organelle, the consumption of ATP by hexokinase (HXK) and phosphofructokinase (PFK) equals the

Downloaded from https://academic.oup.com/femsre/article-abstract/28/5/603/777468 by guest on 02 March 2019

of trypanosomes. The progress in this research has been summarized in a recent review [19]. Whereas glycolysis is the only ATP-generating process in the bloodstream-form of African trypanosomes, this is not the case for the mammalian-intracellular stage (amastigote) of *Leishmania* spp. *Leishmania* amastigotes have a functional mitochondrion with enzymes of the tricarboxylic acid cycle and a respiratory chain [29–32]. Fatty acids and amino acids may play a more important role in the free energy supply of these cells [33]. However, a complete set of enzymes for carbohydrate metabolism is present [30–32,34] and indeed glucose catabolism has been measured albeit at a lower rate than in bloodstream-form *T. brucei* [35] or cultured insect-stage (promastigote) *Leishmania* cells. Moreover, both the presence of glucose and the possession of a functional glucose transporter have been shown to be essential for cultured *Leishmania mexicana* amastigotes [36].

Not much information is as yet available about the energy and carbohydrate metabolism of *T. cruzi* amastigotes. However, it has been shown that cultured amastigote-like forms readily ferment glucose to succinate, have a high activity of some glycolytic enzymes, a low activity of tricarboxylic acid cycle enzymes and a low cytochrome content [37]. The insect stages of all human-pathogenic trypanosomatids (*T. brucei*, *T. cruzi* and *Leishmania* spp.) seem to be capable of metabolizing both amino acids and sugars for their free energy supply; they have a large repertoire of enzymes for carbohydrate metabolism, including the glycolytic pathway, and a well-developed mitochondrion with a respiratory-chain-linked system for oxidative phosphorylation (reviewed in [9,38,39]).

3. Glycosomes, structure and function

Whereas in almost all organisms studied glycolysis is a process that occurs in the cytosol, in Kinetoplastida the major part of the pathway is localized in organelles called glycosomes. The glycosomes of bloodstream-form *T. brucei* contain the enzymes converting glucose into 3-phosphoglycerate; only the last three enzymes of the glycolytic pathway are present in the cytosol (Fig. 1) [5–9]. A consequence of this organization is that, inside the organelle, the consumption of ATP by hexokinase (HXK) and phosphofructokinase (PFK) equals the ATP production by phosphoglycerate kinase (PGK). Net ATP production occurs only in the cytosol, in the reaction catalyzed by pyruvate kinase (PYK). Similarly, the NADH formed in the reaction catalyzed by GAPDH is re-oxidized inside the organelle by a NAD\(^+\)-dependent glycerol-3-phosphate dehydrogenase (GPDH). Subsequently, the electrons are transferred to a mitochondrial glycerol-3-phosphate oxidase via a redox shuttle comprising a putative transporter in the glycosomal membrane which exchanges glycerol 3-phosphate for dihydroxyacetone phosphate (Fig. 1). The transporter remains to be identified, but its existence is inferred from the low permeability of the glycosomal membrane (see below) and the requirement for strict coupling of the fluxes by which the two triosephosphates are exchanged [5,40]. The glycerol-3-phosphate oxidase complex contains a FAD-linked GPDH, ubiquinone and a terminal oxidase, known as the alternative oxidase, that is insensitive to cyanide, but can be inhibited by salicylhydroxamic acid. The redox process catalyzed by the glycerol-3-phosphate oxidase is nonprotonmotive and thus not coupled to ATP formation.

As described above, amastigotes of *Leishmania* spp. and *T. cruzi*, and insect-stage form parasites (called ‘procyclic’ form in *T. brucei*, ‘epimastigote’ form in *T. cruzi*, and ‘promastigote’ form in *Leishmania*) have a more elaborate energy- and carbohydrate-metabolic network. In these cells, other ATP-dependent kinases (e.g., phosphoenolpyruvate carboxykinase and pyruvate phosphate dikinase) may be found in the glycosome, whereas PGK is relocated to the cytosol. These alterations have as apparent consequence that the ATP/ADP balance is maintained. Similarly, the presence of other NAD\(^+\)-dependent oxidoreductases (malate dehydrogenase and fumarate reductase) in the glycosomal matrix of these cells is accompanied by a significant drop in the activity of GPDH, suggesting that also the intraglycosomal production and consumption of NADH remain balanced. Although the enzyme distribution and stoichiometric relationship of metabolites have not yet been studied in much detail in other kinetoplastids than *T. brucei*, the available evidence suggests that the intraglycosomal balance in formation and consumption of both ATP and NADH is a general feature of these organisms and thus possibly important for proper functioning of the organelle.

In bloodstream-form *T. brucei*, glycolysis is virtually the only process occurring in the glycosomes. Glycolytic enzymes may comprise up to 90% of the matrix proteins of the organelles in these cells. In contrast, in other trypanosomatids, or other *T. brucei* life-cycle stages, the concentration of glycolytic enzymes in glycosomes is less important. Other metabolic systems that – at least in part – may be found in the glycosomes are the pentose-phosphate and gluconeogenic pathways, enzymes of purine salvage and pyrimidine biosynthesis, β-oxidation of fatty acids, ether-lipid biosynthesis, isoprenoid biosynthesis and oxidant defense mechanisms [6,8,9]. Despite the presence of this unique variety of metabolic systems in glycosomes, these organelles should be considered as authentic peroxisomes on the basis of the following criteria: (1) glycosomes, like other members of the peroxisome family (peroxisomes, yeast microbodies and plant glyoxysomes), are bounded by a single phospholipid bilayer
membrane, have an electron-dense proteinaceous matrix and do not contain any detectable DNA; (2) although glycolytic enzymes are the most prominent proteins in glycosomes, some other enzymes or enzyme systems are shared with peroxisomes, notably enzymes of peroxide metabolism, fatty-acid oxidation and ether-lipid biosynthesis; (3) glycosomes and other members of the peroxisome family have a similar process of biogenesis involving several cascades of proteins which are homologous in the different organelles [7,41,42]. The process of biogenesis is discussed below. It may be concluded, therefore, that the common ancestor of Kinetoplastida and other eukaryotes already contained peroxisomes and that an ancestral organism in the kinetoplastid lineage additionally acquired the capability of compartmentalizing the major part of the glycolytic pathway in these organelles. How this may have happened has been discussed elsewhere [9,43,44].

Although the presence of glycolysis in peroxisomes is unique, it should be appreciated that the enzymic content of the peroxisomal family may differ considerably between different organisms. Indeed, the astounding metabolic diversity of the members of the peroxisome family has been already identified many years ago and reviewed by de Duve [45] and Borst [46].

4. Proper compartmentation of glycosomal enzymes and intact glycosomes are vital to trypanosomes

As described in Section 2, glycolysis is essential for trypanosomes and is considered as a validated drug target. One could wonder if proper compartmentation of glycolytic enzymes inside glycosomes and possession of intact organelles are equally essential. If so, compounds that will interfere with glycosome biogenesis or integrity may offer an additional possibility for therapy against trypanosomatid-borne diseases.

Indeed, several experiments have provided evidence that mislocalization of glycolytic enzymes is deleterious for the trypanosome. Trypanosomes have two distinct,
but highly similar isoenzymes of PGK, a glycosomal one (gPGK) only expressed in the bloodstream form and a cytosolic one (cPGK) only found in the procyclic trypanosomes growing in the gut of tsetse flies. The signal for glycosome-targeting of gPGK has been located within the C-terminal six residues [47,48]. Bloodstream-form trypanosomes are rapidly killed when a transgene of cPGK or a C-terminally truncated form of gPGK is expressed [49]. The toxicity depends both on enzyme activity – expression of a mutant PGK with a single residue substitution, rendering the enzyme inactive, has no effect – and cytosolic location. Expression of a transgene of gPGK in procyclic trypanosomes had a less dramatic effect; nonetheless, the growth rate was slowed down and the morphology of the cells was changed. Also, expression of cytosolic triosephosphate isomerase (TIM), after transfection of bloodstream-form T. brucei with the gene of the Saccharomyces cerevisiae enzyme, inhibited cell growth [50].

To understand why the expression of an additional copy of a glycolytic enzyme in the wrong compartment is detrimental for trypanosomes, one has to know the function or consequences of this form of metabolic compartmentation. For years, it was thought that the confinement of glycolytic enzymes inside an organelle was an evolutionary adaptation of trypanosomatids to sustain a high glycolytic flux [6]. All glycosomes together constitute about 4% of the total cellular volume of T. brucei, and the enzymes are present at a relatively high concentration. It was therefore assumed that the sequestering of enzymes overcomes a diffusion limitation of glycolytic metabolites and thus enables a high flux. However, this notion is now considered unlikely. First, a relatively high glycolytic flux is only characteristic of the bloodstream-form T. brucei, not of most other trypanosomatids living in environments where sugars are less abundant. There is no reason to assume that the capacity for a high glycolytic flux must have been a selective advantage for the ancestral kinetoplastids [51,52]. Second, it was shown that S. cerevisiae can have a 2-fold higher flux than T. brucei, with a less than 2-fold higher glycolytic protein content that is not compartmentalized [53]. Third, calculations have shown that, even if the enzymes were dispersed over the entire cytosol, the glycolytic flux of bloodstream-form T. brucei would not be limited by diffusion of metabolites from one enzyme to the next, but rather by the catalytic activity of the enzymes [52]. Computer simulation of metabolism of bloodstream-form T. brucei has, however, provided a new clue about the possible function of the glycosome and a possible explanation why proper compartmentation of enzymes may be essential. The explanation provided by this computer analysis is dependent on two premises: first that the glycosomal membrane has a poor permeability for most glycolytic intermediates, adenine nucleotides and nicotinamide adenine nucleotides, and second that the glycolytic enzymes inside the glycosomes lack the usual mechanisms of activity regulation.

Several studies have provided strong evidence that the glycosomal and cytosolic compartments are indeed separated by a permeability barrier. First, the kinetics of incorporation of radioactivity in glycolytic intermediates, upon a pulse of radioactively labelled glucose added to trypanosomes, indicated the existence of two pools of intermediates: a rapidly labelled pool, representing 20–30% of the total cellular metabolites, and a slowly labelled pool representing 70–80% of the total. The rapidly labelled pool is directly involved in glycolysis and was interpreted as being present in the glycosomes, whereas the second pool, in the cytosolic compartment, is not directly involved in glycolysis. Exchange between the two pools occurs only very slowly [54]. Second, under anaerobic conditions, bloodstream-form trypanosomes produce one molecule of pyruvate and one molecule of glycerol per molecule of glucose, instead of two molecules of pyruvate [55]. This is due to the presence of a glycerol kinase present in the glycosome. It catalyses the thermodynamically unfavourable ‘reverse’ reaction with ADP and glycerol 3-phosphate as substrates (Fig. 1). However, this would only be feasible at high concentrations of these substrates. Such conditions would be difficult to attain in the entire cell, but can occur in a small compartment, such as the glycosome, with a membrane impermeable for the substrates. The consequence of this reverse reaction is that glycolysis can proceed under anaerobic reactions, while maintaining the ATP/ADP and NADH/NAD+ balances within the glycosome, but with a net production of only one molecule of ATP per molecule of glucose instead of two ATP under aerobic conditions.

The notion that the glycosomal membrane is poorly permeable is in line with current ideas about the membrane of mammalian and yeast peroxisomes. Although peroxisomes were, for many years, considered leaky for low-molecular weight compounds, a considerable amount of evidence is now available supporting the view that peroxisomes form closed compartments under in vivo conditions (reviewed in [56]). This evidence includes genetic data showing the need for shuttle systems for redox equivalents [57]. The notion that peroxisomes are surrounded by a membrane that is poorly permeable for most metabolites implies that specific transporter molecules should exist for the transmembrane movement of metabolites. Indeed, two classes of transporters in the peroxisomal membrane have been identified: ABC transporters (reviewed in [58]), some of which have been shown to be involved in transport of fatty acids in S. cerevisiae [59] and have
possibly similar functions in other organisms, and transporters homologous to the mitochondrial solute carriers or AAC transporters. Recently, homologues of the ABC transporters involved in fatty-acid transport have been identified in the glycosomal membrane of procyclic T. brucei (C. Yernaux and PM, unpublished).

It should be noted that some researchers working on plant peroxisomes claim that compartmentalization of plant peroxisomal metabolism is in major parts not caused by a boundary function of the membrane but is primarily due to the specific structure of the protein matrix. Transfer of metabolites across the membrane would proceed through relatively unspecific porin-like channels (reviewed in [60]). However, this claim is not supported by available data for peroxisomes from other organisms [56–59].

With regard to the second premise made for the computer analysis of glycosome metabolism, indeed, the regulation of glycolysis in trypanosomes differs largely from that in other organisms. There is an apparent lack of activity regulation of the trypanosomatid HK and PFK [61–64]. In most organisms, the activities of these two glycolytic enzymes are highly regulated by their products, by metabolites further downstream in the metabolism or by other (allosteric) effectors. This regulation serves two purposes. First, it prevents the loss of ATP by futile cycling when glycolysis and gluconeogenesis occur simultaneously. Second, it has been argued, and experimentally confirmed by using S. cerevisiae mutants, that a tight regulation of the first steps of glycolysis is vital [65]. Indeed absence of regulation may lead to unrestricted accumulation of glycolytic intermediates, a situation that will be highly toxic for the cell.

Using available kinetic data of the different glycolytic enzymes, a mathematical model of glycolysis in bloodstream-form T. brucei was developed [66] and used to assess the metabolic consequences of compartmentation [53]. The question addressed was how the functional behaviour of trypanosome glycolysis would change if the pathway were not compartmentalized. According to the computer model this would not significantly affect the steady-state glycolytic flux, in agreement with the conclusions described above. But strikingly, it would lead to toxic accumulation of hexose phosphates upon addition of glucose. Such toxic accumulation was prevented by the existence of distinct glycosomal and cytosolic adenine nucleotide pools. The kinases in the beginning of the pathway respond to the glycosomal ATP/ADP ratio, not the cytosolic one. The glycosomal ATP/ADP ratio, which – according to the computer simulation – is usually low, controls the activity of HK and PFK and maintains the concentration of the hexose phosphates constrained within a narrow range. If the enzymes would sense the higher cytosolic ATP/ADP ratio, their activity would not be restrained and the products would accumulate. Therefore, the computer analysis supports the notion that the compartmentation of glycolysis serves a regulatory function that compensates for the lack of activity regulation of its enzymes, and that compartmentation of the process within a membrane that is poorly permeable for solutes is essential for the parasite. Would such a function also provide an explanation why the additional presence of a glycosomal glycolytic enzyme in the cytosol is highly detrimental to the cell?

It is essential that a stoichiometric relationship of all metabolites – phosphorylated glycolytic intermediates, adenine nucleotides and nicotinamide adenine nucleotides – participating in the reactions inside the glycosome is maintained, to allow glycolysis to proceed [66]. If such a metabolite would leave the glycosome, it would affect the stoichiometric relationship and cause arrest of glycolysis by accumulation of another metabolite and/or disruption of the ATP/ADP and NADH/NAD+ balance. For some metabolites transport mechanisms should exist across the membrane, for example for glycerol 3-phosphate and dihydroxyacetone phosphate, which are involved in the redox shuttle to the mitochondrion. Furthermore, the pulse-labelling experiments mentioned above [54], in conjunction with measurements of intracellular metabolite concentrations [67] indicated, under steady-state conditions, similar concentrations of glycolytic intermediates in the glycosomal and cytosolic pools, despite the slow equilibration across the membrane. It is conceivable that the partial mislocalization of a glycosomal enzyme to the cytosol would catalyze a reaction between the cytosolic intermediates, causing a drain on glycosomal metabolites that are capable of crossing the membrane and thus disturb the stoichiometry of the intraglycosomal reactions. This effect would even be reinforced if an extraglycosomal shunt of part of the glycolytic pathway could be formed [49,50]. This is feasible, because some enzymes have a constitutively dual subcellular localization: part is in the glycosomes, another part in the cytosol (reviewed in [68]). Trypanosomatids have two different isoenzymes of GAPDH. The metabolic role of the cytosolic enzyme is not known yet, but it may be involved in redox buffering of the cytosol. In addition, the parasites have only one isoenzyme of glucose-6-phosphate isomerase, but part of it is permanently found in the cytosol. Possibly, this cytosolic activity is involved in the pentose-phosphate pathway that also has a dual localization: it may isomerize fructose 6-phosphate, one of the products of the pathway’s activity, into the substrate glucose 6-phosphate. Thus, the incorrect compartmentation of even a low amount of a glycosomal enzyme may thus have a highly detrimental effect on trypanosomes, as most clearly shown for PGK [49].
5. Overview of peroxisome biogenesis

Peroxisomes, like mitochondria and chloroplasts, multiply by growth and division, and segregate at cell division. However, contrary to these other organelles, peroxisomes do not contain any DNA. All their proteins are encoded by nuclear genes, are synthesized in the cytosol on free ribosomes and post-translationally imported into pre-existing organelles. Moreover, the mechanism of peroxisome biogenesis is fundamentally different from that of mitochondria and chloroplasts, as revealed by studies performed during the last two decades, mainly on various yeasts and mammalian cells (reviewed in [69,70]). Conceptually, the biogenesis of peroxisomes can be considered as a number of different processes: recruitment of lipids for membrane formation, synthesis and insertion of membrane proteins, synthesis and import of matrix proteins, and organelle fission and inheritance. To date, 32 proteins, called peroxins (or with the acronym PEX) [71], have been identified as required for the different aspects of peroxisome biogenesis (Table 1). The first peroxins were discovered in yeasts by screens of mutant cells presenting defects in the assembly of peroxisomes [72]. These cells were unable to grow on oleate as unique source of carbon. This substrate is known to induce the proliferation of peroxisomes [73]. Similarly, Chinese Hamster Ovary (CHO) mutant cells affected in plasmalogon synthesis turned out to be also impaired in peroxisome biogenesis [74]. Complementation studies of these two types of mutant cells with genomic and cDNA libraries, respectively, led to the discovery of proteins involved in the biogenesis of the organelles. Subsequently, other peroxins were identified by additional genetic studies, and upon the detection of interactions between known peroxins and other ones, as revealed by yeast or bacterial two-hybrid studies and coprecipitation analyses and afterwards confirmed by in vitro binding studies.

Table 1
Identified peroxins

<table>
<thead>
<tr>
<th>Peroxin</th>
<th>Organisms</th>
<th>Yeasts/fungi</th>
<th>Mammals</th>
<th>Plants</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEX1</td>
<td>Sc</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX2</td>
<td>Y.l.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX3</td>
<td>D.m./C.e.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX4</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX5</td>
<td>X</td>
<td>X</td>
<td>D.m./C.e.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX6</td>
<td>X</td>
<td>X</td>
<td>C.e.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX7</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX8</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX9</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>?</td>
<td>X</td>
</tr>
<tr>
<td>PEX10</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX11</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX12</td>
<td>X</td>
<td>X</td>
<td>C.e.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX13</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX14</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX15</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX16</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX17</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX18</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX19</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>C.e.</td>
<td>X</td>
</tr>
<tr>
<td>PEX20</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX21</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX22</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX23</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX24</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX25</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX26</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX27</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX28</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX29</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX30</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX31</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PEX32</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Peroxin function and Localization:

- Matrix protein import – receptor recycling or vesicle fusion
- Matrix protein import – translocation
- PMP membrane insertion
- Matrix protein import – PTS1 receptor
- Matrix protein import – receptor recycling or vesicle fusion
- Matrix protein import – PTS2 receptor
- Matrix protein import
- Matrix protein import – receptor recycling or vesicle fusion
- Matrix protein import – PTS2 receptor
- Matrix protein import
- Matrix protein import – docking
- Matrix protein import – docking
- Recruitment PEX6 (and PEX1) to peroxisomal membrane
- PMP membrane insertion
- Matrix protein import – docking
- Matrix protein import – docking
- Recruitment PEX6 (and PEX1) to peroxisomal membrane
- PMP membrane insertion
- Matrix protein import – docking
- Recruitment PEX6 (and PEX1) to peroxisomal membrane
- Matrix protein import – docking
- Recruitment PEX6 (and PEX1) to peroxisomal membrane
- PMP membrane insertion
- Recruitment PEX6 (and PEX1) to peroxisomal membrane
- Matrix protein import – docking
- Recruitment PEX6 (and PEX1) to peroxisomal membrane
- PMP membrane insertion
- Recruitment PEX6 (and PEX1) to peroxisomal membrane
- PMP membrane insertion
- Recruitment PEX6 (and PEX1) to peroxisomal membrane
- PMP membrane insertion

Abbreviations: S.c., Saccharomyces cerevisiae; Y.l., Yarrowia lipolytica; D.m., Drosophila melanogaster; C.e., Caenorhabditis elegans; cyt, cytosol; MA, membrane associated; IMP, integral membrane protein.

* Other yeasts/fungi: Pichia pastoris, Hansenula polymorpha and Neurospora crassa.
Results from a large number of genetic, molecular and cell-biological studies show convincingly that the general principles of peroxisome biogenesis have been maintained in all organisms analyzed and are similar for the different members of the organelle family (peroxisomes, microbodies, glyoxysomes, glycosomes). Models for various steps of peroxisome assembly emerged from these studies and will be discussed in the following sections. However, it will also be shown that differences have been found in details of the process and that the molecules (peroxins) involved may differ considerably between species. In some cases, different species seem to use non-homologous proteins to perform corresponding steps in the organelle-assembly process, or homologous proteins could not (yet) be identified. In Section 6, available data about the biogenesis of trypanosomatid glycosomes will be discussed in much detail. The presently known differences between glycosome biogenesis in the parasites and peroxisome biogenesis in humans will be highlighted and the possible exploitation of these differences for the design of trypanocidal drugs will be discussed.

5.1. Import of matrix proteins; receptor recognition

The majority of the peroxins discovered until now are involved in the routing of newly synthesized proteins to the matrix of peroxisomes (Fig. 2). These matrix proteins are synthesized on free ribosomes in the cytosol where they interact with a cytosolic receptor, either PEX5 or PEX7, dependent on their peroxisome-targeting signal (PTS).

The principal peroxisomal matrix protein–import pathway involves the proteins harbouring a peroxisome-targeting signal of type 1 (PTS1) which is recognized by the receptor PEX5. Early studies on PTS1 revealed that the signal is a C-terminal tripeptide with the consensus sequence (S/A/C)-(K/R/H)-L [75]. Later, it was shown that this motif may vary considerably more regarding to its length and sequence, as also the specificity of the signal may vary in different organisms [76–78]. Recently, Neuberger et al. [79] reanalyzed all published data on PTS1–PEX5 interaction and arrived at the conclusion that at least the 12 C-terminal residues of a given protein are implicated in PTS1 recognition. This 12 amino-acid motif can be structurally and functionally divided into three regions: (1) the C-terminal tripeptide, (2) a region interacting with the surface of PEX5 (about four residues further upstream) and (3) a polar-solvent accessible and unstructured region with linker function (the five remaining residues). Gatto et al. [80] studied in great detail the affinities of variant PTS1-containing peptides to investigate the sequence requirement for receptor binding. Their results suggested that there is a threshold affinity of the peptide for the PTS1-binding domain of PEX5. If its affinity

![Fig. 2. General model of peroxisomal matrix protein import. Matrix proteins harbouring a PTS1 or a PTS2 are recognized in the cytosol by the receptors, PEX5 and PEX7, respectively. A cascade of protein–protein interactions occurs in the consecutive steps of peroxisomal protein import which are: docking, translocation, cargo release and receptor recycling. For details, see text.](https://academic.oup.com/femsre/article-abstract/28/5/603/777468)
for the receptor is tighter than this threshold, a given sequence will function as a peroxisome-targeting signal.

The PTS1 receptor, PEX5, is one of the most intensively studied and better understood peroxins. First identified in the yeast *Pichia pastoris* [81], it has then been cloned in numerous organisms: mammals, other yeast species, *Caenorhabditis elegans*, plants [82–90] and the trypanosomatids *T. brucei* and *Leishmania donovani* (see Section 6). Subcellular localization experiments indicated that PEX5 is mostly cytosolic with a smaller proportion associated with peroxisomes [81,83,87,90–93]. But Szilard et al. [86] observed that PEX5 was entirely localized inside the peroxisome of the yeast *Yarrowia lipolytica* and Terlecky et al. [94] found PEX5 tightly associated with the matrix face of the peroxisomal membrane of *P. pastoris*. The different localizations of PEX5 are in agreement with results of recent studies in which PEX5 (and PEX7, another PTS receptor – see below) were found to undergo multiple rounds of entry into the peroxisome lumen. This gave rise to the ‘extended shuttle’ hypothesis in which both receptors actually enter the peroxisome together with their bound cargo molecules and release them in the lumen before shuttling back to the cytosol (see Section 5.4).

The C-terminal half of PEX5 contains a region of 6–7 tetratricopeptide repeats (TPRs), a 34 amino-acid degenerate motif (Fig. 3). These TPRs have been shown to be essential for the interaction between PEX5 and the PTS1 [83,84,94,95]. The three-dimensional structure of human PEX5, complexed with a PTS1-containing pentapeptide, has been resolved by X-ray crystallography unravelling the nature of the interaction between PEX5 and the PTS1 [96] (Fig. 4). In addition, random

![Diagram](https://example.com/diagram.png)

**Fig. 3.** Scheme of known interactions between proteins to be imported into the peroxisomal matrix, their respective receptors and the docking-complex peroxins. Yeast and mammalian PEX5 possess, in their C-terminal part, seven TPR motifs (indicated as white boxes) responsible for the recognition of the PTS1. PTS2 proteins are recognised by PEX7. (a) In mammals, a proline-rich motif (PXXP) in PEX14 seems to be responsible for interaction of this peroxin with the SH3 domain of PEX13. The N-terminal part of PEX13 binds to the WXXXF/Y motifs (indicated as small black bars) no. 2 to 4 of PEX5, whereas other WXXXF/Y motifs are involved in binding the N-terminal part of PEX14. The longer version of PEX5, PEX5L, possesses a putative PEX7 binding box. (b) Also in the yeast *S. cerevisiae*, the proline-rich PXXP sequence of PEX14 allows the binding of this peroxin to the SH3 domain of PEX13. Moreover, this PEX13 SH3 domain interacts also with the first of only two WXXXF/Y motifs present in *S. cerevisiae* PEX5. PEX18 and PEX21 are yeast specific peroxins which bind to PEX7, and are necessary for PTS2 import.
and site-directed mutagenesis of *S. cerevisiae* PEX5 (ScPEX5) confirmed the importance of the TPR domain for PTS1 binding [97]. These studies showed that the seven TPR segments of PEX5 form two clusters (TPR 1–3 and TPR 5–7) with TPR 4 acting as a hinge. The PTS1 sequence binds to the groove between these two TPR clusters.

In mammals, two isoforms of PEX5 are produced by alternative splicing [98,99]. The isoforms are called short-form and long-form of PEX5, PEX5S and PEX5L, respectively. This longer form has a 37 amino-acid insertion located upstream of the TPR domain. It has been shown that PEX5L, but not PEX5S, interacts with PEX7, the cytosolic receptor of PTS2 proteins. In these organisms, PEX7 does not bind directly to PEX14, a protein of the docking complex at the peroxisomal membrane (see also Section 5.2). Therefore, it has been proposed that PEX5L interacts with PEX7 in order to enable the docking of the PTS2 receptor–ligand complex to the membrane of the peroxisome. Thus PEX5 seems in these organisms not only responsible for the PTS1 import pathway but also involved in the PTS2 pathway.

Protein–protein interaction studies demonstrated that PEX5 binds to proteins forming the docking complex at the membrane: PEX13 [92,93,100–105] and PEX14 [101,104–111] (Figs. 2 and 3). The interactions with these two peroxins, PEX13 and PEX14, are mediated via pentapeptides WXXXF/Y localized in the N-terminal domain of PEX5. The WXXXXF/Y motifs are present in two or more copies depending on the species: two in *S. cerevisiae*, seven in human or 11 in tobacco. PEX5 has also been shown to interact with PEX10 and PEX12 [112–114], peroxins involved in the translocation step, and to PEX8 [115] leading to an extensive network of interactions. A scheme of the different mapped regions of PEX5 is shown in Fig. 3. These interactions will be described in more detail in Sections 5.2 and 5.3.

The second import pathway involves the other known cytosolic receptor, PEX7, and proteins harbouring the peroxisomal-targeting signal of type 2 (PTS2). This PTS2 is a relatively rare topogenic signal and the PTS2–PEX7 interaction as well as the PTS2–protein import mechanism are far less understood. The first PTS2 to be identified was the one of rat thiolase [116]. Despite the fact that only a few proteins employ the PTS2 pathway, their import is absolutely essential. In human, an inability to import PTS2 proteins leads to the disease rhizomelic chondrodysplasia punctata that causes death of the patient within two years of life [117,118]. The PTS2 consists of a nonapeptide (R/K)(L/I/V)(L/V/I)X(R/H)(L/A) close to the N-terminus of the protein [119,120]. The five amino-acid linker between the conserved residues is crucial for recognition by the PTS2 receptor. However, it seems that the sequence required for import is somewhat different between species and therefore the above consensus may be an oversimplification [121–123]. In the cytosol, PTS2 proteins are recognized by PEX7. This receptor has been found in mammals and yeasts [119,124–129], but is apparently absent from *C. elegans*. Indeed *C. elegans*, and perhaps even more species, does not utilize this targeting pathway. Every known peroxisomal matrix protein from this organism possesses a PTS1 sequence and is imported via PEX5 [130].

The protein PEX7 consists of six or seven WD repeats, which are approximately 40 amino acids long and contain a central tryptophan (W)-aspartate (D) motif. These repeats, constituting almost the entire protein, are preceded by a distinct N-terminal region of approximately 60 amino acids. The subcellular localization of PEX7 has been described as being in the cytosol but the protein has also been found in the lumen of the peroxisome [119,125,126,129–133]. This intraperoxisomal localization found for PEX7 and PEX5 (see above) gave evidence for the ‘extended shuttle’ model that invokes entry of the receptor–cargo complex into the peroxisomal matrix (see Section 5.4). Little is known about the critical regions of PEX7 involved in protein–protein interactions. However, PEX7 has been shown to interact with several peroxins. These peroxins are PEX14 [106,107,109], PEX13 [134,135] and, only observed in *S. cerevisiae*, both PEX18 and PEX21 [136].
As mentioned above, it has been shown that in human cells PEX7 interacts with the longer isofrom, of PEX5, PEX5L, which has a 37-amino-acid insertion located upstream of the TPR domain, but does not bind to PEX14 [98,99]. Initially, this additional sequence was thought to be responsible for the interaction of PEX5L with PEX7. Later, studies to unravel this interaction demonstrated that a region upstream of the 37 residues is also crucial for the interaction with PEX7. Moreover, the C-terminal 18–21 amino acids of the 37-residues insert can be removed without significantly reducing the ability of PEX5L to mediate PTS2 protein import. However, a sequence of at least eight residues is needed in order to bind PEX7 [99,131]. Recently, two-hybrid analysis of Arabidopsis thaliana peroxins showed that PEX14 interacts with PEX5, but not with PEX7. Instead, PEX7 binds PEX5 indicating that, in Arabidopsis, the PTS2 pathway depends also on the PTS1 pathway. The interaction analysis showed that the binding domain of AtPEX7 was restricted to the very C-terminal region which in Arabidopsis does not include the WD40 repeats. Indeed, the WD repeats were identified as the binding domain for PTS2 proteins. And an N-terminal region of AtPEX5, which does not overlap with the binding domain for AtPEX14, is responsible for the interaction with AtPEX7 [137]. In that perspective, the PEX7 binding partner in Arabidopsis is similar to that of long-form PEX5 in human, even though AtPEX5 is not found as differently spliced isoforms.

PEX8 and PEX21 are two structurally related peroxins involved in PTS2 protein import in S. cerevisiae. Comparison of both protein sequences revealed a weak but significant level of similarity distributed throughout the protein sequences (23% identity). The level of identity is most prominent at the C-terminus (35% for the last 60 amino acids). Absence of both peroxins abolishes the PTS2 branch of peroxisome biogenesis without affecting the PTS1 branch [136]. They are found mostly in the cytosol where they bind PEX7 through their C-terminal regions. It seems that either PEX18 or PEX21 is required prior to the docking event for the formation of a PTS2 protein-containing import-competent complex. So in the cytosol, PEX18 and PEX21 would promote the formation of an oligomeric complex formed by PEX7-PTS2 ligand and PEX18/PEX21 [132,135]. No orthologues of PEX18 or PEX21 have ever been found in higher eukaryotes. PEX7 was restricted to the very terminal of the former [145,146]. Strong support for ‘piggy-backing’ as a natural mechanism, a process that would involve the formation of a heteromeric complex of the protein harbouring no consensus targeting motif with a PTS1- or PTS2-containing protein to enable the import of the former [145,146]. Strong support for ‘piggy-backing’ targeting has been obtained for a yeast enzyme involved in β-oxidation of unsaturated fatty acids (\(\Delta^1\)-cis:\(\Delta^2\)-trans-enoyl-CoA isomerase) that has only a weakly functioning PTS1 but is imported due to interaction with a related protein having an efficient PTS1 [147]. It is therefore conceivable that an I-PTS could serve as an anchor enabling a protein harbouring such a sequence to form a complex with a PTS1 or PTS2 protein.

crass can also substitute for ScPEX18/PEX21 [140]. Therefore, it is likely that YlPEX20 and HsPEX20 replace PEX18 and PEX21 in forming an import-competent complex and function together with PEX7 in PTS2 import. Recently, a 21-amino-acids long peptide motif of mammalian PEX5L, shared by S. cerevisiae PEX18 and PEX21 and Y. lipolytica PEX20, was identified that overlaps the regions shown to be essential for PTS2 protein import [131,139]. Thus HsPEX5L, ScPEX18/ScPEX21 and YlPEX20 all contain a conserved sequence which most likely represents a common PEX7 binding site.

A few proteins lack both PTS1 and PTS2 but are nevertheless targeted to the peroxisome. The first example of such a peroxisomal enzyme studied was Candida tropicalis acyl-CoA oxidase, which was shown to be targeted to the peroxisome by a polypeptide-internal sequence (I-PTS) [141]. Recently, Klein et al. [142] demonstrated the existence of a novel, PEX5-dependent non-PTS1 pathway for S. cerevisiae acyl-CoA oxidase, POX1, another I-PTS protein. POX1 and PEX5 can interact directly with each other without support of other proteins, but the binding involves novel contact sites in both proteins. The binding region in PEX5 is located in a defined area of the N-terminal part of the protein, thus outside the C-terminal TPRs constituting the PTS1-recognition domain, and the interaction site of the POX1 polypeptide is indeed located internally. Internal peroxisome-targeting signals might thus be recognized by the cytosolic receptors, PEX5 or PEX7, or by another yet unknown receptor.

However, a direct interaction between a cytosolic receptor and the I-PTS proteins is not the only import mechanism considered nowadays. In contrast to mitochondria, peroxisomes can import folded and even oligomeric proteins (see Section 5.5). Various studies showed that, for both PTS1 and PTS2 proteins, subunits containing the import signal could assemble with subunits from which the motif had been removed and be transported as multimeric molecules into the peroxisomes of yeast, plant and mammalian cells [121,143,144]. This is suggestive for import by ‘piggy backing’ as a natural mechanism, a process that would involve the formation of a heteromeric complex of the protein harbouring no consensus targeting motif with a PTS1- or PTS2-containing protein to enable the import of the former [145,146]. Strong support for ‘piggy-backing’ targeting has been obtained for a yeast enzyme involved in β-oxidation of unsaturated fatty acids (\(\Delta^1\)-cis:\(\Delta^2\)-trans-enoyl-CoA isomerase) that has only a weakly functioning PTS1 but is imported due to interaction with a related protein having an efficient PTS1 [147]. It is therefore conceivable that an I-PTS could serve as an anchor enabling a protein harbouring such a sequence to form a complex with a PTS1 or PTS2 protein.
5.2. The docking complex for import of matrix proteins

Following the recognition in the cytosol, the receptor with its bound cargo docks on the peroxisomal membrane at a complex formed by PEX14, PEX13 and PEX17; this latter peroxin is only found in yeasts. PEX14 has been cloned from different species: mammals [108–110], various yeast species (S. cerevisiae, P. pastoris and Hansenula polymorpha) [104,106,107,148,149], plants [150,151], T. brucei and L. donovani (see Section 6).

In S. cerevisiae and P. pastoris, PEX14 has been shown to be a peripheral peroxisomal membrane protein tightly bound to the cytosolic face [106,148,149]. Despite the lack of a membrane-spanning domain, PEX14 presents short hydrophobic regions that could possibly interact with membrane proteins and lipids and may therefore be responsible for the protein’s association with the membrane. Contrary to yeast, mammalian PEX14 behaves like an integral membrane protein, with its C-terminal two-thirds largely exposed to the cytosol [109,110,152]. By a very complete protease protection assay, it was shown that a small domain of 16 kDa is embedded in the peroxisomal membrane. So this domain alone is responsible for the fact that this PEX14 displays the characteristics of an intrinsic membrane protein [152]. As in mammals, plant PEX14 behaves as an integral membrane protein having at least two hydrophobic segments and part of the polypeptide located in the cytosol [150,151].

Again, extensive studies have been performed on the interacting partners of PEX14. In addition to its interaction with PEX5 and PEX7, it was shown to bind PEX13 [106,107] and PEX17 [153,154]. Some of these studies also showed that PEX14 interacts with itself [106,107], and would exist predominantly as a dimer in vivo [155]. A predicted coiled-coil region was found in the mammalian and yeast peroxin and may be responsible for dimerization of the protein [108–110,148,149,152].

The interaction of PEX14 with both the PTS1 and PTS2 receptor, PEX5 and PEX7, in yeasts, and with PEX5 in mammals and plants, was shown using two-hybrid and co-immunoprecipitation studies [104–108,110,151]. Mammalian PEX5 possesses multiple high-affinity binding sites for PEX14, which appear to be distributed throughout its N-terminal half. Pentapeptide motifs WXXY/F (present seven times in the protein) were proposed to be responsible for the interaction with PEX14 [101]. Later it was demonstrated that each of these PEX5 motifs are capable of forming a single, high-affinity binding site for the N-terminal PEX5-binding domain of PEX14. Indeed, both aromatic amino acids tryptophan and phenylalanine/tyrosine at positions 1 and 5, respectively, are indispensable for the high affinity interaction with PEX14. The authors hypothesized that each WXXY/F motif, supposed to be part of an amphipathic α-helix, binds directly to a hydrophobic region of PEX14. Anyhow, it was concluded from these studies that the pentapeptide motifs form multiple PEX14-binding sites, in a mutually independent manner [111,115]. In S. cerevisiae and P. pastoris, PEX5 possesses, respectively, two and three WXXY/F motifs. However, in both yeasts, no interaction between PEX5 fragments containing one or more WXXY/F motifs and PEX14 could be detected using overlay and pull-down assays [102–104]. In contrast, some of these motifs have been shown to be essential for the association of PEX5 with PEX13 (see below). In A. thaliana, two distinct regions in the N-terminal half of PEX14, i.e., the segments I58-L65 and R78-R97, interact with PEX5. Therefore, it was proposed that these regions might form a binding pocket for PEX5 [137].

PEX14 has been postulated to be the initial docking site for cargo-loaded PEX5 and PEX7 and would be therefore the point of convergence of the PTS1 and PTS2 pathways. In mammalian cells, overexpression of PEX14 led to the accumulation of PEX5 at the peroxisomal membrane. In contrast, overexpression of PEX13, PEX10 or PEX12 did not cause apparent PEX5 accumulation at the peroxisomal membrane, and import of PTS1 and PTS2 proteins remained unaffected. Moreover, in mutant cells lacking PEX14, PEX5 is found in the cytosol whereas it is found at the membrane in cells lacking PEX13, PEX10 or PEX12 [156]. Later, it was shown that PEX5 loaded with PTS1-containing protein interacted only with PEX14, while unloaded PEX5 bound to PEX13. This implies that PEX13 functions at a step following the detachment of the cargo-protein from the receptor. The authors proposed a model for peroxisomal protein import mediated by the receptor PEX5 and the import machinery comprising the initial PEX5-anchoring site, PEX14, and PEX13 [156].

As mentioned above, PEX14 interacts also with another component of the docking complex, PEX13. This interaction has been well studied in both yeasts (S. cerevisiae and P. pastoris) and mammalian cells. Contrary to PEX14, it has been well established that PEX13 is an integral peroxisomal membrane protein and its termini are both oriented toward the cytosol [134]. PEX13 possesses, in its C-terminal part, a SH3 domain which is involved in protein–protein interactions. In S. cerevisiae, both PEX5 and PEX14 are able to bind PEX13 via its SH3 domain [92,93,100,106]. Competition assays demonstrated that the two ligands can bind simultaneously to this domain [157]. In this organism, PEX14 is the classical SH3 ligand as it presents a PXXP motif in the N-terminal part of the protein [134]. Indeed, the PXXP motif has been shown to interact directly with SH3 domains (reviewed in [158]). Surprisingly, PEX5 does not possess such a PXXP motif in its amino-acid sequence but still interacts directly with...
PEX13. Instead, PEX5 presents an α-helical element that binds to a novel interaction site on the SH3 domain which is distinct from the classical PXXP binding cleft [102]. In *P. pastoris*, one identifiable feature of the SH3 binding site in PEX5 is the presence of three pentapeptide repeat motifs [104]. Bottger et al. [103] reported that the first WXXxF/Y motif of *S. cerevisiae* PEX5 is important for the interaction with PEX13 whereas the second motif does not present any affinity for either PEX13 or PEX14. The residues Trp and Phe of the first pentapeptide, as well as a negatively charged residue outside this motif, were shown to be indispensable for the binding of PEX5 to the SH3 domain of PEX13 [102,103,157]. Later, it was proven that the two ligands, ScPEX14 and ScPEX5, bind in unrelated conformation to patches located at opposite surfaces of this SH3 domain [159]. In human, several PXXP motifs in PEX14 are likely to be responsible for the interaction with PEX13, as reported for *S. cerevisiae*. However, human PEX5 does not bind to the SH3 domain of PEX13 but rather to the N-terminal part of the protein. This interaction occurs via the WXXxF/Y motifs 2–4 of PEX5 [156]. Therefore, it is most likely that PEX5 interacts with PEX13 differently in mammalian and yeast cells.

The third component attributed to the docking complex is PEX17. It is only found in yeast species, *Y. lipolytica*, *S. cerevisiae* and *P. pastoris*, where it is peripherally but tightly associated with the peroxisomal membrane [153,154,160]. In *S. cerevisiae*, PEX17 was shown to interact with PEX14 and therefore proposed to be part of the docking complex [153]. Nevertheless, the presence of PEX17 is not a prerequisite for the docking of PEX5 and PEX7 to the peroxisomal membrane [134]. Based on studies with Δpex17 mutants of *P. pastoris* and *S. cerevisiae* it was suggested that PEX17 may also be involved in the insertion of PMPs into the peroxisomal membrane [153,154,161,162]. However, such mutant cells often displayed only slight effects on PMP abundance and localization, similar to the changes sometimes observed in Δpex5 mutants. Therefore, it is unlikely that PEX17 plays an important role in the insertion of PMPs into the membrane [162].

### 5.3. Import of matrix proteins; cargo–receptor complex translocation and receptor recycling

Following the docking step, the cargo–receptor complex will enter the peroxisome where the receptor will release its cargo protein before being recycled back to the cytosol (Fig. 2). These last processes are far less well understood and questions still arise about many aspects. PEX2, PEX10 and PEX12 are thought to be involved in the translocation step that follows the docking of PEX5 and PEX7 together with their respective ligand on the membrane. The import thus results from a cascade of events ending eventually in the internalization of the matrix proteins. The hypothesis is that PEX2, PEX10 and PEX12 form another protein complex in the membrane distinct from the docking machinery. These three peroxins are integral peroxisomal membrane proteins, possessing a ‘RING finger’ domain exposed to the cytosol [112,113,163–165]. The RING domains of PEX10 and PEX12 enable the interaction between the two proteins as well as with PEX5 [112,113]. In vitro, PEX10 also interacts with PEX2, a peroxin that has been suggested to act downstream of PEX10 and PEX12 [113]. Furthermore, PEX2 is also essential for PTS2 protein import. However, a single substitution in the zinc-binding RING motif results in loss of the PTS1 but not PTS2 import implying that the protein possesses two distinct domains necessary for the PTS1 and PTS2 pathways, respectively [166]. The main reason to group PEX2, PEX10 and PEX12 in the translocation machinery, and not in the docking complex, came from the observation that human cells defective in PEX2, PEX10 and PEX12 are still able to recruit PEX5 to the peroxisomal membrane, presumably through PEX14 [105,112]. Yet, Reguenga et al. [114] demonstrated that PEX2, PEX12, PEX5 and PEX14 are constituents of the same protein assembly. This again shows the complexity of the import mechanism in which each step interconnects with the other in an intimate manner. On the other hand, Dodt and Gould [91] isolated cells from a patient suffering from a disorder caused by a mutation in *PEX12* and found PEX5 accumulating inside the peroxisome. From this observation and later studies, it was concluded that PEX12 actually appears to have a role in recycling PEX5 across the membrane [112]. Furthermore, Eckert and Erdmann [70] state that, based on the available evidence, it is not impossible that PEX14, PEX13 and PEX17 could also act in the translocation event in addition to their role in the docking step.

The identification and characterization of a *T. brucei* and *L. donovani* PEX2 has also been reported, and the importance of this peroxin glycosome biogenesis has been confirmed (see further details in Section 6.2).

PEX4 is a member of the family of ubiquitin-conjugating enzymes and can bind to ubiquitin in vitro [167,168]. In *S. cerevisiae* and *P. pastoris*, deletion of the *PEX4* gene disrupts targeting via both the PTS1 and PTS2 pathways [167,169]. The protein is anchored to the cytosolic face of the peroxisome by the integral membrane protein PEX22. The model predicts that PEX4 and PEX22 act together in the import of peroxisomal matrix proteins [170]. Recent studies strongly suggested that the complex of PEX4 and PEX22 is linked to the protein import apparatus via a unique binding to PEX10 [171]. The exact role of the two peroxins PEX4 and PEX22 in the import process remains unclear. In
P. pastoris Δpex4 and Δpex22 mutants, the abundance of PEX5 is drastically decreased. Although the phenomenon is not understood [170], it has been used in an epistasis analysis which revealed that PEX13, PEX14, PEX10, PEX12, PEX2 and PEX8 function at the initial stage of PTS1 protein import, followed by PEX1 and PEX6 and subsequently PEX22 and PEX4 [172]. It has also been suggested that H. polymorpha PEX4 may be involved in PEX5 recycling to the cytosol [173]. In addition, an entirely different role has been proposed for PEX4. Several studies support the idea that degradation of peroxins occurs in proteasomes and that this process would be dependent on PEX4 and its ubiquitination. Purdue and Lazarow [174] showed that in wild-type S. cerevisiae PEX18 undergoes a rapid degradation with a half-time of less than 10 min. In these cells, mono- and di-ubiquitinated forms of PEX18 could be detected during peroxisome biogenesis. However, a role for PEX4 in this PEX18 ubiquitination remains hypothetical and more experiments will be required to investigate whether this might be the case.

Another peroxin required for both the PTS1 and PTS2 import pathways is PEX8 whose role remains to be elucidated. It is localized at the matrix side of the peroxisomal membrane but nevertheless interacts with PEX5 in S. cerevisiae [115] and with PEX20 in Y. lipolytica [175]. In S. cerevisiae, Agne et al. [176] have defined two core complexes of the peroxisomal import machinery: the docking complex comprising PEX14 and PEX17, with the loosely associated PEX13 on one side and the RING finger complex containing PEX2, PEX10 and PEX12 on the other side. They conclude that PEX8 organizes the formation of a larger import complex from the matrix side of the peroxisome membrane and thus might enable functional communication between both sides of the membrane. But how PEX8 promotes the association of the docking and the RING finger complex from the inner face of the membrane will be a challenging question to answer.

PEX1 and PEX6 are membrane-bound peroxins belonging to the family of AAA (ATPase-associated activity) proteins [177,178] and have been shown to interact with each other in an ATP-dependent manner [179–181]. Their two characteristic conserved AAA domains (denoted D1 and D2), containing Walker ATP-binding sequences, are responsible for the ATPase activity. Studies on the localisation of PEX1 and PEX6 showed striking differences among species. It runs from peroxisomal [181], to peripherally associated with peroxisomes [182], to cytosolic when overexpressed [183,184], and even to be present in vesicular structures distinct from mature peroxisomes [179,185,186]. The role of these peroxins is still a matter of debate. In P. pastoris, PEX1 and PEX6 have been proposed to be involved in matrix protein import. They were shown to be part of the translocation machinery where they act downstream of the docking and translocation steps but upstream of PEX4 and PEX22 [172]. The results of this study suggest, therefore, that these peroxins may play a role in the recycling of PEX5 [187] (Fig. 2). On the other hand, it is also tempting to speculate that these two ATPase peroxins could provide the energy needed for matrix protein import from ATP hydrolysis. These two functions are not necessarily contradictory. It has been reported in many publications that peroxisomal matrix protein import is known to be dependent on hydrolysis of ATP (e.g., see [188,189]). However, Oliveira et al. [190], performing in vitro import assays with rat liver peroxisomes, showed that the ATP hydrolysis is not required for the insertion of PEX5 into the peroxisomal membrane. They argued that protein translocation across the peroxisomal membrane may be driven by both energy released during the interactions of peroxins (e.g., PEX5 and PEX14) and ATP hydrolysis downstream of the translocation step to reset the PEX5-mediated transport system (e.g., recycling of the receptor by PEX1 and PEX6). Also other functions than energization of receptor recycling have been proposed for PEX1 and PEX6, such as mediating membrane fusion in peroxisomal membrane formation in Y. lipolytica and P. pastoris [179,185,186,191]. Clearly, further studies will be essential to elucidate the precise role of PEX1 and PEX6 in peroxisome biogenesis in different organisms.

Out of the 32 known peroxins, some remain with an as yet undefined or unclear role. PEX9 and PEX23 are integral membrane proteins in Y. lipolytica, with unknown function in peroxisome biogenesis [192,193]. Recently, a Blast search of the genome in N. crassa revealed the presence of a putative PEX23 [140]. PEX24 is a novel integral membrane protein of Y. lipolytica [194] which is also present in N. crassa [140]. Studies of a Y. lipolytica pex24 mutant revealed that these cells are defective in matrix and membrane protein import. Furthermore, PEX6 was identified using a mammalian CHO mutant cell line [195,196]. It is a peroxisomal membrane protein. In immunoprecipitation, PEX1 and PEX6 co-precipitate with PEX26. Epitope-tagged PEX1 and PEX6 could not be detected as puncta in PEX26 defective cells, but PEX26 expression re-established colocalization, in a PEX6-dependent manner. PEX26 appeared to recruit PEX6–PEX1 complexes to the peroxisomes; PEX26 anchors PEX6 and PEX1 through PEX6–PEX6 and PEX6–PEX1 interactions, respectively.

Interestingly, these observations made on PEX26 are similar to those recently made on PEX15 in S. cerevisiae [197]. This latter membrane-bound peroxin had previously been identified in this yeast, but its function remained unknown. Now it has been shown that it is an integral membrane protein with most of its sequence,
including the N-terminus, facing the cytosol. It binds, both in vivo and in vitro, to the AAA peroxin PEX6. This suggests that PEX15 in *S. cerevisiae* exerts a similar function as PEX26 in mammalian cells, i.e., the recruitment of PEX6 to the peroxisomal membrane. The N-terminal part of PEX6 contains the binding site for PEX15. Mutagenesis of the Walker motifs of the two AAA cassettes D1 and D2 of PEX6 show that D1 and D2 have opposite effects on the PEX6-PEX15 interaction. The cycle of recruitment and release seems dependent on ATP hydrolysis by PEX6.

### 5.4. Receptor recycling and the ‘extended shuttle’ model for the import of matrix proteins

Initially, it was proposed by Marzioch et al. [124] that the receptors first recognize their cargo proteins in the cytosol. Subsequently, the receptor–cargo complexes would dock at the peroxisomal membrane where the receptors deliver their cargo without entering the peroxisome. After release of the cargo proteins, the free receptors would then return to the cytosol. This model became known as the ‘simple shuttle’ mechanism. Later, experimental evidence that PEX5 actually enters the peroxisome, subsequently releases its PTS1 protein and recycles back to the cytosol, was provided by Dodt and Gould [91]. Reversible accumulation of PEX5 in the peroxisomal compartment could be demonstrated in vivo by manipulating the ATP level and temperature in cultured human fibroblasts and it seemed that PEX5 exits the peroxisomal compartment by a process that requires ATP. More recently, Dammai and Subramani [198] elegantly showed that PEX5 enters human peroxisomes during the course of its normal function and reemerges in the cytosol to carry out further rounds of import. This led to the ‘extended shuttle’ model in which both receptors enter the peroxisome together with their cargo and release them in the lumen before shuttling back to the cytosol (Fig. 2). However, nothing is known yet about a protein-export machinery or PEX5 and PEX7 sorting machinery [199]. In this context, it is tempting to speculate that some peroxins proposed to play a role in the import of proteins might in fact act in receptor export.

---

**Fig. 5.** Three models for import of folded and oligomeric proteins across the peroxisomal membrane (b, c and d) compared with the general model for transmembrane transport of unfolded proteins (a). (a) Chaperonins, indicated by green-coloured dots, keep the newly synthesized protein unfolded at the *cis* side of the membrane, where it is delivered to a receptor. Subsequently, it is translocated through a channel and, at the *trans* side, properly folded with the assistance of chaperonins. (b) Proteins are folded and form oligomers in the cytosol, prior to their import through a regulated pore in the peroxisomal membrane. (c) Cytosolic oligomeric proteins are wrapped by a newly formed double membrane; the outer membrane of the protein-containing vesicle fuses with the peroxisomal membrane and subsequently the inner membrane is degraded, releasing the protein in the peroxisomal matrix. (d) The oligomeric protein in the cytosol is surrounded by an invaginating peroxisomal membrane and internalized by a process reminiscent of endocytosis. At the matrix side, the membrane is degraded to release the translocated protein. (Figure adapted from [200]).
5.5. Three models for the translocation mechanism of matrix proteins

What happens after the docking event still remains largely unknown. Whereas several peroxins are known to be required for the translocation step as discussed above, the mechanism of internalisation of matrix proteins itself has not yet been unravelled. Several data suggest that protein unfolding is not a requirement for the import of matrix proteins as it is the case for protein translocation across the membranes of the endoplasmic reticulum (ER), mitochondria and chloroplasts (reviewed in [145,200]). Indeed, it has been demonstrated that peroxisomes can import folded and oligomeric proteins [121,143,201–203]. However, import of oligomeric proteins is not a general feature of peroxisomal matrix proteins [204]. A recent study demonstrated that proteins are imported into the peroxisome either folded or in unfolded state [205]. This study supports the idea that ternary structure has little or no effect on protein import into the peroxisomal matrix. Misfolded reporter proteins bearing a PTS1 appeared to be imported in association with chaperones, molecules which have never been described within the matrix of the mammalian peroxisomes. However this process might not be functionally relevant under normal conditions.

Three possible mechanisms for transport of oligomeric proteins across the peroxisomal membrane have been considered over the years (Fig. 5). Because large complexes of oligomeric proteins and even 4–9 nm gold beads coated with PTS1-containing protein can be imported into the peroxisome [206], the existence of a pore has been postulated as is the case in the nuclear membrane [200,207]. But freeze-fracture studies have not given any indications of the presence of such a structure in the peroxisomal membrane. Along the same range of ideas, the existence of a channel assembled transiently in response to cargo delivery to the membrane has been postulated. Again, no clear indications for the presence of such a channel have been provided. Another mode of internalization would be through invagination of a portion of the peroxisomal

![Diagram](https://example.com/diagram.png)
membrane where the cargo proteins are docked [143]. After the internalization of the vesicle by a process reminiscent to endocytosis, the membrane surrounding the cargo proteins must be degraded in order to deliver its content in the peroxisomal matrix. Finally, the last mode of transport hypothesized is the wrapping of the cargo proteins by a double membrane [70]. The first step of this process is the formation of a vesicle around the matrix proteins. The outer membrane of the vesicle will then fuse to the peroxisomal membrane. In this process, the inner membrane has also to be degraded in order to release the vesicle content inside the peroxisomal lumen. There is little evidence for any of these three pathways in peroxisome biogenesis and speculations on this subject are still numerous.

5.6. Preimplex hypothesis for the import of matrix proteins

As mentioned before, folded and oligomeric proteins can be imported into the peroxisome. Gould and Collins [208] have proposed the ‘preimplex’ hypothesis that serves to explain this unique feature of peroxisomal matrix protein import. It describes a mechanistic model for the early steps of the import. The authors hypothesize that newly synthesized matrix proteins, together with the PTS receptors, form large protein complexes shortly after their synthesis but before their import (Fig. 6). These pre-import complexes are referred to as preimplexes. It has been shown that one PEX5 monomer binds one PTS1 protein. Because PEX5 seems to exist in vivo as a tetramer [101], it is potentially able to bind several PTS1 proteins at the same time and so assemble the preimplex. Moreover PEX14, the primary PEX5-docking site at the peroxisomal membrane, is a dimeric protein with at least one PEX5 binding site per monomer [101,111]. Preimplex expansion may be further promoted by the fact that each PEX5 monomer has several binding sites for PEX14. Eckert and Erdmann [70] proposed that PEX7 together with PEX18/PEX21 would contribute to forming preimplexes in a comparable manner as PEX5.

What are the benefits of assembling large complexes at the membrane before the translocation step? It allows the rapid and specific delivery of matrix proteins using only diffusion. It also generates an extremely high concentration of matrix proteins near the docking machinery, thus creating a situation that is likely to promote the subsequent translocation of the proteins. Bellion and Goodman [209] indeed observed that several peroxisomal enzymes enter into an extremely large protein complex immediately after their synthesis but before their translocation into the peroxisomal lumen. The preimplex model may nicely accommodate the ‘piggy-backing’ hypothesis and thus offers an explanation to the mode of I-PTS protein import. Finally, the internalization mechanism itself needs to be able to accommodate these large preimplexes. Some models for this mechanism, all compatible with the existence of preimplexes, have already been discussed in Section 5.5. In conclusion, the preimplex hypothesis offers explanations for different aspects of matrix protein import and proposes a possible mechanism of how the matrix proteins are translocated.

5.7. Insertion of peroxisomal membrane proteins

Matrix protein import and peroxisomal membrane proteins (PMPs) insertion occur via two distinct pathways using different sets of peroxins. Evidence for separate processes came from the observation that in cells of Zellweger syndrome patients, empty peroxisomal membranes called ghosts were found; such ghosts still contain peroxisomal membrane proteins but lack matrix proteins [210–212]. Later, it was demonstrated that depletion of essential peroxins for the import of matrix proteins did not affect the topogenesis of peroxisomal membrane proteins [93,106,213,214]. It indicates that indeed two different pathways are involved in the insertion of PMPs into the membrane and the import of matrix proteins.

It has been accepted for a long time that peroxisomal membrane proteins are synthesized on free ribosomes in the cytosol and post-translationally inserted into the membrane of peroxisomes [215-218]. Some peroxins, including integral membrane proteins, have been identified that are responsible for the correct insertion of PMPs in the peroxisomal membrane, but their individual contribution is still not fully understood. Moreover, a recent publication is challenging the old view on the origin and the formation of peroxisomes, therefore reevaluating the mechanism of the PMPs insertion process [219]. The hypothesis presented in this paper will be discussed in Section 5.10.

Peroxisomal membrane proteins contain a peroxisomal membrane-targeting signal (mPTS), which is presumably recognized by a mPTS-receptor that targets the proteins to the peroxisomal membrane, where they are inserted [220,221]. But so far, no general consensus for a mPTS has been identified. The first mPTS, described for Candida boidinii, is present in a hydrophobic loop within PMP47 [220]. It was shown to be necessary and sufficient for targeting the protein to the peroxisomal membrane. Later, peptides containing a mPTS were also defined for PMP34 (the human homologue of CbPMP47), PEX3, PMP22, PEX13 and PEX16. These PMPs do not possess mPTS with a consensus sequence apart from a common short stretch of basic amino acids associated with transmembrane domains. Table 2 summarizes the information presently available about these mPTS. The conclusion that can be drawn from the data in this table is that a mPTS consists of a hydrophilic...
the role of PEX3 as proposed by Hazra et al. [229] is necessary for matrix protein import. It should be noted that PEX3 likely plays a role in the docking and translocation subcomplexes, is similar to that proposed by Agne et al. [176] for PEX8 in the context of their studies with PEX19 act at the earliest stage of peroxisome biogenesis, which is the insertion of PMPs into the membrane [231,227,228,229]. It also interacts with a range of other PMPs [233,234] including PEX10 which requires its C-terminus for this interaction [231]. In several organisms (e.g., human, S. cerevisiae) PEX19 appears farnesylated at its C-terminus, but such modification was not observed in P. pastoris PEX19 [231]. The role of farnesylation is still under debate, but it has been shown for human PEX19 that it is an important determinant in its affinity for other peroxins [235]. Because of its extensive network of interactions, PEX19 has been proposed to act as the cytosolic receptor of PMPs, subsequently targeting them to PEX3 [233]. However this is still a matter of debate. Indeed, other authors suggested that PEX19 could act as a chaperone. This idea was based on the observation that PEX19, in some cases, interacts with regions of PMPs which are different from the sequences required for their targeting to the peroxisomal membranes, defined as the mPTS [234,235]. It is however clear that PEX3 and PEX19 act at the earliest stage of peroxisome biogenesis which is the insertion of PMPs into the membrane [231,227,236,237].

In human, PEX16 is an integral membrane protein with both termini exposed to the cytosol [238–240]. The complementation of Δpex3 mutant fibroblasts by PEX16 expression, restored the biogenesis of peroxisomal membranes, subsequently followed by peroxisomal matrix protein import restoration [238,239]. The participation of PEX16 in the biogenesis of peroxisome membranes is therefore obvious but its precise role in the import of PMPs or formation of the lipid bilayer is unclear. Recent studies on Δpex3 mutant cells showed that upon expression of PEX3, the co-expression of the cytosolic C-terminal part of PEX16 in these cells led to an abolition of peroxisome assembly restoration. These results imply that most likely PEX16 functions upstream of PEX3 in peroxisome membrane assembly [240]. In contrast, Y. lipolytica PEX16 is a peripheral membrane protein that does not seem to participate in the biogenesis of the peroxisomal membrane [241] and it appears to be absent from S. cerevisiae [242].

Table 2
Peroxisomal membrane targeting signals

<table>
<thead>
<tr>
<th>Protein</th>
<th>Organism</th>
<th>mPTS</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMP47</td>
<td>Ch, Sc</td>
<td>Cytoplasmic basic cluster of amino acids adjacent to a membrane anchoring transmembrane domain</td>
<td>[303,304]</td>
</tr>
<tr>
<td>PMP34</td>
<td>Hs</td>
<td>Basic loop + three transmembrane domains. Existence of two independent mPTS</td>
<td>[305,306]</td>
</tr>
<tr>
<td>PEX13</td>
<td>Hs</td>
<td>Existence of two independent mPTS</td>
<td>[306]</td>
</tr>
<tr>
<td>PMP22</td>
<td>Rn</td>
<td>Two transmembrane domains adjacent to a loop containing a basic cluster of amino acids</td>
<td>[307,308]</td>
</tr>
<tr>
<td>PEX16</td>
<td>Hs</td>
<td>Basic cluster of amino acids + a transmembrane segment</td>
<td>[240]</td>
</tr>
<tr>
<td>PEX3</td>
<td>Sc, Pp, Hs, Rn</td>
<td>N-terminal block of positively charged amino acids</td>
<td>[225,224,236,309]</td>
</tr>
</tbody>
</table>

Proteins PMP47, PMP34, PEX13, PMP22, PEX16 and PEX3 from different species; Ch, C. boidinii; Hs, Homo sapiens; Rn, Rattus norvegicus; Sc, S. cerevisiae; Pp, P. pastoris.
5.8. Proliferation of peroxisomes and the formation of their membrane

PEX11 has been postulated to play a role in the regulation of peroxisome proliferation since cells lacking PEX11 contain few large peroxisomes [213,243,244] and overexpression of the peroxin led to hyperproliferation of small peroxisomes [245,246]. However, in S. cerevisiae, PEX11 has been proposed to be indirectly involved in peroxisome proliferation by mediating the production of a signaling molecule that is responsible for modulating peroxisome abundance [257]. In S. cerevisiae, PEX11 was shown to form homodimers [244]. When present in monomeric conformation, S. cerevisiae PEX11 is active in stimulating proliferation. In contrast, PEX11 dimerization leads to inhibition of organelle division [244].

Mammalian cells contain up to three different PEX11 isoforms, α, β and γ [248,249]. The PEX11β gene is constitutively expressed. Mice lacking this PEX11β have no apparent lack in peroxisomal import and only mild effects in peroxisomal metabolic functions. However, they show neuropathological features characteristic of the known peroxisomal disorder ‘Zellweger syndrome’. Expression of the PEX11α gene is inducible by peroxisomal proliferators agents. PEX11α overexpression promotes peroxisome division. However, pex11α-deficient mice have normal peroxisome abundance. The role of PEX11γ remains to be studied.

Recently, several additional peroxins have been identified which appear to be involved in peroxisome proliferation. First, S. cerevisiae cells lacking a protein subsequently denoted PEX25 have fewer but larger peroxisomes [250]. Subsequently, a PEX27 was identified in this yeast species [251,252]. PEX25 and PEX27 are a homologous pair of proteins whose C-termini are similar to the entire PEX11. All three proteins localize to the peroxisomal membrane. Deletion of any of these PEX genes resulted in enlarged peroxisomes. Cells deleted for PEX27 show few enlarged peroxisomes, cells harbouring double deletions of PEX27 and PEX25, PEX27 and PEX11, and PEX25 and PEX11 show greatly enlarged peroxisomes. A severe peroxisomal import defect was observed in a triple deletion mutant strain. Overexpression of PEX25 in the triple mutant strain resulted in a massive proliferation of lamellar membranes and peroxisomes. Overexpression of PEX11, PEX25 or PEX27 led to peroxisome proliferation and the formation of small peroxisomes. Two-hybrid analysis showed that PEX27 interacts with PEX25 and itself, PEX25 interacts with PEX27 and itself and PEX11 interacts only with itself. Together, the data indicate that these three peroxins build a family of proteins whose members are required for peroxisome biogenesis and play a role in the regulation of peroxisome size and number.

The existence of a family of PEX11-like proteins is not limited to mammalian cells (PEX11α, β and γ) and S. cerevisiae (PEX11, PEX25 and PEX27), but extends to trypanosomatids, where recently several PEX11 homologues have been identified: PEX11, GIM5A and GIM5B, as discussed in detail in Section 6.2.

In recent years, PEX24 has been reported for Y. lipolytica and N. crassa. It is an integral peroxisomal membrane protein with unknown function. Mutant cells are defective in matrix and membrane protein import. Recently, two PEX24 homologues have been identified in the S. cerevisiae genome [253]. The encoded proteins are called PEX28 and PEX29. They are constitutively expressed integral membrane proteins. Peroxisomes of cells deleted for either or both the PEX28 and PEX29 genes are increased in number, exhibit extensive clustering and are smaller in area than peroxisomes of wild-type cells. They also often show membrane thickening between adjacent peroxisomes in a cluster. The buoyant density of peroxisomes from double deletion cells is decreased. Overexpression of PEX25 (not PEX11) can restore the wild-type phenotype in both single and double mutant cells. These data suggest that PEX28 and PEX29, together with PEX25, regulate peroxisome number, size and distribution in S. cerevisiae.

Furthermore, in the S. cerevisiae genome, three genes were found coding for proteins with high similarity to PEX23, also with unknown function so far, of Y. lipolytica [254]. These three proteins, called PEX30, PEX31 and PEX32, are integral to the peroxisomal membrane. Cells deleted for PEX30 exhibit increased numbers of peroxisomes, whereas cells deleted for PEX31 and PEX32 exhibit enlarged peroxisomes. PEX30 and PEX32 cannot functionally substitute for one another or for PEX31, whereas PEX31 shows partial functional redundancy with PEX30 and PEX32. PEX30 and PEX31 and PEX32 interact with each other and with PEX28 and PEX29. Systematic deletion studies demonstrated that PEX28 and PEX29 function upstream of PEX30, PEX31 and PEX32 in the regulation of peroxisome proliferation. These data suggest a role for PEX30, PEX31 and PEX32, together with PEX28 and PEX29, in controlling peroxisome size and proliferation in S. cerevisiae.

5.9. Growth and division model versus de novo formation model

Over the years, the origin of peroxisomes was argued and two points of view arose. The earliest one proposed that peroxisomes proliferate and multiply by growth and division [255]. In this model, preexisting peroxisomes posttranslationally import new matrix and membrane proteins and subsequently divide into new peroxisomes. These preexisting peroxisomes are small and of low density and will mature into large peroxi-
some of higher density upon import of these new proteins [256]. Recently, in vivo studies demonstrated that peroxisomes have indeed the ability to multiply by division [257].

In the second half of the 1990s, various lines of evidence for de novo synthesis of peroxisomes, or formation of new peroxisomes in the absence of preexisting ones, were brought to light. Early analysis of $\Delta$peX3 and $\Delta$peX19 mutants of various yeasts and human cells, showed that neither of these cells contained detectable peroxisomes nor peroxisomal remnants (e.g., ghosts) and mislocalize PMPs to the cytosol [162,222,224,228,230,232,233,258,259]. When mutant cells were complemented with their respective proteins, PEX3 or PEX19, the peroxisomal activity was restored. This was considered to be strong evidence for de novo formation of peroxisomes. But doubt has been shed on this hypothesis by other studies. The detection of these particular membrane structures was performed using electron microscopic observation, immunofluorescence or immunoelectron microscopy with antibodies against, e.g., PEX10, PEX14 or also PEX11 [162,222,224]. These methods could have easily missed very small peroxisomal remnants [229,231]. Indeed, in recent studies on $\Delta$peX3 and $\Delta$peX19 cells of P. pastoris and Y. lipolytica, small vesicles and tubules, resembling early peroxisomes, were observed containing PMPs or peroxins [227,229,231,260]. Scientists in favour of de novo formation, argued that depletion of PEX19 and PEX3 could lead to a mistargeting of PMPs to other structures, incorrectly interpreted as preperoxisomes, in the absence of normal target membranes. So there is still an intense debate today whether de novo formation of peroxisomes indeed exists.

For Y. lipolytica, evidence has been presented for the existence of a dynamic population of organelles that differ in various aspects such as protein import competence [185,261,262]. Titorenko and Rachubinski [191,263] presented a model for a multistep peroxisome assembly pathway. Initially pre-peroxisomal structures, thought to arise from a subdomain of the ER (see also Section 5.10), transform into nascent peroxisomes as a result of the uptake of the membrane peroxins PEX3 and PEX16. This second population of peroxisomes is competent to import new matrix and membrane proteins and finally grows into mature peroxisomes which are fully active.

Some authors tried to reconcile both models (Fig. 7). This led to the creation of a two-way model for the biogenesis of peroxisomes [70,218]. The first pathway includes the ‘growth and division’ model in which early peroxisomes grow into mature peroxisomes by import
of PMPs and matrix proteins. Subsequently, these mature peroxisomes will divide to create new early peroxisomes competent for newly synthesized protein import. The second half of the pathway is issued from the de novo formation hypothesis. Some pre-peroxisomes originate from an undefined membrane source, e.g., the ER. The newly formed peroxisomes would first acquire PEX3 and PEX16, peroxins involved in the earliest step of peroxisome biogenesis. Subsequently, with the help of PEX19, PEX3 and PEX16 they could import other PMPs involved in matrix protein import, thereby starting the process leading to mature peroxisomes.

5.10. A role for the ER in peroxisomal membrane formation and insertion of peroxisomal membrane proteins?

How the peroxisomal membranes acquire their lipid components is still largely unknown. Phosphatidyl choline and phosphatidyl ethanolamine are the main phospholipids of the lipid bilayer [264]. The site of synthesis of these phospholipids has been shown to be the ER [255]. The transport from the ER to the peroxisome may occur via specific carrier proteins, transport vesicles or through intermittent connections between the ER and the peroxisomal membrane [69]. This subject is still very controversial because there is no clear evidence of any phospholipid transport mechanism from the ER to the peroxisomes.

In de novo formation, the obvious source of new membranes is the ER. Newly synthesized PMPs could transit via the ER in order to be incorporated into membranes and subsequently transported to the peroxisomal membranes. Elgersma et al. [221] showed that upon overexpression of PEX15, this peroxin was indeed targeted to ER membranes. But later, it was shown that the insertion of PEX15 into the ER membranes should be interpreted as mistargeting resulting from the overexpression of the protein [162]. Such artifact could be due to the protein’s hydrophobic segments which would insert the protein into the membranes of the ER or other compartments if peroxisomal membranes are not available.

A very recent study shed new light on this controversial question of the ER’s involvement in peroxisome biogenesis [219,265]. The authors presented evidence that peroxisomal membranes are derived from the ER in non-manipulated mouse dendritic cells. Using immuno-electronmicroscopy and electron tomography, they observed special lamellar structures that were directly connected to the ER and that contained PEX13 and the peroxisomal transporter protein PMP70. These lamellar structures, named ‘lamellae’, present a special density and striping and were observed to encircle peroxisomal clusters. Moreover, by manually tracing the membranes in tomographic slices, they reconstructed the three-dimensional volume of a portion of a cell. In these slices, they showed the membrane continuity of the ER with the lamellae and the connection between a peroxisome with the lamellar reticulum. Based on these observations, the authors proposed the following model (Fig. 8).

![Fig. 8. ER-peroxisome connection and protein trafficking between the different compartments. Some integral membrane proteins [PEX(X)] are inserted into the ER membranes which become the lamellae. Insertion of additional proteins [PEX(Y), PEX(Z)] which will be components of the protein import machinery or involved in peroxisomal metabolism, occurs in a second step. Later, matrix proteins are imported and the lamellae give rise to mature peroxisomes. (Figure adapted from [219]).](https://academic.oup.com/femsre/article-abstract/28/5/603/777468)
proteins enter the ER giving rise to the formation of the specialized ER, the lamellae. Subsequently, additional integral and peripheral proteins are recruited to assemble a competent protein import machinery. At that moment, matrix proteins start to be imported and globular, mature peroxisomes can bud off the lamellar structures. This model explains nicely how peroxisomal membranes are formed and also accommodates previous observations in favour of the division model. However, it remains to be determined how the very first peroxisomal membrane peroxins, necessary for the subsequent insertion of other PMPs in the pre-peroxisomes, are introduced into the ER membrane. Moreover, the results of this study are only based on morphological observations and therefore, it will be crucial to confirm these data by genetic and biochemical analysis.

6. Aspects of glycosome biogenesis

Like peroxisomes in other organisms, no DNA is detectable in trypanosomatid glycosomes. Glycosomal enzymes are synthesized on free polysomes in the cytosol, and then transferred into glycosomes [266]. The polypeptides are synthesized at their mature size and do not undergo any detectable proteolytic cleavage or secondary modification, such as phosphorylation, glycosylation or covalent linkage of lipids, upon transfer into the glycosomes. The half-life in the cytosol before being imported into the glycosomes has been measured for various glycolytic enzymes of procyclic T. brucei and was found to be short, between 1.0 and 3.4 min [266,267], indicating that import into glycosomes can be a rapid process. These values are comparable to the fastest translocation rates reported for rat-liver and yeast peroxisomal proteins. The glycosomal enzymes of both bloodstream-form and procyclic T. brucei are very stable; no significant turnover could be measured during a 3-h period [267,268].

6.1. Targeting and import of glycosomal matrix proteins

Most of the glycosomal enzymes analyzed to date have a PTS1, some of them a PTS2, and in a small subset of glycosomal enzymes no consensus motif that might be responsible for protein targeting to the glycosome could be detected. Table 3 presents the PTS motifs of glycosomal enzymes involved in carbohydrate metabolism in the various life-cycle stages of T. brucei and some Leishmania species.

Analysis of the trypanosomatid genome projects databases revealed the presence of about 250 predicted

| Table 3 Glycosome-targeting signals of Trypanosoma brucei and Leishmania spp. enzymes involved in carbohydrate metabolism |
|-----------------|-----------------|-----------------|
| **Type of signal** | **Enzyme** | **T. brucei** | **Leishmania spp.** |
| **PTS1** | Extended glycolysis | PGI | -SHL | -AHL |
| | | PFK | -AKL | -SKV |
| | | GAPDH | -AKL | -SKM |
| | | PGK-C | -SSL | - |
| | | GPDH | -SKM | -SKL |
| | | GK | -AKL | -AKL |
| | | PEPCK | -SRL | -AHL |
| | | MDH | -SKL | -SHL |
| | | FR | -SKI | -AKI |
| | | PPDK | -AKL | -SKL |
| | | G6PDH | - | -NKL |
| **PPP and gluconeogenesis** | 6PGL | -AKF | - |
| | FBPase | -SKL | -SKL |
| | SBPase | -SKL | enzyme absent |
| **PTS2** | Extended glycolysis | HKX | -RL(X)2HI | -RL(X)2HI |
| | | ALD | -RV(X)2QL | -RV(X)2QL |
| **PPP and gluconeogenesis** | 6PGL | - | -KV(X)2KI/2HI |
| | FBPase | -RV(X)2QF | - |
| **I-PTS** | Extended glycolysis | TIM | amino acids 140 to 161 | ? |
| | | PGK-A | amino acids 24 to 91 | enzyme absent |

Abbreviations: PPP, pentose-phosphate pathway. Enzymes: ALD, aldolase; FBPase, fructose-1,6-bisphosphatase; FR, fumarate reductase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GK, glycerol kinase; GPDH, glycerol-3-phosphate dehydrogenase; G6PDH, glucose-6-phosphate dehydrogenase; HKX, hexokinase; TIM, triosephosphate isomerase; PEPCK, phosphoenolpyruvate carboxykinase; 6PGL, 6-phosphogluconolactonase; PGI, glucose-6-phosphate isomerase; PFK, phosphofructokinase; PGK, phosphoglycerate kinase (isoenzyme A or C); PPDK, pyruvate phosphate dikinase; SBPase, sedoheptulose-1,7-bisphosphatase. –, The consensus motif of the PTS is not present in the available amino-acid sequence; ?, the targeting signal has not yet been identified in the protein.
proteins with a candidate PTS motif (F. Opperdoes, unpublished). Approximately two-third of these proteins have a putative PTS1, one-third a PTS2. For about 50% of these predicted PTS-containing proteins no function is known as yet, but the remainder comprises enzymes of known glycosomal processes (see Section 3): glycolysis, pentose-phosphate pathway, purine salvage, pyrimidine biosynthesis, \( \beta \)-oxidation of fatty acids, fatty-acid elongation, ether-lipid biosynthesis, several reactions of the isoprenoid biosynthetic pathway and proteins involved in protection against oxidant stress.

Detailed studies have been done with stably transfected bloodstream-form \( T. brucei \) expressing reporter proteins to determine which variations of the C-terminal tripeptide (PTS1) can route a reporter protein to glycosomes [47, 48, 269, 270]. These analyses, as well as inspection of the sequences of glycosomal enzymes, show a spectrum of variants of the canonical PTS1, -SKL, that are functional, albeit at different efficacies. The variety of functional PTS1 permutations is markedly larger in \( T. brucei \) than observed for mammalian cells [76, 79].

Two proteins of glycolysis of trypanosomatids have a PTS2 motif, hexokinase and aldolase. In addition, a few other proteins involved in core carbohydrate metabolism have such a signal (Table 3). Indeed, the PTS2 motif-containing N-terminal sequence of aldolase was capable of directing a reporter protein to glycosomes [271]. The presence of an N-terminal PTS2 in aldolase, instead of a C-terminal PTS1, can be rationalized because a free C-terminal tyrosine residue is essential for the catalytic activity of this enzyme.

For two glycolytic enzymes of \( T. brucei \), the involvement of an I-PTS in glycosome targeting was shown. A minor glycosomal PGK isoenzyme (PGK-A) is targeted to the glycosomes by an internal sequence located between amino acids 24 and 91 [272]. A second glycosomal enzyme from \( T. brucei \), TIM, was found to be imported via a 22 amino-acid internal sequence (de Walque, S., Chevalier, N. and PM, unpublished) (Table 3). A peptide comprising residues 140–161 of the 250 amino acids long TIM polypeptide was shown to be sufficient and necessary to direct the reporter protein chlorramphenicol acetyl transferase to the glycosomes when fused to its N-terminus. The corresponding peptide of \( S. cerevisiae \) TIM could not do so. Indeed, the trypanosomatid (\( T. brucei, T. cruzi, L. mexicana \)) TIM peptide has some specific residues (residues only conserved in trypanosomatids and a two-residue insertion) that may be responsible for the targeting. We hypothesize that this peptide, which in the three-dimensional TIM structure contributes to a surface-exposed loop [273], is responsible for interaction with a different protein. This other protein may be a peroxin or a different glycosomal matrix protein enabling TIM import by piggy-backing (see also Section 5.1).

Interestingly, a considerable number of enzyme activities of kinetoplastids have a dual localization; part of the enzyme activity is found in the glycosome, another part in the cytosol (reviewed in [68]). The extent of compartmentation may differ between the various kinetoplastid organisms and between cells from different life cycle stages. Some activities with a dual localization are represented in each compartment by different isoforms, encoded by distinct genes (e.g., PGK and GAPDH in \( T. brucei \) [274, 275]). In most cases studied, however, no differences could be detected between the enzyme present in the glycosomes and cytosol, and the differently located activities had to be attributed to the product encoded by a single gene. Examples are glucose-6-phosphate isomerase, glucose-6-phosphate dehydrogenase and 6-phosphogluconolactonase [276, 277] (see also Section 4). How such enzymes are distributed in the kinetoplastid cell in a manner that is physiologically appropriate for each species and its life cycle stage is not clear yet. Several possible mechanisms, separately or acting together, could be considered, such as different efficiencies of the targeting signal or the structural context in which the targeting signal is present in the protein, not yet detected posttranslational mechanisms for selective retention of part of a protein in a compartment, or a kinetic control determined by different rates of biosynthesis, import and turnover of different proteins. These and other possible mechanisms have been discussed in detail by Hannaert et al. [68].

Also for glycosomes, protein unfolding is not a prerequisite for import. One of the first, and most convincing examples that proteins in folded form can enter peroxisomes was reported for \( T. brucei \) glycosomes. Häusler et al. [201] employed a methodology previously applied to show that a polypeptide during its transit into mitochondria must be in an unfolded state. Methotrexate, a folate analogue, binds with high affinity to dihydrofolate reductase (DHFR), stabilizing its folded state and thereby inhibiting unfolding [278]. DHFR coupled to a mitochondrial signal peptide can be imported into the organelle, but upon addition of methotrexate or aminopterin, a membrane-permeable analogue, the stably folded complex gets stuck in the import channel. When a similar experiment was done in trypanosomes, targeting of the DHFR fusion protein to the mitochondrion was also inhibited. However, import of a fusion protein of DHFR and the C-terminus of gPGK containing a PTS1, into glycosomes was not inhibited upon addition of the same DHFR inhibitors.

6.2. Peroxins involved in glycosome biogenesis

Various homologues of mammalian and yeast peroxin genes have been identified in the genomes of \( T. brucei, T. cruzi \) and \( Leishmania \) spp (Table 4A). These genes were identified by different methods. Flaspohler et al.
### Table 4
Peroxin homologues in Trypanosomatidae

<table>
<thead>
<tr>
<th>Peroxin</th>
<th><em>Trypanosoma brucei</em></th>
<th></th>
<th></th>
<th><em>Leishmania</em> spp.</th>
<th></th>
<th></th>
<th></th>
<th><em>Trypanozoma cruzi</em></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Homologue identified</td>
<td>Peroxin identity confirmed</td>
<td>Essentiality confirmed</td>
<td>Homologue identified</td>
<td>Peroxin identity confirmed</td>
<td>Essentiality confirmed</td>
<td>Homologue identified</td>
<td>Essentiality confirmed</td>
<td></td>
</tr>
<tr>
<td>PEX1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX2</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX5</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX6</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX7</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX10</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX11/GIM5A</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX11/GIM5B</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX12</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX14</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX19</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**A. Identification, confirmation of identity and essentiality of the peroxins for the parasites**

- PEX1, PEX2, PEX5, PEX6, PEX7, PEX10, PEX11, PEX12, PEX14, PEX19 are confirmed to be essential for the parasites.
- PEX3, PEX4, PEX8, PEX9, PEX13, PEX15, PEX16 are not confirmed.

**B. Percentages of identity between peroxin amino-acid sequences**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PEX1</td>
<td>60</td>
<td>49</td>
<td>49</td>
<td>27</td>
<td>28</td>
<td>27</td>
<td>23</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>PEX2</td>
<td>65</td>
<td>51</td>
<td>56</td>
<td>24</td>
<td>26</td>
<td>26</td>
<td>15</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>PEX5</td>
<td>64</td>
<td>52</td>
<td>52</td>
<td>26</td>
<td>27</td>
<td>26</td>
<td>22</td>
<td>22</td>
<td>21</td>
</tr>
<tr>
<td>PEX6</td>
<td>76</td>
<td>65</td>
<td>68</td>
<td>32</td>
<td>34</td>
<td>33</td>
<td>34</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>PEX7</td>
<td>76</td>
<td>65</td>
<td>68</td>
<td>32</td>
<td>34</td>
<td>33</td>
<td>34</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>PEX10</td>
<td>65</td>
<td>53</td>
<td>54</td>
<td>27</td>
<td>26</td>
<td>28</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>PEX11</td>
<td>61</td>
<td>59</td>
<td>62</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>12-17</td>
<td>12-16</td>
<td>12-19</td>
</tr>
<tr>
<td>PEX11/GIM5A</td>
<td>66^a</td>
<td>56^a</td>
<td>56^a</td>
<td>12-17^b</td>
<td>12-16^b</td>
<td>12-19^b</td>
<td>11-13^b</td>
<td>12-15^b</td>
<td>12-13^b</td>
</tr>
<tr>
<td>PEX11/GIM5B</td>
<td>56^a</td>
<td>56^a</td>
<td>56^a</td>
<td>12-17^b</td>
<td>12-16^b</td>
<td>12-19^b</td>
<td>11-13^b</td>
<td>12-15^b</td>
<td>12-13^b</td>
</tr>
<tr>
<td>PEX12</td>
<td>57</td>
<td>40</td>
<td>43</td>
<td>20</td>
<td>20</td>
<td>18</td>
<td>17</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>PEX14</td>
<td>45</td>
<td>36</td>
<td>34</td>
<td>16</td>
<td>15</td>
<td>14</td>
<td>13</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>PEX19</td>
<td>36</td>
<td>27</td>
<td>26</td>
<td>19</td>
<td>13</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
</tbody>
</table>

References for the trypanosomatid peroxins experimentally analyzed are given in the text. Confirmation of essentiality of peroxins is based on published experiments described in the text and unpublished RNAi experiments (for *T. brucei* bloodstream-form, PEX6, PEX10 and PEX12; H. Krazy and PM, unpublished). The identification of putative peroxins was done by searches in the GeneDB database of the trypanosomatid genome projects (http://www.geneDB.org). *Leishmania* sequences used for the data in Table 4 B are either of *L. donovani* (PEX2, PEX5, PEX14), if characterized and described in the literature, and otherwise of *L. major*, as identified in the genome database. Human and *S. cerevisiae* sequences were taken from the SwissProt database (http://www.expasy.org/sprot/) or GenBank (http://www.ncbi.nlm.nih.gov/). A multiple alignment of the amino-acid sequences of each of the peroxins of human, *S. cerevisiae* and the three trypanosomatid species was prepared with the program ClustalX [310] and used to calculate the percentages of identity given in Table 4 B. Abbreviations: GIM, glycosomal integral membrane protein; PEX, peroxin; *L.* spp, *Leishmania* species; *T. b.*, *T. brucei* ; *T. c.*, *T. cruzi* ; *S. c.*, *S. cerevisiae.*

^a average of identities of GIM5A and GIM5B homologues in the trypanosomatid species indicated.

^b range of the identities between the different trypanosomatid PEX11 isoforms (PEX11, GIM5A and GIM5B) and the three isoforms of human PEX11: PEX11α, β and γ.
[279,280] identified a *L. donovani* PEX2 homologue by a functional screen. Promastigotes were first transfected with a gene coding for a PTS1-containing protein rendering cells resistant to the drug bleomycin. The rationale of the screen is that the cells remain susceptible to bleomycin or phleomycin when the protein is sequestered inside the glycosomes, but become resistant when drug uptake into the organelles is abolished. Upon random mutagenesis, resistant cells were obtained which showed a defect in the import of glycosomal protein. A clone thus obtained (called *gim1*-1, for ‘glycosome import mutant’) showed indeed a cytosolic localization of the drug resistance protein and a partial mislocalization of a glycosomal protein involved in purine salvage, hypoxanthine-guanine phosphoribosyltransferase (HG-PRT), a PTS1-protein (-SKV). Other glycosomal enzymes, GAPDH and PGK-C (having a PTS1 and an unknown targeting signal, respectively), remained correctly compartmentalized and glycosome morphology and number were not significantly affected [279]. Strikingly, the mutant cells lacked lipid bodies, which have been postulated to serve as an energy reserve. The *gim1*-1 mutant gene was identified as a PEX2 homologue. The phenotype was due to the dominant-negative nonsense mutation in one of the two PEX2 alleles of the diploid organism [280], resulting in a protein lacking one transmembrane domain and the RING finger. The same methodology has subsequently yielded two additional mutant clones (called *gim2* and *gim4*) that have partial defects in glycosomal import of HG-PRT and TIM, but not GAPDH and PGK-C, but no obvious glycosomal or other structural changes compared to wild-type cells [281]. The *gim2* and *gim4* genes remain to be characterized.

The *T. brucei* PEX2 was identified by homology search of the sequence database [282]. It has 51% amino-acid sequence identity with *L. donovani* PEX2, and about 15-25% with all other known PEX2 sequences (Table 4B). Northern and Western blotting experiments suggest that PEX2 is expressed at low abundance in trypanosomes, and overexpression seemed not possible. Expression of an additional, truncated PEX2, similar to the *Leishmania* mutant protein, appeared lethal in bloodstream-form trypanosomes. When PEX2 expression was silenced by RNA interference (RNAi), bloodstream-form trypanosomes died rapidly upon PEX2 depletion. PEX2 knockdown in procyclic forms led to relocation of part of the tested glycosomal matrix proteins (PFK and GAPDH, both PTS1 proteins, and aldolase, a PTS2 protein) to the cytosol and eventually also these cells died.

Purified glycosomal membrane fractions of *T. brucei* contain two predominant protein species of 24 and 26 kDa [283]. Both proteins were purified, subjected to peptide sequencing, and the genes cloned [284,285]. The protein encoded by the 24 kDa protein has low sequence identity with PEX11 from other organisms (less than 20%, see Table 4B), but nevertheless was identified as an authentic PEX11 by complementation of a *S. cerevisiae Δpex11* mutant which is defective in peroxisome proliferation [284]. Most conservation is found in the N- and C-terminal domains of this integral membrane protein; both domains are oriented towards the cytosol. The protein can form homodimers, like its yeast counterparts. The *TbPEX11* gene appeared essential in trypanosomes; by inducible knockout in bloodstream-form and procyclic trypanosomes, the expression could be reduced to 10% of the wild-type level, but not lower. This reduced expression resulted in a decrease in the number of glycosomes, with corresponding increase in size. However, this reduction did initially not result in slowed growth. *TbPEX11* overexpression caused growth arrest and transformed the globular glycosomes into clusters of long tubules filling a large part of the cytoplasm. Therefore, it seems that in trypanosomes *TbPEX11* plays a role homologous to that of PEX11 in yeast and mammalian cells, i.e., an involvement in peroxisome proliferation (see Section 5.8) [247]. Two related genes, called *GIM5A* and *GIM5B* (GIM stands for glycosomal integral membrane protein), encode the 26 kDa glycosomal membrane protein [285]. GIM5A and GIM5B are isoforms of PEX11, albeit with a low level of sequence identity (both 15% compared to *T. brucei* PEX11). They have all a similar predicted topology in the glycosomal membrane, but unlike trypanosome PEX11, the GIM5 proteins have no nucleotide-binding motif, nor a potential consensus peroxisomal membrane-targeting signal. GIM5 is essential for bloodstream-form *T. brucei* [285,286]. Deletion of the *GIM5A* gene from the bloodstream-form parasite has no effect on trypanosome growth, but depletion of GIM5B to below 10% of wild-type levels in a *gim5A* null background leads to rapid cell death. In contrast, procyclic trypanosomes survive without GIM5A and with very low levels of GIM5B. Attempts to completely delete these *PEX11*-related genes were unsuccessful, suggesting that they are essential for the survival of the parasite. *Δgim5A* GIM5B depleted procyclic cells have a drastically reduced number of glycosomes and are osmotically fragile. The import of aldolase and GAPDH into glycosomes was not significantly affected in these cells, but some leakage of matrix proteins into the cytosol occurred. Other changes noticed in these cells are an altered mitochondrial morphology and a drastically reduced level of ether lipid species. The bloodstream-form *Δgim5A* cells in which also GIM5B depletion was induced, showed, before GIM5B depletion became critical, a shift in metabolic pattern: no glycerol was produced anymore from glucose, whereas pyruvate production increased. In contrast, glycerol could still be metabolized when given as substrate, indicating that
glycerol uptake and metabolism by glycosomes were not affected.

Overexpression of GIM5B inhibits growth. It was also shown that TbPEX11, GIM5A and GIM5B have non-redundant, but overlapping functions. GIM5A and GIM5B form heterodimers, but at present no other proteins interacting with GIM5 are known. From all data obtained, it was concluded that the GIM5 proteins are required for the maintenance of normal glycosome numbers in T. brucei. The authors proposed that the function of PEX11-family proteins is to determine the overall structure and composition of the peroxisomal membrane [286]. Homologues of L. major GIM5A and GIM5B have also been detected, but not studied in any detail [285].

It should be noted that the PEX11/GIM5 situation in trypanosomatids is reminiscent of that in mammalian cells and S. cerevisiae, where also three PEX11 family members have been detected: the PEX11 isoforms α, β and γ and PEX11, PEX25 and PEX27, respectively (see Section 5.8).

The genes of PEX5, the PTS1 receptor, of T. brucei and L. donovani were both cloned after successful PCR amplification of a gene fragment using primers corresponding to regions conserved in the peroxin of mammalian cells and yeasts [287,288]. The overall identity with other PEX5 sequences is low (21–27%), but the (sub)domain organization of the trypanosomid PEX5 homologue is similar to that of the PTS1 receptors of other organisms, with seven copies of the TPR motif in the C-terminal domain, and a N-terminal half containing three copies of the WXXXF/Y pentapeptide motif. The amino-acid sequences of the two trypanosomatid PEX5 sequences are 52% identical. The identity of the T. brucei protein product as a PTS1 receptor was proven by its specific interaction in vitro with glycosomal PGK-C, with an equilibrium dissociation constant (Kd) of 40 nM [287]. A C-terminally truncated PGK-C, without PTS1 did not bind to purified trypanosome PEX5, produced in E. coli. Similarly, the L. donovani protein displayed specific binding of the PTS1 containing xanthine phosphoribosyltransferase (XPRT), with a Kd of 4.2 nM, but not of XPRT in which the PTS1 had been deleted [288]. Using biochemical, immuno-cytochemical and immuno-fluorescence techniques, a predominantly cytosolic localization was found for both TbPEX5 and LdPEX5, in agreement with findings for the PTS1 receptor in other organisms. The repeated inability to generate a ΔPEX5 null mutant by targeted gene replacement from ΔpeX5/PEX5 heterozygotes suggested that the PEX5 is essential for L. donovani promastigotes. In both the bloodstream and procyclic form of T. brucei PEX5 was shown to be essential by applying RNAi (JM, N. Galland and PM, unpublished).

A fragment of the TbPEX5 containing the first three TPR motifs has been crystallized and its structure determined at 1.6 Å resolution [289]. The two first of the three TPR motifs appear to adopt the canonical antiparallel helix hairpin structure of TPRs. In contrast, the third TPR motif has a dramatically different conformation: the two helices that usually form a hairpin are folded into a single, long continuous helix never observed before for a TPR motif (Fig. 9). The potential functional importance of this different conformation has not yet been established.

Also the gene of a T. brucei homologue of PEX7, the PTS2 receptor, has been cloned and sequenced (V. Hannaert, JM and PM, unpublished). This was possible after recognition of an EST coding for a fragment with a typical PEX7 sequence in the T. cruzi sequence database. The putative T. brucei PEX7, as well as the corresponding sequence recently revealed in the L. major and T. cruzi databases, have 32–36% amino-acid sequence identity with human and yeast PEX7 (Table 4B), and 40% with A. thaliana PEX7. The T. brucei, T. cruzi and L. major sequences are 65–76% identical. The trypanosomatid sequences have a similar organization as the PEX7 sequences from other organisms: a core of six predicted WD motifs with an N-terminal region of about 60 residues. Surprisingly, they have an approximately 40 residues long extension at the C-terminus, whereas other PEX7s have only 5-10 residues. Furthermore, this C-terminal extension is very rich in proline residues. The identity of these sequences as the trypanosomatid PTS2 receptors is very likely, but remains to be proven.

Short potential PEX14 fragments were recognized in the databases of T. brucei and L. major, by BLAST searches with yeast and mammalian PEX14 sequences. This led to the cloning and sequencing of T. brucei [290,291] and L. donovani [292] PEX14 homologues. The overall identity with PEX14 sequences of other organisms is only around 15–25%; the T. brucei and
also the I-PTS protein TIM[291]. The partial mislocalization of matrix proteins to the cytosol was observed for different classes of matrix proteins: glycosomal matrix proteins import pathways. To our knowledge, no evidence for the involvement of PEX14 in I-PTS protein import has been reported so far for peroxins of the trypanosomatids and their human homologues, with glycosomal sequences share 34% identity. The protein was found to be associated with glycosomes by immunofluorescence microscopy. Indeed, the sequence displays a stretch of 19 consecutive hydrophobic residues extending from residue 148 to 166, with a similar hydrophobic nature as the equivalent region in other PEX14 sequences, and that is presumably responsible for its anchoring to the membrane of the organelle [291]. The N-terminal domain, upstream of this hydrophobic region, was shown to be responsible for specific binding to TbPEX5 and LdPEX5, respectively [291,292], proving the identity of these trypanosomatid proteins as PEX14, a constituent of the receptor docking complex. Moreover, detailed analysis of the interaction of the L. donovani protein with the glycosomal membrane indicated that LdPEX14 is tightly associated to the outer side of the membrane rather than being an integral membrane protein. The N-terminal region also has the PXXP box typically for SH3-domain binding. The equivalent region in PEX14 of other organisms has been shown to be responsible for interaction with PEX13 (see Section 5.2).

Depletion of PEX14 by RNAi showed that the protein is essential to both procyclic and bloodstream-form T. brucei [290,291]. RNAi results, in cells of both life-cycle stages, in a partial mislocalization of glycosomal matrix proteins to the cytosol. The mislocalization was observed for different classes of matrix proteins: PTS1 (GAPDH), PTS-2 (aldolase), and interestingly was also the I-PTS protein TIM [291]. The partial mislocalization was interpreted as the retention of newly synthesized proteins in the cytosol due to the inability to import matrix proteins by the fraction of growing and dividing glycosomes. The miscompartmentation of both PTS-1 and PTS-2 proteins strongly suggests that also in T. brucei, PEX14 is the point of convergence of both matrix proteins import pathways. To our knowledge, no evidence for the involvement of PEX14 in I-PTS protein import has been reported so far for peroxins of any organism. This observation provides support for the notion that TbTIM enters glycosomes by “piggy backing”, involving the formation of a heteromeric complex with a PTS-1 or PTS-2 protein (see Section 5.1).

The reduction of growth rate upon induction of RNAi appears in the bloodstream-form cells already after 24 hours, whereas the effect on growth of the procyclic form is seen only after 4 days of treatment. Bloodstream-form trypanosomes are entirely dependent on glycolysis and glycosomes for their ATP supply. As discussed in Section 4, it was concluded from computer simulations of the metabolism that an intact glycosomal membrane and proper enzyme compartmentation are of vital importance for these cells to compensate for a lack of activity regulation of hexokinase and phosphofructokinase by feed-back loops [53]. In contrast, procyclic cells have a highly branched metabolic network by which the harmful effects of the unregulated kinases may be alleviated. The presence of an active pentose-phosphate pathway in procyclics, with its enzymes present not only in the glycosomes but also in the cytosol [277,293], may provide an extra-glycosomal shunt for glucose catabolism that partially overcomes the problems caused by the miscompartmentation of glycolytic enzymes. Thus a mistargeting of glycosomal enzymes could result in a more severe phenotype in bloodstream-form trypanosomes than in procyclics. Similar differences between bloodstream-form and procyclic T. brucei have been observed when PEX2 was targeted by RNAi [282] (see above).

Interestingly, Furuya et al. [290] demonstrated that removal of simple sugars from the medium allowed the procyclic trypanosomes in which PEX14 was depleted to survive. We interpret these data as additional support for the hypothesis that proper compartmentation is particularly important during growth on sugars (as occurs in the bloodstream form) to prevent the unrestrained accumulation of sugar-phosphates that would otherwise result from the lack of regulation of sugar kinases, as mentioned above [53]. During growth on other carbon sources such as amino acids, trypanosomes do not rely on glycosomes for catabolism and, consequently, lesser harm may occur to the cell.

Several other potential peroxins have been identified in the T. brucei, T. cruzi and L. major genome databases by Blast searches with mammalian or yeast sequences: PEX1, PEX6, PEX10, PEX12 and PEX19 (Table 4). Like the other peroxins, their amino-acid sequences display low identities with that of their human homologues, in the range of 15–35%. The putative T. brucei PEX6, PEX10 and PEX12 genes have been cloned and expressed in E. coli as recombinant proteins (H. Krazy and PM, unpublished). The predominant association of the PEX6 and PEX10 homologues with glycosomes has been confirmed by cell fractionation through differential centrifugation followed by immunoblot analysis. RNAi was used to show that the putative PEX6, PEX10 and PEX12 are essential proteins in both bloodstream-form and procyclic T. brucei.

6.3. Protein interactions in glycosome biogenesis and possibilities for drug design

As described in Section 5, the biogenesis of peroxins and glycosomes occurs via a multi-step process involving a large number of proteins that form, transiently or permanently, specific interactions with each other. Most of these steps require the formation of complexes of several proteins. Because of the vital importance of proper glycosome biogenesis to the parasite (see Section 4), and the low sequence identities between the peroxins of the trypanosomatids and their human
host (Table 4B), these peroxins may be considered as highly promising anti-parasite drug targets. Targeting peroxins with compounds that should interfere with the function of these proteins in glycosome biogenesis implies targeting in a selective way the interactions between the trypanosome peroxins, without disrupting the interactions between the corresponding human peroxins. Although this research is still in a preliminary stage, important information about the interactions of some trypanosomatid proteins involved in glycosome biogenesis has already been obtained, providing clues for strategies of future design of selective inhibitors.

6.3.1. PTS1 protein–PEX5 interaction

No crystal structure is as yet available for the complete TPR domain, responsible for PTS1 binding, of a trypanosomatid PEX5. However, the structure of a fragment of human PEX5 that includes all seven TPR motifs in complex with a pentapeptide containing a PTS1 sequence has been reported (Fig. 4) [96]. This structure, in conjunction with binding studies with fluorescent peptides bearing PTS1 variants [80], revealed the molecular basis for PTS1 recognition. The low percentage of identity between the human and trypanosomatid PEX5 (36% for the TPR domain) and the markedly larger variety of

Fig. 10. The PTS2 dimer of T. brucei aldolase. Panels a and b show, from two different angles, the aldolase surface comprising the PTS2 dimer formed by the intertwined N-termini of subunits A and B. Subunits A and B are coloured green and blue, respectively, and their N-terminal residues comprising the PTS2 (SKRVEVLLTQLPAY; residues in bold correspond to the consensus PTS2 motif) are labeled. The numbers indicate the position of each residue in the polypeptide (starting from the N-terminal Ser). The flanking subunits C and D of the tetrameric aldolase are coloured in purple. Notice the deep grooves between the PTS2 dimer and the subunits C and D, and which might be targets for drug design, by exploiting structural features of both the PTS2 dimer and flanking subunits C and D, a combination that is unique to the aldolase of the parasite. Such drugs would prevent interaction of trypanosome aldolase with PEX7, either directly by steric occlusion, or by allosteric inhibition if a change of PTS2 conformation is needed for its recognition by PEX7. The T. brucei aldolase crystal structure was determined by Chudzik et al. [296]; the coordinates are available from the Protein Data Bank via accession code 1F27.
functional PTS1 permutations observed for *T. brucei* compared to mammalian cells (see Section 6.1) suggest that the PTS1 binding groove of the parasites’ receptors may be targeted by selective inhibitors. The finding by Gatto et al. [80] that a binding energy threshold may determine the functionality of PTS1 sequences made these authors suggest that inhibitors should be designed that interfere with the binding to the trypanosomatid PEX5 but do not cross the functional threshold in humans.

### 6.3.2. PTS2 protein–PEX7 interaction

Very little is known about the mode of interaction of PTS2 with PEX7. This analysis is hampered by the apparent difficulty to overexpress and purify PEX7 in soluble form to perform in vitro binding studies. PEX7 belongs to the family of WD-repeat proteins, which are known to be involved in interactions with other proteins. Crystal structures of two WD-repeat proteins have been reported—the β subunit of the G-protein involved in signaling in mammalian cells [294] and an actin-interacting protein from *S. cerevisiae* [295]—and shown to have a circularized seven-bladed β-propeller structure. All WD-repeat proteins are expected to have such a β-propeller structure, but defining the binding domains remains a challenge. Although no structure is yet known of the PTS2 receptor, crystallographic analysis of *T. brucei* and *L. mexicana* aldolase has provided information about the PTS2-structure. Aldolase is a tetrameric enzyme. In both trypanosomatid enzymes, the PTS2s from two different subunits form two closely intertwined structures (Fig. 10) [296]. The finding or this specific, ordered structure offers promise for the design of trypanocidal drugs. Any compound binding tightly to this structure would most likely interfere with PTS2-PEX7 binding and thus prevent aldolase from entering the glycosome. Human aldolase is a cytosolic enzyme, not containing a PTS2. Selectivity of an inhibitor, to prevent interference with the import of any host PTS2 protein could be obtained by designing compounds that also recognize features specific of the trypanosomatid aldolase. Indeed, flanking the PTS2 dimer is a groove at the aldolase surface containing a number of unique residues that could be exploited (Fig. 10). A similar approach is possibly feasible to inhibit the import of other PTS2-containing enzymes into the glycosomes.

### 6.3.3. PEX5–PEX14 interaction

The interaction of trypanosomatid (*T. brucei* and *L. donovani*) PEX5 and PEX14 has been studied in considerable detail and compared with that of the corresponding human and yeast peroxins [291, 292, 297]. Using purified recombinant forms of PEX5 and PEX14 specific interactions were shown in vitro, using different techniques, both for the *Trypanosoma* and *Leishmania* proteins. A $K_d$ of 2.75 μM was determined for the *L. donovani* peroxins [292]. When the *LdPEX5* was preloaded with a PTS1 peptide, the affinity increased to 0.40 μM. The N-terminal halves of both peroxins were shown to be responsible for the interaction. The binding sites, and the strength of the interactions between the two peroxins were characterized in detail in *T. brucei* by a fluorescence assay, using a panel of N-terminal regions of *TbPEX14* variants and a series of different peptides derived from *TbPEX5* [297]. On the PEX5 side, WXXXF/Y motifs in the N-terminal domain were shown to be responsible for PEX14 binding, similar to earlier findings in human and yeast cells. However, only two of these pentapeptide motifs showed affinity for PEX14, the second motif peptide did not. Studies using different variants of the peptides suggested that the propensity to form an α-helix, as well as the nature of residues outside the pentapeptide, notably the absence of a negative charge at the 9th residue counting from the conserved tryptophan, are important for binding. Indeed, peptides with a length of 13 or 16 residues show comparable affinity, in the low nanomolar range, whereas the affinity of shorter peptides (10- and 7-mers) for PEX14 is considerably (10- to 100-fold) reduced. The non-binding second motif has a low, predicted, tendency to form an α-helical conformation and contains a negative charge on the ninth residue (E instead of Q in motifs 1 and 3). Interestingly, WXXXF/Y motifs with a negatively charged residue at the ninth position have been shown to be responsible for PEX5 interaction with PEX13, another component of the docking complex, in *P. pastoris*, *S. cerevisiae* and Chinese hamster [102, 103, 156] (see Section 5.2). On the *T. brucei* PEX14 side, the region for binding was specified within the N-terminal 84 residues. Site-directed mutagenesis studies identified a number of specific residues, most of them hydrophobic and conserved in all known PEX14 sequences, which seem important for binding. Similar findings have been reported for *A. thaliana* [137]. Particularly two phenylalanines (F35 and F52 in *T. brucei*) are critical. The results together suggest that hydrophobic interactions are important for PEX5-PEX14 interaction. The likely α-helical conformation of the PEX5 peptides containing the WXXXF/Y motif implies that the Trp and Phe/Tyr residues are facing the same side. They may interact with the hydrophobic residues in PEX14. Other residues in the vicinity also play an important role in determining the affinity. Interestingly, not only the overall identities of PEX5 and PEX14 between parasite and its human host are low, but there exists also considerable variation in the residues around the critical residues involved in transient PEX5:PEX14 interactions. This may have great potential for the design of selective, peptidomimetic inhibitors.
6.3.4. Putative interactions of I-PTS proteins

As described in Sections 5.1 and 6.2, there is ample support for the notion that some enzymes enter the peroxisomal/glycosomal matrix using a piggy-backing strategy, by forming a heteromeric complex with a PTS1 or PTS2 containing protein. Such a mechanism opens additional possibilities to design compounds that may selectively interfere with the biogenesis of the organelle of the parasite, because it is highly unlikely that the areas responsible for interaction of both proteins are conserved in host and parasite. For example, the 22 amino-acid sequence, present on the surface of *T. brucei* TIM and that was identified as responsible for the targeting of the protein to the glycosome (see Section 6.1), is very different from the corresponding region in the enzyme of non-trypanosomatids. This sequence has also not been found in any other glycosomal protein. Presumably, it interacts with an equally unique area of another glycosomal matrix protein. Therefore, it may be feasible to develop a compound with specificity to prevent or disrupt this interaction and thus to prevent glycosome import of TIM. Similar approaches may be feasible to prevent import of other I-PTS proteins.

7. Feasibility to target transient protein–protein interactions

Designing inhibitors that will interfere with glycosome biogenesis involves other challenges than designing inhibitors of (glycosomal or cytosolic) enzymes. Many enzyme inhibitors are designed or selected to bind in pockets, the binding sites for natural ligands, of the proteins where they exploit the enlarged surface with its specific conformation and presence of reactive residues. This enables the development of inhibitors that interact with very high affinity or irreversibly, and with high specificity to the target. In contrast, compounds meant to disrupt glycosomal biogenesis will target protein–protein interactions. For some time, the design of drugs that act through disrupting protein–protein interactions has been a controversial issue in pharmaceutics. The fact that small molecules should insert themselves between proteins in which large, complementary surfaces are responsible for interactions often with high affinities was considered as a formidable task. Usually, the interfaces are relatively flat, thus not providing a defined binding pocket for an antagonistic small molecule. However, drugs that target protein–protein interactions have already been in use for many years. For example vinca alkaloids, used in treatment of cancer, have been identified in a natural screening program about 50 years ago. But only in 1996 it was shown that vinblastin acts by interfering with the polymerization of the α–β heterodimer of tubulin and consequently with the formation of the mitotic spindle. In the last decade, considerable progress has been made in developing antagonists of a variety of inter-protein interactions. Ribonucleotide reductase antagonists have been obtained which prevent assembly of the two components of Herpes simplex virus enzyme with a *K*$_i$ of 0.3 nM [298]. Another interesting example involves the trisoxazole macrolide toxins which target actins with high affinity and specificity and have promising pharmacological properties. Recently, the binding site of these compounds to actin has been revealed; the results suggested that these macrolides may be viewed as small molecule biomimetics of actin-binding proteins [299].

A number of small molecule antagonists of protein–protein interactions have reached the market or are in various stages of trials for use against all kinds of different diseases, including cancers, vascular and coronary diseases and viral infections (cited in [300]). All these cases do not involve tightly bound proteins as found in stable homodimeric proteins, but heteromeric complexes of which the units are in dynamic equilibrium. It appears to be perfectly possible to disrupt such dynamic protein complexes with small molecules, by peptidomimics or inhibitors (reviewed in [300,301]). The weaker such critical protein–protein interactions are, and the more convoluted the interface, the larger the chance that successful small molecule inhibitors will be developed.

The process of peroxisome and glycosome biogenesis involves transient interactions of a large number of different proteins. In view of the success achieved in developing antagonists of transient protein–protein interactions as drugs and/or lead compounds for other diseases, success with a similar approach for developing anti-trypanosome drugs seems realistic.

8. Conclusions

A major part of the core carbohydrate metabolism, most notably glycolysis, of Trypanosomatidae occurs in glycosomes, remarkable trypanosomatid organelles related to peroxisomes. It has been shown that the correct compartmentalization of glycosomal enzymes involved in glycolysis in bloodstream-form *T. brucei* and the integrity of the organelles are essential to the parasites. Therefore, compounds that would interfere with glycosome biogenesis would potentially be effective drugs.

In the last two decades, much research has been performed to unravel the routes of peroxisome biogenesis. At present, 32 peroxins involved in different aspects of the organelle’s formation have been identified and it has been shown that many of these proteins exert their function by transient interactions with one or multiple other peroxins. These studies have provided a general
picture of peroxisome biogenesis, conserved among eukaryotes, but at the same time also revealed significant differences between organisms such as mammals and yeasts. However, almost all details of the molecular mechanisms of processes such as matrix protein import, formation of the lipid membrane, insertion of proteins in the membrane and energetics of the organelle formation are currently still poorly understood.

Biogenesis of trypanosomatid glycosomes appears to be, in general, similar to that of peroxisomes. Several peroxins have been identified in different human-pathogenic Trypanosomatidae; they are homologous to their counterparts in mammals, yeasts and plants, however with very low percentages of amino-acid sequence identity. The presently available data on the assembly of the different members of the peroxisome family and the sequence and properties of peroxins do not permit us to conclude that glycosome biogenesis more closely resembles the corresponding process in yeasts, plants or mammals.

For various trypanosomatid peroxins it has been shown that decreasing their expression by reverse genetics methodologies leads to growth arrest and death of the parasites. The fact that the peroxins responsible for glycosome biogenesis are essential for the viability of trypanosomes, and the low degree of sequence conservation of peroxins, render these proteins promising targets for the design of anti-parasite drugs that act through preventing the formation of functional organelles and proper metabolic compartmentation. Current research is devoted to determining the transient interactions between different sets of proteins involved in glycosome biogenesis at the atomic and molecular level. This will provide a solid basis for future design of peptidomimetic and other compounds that will selectively interfere with these interactions without affecting protein–protein interactions in the human host.

Acknowledgements

The research on glycosome biogenesis and anti-trypanosomatid drug design in the laboratories of the authors is financially supported through the following grants: to PM from the Belgian ‘Fonds de la Recherche Scientifique Médicale’ (FRSM) and Interuniversity Attraction Poles — Federal Office for Scientific Technical and Cultural Affairs, to EF from the W.M. Keck Foundation Center for Microbial Pathogens at the University of Washington. JM acknowledges a PhD scholarship from the ‘Fonds à la Recherche dans l’Industrie et l’Agriculture (FRIA, Belgium). We are grateful to our coworkers Véronique Hannaeft, Hanane Krazy, Nathalie Galland, Cédric Yernaux, Nathalie Chevalier and Stéphane de Walque (ICP, Brussels) for their contribution to the work cited in this review and sharing unpublished data. We would like to thank Claudia Roach, Stephen Suresh, Abhinav Kumar and David Chudzik (UoW, Seattle) for their contributions to structural biology of trypanosomatid glycosomal proteins. We thank Dr Marc Fransen (‘Katholieke Universiteit Leuven’, Belgium) for stimulating discussions and comments on preliminary versions of the manuscript and Nathalie Chevalier for assistance in the preparation of figures and tables.

References


[219] Baerends, R.J.S., Rasmussen, S.W., Hilbrands, R.E., Van der Heide, M., Faber, K.N., Reuvekamp, P.T.W., Kiel, J.A.K.W.,


