

# The N-terminal Half of the Peroxisomal Cycling Receptor Pex5p is a Natively Unfolded Domain

Andreia F. Carvalho<sup>1,2</sup>, João Costa-Rodrigues<sup>1,2</sup>, Isabel Correia<sup>3</sup>  
João Costa Pessoa<sup>3</sup>, Tiago Q. Faria<sup>4</sup>, Cristina L. Martins<sup>5</sup>, Marc Fransen<sup>6</sup>  
Clara Sá-Miranda<sup>1</sup> and Jorge E. Azevedo<sup>1,2\*</sup>

<sup>1</sup>*Instituto de Biologia Molecular e Celular, Rua do Campo Alegre 823, 4150-180 Porto, Portugal*

<sup>2</sup>*Instituto de Ciências Biomédicas de Abel Salazar Universidade do Porto, Largo do Professor Abel Salazar, 2 4099-003 Porto, Portugal*

<sup>3</sup>*Instituto Superior Técnico Centro de Química Estrutural 1049-001 Lisboa, Portugal*

<sup>4</sup>*Instituto de Tecnologia Química e Biológica, Rua Quinta Grande, 6, 2780-156 Oeiras, Portugal*

<sup>5</sup>*Instituto de Engenharia Biomédica, Universidade do Porto, Laboratório de Biomateriais, Rua do Campo Alegre, 823, 4150-180 Porto Portugal*

<sup>6</sup>*Katholieke Universiteit Leuven Faculteit Geneeskunde, Departement Moleculaire Celbiologie, Herestraat 49 B-3000 Leuven, Belgium*

Targeting of most newly synthesised peroxisomal matrix proteins to the organelle requires Pex5p, the so-called PTS1 receptor. According to current models of peroxisomal biogenesis, Pex5p interacts with these proteins in the cytosol, transports them to the peroxisomal membrane and catalyses their translocation across the membrane. Presently, our knowledge on the structural details behind the interaction of Pex5p with the cargo proteins is reasonably complete. In contrast, information regarding the structure of the Pex5p N-terminal half (a region containing its peroxisomal targeting domain) is still limited. We have recently observed that the Stokes radius of this Pex5p domain is anomalously large, suggesting that this portion of the protein is either a structured elongated domain or that it adopts a low compactness conformation. Here, we address this issue using a combination of biophysical and biochemical approaches. Our results indicate that the N-terminal half of Pex5p is best described as a natively unfolded pre-molten globule-like domain. The implications of these findings on the mechanism of protein import into the peroxisome are discussed.

© 2005 Elsevier Ltd. All rights reserved.

\*Corresponding author

**Keywords:** Pex5p; PTS1-receptor; natively unfolded; peroxisomal protein import

## Introduction

Peroxisomal matrix proteins are synthesized in cytosolic ribosomes and post-translationally imported into the organelle. Two soluble receptor

proteins, Pex5p and Pex7p, promote the specific targeting of these newly synthesized proteins into the peroxisome.<sup>1–5</sup> By far the majority of peroxisomal matrix proteins use Pex5p as their cognate receptor. These proteins possess the so-called peroxisomal targeting signal (PTS) type 1, a tripeptide located at their extreme C terminus with the sequence S-K-L or a variant.<sup>6–8</sup> The molecular basis behind the interaction of Pex5p with PTS1-containing cargo proteins has been extensively characterized. Indeed, a variety of biochemical studies have shown that the C-terminal

Abbreviations used: PTS, peroxisomal targeting sequence; TPRs, tetratricopeptide repeats; DSC, differential scanning calorimetry; ATR-FTIR, attenuated total reflection Fourier transform infrared spectroscopy.

E-mail address of the corresponding author: [jazevedo@ibmc.up.pt](mailto:jazevedo@ibmc.up.pt)

half of Pex5p, a region containing several TPR motifs, interacts with the PTS1 sequence.<sup>7–14</sup> Crystallization of this domain of Pex5p in complex with a PTS1-containing peptide provided the details of this interaction at the atomic level.<sup>15</sup>

According to the current model of protein import into the peroxisome,<sup>1–5</sup> the next step in the Pex5p-mediated protein import pathway is the recognition of the cytosolic Pex5p–cargo protein complex by the so-called peroxisomal docking/translocation machinery. This is a transmembrane multisubunit protein complex comprising Pex2p, Pex10p, Pex12p, Pex13p, Pex14p, Pex8p and Pex17p (these two latter peroxins were only found in lower eukaryotes<sup>16–23</sup>). Somewhere during or after this recognition event, cargo proteins are translocated across the peroxisomal membrane, the third step in this protein import pathway. Although the nature of this step remains to be determined there are now several data suggesting that Pex5p is actively involved in the translocation step (see Discussion). Finally, Pex5p is recycled back to the cytosol in order to catalyse further rounds of protein import.<sup>24–27</sup>

Although this model provides a global perspective on the process of protein import into the peroxisome, it still lacks the required level of detail in order to describe accurately the mechanism of this unique process. Indeed, besides the uncertainties described above concerning the nature of the translocation step and its relationship to the docking (recognition) event, there are numerous aspects that remain enigmatic. The fact that the peroxisome is capable of importing already folded oligomeric proteins, despite the absence of stable large protein pores on its membrane, is perhaps the most puzzling one.<sup>28–31</sup> Clearly, now that most components of this import machinery have already been identified, data on their structure/architecture are needed in order to explain how they function.

Being one of the central components of the peroxisomal protein import machinery, Pex5p has been the target of particular attention. Besides the characterization studies on the Pex5p cargo protein-binding domain mentioned above, intensive research has been focused on the structural/functional relationships in the N-terminal half of this peroxin. Using biochemical and genetic approaches it has been demonstrated that this region of the human PTS1 receptor contains seven binding sites for Pex14p.<sup>32,33</sup> Pex13p, a component of the docking/translocation machinery possibly involved in the recycling step of Pex5p or in the release of cargo proteins from Pex5p, also interacts with this domain of Pex5p.<sup>13,34</sup> In addition, Pex7p, the receptor for PTS2-containing proteins,<sup>35–37</sup> binds to the N-terminal half of mammalian Pex5p, an interaction that is essential for the targeting of these proteins to the organelle.<sup>38–40</sup> Furthermore, both *in vivo* and *in vitro* experiments have shown that this domain of Pex5p alone is sufficient for specific targeting to the peroxisome.<sup>38,41</sup> Data suggesting that C-terminally truncated versions of Pex5p (as the full-length version of the

peroxin) are correctly inserted into the peroxisomal docking/translocation machinery have also been provided.<sup>41</sup> Finally, a role for the first N-terminal amino acid residues of Pex5p in the ATP-dependent export step of this peroxin from the peroxisomal membrane into the cytosol was recently suggested.<sup>24</sup> Thus, all these data point to the possibility that this domain of the PTS1 receptor is crucial not only for the docking/recognition step but also for all the subsequent events occurring at the peroxisomal membrane level.

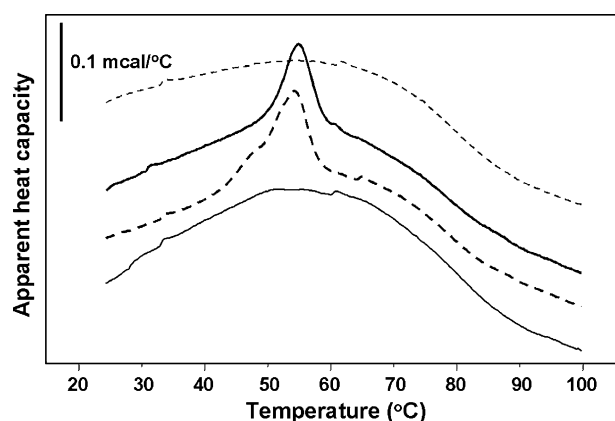
In recent work on the hydrodynamic characterization of human Pex5p it was noticed that the N-terminal half of this protein displays an unusually large Stokes radius.<sup>34</sup> This observation indicates that this fragment of the peroxin is either a structured elongated domain or that it adopts a low compactness conformation. Here, we have addressed this issue using a combination of biophysical and biochemical techniques. Our results indicate that the N-terminal half of Pex5p is a natively unfolded pre-molten globule-like domain. The implications of these findings on the mechanism of protein import into the peroxisome are discussed.

## Results

### Differential scanning calorimetry reveals that the N-terminal half of Pex5p lacks stable tertiary structure

We have shown previously that the N-terminal half of Pex5p (residues 1–324) displays an anomalous high Stokes radius/frictional ratio. Although this observation suggests that this fragment of the protein is not a compact globular/spherical domain, it does not provide any more clues regarding its structure. *A priori*, there are two possibilities (not necessarily mutually exclusive) to explain a larger than expected Stokes radius. The protein may possess a structured but elongated shape. One of the best examples of this type of structure is provided by myosin,<sup>42</sup> a protein in which a major fraction of its polypeptide chain adopts an alpha-helical conformation. Alternatively, the protein may display a large Stokes radius simply because it does not exist in solution as a compact structure. Natively unfolded (or intrinsically unstructured) proteins as the Nup2p protein<sup>43</sup> or the C-terminal domain of caldesmon<sup>44</sup> are extreme examples of this family of proteins.<sup>45</sup>

To discriminate between these two possibilities, we used a combination of biophysical and biochemical approaches to characterize the secondary/tertiary structural properties of human Pex5p. We started these studies by looking at the unfolding behaviour of Pex5p as determined by differential scanning calorimetry (DSC). Besides the full-length version of this peroxin, we included in this analysis two additional recombinant proteins, one corresponding to its N-terminal half (comprising amino acid residues 1–324;  $\Delta$ C1-Pex5p) and



**Figure 1.** DSC thermograms of Pex5p, TPRs-Pex5p and  $\Delta$ C1-Pex5p. Raw data of the thermal unfolding of Pex5p (thick continuous line), TPRs-Pex5p (thick broken line) and  $\Delta$ C1-Pex5p (thin continuous line) are presented. The baseline obtained using the same experimental conditions is also shown (thin broken line). Protein concentrations were 7.36  $\mu$ M for Pex5p, 8.28  $\mu$ M for TPRs-Pex5p and 21.1  $\mu$ M for the  $\Delta$ C1-Pex5p.

the other containing the tetratricopeptide repeat (TPR) domains of Pex5p (the cargo protein-binding domain), corresponding to its C-terminal half (amino acid residues 315–639; TPRs-Pex5p). As shown in Figure 1, analysis of Pex5p by DSC results in an endotherm displaying a single well-defined peak. This observation indicates the existence of cooperativity in the unfolding process of Pex5p, implying that this protein, or at least some of its domain(s) has/have some degree of globularity. The thermodynamic parameters of unfolding obtained from the Pex5p thermogram were estimated (see Materials and Methods). Values of 54.7 ( $\pm$ 0.1)  $^{\circ}$ C, 146 ( $\pm$ 2) kcal/mol and 126 ( $\pm$ 2) kcal/mol were obtained for the melting temperature ( $T_m$ ), calorimetric enthalpy ( $\Delta H_{cal}$ ) and the van't Hoff enthalpy ( $\Delta H_{vh}$ ), respectively. Data suggesting that the major contribution for the endothermic peak observed with Pex5p comes from the unfolding of its C-terminal half were obtained when TPRs-Pex5p was analysed. Indeed, the thermograms obtained with Pex5p and TPRs-Pex5p are quite similar, with the exception that the main endothermic peak detected for TPRs-Pex5p seems to present a small shoulder at around 49  $^{\circ}$ C. This observation could suggest the existence of two transitions during the unfolding process of TPRs-Pex5p although further work is necessary to corroborate this finding. The estimated thermodynamic parameters of unfolding obtained for TPRs-Pex5p (assuming only one transition) are 53.1 ( $\pm$ 0.1)  $^{\circ}$ C, 157 ( $\pm$ 6) kcal/mol and 98 ( $\pm$ 4) kcal/mol for  $T_m$ ,  $\Delta H_{cal}$  and  $\Delta H_{vh}$ , respectively. It should be noted that the TPRs-Pex5p recombinant protein used in this work is known to be fully functional in a PTS1-binding assay (unpublished results). Thus, the differences between Pex5p and TPRs-Pex5p

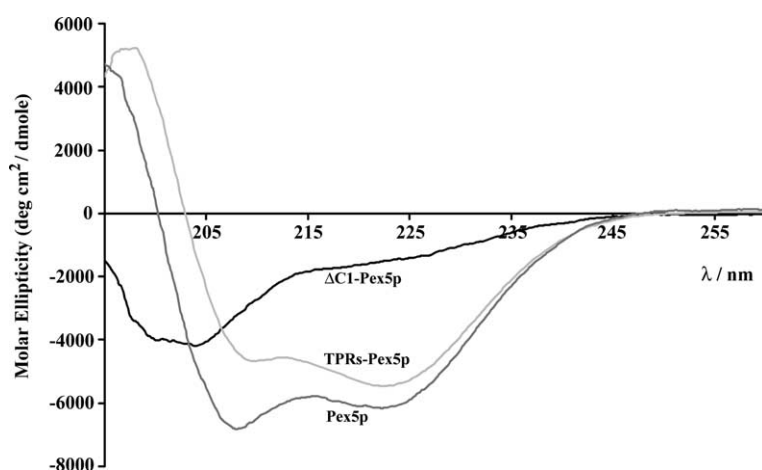
detected in these assays (if any) most likely reflect a context-dependent structural stability of the C-terminal half of Pex5p.

The DSC analysis of  $\Delta$ C1-Pex5p is also presented in Figure 1. Remarkably, the thermogram of this protein corresponds basically to the baseline of the calorimeter. This observation indicates that the N-terminal half of Pex5p lacks any stable tertiary structure.

#### Far-UV CD and ATR-FTIR analysis of Pex5p, TPRs-Pex5p and $\Delta$ C1-Pex5p

The absence of a stable tertiary structure is one of the requisites to classify a given protein as natively unfolded. Another important parameter that should be determined when characterizing a natively unfolded protein is its content in ordered secondary structure. According to Uversky<sup>45</sup> there are two types of natively unfolded proteins. The so-called intrinsic coiled proteins do not possess any (or almost any) ordered secondary structure. Natively unfolded pre-molten globule-like proteins, on the other hand, have some ordered secondary structure. The amount of ordered secondary structure detected in this sub-family of natively unfolded proteins is highly variable but values of up to 50% of alpha-helical plus beta-sheet structure have been reported (see Lydakakis-Simantiris *et al.*<sup>46</sup> and references cited therein).

In order to assess the type and amount of secondary structure in Pex5p, the complete version of this peroxin, TPRs-Pex5p and  $\Delta$ C1-Pex5p were analysed by CD and attenuated total reflection Fourier transform infrared spectroscopy (ATR-FTIR). The far UV CD spectra obtained with these proteins are shown in Figure 2. TPRs-Pex5p displays two maxima of negative ellipticity: one at 210 nm and the other at 222 nm. Analysis of this spectrum using the CDPPro program<sup>47</sup> allowed us to estimate the amount of the different types of secondary structure present in TPRs-Pex5p (see Table 1). As expected, alpha-helix is, by far, the most abundant type of secondary structure present in this domain of Pex5p: a value of 52% was obtained. This value is 10% lower than the one obtained from the three-dimensional structure of residues 318–639 of human Pex5p in complex with a PTS1-containing peptide.<sup>15,48</sup> Whether the CD-derived data reflect a different secondary structure content of the unliganded Pex5p domain or simply experimental variation remains to be determined. The CD spectrum obtained with Pex5p also shows two maxima of negative ellipticity. In this case, however, the first maximum is shifted to 208 nm and is clearly more intense than the one at 222 nm. Analysis using the CDPPro program indicates that Pex5p contains less alpha-helical structure (on a percent basis) than its C-terminal half (33% versus 52%, respectively). This decrease in alpha-helical content is compensated by an increase in other types of secondary structure, particularly in unordered structure (38% versus 23% in Pex5p and TPRs-Pex5p, respectively).



**Figure 2.** Far-UV CD spectra of Pex5p, TPRs-Pex5p and  $\Delta$ C1-Pex5p. The CD spectra of Pex5p (dark-grey line), TPRs-Pex5p (light-grey line) and  $\Delta$ C1-Pex5p (black line) are presented. The molar ellipticities are shown.

Analysis of the CD spectrum obtained for  $\Delta$ C1-Pex5p provides the explanation for this increase. This protein displays a maximum of negative ellipticity at 203 nm. Intense negative ellipticity in the vicinity of 200 nm is a property of unfolded polypeptides.<sup>45</sup> Accordingly, spectrum analysis predicts that unordered structure is the most prevalent type of structure in  $\Delta$ C1-Pex5p (57%), with beta-sheet structure presenting the second highest percentage.

Qualitatively similar conclusions regarding the type of secondary structure present in these recombinant proteins were obtained by attenuated total reflection fourier transform infrared spectroscopy (ATR-FTIR). As shown in Figure 3, from all the three analysed proteins, TPRs-Pex5p is the one presenting the highest amount of alpha-helical structure. On the other hand,  $\Delta$ C1-Pex5p is the most enriched in random coil structure. Estimates for the amounts of the different types of secondary structure present in these recombinant proteins are provided in Table 1. It is clear that the obtained values are considerably different from those derived from CD spectroscopy. The fact that the FTIR spectra were obtained with partially dehydrated adsorbed protein films (a technique that may lead to some secondary structure changes in the analysed proteins<sup>49</sup>) together with the inherent difficulties in the deconvolution of this type of spectra (for a thorough discussion see Nevea *et al.*<sup>50</sup>) may explain these discrepancies. Regardless of the reasons behind these differences, the important conclusion to be drawn from these experiments is that the structure of the N-terminal half of Pex5p, although largely unordered, does contain elements with an ordered secondary structure. Thus, this domain of Pex5p cannot be considered a natively unfolded intrinsic coil.

#### Size-exclusion chromatography analysis of C-terminal truncated versions of Pex5p

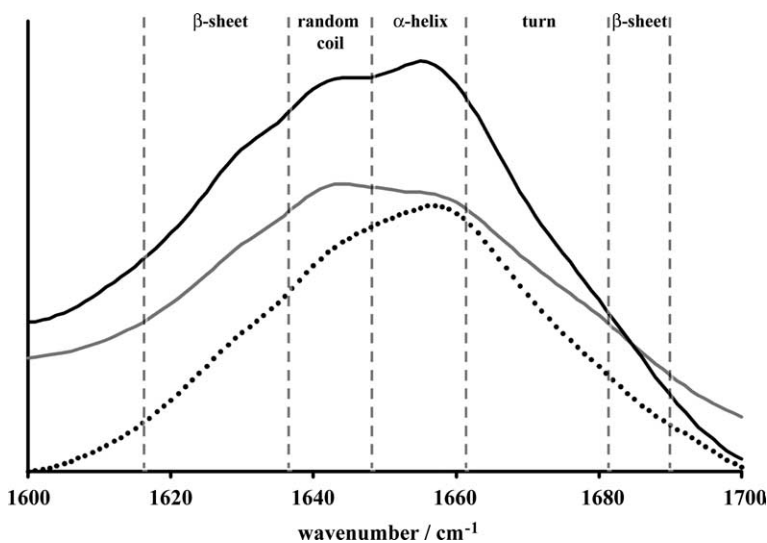
One parameter frequently used to classify the conformational state of a given protein is its

hydrodynamic dimension.<sup>45</sup> Indeed, clear empirical relationships between the Stokes radii and the molecular masses of proteins adopting different conformational states have been described.<sup>44,45</sup> Of relevance here, are the relationships describing how the Stokes radius of a protein increases with its molecular mass assuming a natively unfolded intrinsic coil conformation, a natively unfolded pre-molten globule-like conformation or a globular (compact) conformation. These relationships are described by straight lines when the logarithm of the Stokes radius is plotted against the logarithm of the molecular mass.<sup>44</sup> We have recently described that the Stokes radius of an *in vitro* synthesized protein fragment corresponding to amino residues 1–324 of human Pex5p is  $5(\pm 0.25)$  nm.<sup>34</sup> As shown in Figure 4(a) and (b), a similar value (4.6 nm) was obtained when recombinant  $\Delta$ C1-Pex5p was analysed by size-exclusion chromatography. Plotting the molecular mass and the Stokes radius of this protein into the graphic presented in Figure 4(c), reveals that  $\Delta$ C1-Pex5p falls in a region between the two straight lines corresponding to natively unfolded intrinsic coils and natively unfolded pre-molten globules. A similar conclusion is obtained when shorter C-terminal truncated versions of Pex5p are included in this analysis. As shown in Figure 4, the Stokes radii of polypeptides comprising amino acid residues 1–110, 1–197 or 1–268 of Pex5p are compatible with the existence of a natively unfolded conformation in these regions of the protein. These results also suggest that

**Table 1.** Secondary structure content of Pex5p, TPRs-Pex5p and  $\Delta$ C1-Pex5p as determined by CD and ATR-FTIR spectroscopies

	CD				ATR-FTIR			
	% $\alpha^a$	% $\beta$	% T	% U	% $\alpha$	% $\beta$	% T	% C
Pex5p	33	12	17	38	49	10	19	22
TPRs-Pex5p	52	7	18	23	53	16	14	17
$\Delta$ C1-Pex5p	12	17	14	57	36	8	14	42

<sup>a</sup>  $\alpha$ , Alpha-helix;  $\beta$ , beta-sheet; T, turn; U, unordered; C, coil.



**Figure 3.** ATR-FTIR spectra of Pex5p,  $\Delta$ C1-Pex5p and TPRs-Pex5p. The amide I region of the FTIR spectra obtained for Pex5p (black line),  $\Delta$ C1-Pex5p (grey line) and TPRs-Pex5p (dotted line) are presented. The regions delimited by vertical broken lines correspond to the different types of secondary structure (indicated at the top of the Figure) as described by Vigano *et al.*<sup>65</sup>

the secondary structure elements present in the N-terminal half of Pex5p do not give rise to globular domains, corroborating the results obtained by DSC. The fact that the coordinates of the proteins analysed in these experiments do not coincide exactly with the straight lines corresponding to the two families of natively unfolded proteins is not unexpected. Indeed, as stated above, the secondary structure content of pre-molten globule natively unfolded proteins is highly variable. Thus, it is plausible that these two families represent the extremes on a continuum rather than specific classes.

**The N-terminal half of Pex5p is highly sensitive to mild proteolysis, is resistant to heat-treatment and has an amino acid composition typical of natively unfolded proteins**

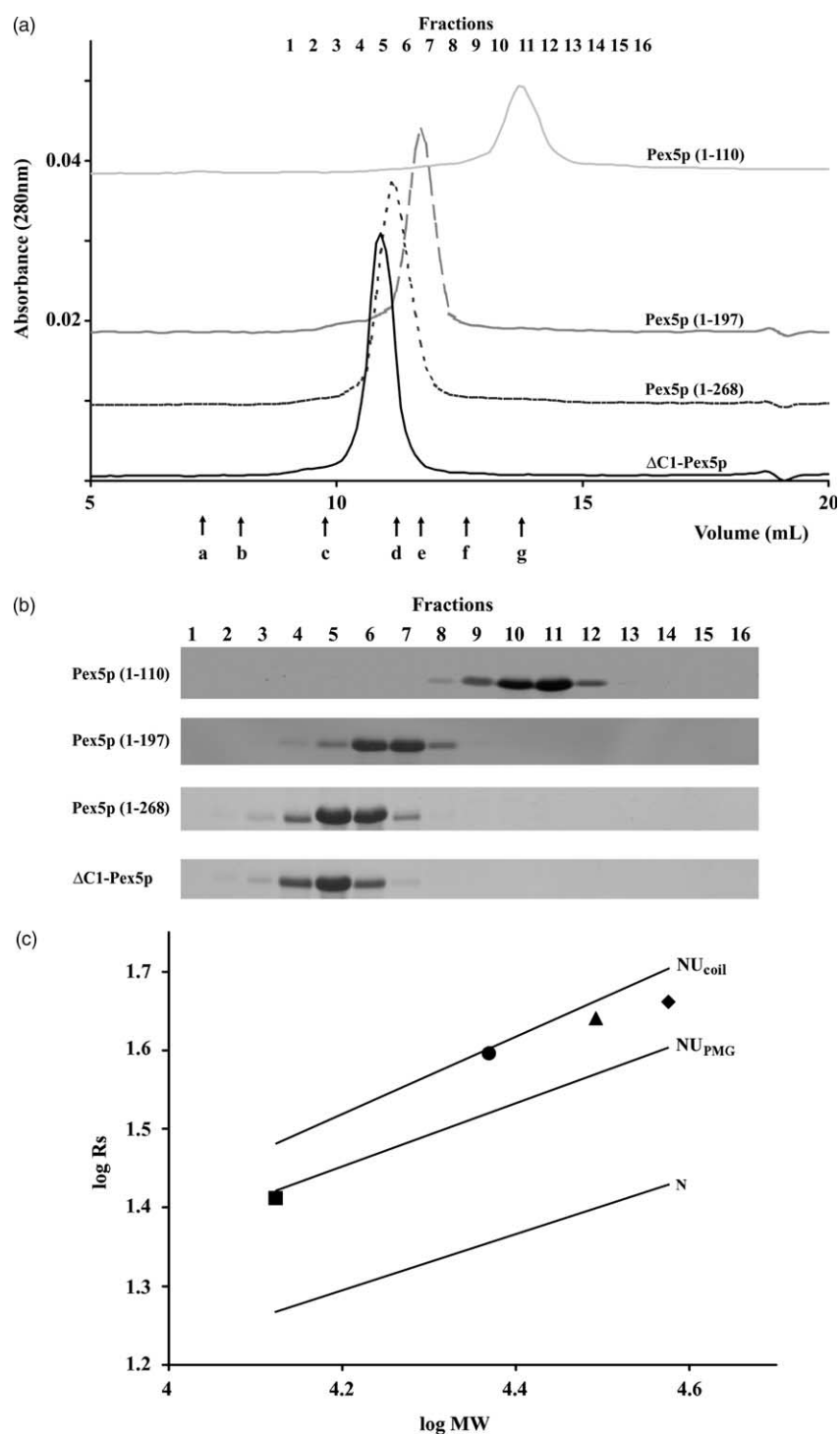
Due to the flexibility of their polypeptide chains, natively unfolded proteins are easily hydrolysed when incubated with low concentrations of proteases. As shown in Figure 5(a) treatment of  $\Delta$ C1-Pex5p with only 125 ng/ml of proteinase K results in extensive hydrolysis of the recombinant protein (note the Coomassie blue-stained material running at the electrophoretic front in lane 7 but not in lane 6). Under the same conditions TPRs-Pex5p is only partially degraded (Figure 5(a), compare lanes 11 and 12). Naturally, the degree of proteolysis increases when the protease concentration is raised but large protein fragments (apparent molecular masses of 14–20 kDa) are still visible even at the highest protease concentration used in this experiment (lane 15; 1  $\mu$ g/ml). It should be noted that the three-dimensional structure of the C-terminal half of Pex5p suggests the presence of several unstructured regions<sup>15,48</sup> (see also below). Thus, a complete resistance to proteolysis as the one observed, for example, when albumin is used in these assays (data not shown) was not expected for recombinant TPRs-Pex5p. The proteolytic profile obtained for Pex5p is, perhaps, the most informative. At low proteinase K concentrations (e.g. 125 ng/ml of

protease; lane 2) numerous protein bands displaying apparent molecular masses of 33–90 kDa are detected. When the protease concentration is increased (lanes 3–5) most of these bands disappear while the amounts of two polypeptides displaying apparent molecular masses of 33 kDa and 35 kDa increase. The 35 kDa polypeptide was shown before to correspond to the C-terminal half of Pex5p.<sup>34</sup> Edman degradation of the 33 kDa polypeptide provided essentially the same conclusion (unpublished results). We conclude that the N-terminal half of Pex5p is highly sensitive to proteolytic degradation. Interestingly, the protease sensitivities of the two Pex5p domains characterized in this experiment decrease when the complete Pex5p protein is analysed. Steric hindrance probably explains this difference, i.e. one domain partially protects the other from proteolytic attack and thus higher protease concentrations have to be used with Pex5p in order to obtain the degree of proteolysis observed with the isolated N and C-terminal halves of the peroxin.

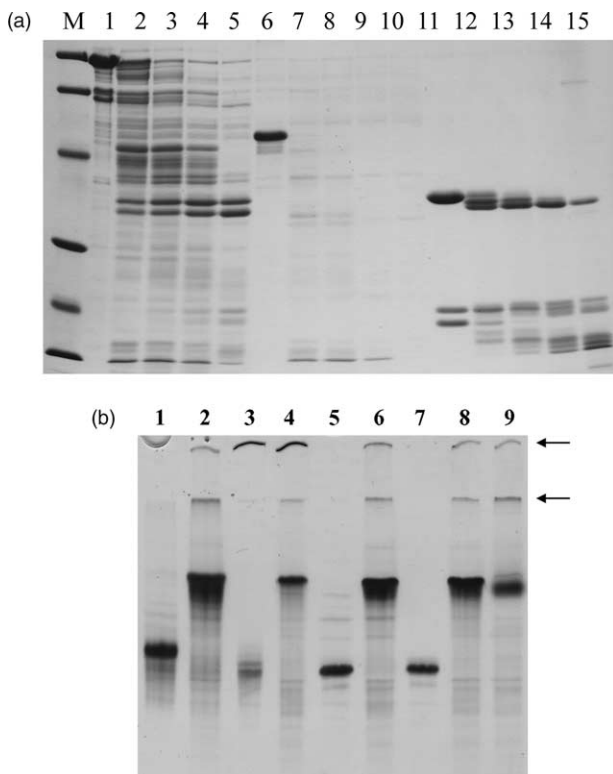
Another characteristic of natively unfolded proteins is their resistance to heat treatment. In order to extend the findings described above we incubated  $\Delta$ C1-Pex5p and Pex5p at 95 °C for 30 min and analysed the electrophoretic behaviour of these proteins as well as their capacity to interact with the Pex5p-binding domain of human Pex14p<sup>33</sup> by native PAGE. As shown in Figure 5(b), heat-treated Pex5p forms high molecular mass aggregates remaining at the wells of the gel. In contrast, both the electrophoretic mobility of  $\Delta$ C1-Pex5p (a function of both its charge/mass ratio and Stokes radius) as well as its capacity to interact with the Pex5p-binding domain of human Pex14p remain unchanged by the heat-treatment.

Finally, we have analysed the primary structure of human Pex5p using the several predictors of natural disordered regions available†. Analysis of

† [www.pondr.com](http://www.pondr.com)



**Figure 4.** Size-exclusion chromatography of N-terminal fragments of human Pex5p. (a)  $\Delta$ C1-Pex5p and recombinant proteins comprising amino acid residues 1–110, 1–197 and 1–268 of human Pex5p (Pex5p(1–110), Pex5p(1–197) and Pex5p(1–268), respectively) were analyzed by size-exclusion chromatography. Numbers at the top of the chromatogram indicate the fractions collected and analysed by SDS-PAGE. The Coomassie blue-stained gels are shown in (b). The void volume in (a) and the elution volumes of tyroglobulin in (b), ferritin in (c), aldolase in (d), bovine serum albumin in (e), ovalbumin in (f) and soybean trypsin inhibitor in (g) are indicated. The determined Stokes radii were 2.6 nm for Pex5p(1–110), 3.9 nm for Pex5p(1–197), 4.4 nm for Pex5p(1–268) and 4.6 nm for  $\Delta$ C1-Pex5p. The logarithms of these values ( $R_s$ ) versus the logarithms of the molecular masses (MW) of the corresponding proteins are plotted into the graphic shown in (c). Pex5p(1–110), Pex5p(1–197), Pex5p(1–268) and  $\Delta$ C1-Pex5p correspond to the square, circle, triangle and rhombus, respectively. The straight lines show how the Stokes radius increases with the molecular mass of proteins assuming a natively unfolded intrinsic coil conformation ( $NU_{coil}$ ), a natively unfolded pre-molten globule-like conformation ( $NU_{PMG}$ ) or a globular conformation (N). The equations of these lines were taken from Permyakov *et al.*<sup>44</sup>



**Figure 5.** The N-terminal half of Pex5p is highly susceptible to proteolysis and resistant to heat treatment. (a) Five micrograms of Pex5p (lanes 1–5),  $\Delta$ C1-Pex5p (lanes 6–10) or TPRs-Pex5p (lanes 11–15) were incubated on ice for 30 min in the absence (lanes 1, 6 and 11) or in the presence of proteinase K at final concentrations of 0.125  $\mu$ g/ml (lanes 2, 7 and 12), 0.25  $\mu$ g/ml (lanes 3, 8 and 13), 0.5  $\mu$ g/ml (lanes 4, 9 and 14) or 1  $\mu$ g/ml (lanes 5, 10 and 15). After inactivation of the protease with phenylmethylsulphonyl fluoride the samples were precipitated with trichloroacetic acid and analysed by SDS-PAGE. M, molecular mass standards (in kDa; from top to bottom): 97, 66.2, 45.0, 31.0, 21.5 and 14.4. Note that a small fraction of the TPRs-Pex5p protein used in this experiment is slightly degraded (polypeptides displaying apparent molecular masses of 18 kDa and 20 kDa present in lane 11). (b). Recombinant Pex5p (lanes 1–4) and  $\Delta$ C1-Pex5p (lanes 5–8) were either kept on ice (lanes 1, 2, 5 and 6) or heated at 95 °C for 30 min and then put on ice (lanes 3, 4, 7 and 8). Five micrograms of a recombinant protein comprising the first 80 amino acid residues of human Pex14p (lane 9) were added to the samples of lanes 2, 4, 6 and 8. The other samples received the corresponding amount of buffer. All samples were incubated for 10 min at 37 °C before loading onto the gel. Formation of a protein complex involving Pex5p or  $\Delta$ C1-Pex5p on one side and the Pex14p protein fragment on the other results in intense and diffuse bands appearing slightly above the Pex14p fragment (compare lanes 2, 4, 6 and 8 with lane 9). Note that heat-treated Pex5p (lanes 3 and 4) aggregates and remains at the wells of the gel (indicated by the first arrow from the top at the right-hand side of the gel). Our preparations of recombinant Pex5p frequently contain small amounts of C-terminally truncated Pex5p (see Costa-Rodrigues *et al.*<sup>34</sup>). These truncated Pex5p forms are visible in lane 3 (i.e. they also display heat-resistance) and explain the small amount of Pex14p-containing protein complex visible in lane 4. The Pex14p recombinant protein used in this experiment contains some

the first 324 amino acid residues of human Pex5p (corresponding to our  $\Delta$ C1-Pex5p protein) with the charge-hydrophathy method<sup>45</sup> predicts that this domain is a natively unfolded protein. In contrast, amino acid residues 325–639 of Pex5p are predicted as ordered. When the VSL1 algorithm<sup>51</sup> was used, 68% of the first 324 amino acid residues of Pex5p were classified as disordered. In contrast, from the remaining 314 residues of Pex5p (residues 325–639) only 13% are predicted as disordered. Significantly, three of the four small regions in the C-terminal half of Pex5p classified as disordered by this algorithm correspond to regions of missing electronic density in the crystal structure of this domain of Pex5p, illustrating the predictive value of this program.

In conclusion, the data described above indicate that the N-terminal half of human Pex5p is best described as a natively unfolded pre-molten globule.

## Discussion

Here we describe the structural properties of the N-terminal half of human Pex5p, the peroxisomal cycling receptor. Using a combination of biophysical and biochemical approaches it is shown that this domain of the PTS1 receptor displays the characteristics of a natively unfolded pre-molten globule-like protein. Thus, our data provide the molecular basis to explain why the Stokes radius of Pex5p is abnormally large, an observation that led so many authors to propose a tetrameric structure for this peroxin (e.g. Gould & Collins<sup>52</sup>).

Several independent data have shown that this domain of Pex5p mediates the interaction of Pex5p with the peroxisomal membrane,<sup>13,38,41</sup> a property that derives from the existence in this portion of the PTS1 receptor of several binding sites for Pex14p and Pex13p, two components of the peroxisomal docking/translocation machinery.<sup>13,32–34</sup> Thus, the findings reported here have implications on the mechanism of protein targeting into the peroxisomal compartment.

One implication of our data regards the association event of cytosolic Pex5p–cargo protein complexes with the docking/translocation machinery. Natively unfolded proteins have major advantages over globular proteins in establishing protein–protein interactions. First, a flexible unstructured polypeptide has a greater capture radius for a specific binding site than a globular rigid protein, which in turn increases the rate of protein–protein interaction.<sup>53,54</sup> Applied to the Pex5p-mediated protein transport process, this concept translates into faster protein import rates. Secondly, although many native unfolded proteins display disorder–order transitions (see below), they tend to preserve large physical

material staying at the wells of the gel and at the interface of the stacking and separation gels (second arrow from the top at the right-hand side of the gel).

dimensions when they complex with their partners, thus providing "for a disproportionately large binding surface and multiple contact points for a protein of the given size".<sup>55</sup> Such property explains why a relatively short protein domain of Pex5p is capable of interacting simultaneously with seven Pex14p molecules (see discussion in Schliebs *et al.*<sup>33</sup>).

The finding that the N-terminal half of Pex5p is natively unfolded may also have impact on the mechanism of protein translocation across the peroxisomal membrane. Indeed, several data suggesting that Pex5p itself promotes the translocation step of PTS1-containing proteins have been described. These include the membrane topology of Pex5p transiently associated with the peroxisomal membrane,<sup>26,56</sup> the demonstration that Pex5p becomes exposed to the peroxisomal lumen during its protein transport cycle *in vivo*,<sup>57</sup> the observation that cargo-dependent insertion of Pex5p into the peroxisomal membrane does not require energy in the form of ATP or a membrane potential<sup>27,58,59</sup> and the fact that besides Pex5p no other peroxin is capable of interacting directly with PTS1-containing proteins (e.g. Gould & Collins<sup>52</sup>). Based on these findings it was recently proposed that one possible mechanism to explain how Pex5p may accomplish this task is to assume that its N-terminal half displays the properties of a lever capable of coupling the force generated during its association with the components of the docking/translocation machinery to the membrane translocation of its C-terminal half.<sup>60</sup> In the light of the results described here it may seem difficult to envisage how a flexible disordered protein domain is capable of transmitting mechanical force to the cargo-loaded C-terminal half of Pex5p. There is, however, an important property of many native unfolded proteins that may be relevant here. This is their capacity to oscillate between a disordered and an ordered state upon interacting with their binding partners (see Uversky<sup>45</sup> and Tompa<sup>55</sup> and references cited therein). Although such capacity remains to be shown for human Pex5p, available data for the *Hansenula polymorpha* PTS1-receptor makes this possibility highly plausible. Indeed, van der Klei and co-workers reported that the hydrodynamic radius of a Pex5p–Pex8p protein complex is smaller than the one displayed by Pex5p alone.<sup>61</sup> As recently discussed,<sup>34</sup> this observation strongly suggests the existence of a drastic conformational alteration in *H. polymorpha* Pex5p upon interacting with Pex8p. In the light of our present results this conformational alteration necessarily implies a disorder–order transition in Pex5p. Thus, it remains possible that the N-terminal half of Pex5p acquires rigidity concomitantly with its interaction with the peroxisomal docking/translocation machinery, thus providing the lever-like action necessary for protein translocation across the organelle membrane. Testing this hypothesis and, if true, determining the details of such transition will be demanding tasks that must be performed if we want to understand the process of peroxisomal protein import.

## Materials and Methods

### Recombinant proteins

The recombinant His-tagged human Pex5p was obtained as described.<sup>34</sup> For the production of His-tagged C-terminal truncated versions of human Pex5p comprising amino acid residues 1–324 ( $\Delta$ C1-Pex5p), 1–268, and 1–197 the plasmid pQE30-Pex5<sup>34</sup> was used as a template in PCR reactions using the forward primer 5'-GGGAGAGCCAC-CATGAGAGGATCGCATCAC-3' and the reverse primers 5'-CGCCGTGGTACCTCATTAGTACCCCTTATCATA-3' (for  $\Delta$ C1-Pex5p), 5'-GCGGCGGGTACCTCAAGATGT-GTTACTGGTCTT-3' (for Pex5p residues 1–268), and 5'-GCGGCGGGTACCTCATTACGTGTGCTGCAGAT-3' (for Pex5p residues 1–197). The amplified fragments were digested with BamHI and KpnI, and inserted into pQE-30 (QIAGEN) according to the manufacturer's instructions. The cDNA encoding the first 110 amino acid residues of Pex5p was obtained from pGEM-4-Pex5p<sup>62</sup> by PCR using the forward primer 5'-CCGGCATGCGCAATGCGG-GAGCTGGTGG and the reverse primer 5'-CCGGTC-GACGCCACACCAGGGCTCTCTGG-3'. The fragment was cloned into the pGEM<sup>®</sup>-T easy vector according to the manufacturer's instructions (Promega). The recombinant plasmid was digested with SphI and Sall, and the insert was cloned into pQE-30. Expression of these Pex5p truncated versions was performed in the XL1 strain of *Escherichia coli*. For the production of His-tagged C-terminal half of Pex5p containing amino acid residues 315–639 of this protein (TPRs-Pex5p) the plasmid pGEM-4-Pex5p was used in a PCR reaction using the forward primer 5'-GCGGCATGCGACCTTACGTCAGCTACCTAT-GAT-3' and the reverse primer 5'-GCGGTCGACT-CACTGGGGCAGGCCAAACAT-3'. The amplified DNA fragment was cloned into the pGEM<sup>®</sup>-T easy vector, which was then digested with SphI and Sall and the insert was cloned into pQE-30. Expression was performed in the M15 strain of *E. coli* as described.<sup>15</sup> For the production of the His-tagged fusion protein containing the first 80 amino acid residues of human Pex14, the plasmid pMF101<sup>63</sup> was subjected to a PCR reaction using the forward primer 5'-GGGGGATCCATGGCGTCTCGGAGCAG-3' and the lower primer 5'-GGGCTGCAGCAAGGACGA-AGGTCATCGGC-3'. The amplified fragment was digested with BamHI and PstI, and cloned into the pQE-30 expression plasmid.

Pelleted cells from induced cultures were cooled on ice and lysed by sonication in a buffer containing 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA-NaHO (pH 8.0), 1 mM DTT, 0.1% (w/v) Triton X-100 and 0.1 mg/ml of phenylmethylsulfonyl fluoride. All the proteins were purified by affinity chromatography using HIS-Select<sup>™</sup> Nickel Affinity Gel according to the manufacturer instructions (SIGMA). When necessary proteins were concentrated and the buffer exchanged by six cycles of centrifugation and dilution with the relevant buffer using a Vivaspin 10,000 MWCO PES concentrator (VIVASCIENCE), as described.<sup>34</sup>

### Differential scanning calorimetry (DSC)

DSC scans were performed on a MicroCal VP-DSC MicroCalorimeter controlled by the VP-viewer program and equipped with 0.51 ml cells. Calibration of temperature and heat-flow were carried out according to MicroCal instructions. Good functioning of the calorimeter was attested from the very high reproducibility of



baselines obtained. Protein solutions (in 10 mM sodium phosphate buffer, pH 7.5) were degassed for 8 min before the calorimetric acquisitions. DSC scans were run at a constant heating rate of 1 deg.C/min from 20 °C to 100 °C and an overpressure of about 30 psi was applied to the calorimeter cells to prevent bubble formation during heating. Thermogram analysis was performed using the software supplied with the calorimeter. Raw calorimetric data were converted to the excess heat capacity of unfolding by subtracting the instrumental baselines determined under identical conditions and dividing it by the scan rate and by the sample protein concentration. The protein melting temperature ( $T_m$ ) the calorimetric ( $\Delta H_{cal}$ ) and the van't Hoff ( $\Delta H_{vh}$ ) enthalpies were calculated. Calorimetric enthalpy is the experimental value of the unfolding energy process determined by numerical integration of the area below the endothermic peak, and the van't Hoff enthalpy is based on the endothermic peak shape and reflects the dependence of the equilibrium constant with the temperature determined by the van't Hoff equation.

### Circular dichroism (CD)

The spectra were measured using a cylindrical 1 mm Hellma Quartz Suprasil cuvette, on a Jasco J-720 spectropolarimeter in the wavelength range 185–260 nm (whenever possible) in 10 mM sodium phosphate buffer (pH 7.5) at room temperature. The measuring conditions were: band width, 0.5 nm; response, 4 s; data pitch, 0.2 nm; scanning speed, 50 nm/min; number of accumulations, 5. The analysis of the circular dichroism spectra for the determination of protein secondary structure was carried out using the CDPro suite of programs.<sup>47</sup> The calculations were made using the three programs included in this software: SELCON3, CONTINLL and CDSSTR. The calculations for the TPRs-Pex5p and  $\Delta C1$ -Pex5p proteins were done with 51 CD points between 190 nm and 240 nm. For the Pex5p protein only the 46 CD points in the range 195–240 nm were used. The reference basis set of proteins SDP48 provided by the CDPro software was used. The SDP48 basis set comprises CD and structural data on 43 soluble proteins and five denatured/disordered proteins. The secondary structure of the denatured proteins included in this basis set was estimated to be 90% unordered and 2% of each of the remaining secondary structures,<sup>64</sup> since structures for these proteins are not available. The SDP48 set is expected to be particularly helpful for the determination of secondary structure when the proteins have large unordered components.<sup>64</sup> In the CDPro software the alpha-helical and beta-sheet structures are split into regular and distorted classes, so six secondary structural classes are considered: regular alpha-helix, distorted alpha-helix, regular beta-strand, distorted beta-strand, turns and unordered. In our results combined values of regular and distorted structures are presented. The reported values are the averages obtained with the three programs of CDPro.

### Attenuated total reflection Fourier transform infrared spectroscopy (ATR-FTIR)

ATR-FTIR spectra were performed, at room temperature, on an FTIR spectrophotometer from Perkin Elmer (model 2000) equipped with a liquid nitrogen-cooled mercury-cadmium-telluride detector and coupled with an ATR cell (Graseby SPECAC) with a ZnSe crystal with

an aperture angle of 45°. The proteins (100  $\mu$ l of a 1 mg/ml solution in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl) were applied onto the surface of the crystal and the solvent evaporated under ambient conditions as described.<sup>49</sup> For each spectrum, 1000 scan cycles at a nominal resolution of 4  $\text{cm}^{-1}$  were averaged. The determination of the secondary structure was based on the vibrational bands of the protein in the amide I band region (1600–1700  $\text{cm}^{-1}$ ), which is sensitive to the secondary structure. The spectra were corrected for background and buffer as described<sup>49</sup> and the second derivatives were used to locate the frequency of the main adsorbing components. Deconvolution of spectra was performed using the program PeakFit® (Systat Software Inc.). The frequency limits for each type of secondary structure were assigned according to Vigano *et al.*<sup>65</sup>

### Size exclusion chromatography

Pex5p truncated versions (50  $\mu$ g of protein) in 200  $\mu$ l of a buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA-NaOH (pH 7.5) and 1 mM DTT were analysed on a Superose 12 column HR 10/30 (Amersham Biosciences) running with the same buffer at a flow rate of 0.5 ml/min at 20 °C. Standards used to calibrate the column were (numbers in parentheses are the Stokes radii): thyroglobulin (8.5 nm), ferritin (6.1 nm), aldolase (4.81 nm), bovine serum albumin (3.6 nm), ovalbumin (3.05 nm) and soybean trypsin inhibitor (2.26 nm). Fractions of 500  $\mu$ l were collected, subjected to TCA-precipitation and analysed by SDS-PAGE. The Stokes radii of the Pex5p truncated proteins were determined as described.<sup>66</sup>

### Native polyacrylamide gel electrophoresis

For native gels the amount of protein used was 5  $\mu$ g for each of the Pex5p proteins and for the Pex14p fragment. The proteins were incubated for 10 min at 37 °C in 50 mM Tris-HCl (pH 8.0), 1 mM EDTA-NaOH (pH 8.0), 1 mM DTT and then applied onto the gel. Tris non-denaturing discontinuous 9% polyacrylamide gel electrophoresis was carried out as described.<sup>67</sup> The gels were run at 250 V for 3 h at 4 °C and stained with Coomassie brilliant blue R-250.

### Miscellaneous

Pex5p protein concentrations were determined from the absorbance at 280 nm in buffer containing 6 M guanidine, 20 mM phosphate (pH 6.5), and using the extinction coefficients calculated by ProtParam<sup>†</sup>. Protease assays were performed exactly as described.<sup>34</sup>

---

## Acknowledgements

This work was supported by Fundação para a Ciência e Tecnologia and FEDER funds. J. C.-R. and A. F. C. thank F. C. T., Portugal, for financial support.

## References

1. Brown, L. A. & Baker, A. (2003). Peroxisome biogenesis and the role of protein import. *J. Cell. Mol. Med.* **7**, 388–400.
2. Heiland, I. & Erdmann, R. (2005). Biogenesis of peroxisomes. Topogenesis of the peroxisomal membrane and matrix proteins. *FEBS J.* **272**, 2362–2372.
3. Lazarow, P. B. (2003). Peroxisome biogenesis: advances and conundrums. *Curr. Opin. Cell Biol.* **15**, 489–497.
4. Moyersoen, J., Choe, J., Fan, E., Hol, W. G. & Michels, P. A. (2004). Biogenesis of peroxisomes and glycosomes: trypanosomatid glycosome assembly is a promising new drug target. *FEMS Microbiol. Rev.* **28**, 603–643.
5. Weller, S., Gould, S. J. & Valle, D. (2003). Peroxisome biogenesis disorders. *Annu. Rev. Genomics Hum. Genet.* **4**, 165–211.
6. Gould, S. J., Keller, G. A., Hosken, N., Wilkinson, J. & Subramani, S. (1989). A conserved tripeptide sorts proteins to peroxisomes. *J. Cell Biol.* **108**, 1657–1664.
7. Klein, A. T., Barnett, P., Bottger, G., Konings, D., Tabak, H. F. & Distel, B. (2001). Recognition of peroxisomal targeting signal type 1 by the import receptor Pex5p. *J. Biol. Chem.* **276**, 15034–15041.
8. Lametschwandtner, G., Brocard, C., Fransen, M., Van Veldhoven, P., Berger, J. & Hartig, A. (1998). The difference in recognition of terminal tripeptides as peroxisomal targeting signal 1 between yeast and human is due to different affinities of their receptor Pex5p to the cognate signal and to residues adjacent to it. *J. Biol. Chem.* **273**, 33635–33643.
9. Brocard, C., Kragler, F., Simon, M. M., Schuster, T. & Hartig, A. (1994). The tetratricopeptide repeat-domain of the PAS10 protein of *Saccharomyces cerevisiae* is essential for binding the peroxisomal targeting signal-SKL. *Biochem. Biophys. Res. Commun.* **204**, 1016–1022.
10. Dodt, G., Braverman, N., Wong, C., Moser, A., Moser, H. W., Watkins, P. *et al.* (1995). Mutations in the PTS1 receptor gene, PXR1, define complementation group 2 of the peroxisome biogenesis disorders. *Nature Genet.* **9**, 115–125.
11. Fransen, M., Brees, C., Baumgart, E., Vanhooren, J. C., Baes, M., Mannaerts, G. P. & Van Veldhoven, P. P. (1995). Identification and characterization of the putative human peroxisomal C-terminal targeting signal import receptor. *J. Biol. Chem.* **270**, 7731–7736.
12. Gatto, G. J., Jr, Geisbrecht, B. V., Gould, S. J. & Berg, J. M. (2000). A proposed model for the PEX5-peroxisomal targeting signal-1 recognition complex. *Proteins: Struct. Funct. Genet.* **38**, 241–246.
13. Otera, H., Setoguchi, K., Hamasaki, M., Kumashiro, T., Shimizu, N. & Fujiki, Y. (2002). Peroxisomal targeting signal receptor Pex5p interacts with cargoes and import machinery components in a spatiotemporally differentiated manner: conserved Pex5p WXXXF/Y motifs are critical for matrix protein import. *Mol. Cell. Biol.* **22**, 1639–1655.
14. Terlecky, S. R., Nuttley, W. M., McCollum, D., Sock, E. & Subramani, S. (1995). The *Pichia pastoris* peroxisomal protein PAS8p is the receptor for the C-terminal tripeptide peroxisomal targeting signal. *EMBO J.* **14**, 3627–3634.
15. Gatto, G. J., Jr, Geisbrecht, B. V., Gould, S. J. & Berg, J. M. (2000). Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. *Nature Struct. Biol.* **7**, 1091–1095.
16. Agne, B., Meindl, N. M., Niederhoff, K., Einwachter, H., Rehling, P., Sickmann, A. *et al.* (2003). Pex8p: an intraperoxisomal organizer of the peroxisomal import machinery. *Mol. Cell.* **11**, 635–646.
17. Albertini, M., Girzalsky, W., Veenhuis, M. & Kunau, W. H. (2001). Pex12p of *Saccharomyces cerevisiae* is a component of a multi-protein complex essential for peroxisomal matrix protein import. *Eur. J. Cell Biol.* **80**, 257–270.
18. Reguenga, C., Oliveira, M. E., Gouveia, A. M., Sa-Miranda, C. & Azevedo, J. E. (2001). Characterization of the mammalian peroxisomal import machinery: Pex2p, Pex5p, Pex12p, and Pex14p are subunits of the same protein assembly. *J. Biol. Chem.* **276**, 29935–29942.
19. Huhse, B., Rehling, P., Albertini, M., Blank, L., Meller, K. & Kunau, W.-H. (1998). Pex17p of *Saccharomyces cerevisiae* is a novel peroxin and component of the peroxisomal protein translocation machinery. *J. Cell Biol.* **140**, 49–60.
20. Liu, H., Tan, X., Russell, K. A., Veenhuis, M. & Cregg, J. M. (1995). PER3, a gene required for peroxisome biogenesis in *Pichia pastoris*, encodes a peroxisomal membrane protein involved in protein import. *J. Biol. Chem.* **270**, 10940–10951.
21. Rehling, P., Skaletz-Rorowski, A., Girzalsky, W., Voorn-Brouwer, T., Fransen, M. M., Distel, B. *et al.* (2000). Pex8p, an intraperoxisomal peroxin of *Saccharomyces cerevisiae* required for protein transport into peroxisomes binds the PTS1 receptor pex5p. *J. Biol. Chem.* **275**, 3593–3602.
22. Smith, J. J., Szilard, R. K., Marelli, M. & Rachubinski, R. A. (1997). The peroxin Pex17p of the yeast *Yarrowia lipolytica* is associated peripherally with the peroxisomal membrane and is required for the import of a subset of matrix proteins. *Mol. Cell. Biol.* **17**, 2511–2520.
23. Waterham, H. R., Titorenko, V. I., Haima, P., Cregg, J. M., Harder, W. & Veenhuis, M. (1994). The *Hansenula polymorpha* PER1 gene is essential for peroxisome biogenesis and encodes a peroxisomal matrix protein with both carboxy- and amino-terminal targeting signals. *J. Cell Biol.* **127**, 737–749.
24. Costa-Rodrigues, J., Carvalho, A. F., Gouveia, A. M., Fransen, M., Sa-Miranda, C. & Azevedo, J. E. (2004). The N terminus of the peroxisomal cycling receptor, Pex5p, is required for redirecting the peroxisome-associated peroxin back to the cytosol. *J. Biol. Chem.* **279**, 46573–46579.
25. Dodt, G., Gould, S. J. & Multiple, P. E. X. (1996). genes are required for proper subcellular distribution and stability of Pex5p, the PTS1 receptor: evidence that PTS1 protein import is mediated by a cycling receptor. *J. Cell Biol.* **135**, 1763–1774.
26. Gouveia, A. M., Guimaraes, C. P., Oliveira, M. E., Reguenga, C., Sa-Miranda, C. & Azevedo, J. E. (2003). Characterization of the peroxisomal cycling receptor, Pex5p, using a cell-free *in vitro* import system. *J. Biol. Chem.* **278**, 226–232.
27. Oliveira, M. E., Gouveia, A. M., Pinto, R. A., Sa-Miranda, C. & Azevedo, J. E. (2003). The energetics of Pex5p-mediated peroxisomal protein import. *J. Biol. Chem.* **278**, 39483–39488.
28. Glover, J. R., Andrews, D. W. & Rachubinski, R. A. (1994). *Saccharomyces cerevisiae* peroxisomal thiolase is imported as a dimer. *Proc. Natl Acad. Sci. USA*, **91**, 10541–10545.

29. McNew, J. A. & Goodman, J. M. (1994). An oligomeric protein is imported into peroxisomes *in vivo*. *J. Cell Biol.* **127**, 1245–1257.
30. Stewart, M. Q., Esposito, R. D., Gowani, J. & Goodman, J. M. (2001). Alcohol oxidase and dihydroxyacetone synthase, the abundant peroxisomal proteins of methylotrophic yeasts, assemble in different cellular compartments. *J. Cell Sci.* **114**, 2863–2868.
31. Walton, P. A., Hill, P. E. & Subramani, S. (1995). Import of stably folded proteins into peroxisomes. *Mol. Biol. Cell*, **6**, 675–683.
32. Saidowsky, J., Dodt, G., Kirchberg, K., Wegner, A., Nastainczyk, W., Kunau, W. H. & Schliebs, W. (2001). The di-aromatic pentapeptide repeats of the human peroxisome import receptor PEX5 are separate high affinity binding sites for the peroxisomal membrane protein PEX14. *J. Biol. Chem.* **276**, 34524–34529.
33. Schliebs, W., Saidowsky, J., Agianian, B., Dodt, G., Herberg, F. W. & Kunau, W. H. (1999). Recombinant human peroxisomal targeting signal receptor PEX5. Structural basis for interaction of PEX5 with PEX14. *J. Biol. Chem.* **274**, 5666–5673.
34. Costa-Rodrigues, J., Carvalho, A. F., Franssen, M., Hambruch, E., Schliebs, W., Sa-Miranda, C. & Azevedo, J. E. (2005). Pex5p, the peroxisomal cycling receptor, is a monomeric non-globular protein. *J. Biol. Chem.* **280**, 24404–24411.
35. Elgersma, Y., Elgersma-Hooisma, M., Wenzel, T., McCaffery, J. M., Farquhar, M. G. & Subramani, S. (1998). A mobile PTS2 receptor for peroxisomal protein import in *Pichia pastoris*. *J. Cell Biol.* **140**, 807–820.
36. Rehling, P., Marzioch, M., Niesen, F., Wittke, E., Veenhuis, M. & Kunau, W. H. (1996). The import receptor for the peroxisomal targeting signal 2 (PTS2) in *Saccharomyces cerevisiae* is encoded by the PAS7 gene. *EMBO J.* **15**, 2901–2913.
37. Zhang, J. W. & Lazarow, P. B. (1996). Peb1p (Pas7p) is an intraperoxisomal receptor for the NH<sub>2</sub>-terminal, type 2, peroxisomal targeting sequence of thiolase: Peb1p itself is targeted to peroxisomes by an NH<sub>2</sub>-terminal peptide. *J. Cell Biol.* **132**, 325–334.
38. Dodt, G., Warren, D., Becker, E., Rehling, P. & Gould, S. J. (2001). Domain mapping of human PEX5 reveals functional and structural similarities to *Saccharomyces cerevisiae* Pex18p and Pex21p. *J. Biol. Chem.* **276**, 41769–41781.
39. Einwachter, H., Sowinski, S., Kunau, W. H. & Schliebs, W. (2001). *Yarrowia lipolytica* Pex20p *Saccharomyces cerevisiae* Pex18p/Pex21p and mammalian Pex5pL fulfil a common function in the early steps of the peroxisomal PTS2 import pathway. *EMBO Rep.* **2**, 1035–1039.
40. Matsumura, T., Otera, H. & Fujiki, Y. (2000). Disruption of the interaction of the longer isoform of Pex5p, Pex5pL, with Pex7p abolishes peroxisome targeting signal type 2 protein import in mammals. Study with a novel Pex5-impaired Chinese hamster ovary cell mutant. *J. Biol. Chem.* **275**, 21715–21721.
41. Gouveia, A. M., Guimaraes, C. P., Oliveira, M. E., Sa-Miranda, C. & Azevedo, J. E. (2003). Insertion of Pex5p into the peroxisomal membrane is cargo protein-dependent. *J. Biol. Chem.* **278**, 4389–4392.
42. de la Torre, J. G. (1994). Hydrodynamics of segmentally flexible macromolecules. *Eur. Biophys. J.* **23**, 307–322.
43. Denning, D. P., Uversky, V., Patel, S. S., Fink, A. L. & Rexach, M. (2002). The *Saccharomyces cerevisiae* nucleoporin Nup2p is a natively unfolded protein. *J. Biol. Chem.* **277**, 33447–33455.
44. Permyakov, S. E., Millett, I. S., Doniach, S., Permyakov, E. A. & Uversky, V. N. (2003). Natively unfolded C-terminal domain of caldesmon remains substantially unstructured after the effective binding to calmodulin. *Proteins: Struct. Funct. Genet.* **53**, 855–862.
45. Uversky, V. N. (2002). Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* **11**, 739–756.
46. Lydakis-Simantiris, N., Hutchison, R. S., Betts, S. D., Barry, B. A. & Yocum, C. F. (1999). Manganese stabilizing protein of photosystem II is a thermally stable, natively unfolded polypeptide. *Biochemistry*, **38**, 404–414.
47. Sreerama, N. & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–260.
48. Stanley, W. A., Sokolova, A., Brown, A., Clarke, D. T., Wilmanns, M. & Svergun, D. I. (2004). Synergistic use of synchrotron radiation techniques for biological samples in solution: a case study on protein-ligand recognition by the peroxisomal import receptor Pex5p. *J. Synchrotron Radiat.* **11**, 490–496.
49. Oberg, K. A. & Fink, A. L. (1998). A new attenuated total reflectance Fourier transform infrared spectroscopy method for the study of proteins in solution. *Anal. Biochem.* **256**, 92–106.
50. Navea, S., Tauler, R. & de Juan, A. (2005). Application of the local regression method interval partial least-squares to the elucidation of protein secondary structure. *Anal. Biochem.* **336**, 231–242.
51. Obradovic, Z., Peng, K., Vucetic, S., Radijovic, P., Dunker, A. K. (2006). Exploiting heterogeneous sequence properties improves prediction of protein disorder. *Proteins* In the press.
52. Gould, S. J. & Collins, C. S. (2002). Opinion: peroxisomal-protein import: is it really that complex? *Nature Rev. Mol. Cell. Biol.* **3**, 382–389.
53. Shoemaker, B. A., Portman, J. J. & Wolynes, P. G. (2000). Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc. Natl Acad. Sci. USA*, **97**, 8868–8873.
54. Pontius, B. W. (1993). Close encounters: why unstructured, polymeric domains can increase rates of specific macromolecular association. *Trends Biochem. Sci.* **18**, 181–186.
55. Tompa, P. (2002). Intrinsically unstructured proteins. *Trends Biochem. Sci.* **27**, 527–533.
56. Gouveia, A. M., Reguenga, C., Oliveira, M. E., Sa-Miranda, C. & Azevedo, J. E. (2000). Characterization of peroxisomal Pex5p from rat liver. Pex5p in the Pex5p-Pex14p membrane complex is a transmembrane protein. *J. Biol. Chem.* **275**, 32444–32451.
57. Dammai, V. & Subramani, S. (2001). The human peroxisomal targeting signal receptor, Pex5p, is translocated into the peroxisomal matrix and recycled to the cytosol. *Cell*, **105**, 187–196.
58. Imanaka, T., Small, G. M. & Lazarow, P. B. (1987). Translocation of acyl-CoA oxidase into peroxisomes requires ATP hydrolysis but not a membrane potential. *J. Cell Biol.* **105**, 2915–2922.

59. Wendland, M. & Subramani, S. (1993). Cytosol-dependent peroxisomal protein import in a permeabilized cell system. *J. Cell Biol.* **120**, 675–685.
60. Azevedo, J. E., Costa-Rodrigues, J., Guimaraes, C. P., Oliveira, M. E. & Sa-Miranda, C. (2004). Protein translocation across the peroxisomal membrane. *Cell Biochem. Biophys.* **41**, 451–468.
61. Wang, D., Visser, N. V., Veenhuis, M. & van der Klei, I. J. (2003). Physical interactions of the peroxisomal targeting signal 1 receptor pex5p, studied by fluorescence correlation spectroscopy. *J. Biol. Chem.* **278**, 43340–43345.
62. Oliveira, M. E., Reguenga, C., Gouveia, A. M., Guimaraes, C. P., Schliebs, W., Kunau, W. H. *et al.* (2002). Mammalian Pex14p: membrane topology and characterisation of the Pex14p–Pex14p interaction. *Biochim. Biophys. Acta*, **1567**, 13–22.
63. Fransen, M., Wylin, T., Brees, C., Mannaerts, G. P. & Van Veldhoven, P. P. (2001). Human pex19p binds peroxisomal integral membrane proteins at regions distinct from their sorting sequences. *Mol. Cell. Biol.* **21**, 4413–4424.
64. Sreerama, N., Venyaminov, S. Y. & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis. *Anal. Biochem.* **287**, 243–251.
65. Viganò, C., Margolles, A., van Veen, H. W., Konings, W. N. & Ruyschaert, J. M. (2000). Secondary and tertiary structure changes of reconstituted LmrA induced by nucleotide binding or hydrolysis. A Fourier transform attenuated total reflection infrared spectroscopy and tryptophan fluorescence quenching analysis. *J. Biol. Chem.* **275**, 10962–10967.
66. Laurent, T. C. & Killander, J. (1964). Theory of gel filtration + its experimental verification. *J. Chromatog.* **14**, 217–230.
67. Gallagher, S. R. (1999). In (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds), p. 2, Wiley, New York.

*Edited by C. R. Matthews*

(Received 5 September 2005; received in revised form 26 November 2005; accepted 1 December 2005)  
Available online 19 December 2005