THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 37, pp. 25334–25342, September 11, 2009

Solution Structure of Human Pex5·Pex14·PTS1 Protein Complexes Obtained by Small Angle X-ray Scattering[S]

Received for publication, April 3, 2009, and in revised form, June 30, 2009 Published, JBC Papers in Press, July 6, 2009, DOI 10.1074/jbc.M109.002311

Kumiko Shiozawa[1], **Petr V. Konarev**[2], **Christian Neufeld, Matthias Wilmanns, and Dmitri I. Svergun**[2,3]

From the European Molecular Biology Laboratory-Hamburg Outstation, c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany

The Pex5p receptor recognizes newly synthesized peroxisomal matrix proteins which have a C-terminal peroxisomal targeting signal to the peroxisome. After docking to protein complexes on the membrane, these proteins are translocated across the membrane. The docking mechanism remains unclear, as no structural data on the multicomponent docking complex are available. As the interaction of the cargo-loaded Pex5p receptor and the peroxisomal membrane protein Pex14p is the essential primary docking step, we have investigated the solution structure of these complexes by small angle x-ray scattering and static light scattering. Titration studies yielded a 1:6 stoichiometry for the Pex5p·Pex14p complex, and low resolution structural models were reconstructed from the x-ray scattering data. The free full-length human Pex5p is monomeric in solution, with an elongated, partially unfolded N-terminal domain. The model of the complex reveals that the N terminus of Pex5p remains extended in the presence of cargo and Pex14p, the latter proteins being significantly intermingled with the Pex5p moiety. These results suggest that the extended structure of Pex5p may play a role in interactions with other substrates such as lipids and membrane proteins during the formation of functional multiprotein complexes.

Peroxisomes are ubiquitous organelles in eukaryotes which are involved in different metabolic pathways (1). Peroxisomal matrix proteins, which contain a peroxisomal targeting signal (PTS),[4] are imported into the peroxisome by recognition of two different import receptors, Pex5p or Pex7p. These receptors recognize specific signal sequences, PTS1 and PTS2, respectively (1). At the molecular level the C-terminal PTS1 signal is bound in a central cavity of the ring-like structure of the seven tetrapeptide repeat (TPR) domains of the C-terminal part of Pex5p (Pex5p(C)) (2–5). It was recently proposed that some of the structural principles of the Pex5p/cargo interaction may also apply to the PTS2 cargo recognition of the Pex7p receptor (5).

The next step of PTS-protein import, docking of the cargo loaded receptor to the translocon, involves the peroxisomal protein Pex14p (6). Multiple Pex14p binding sites with diaromatic pentapeptide motifs (W*XXX*(F/Y)) were shown to be present in the N terminus of Pex5p (7–9). The number of these motifs, however, varies among species. The human Pex5p receptor, which has been investigated in this contribution, has a total of seven motifs. A recent NMR structure of the N-terminal domain of Pex14p and the first W*XXX*(F/Y) motif of Pex5p reveals an α-helical conformation of the motif (10). Interactions between Pex5p and other proteins and by their association with the peroxisomal membrane possibly lead to dissociation of the PTS-protein from Pex5p (11–13). The exact sequence of events in the import mechanism remains, however, unknown. It is in particular unclear how, in contrast with other organelles, peroxisomes can import folded oligomeric, functional proteins (14).

Previous biophysical work indicated that the N terminus half of Pex5p is unfolded *in vitro* (15, 16). Recent protease sensitivity assays showed that the proteolytic profiles of the full-length receptor Pex5p(F) change in the presence of PTS1 peptide and the Pex13p Src homology 3 domain, which is another docking factor (16, 17), indicating conformational changes of Pex5p upon binding these receptor ligands. Furthermore, it was found that Pex5p may even traverse the peroxisomal membrane, leaving only a small N-terminal fragment in the cytosol while exposing the C-terminal TPR domain to the luminal side of the membrane (11).

Although recognition of many PTS cargos seems to be confined to the C-terminal TPR domains of Pex5p, it has become clear that the N-terminal part of Pex5p is primarily involved in docking of the receptor onto the peroxisomal membrane and other docking factors. Because only poorly diffracting crystals have been purified to date, we investigated its solution structure by small angle x-ray scattering (SAXS) and static light scattering (SLS). Complexes with the PTS1 cargo sterol carrier protein 2 (SCP2), which functions as lipid transfer protein, were also studied as the crystal structure of Pex5p(C)/SCP2 is already known (4). Our results indicate that human Pex5p(F) is a monomer with an extended N terminus. The stoichiometry of Pex5p(F)·Pex14p(N)·PTS1 complex has been assessed by titration with SAXS, SLS, and gel filtration, and a low resolution structural model of the complex has been reconstructed in which Pex5p(F) remains extended upon Pex14p(N) binding.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Fig. S1.

[1] Supported by Integrated Project 3D Repertoire LSHG-CT-2005512028 from the European Commission (to M. W.).
[2] Supported by EU Design Study SAXIER Contract 011934.

[3] To whom correspondence should be addressed. Tel.: 49-40-89902-125; Fax: 49-40-89902-149; E-mail: svergun@embl-hamburg.de.

[4] The abbreviations used are: PTS, peroxisomal targeting signals; DLS, dynamic light scattering; D_{max}, maximum particle dimension; EM, electron microscopy; MM, molar mass; NSD, normalized spatial discrepancy; $p(r)$, distance distribution function; R_g, radius of gyration; SAXS, small angle X-ray scattering; SLS, static light scattering; SA, simulated annealing; TCEP, Tris(2-carboxyethyl)phosphine hydrochloride; TPR, tetratricopeptide repeat; NTA, nitrilotriacetic acid; SCP2, sterol carrier protein 2.

EXPERIMENTAL PROCEDURES

Expression—Full-length human Pex5p (residues 1–639; Pex5p (F)), the N terminus of human Pex5p (residues 1–335; Pex5p(N)), the C terminus of human Pex5p (residues 268-602; Pex5p(C)), and the N terminus of Pex14p (residues 16–80 plus tryptophan; Pex14p(N)) were expressed with an N-terminal His_6 tag in *Escherichia coli* BL21(DE3) cells using either the pET9d or pETM11 plasmid vector (EMBL). SCP2 (residues 22–143) (19) was expressed with an N-terminal glutathione *S*-transferase and His_6 tag in *E. coli* BL21(DE3) cells using the pETM30 plasmid vector (EMBL). *E. coli* cells, containing Pex5p(F), Pex5p(N), Pex5p(C), and Pex14p(N) in expression plasmids were grown at 37 °C and induced by 0.5 mM isopropyl-D-thiogalactopyranoside at $A_{600} = 0.5$ followed by growth at 20 °C for 10 h. *E. coli* cells, containing a SCP2 expression plasmid, were grown and induced under the same conditions followed by growth at 30 °C for 3 h. The cell pellet was resuspended in lysis buffer (50 mM HEPES-KOH (pH 7.5), 150 mM KCl, 4 mM β-mercaptoethanol, 1% Triton X-100, and 5% glycerol) with protease inhibitors (Roche Applied Science).

Purification—All protein purification steps were performed at 4 °C. Pex5p(F) was purified by successive passage over a Ni^{2+}-NTA-agarose (Qiagen) column in binding buffer (50 mM HEPES-KOH (pH 7.5), 150 mM KCl, 4 mM β-mercaptoethanol, 20 mM imidazole, and 1% glycerol) and a Superdex 200 size exclusion column, ending with buffer I containing 50 mM HEPES-KOH (pH 7.5), 100 mM KCl, 1% glycerol, and 2 mM TCEP.

SCP2 was purified by successive passage over Ni^{2+}-NTA-agarose in binding buffer and glutathione-Sepharose 4B (Amersham Biosciences) columns in buffer II, containing 50 mM HEPES-KOH (pH 7.5), 100 mM KCl, and 1 mM TCEP. The glutathione *S*-transferase tag of SCP2 was removed with His-tobacco etch virus protease (30 μg/ml) for 2 h at room temperature. The SCP2 released from the resin was passed through Ni^{2+}-NTA-agarose and finally purified over Superdex 75 in buffer II.

Pex14p(N) was purified by successive passage over Ni^{2+}-NTA-agarose in binding buffer and Superdex 75 columns, ending with buffer II. Fractions containing Pex5p(N) and Pex5p(C) were purified by successive passage over Ni^{2+}-NTA-agarose in binding buffer and Superdex 75 columns, ending with buffer II.

To form the binary Pex5p(F)·Pex14p(N) complex, after harvesting the two sets of *E. coli* cells containing (Pex5p(F) and Pex14p(N)) were mixed on ice for 30 min and resuspended in lysis buffer (see above) with protease inhibitors. The Pex5p(F)·Pex14p(N) complex was purified by successive passage over Ni^{2+}-NTA-agarose in binding buffer and Superdex 200 size exclusion columns, ending with buffer III containing 50 mM HEPES-KOH (pH 7.5), 100 mM KCl, and 2 mM TCEP.

The binary Pex5p(F)·SCP2 complex was formed by incubating a (1:1) mixture of purified Pex5p(F) and SCP2 on ice for 30 min and purified on a Superdex 200 size exclusion column, ending with buffer III. The ternary Pex5p(F)·Pex14p(N)·SCP2 complex was formed by incubating a (1:1) mixture of purified (Pex5p(F)·Pex14p(N)) and SCP2 on ice for 30 min and purified on a Superdex 200 size exclusion column, ending with buffer III.

To form the ternary Pex5p(C)·Pex14p(N)·SCP2 complex, after harvesting, three sets of *E. coli* cells containing Pex5p(C), Pex14p(N), and SCP2, were mixed on ice for 30 min and resuspended in lysis buffer (see above) with protease inhibitors. The complex was purified by successive passage over Ni^{2+}-NTA-agarose in binding buffer and glutathione-Sepharose columns in buffer II. The glutathione *S*-transferase tag of SCP2 was removed with His-tobacco etch virus protease (30 μg/ml) for 2 h at room temperature. The complex released from the resin was passed through Ni^{2+}-NTA and finally purified over Superdex 200 in a buffer II. All constructs were pooled and concentrated for the SAXS measurements using a Vivaspin 20 spin concentrator (Vivascience) with a 3- or 5-kDa molar mass (MM) cutoff.

The purity and homogeneity of all proteins were confirmed by SDS-PAGE, native PAGE, and dynamic light scattering (DLS). The extinction coefficients of His_6-Pex5p(F), His_6-Pex14p(N), and SCP2, used in determining concentrations and/or mixing ratios, were calculated from their sequences to be 100,500, 5,690, and 5,500 $M^{-1}\cdot cm^{-1}$, respectively.

SLS—Pex5p(F), Pex5p(F)·Pex14p(N), Pex5p(F)·SCP2, Pex5p(F)·Pex14p(N)·SCP2, and Pex5p(C)·Pex14p(N)·SCP2 were injected at concentrations of 0.3, 0.7, 0.5, 0.7, and 0.1 mg/ml, respectively, into a Superdex 200 10/300 GL tricorn column (Amersham Biosciences), equilibrated with buffer II at room temperature, and connected with an SLS-detector (mini DAWN Tristar; Wyatt Technology, Santa Barbara, CA). The instrument measures the protein concentration by UV absorbance, the refractive index, and the scattering intensity of the eluted peaks. The ASTRA software (Version 4.90.08; Wyatt Technology, Santa Barbara, CA) was used to compute the MM via the basic light-scattering equation (20). A set of standard proteins (Bio-Rad) was used to calibrate the column. The fractions containing the complexes of interest were analyzed on SDS-PAGE and native gels (Invitrogen).

SLS Titration Experiment—Pex5p(F) and Pex14p(N), at 3.5 and 6.3 mg/ml, respectively, were dialyzed in buffer III. Solutions containing appropriate molar ratios of Pex5p(F) and Pex14p(N) between 1:1 and 1:8 were incubated on ice for 1 h to form complexes and processed as indicated above. The fractions containing the complexes of interest were analyzed on a native gel.

Small Angle X-ray Scattering Measurements—The synchrotron radiation x-ray scattering data were collected on the X33 beamline of the EMBL on the storage ring DORIS III (DESY, Hamburg, Germany) (21). Solutions of Pex5p and its binary and ternary complexes with Pex14p(N) and SCP2 proteins were measured at 15 °C at solute concentrations of 0.5, 1.0, and 2.0 mg/ml. Before the measurements, 20 mM TCEP was added to samples to reduce the radiation damage. The data were recorded using a MAR345 image plate detector at a sample-detector distance of 2.7 m and a wavelength of $\lambda = 0.15$ nm, covering the range of momentum transfer $0.12 < s < 4.5\ nm^{-1}$ ($s = 4\pi \sin\theta/\lambda$, where 2θ is the scattering angle). No measurable radiation damage was detected by comparison of successive time frames with 2-min exposures. The data were averaged after normalization to the intensity of the transmitted beam,

and the scattering of the buffer was subtracted. The difference data were extrapolated to zero solute concentration following standard procedures. All data manipulations were performed using the program package PRIMUS (22).

The forward scattering $I(0)$ and the radius of gyration R_g were evaluated using the Guinier approximation (23) assuming that at very small angles ($s < 1.3/R_g$) the intensity is represented as $I(s) = I(0)\exp(-(sR_g)^2/3)$. These parameters were also computed from the entire scattering patterns using the program GNOM (24), which provides the maximum particle dimensions D_{max} and the distance distribution functions $p(r)$. The molecular masses of the solutes were estimated from the forward scattering by normalization against reference solutions of bovine serum albumin. The excluded (Porod) volumes of hydrated particles were computed as (25)

$$V_p = 2\pi^2 I(0) / \int_0^\infty s^2 I(s)\,ds \qquad \text{(Eq. 1)}$$

Before the calculation, an appropriate constant was subtracted from each data point to force the s^{-4} decay of the intensity at higher angles following Porod's law (25) for homogeneous particles. This procedure yields a "shape scattering" curve corrected for the unwanted scattering contribution from the internal structure. For globular proteins the computed Porod (*i.e.* hydrated) volumes in nm^3 are expected to be about twice the MMs in kDa.

Ab Initio Shape Determination—The low resolution shapes of Pex5p(F) and the ternary Pex5p(F)·Pex14p(N)·SCP2 complex were reconstructed by the program DAMMIN (26) and its multiphase version MONSA (27). These programs represent the particle as a collection of $M \gg 1$ densely packed beads inside a sphere with diameter D_{max}. In DAMMIN, each bead is assigned either to the solvent or to the particle, and the latter is represented by a simple "phase" (non-solvent beads). In MONSA each bead is assigned either to the solvent (index = 0) or to one of the parts of the complex (index = 1 corresponding to Pex5p(N) (residues 1–316), index = 2 for Pex5p(C) (residues 317–639), index = 3 for the first Pex14p(N) protein binding to Pex5p(F) together with SCP2 protein, and index = 4 for all other binding Pex14p(N) proteins). The particle is, therefore, represented at low resolution by four phases (portions of the complex), and the overall model is described by a string of length M containing the phase index for each bead (0 = solvent or 1, 2, 3, 4 for any of the parts of the complex). Starting from a random string, simulated annealing (SA) is employed in both DAMMIN and MONSA to search for a model composed of interconnected compact phases, fitting single (DAMMIN), or multiple shape (MONSA) curves to minimize the overall discrepancy,

$$\chi^2 = \sum_k \frac{1}{N_k - 1} \sum_j \left[\frac{I_k(s_j) - c_k I_k^{calc}(s_j)}{\sigma_k(s_j)} \right]^2 \qquad \text{(Eq. 2)}$$

where $I_k(s)$ is the scattering intensity from a construct k, the index k runs over the scattering curves, N_k are the numbers of experimental points in each curve, c_k are scaling factors, and $I_{calc}(s)$ and $\sigma(s_j)$ are the intensities calculated from the subsets of the beads belonging to the appropriate phases and the experimental errors at the momentum transfer s_j, respectively.

In MONSA, the *ab initio* procedure was applied to simultaneously fit five scattering curves (one curve from Pex5p(F), two curves from binary complexes of Pex5p(F)·Pex14p(N) and Pex5p(F)·SCP2, and two from ternary constructs of Pex5p(F)·Pex14p(N)·SCP2 and Pex5p(C)·Pex14p(N)·SCP2). Given the assumption that each phase has a uniform density, the shape-scattering curves after appropriate constant subtraction (see above) were fitted up to the resolution of about 3 nm (*i.e.* in the range of the scattering vectors up to $s = 2\ nm^{-1}$). The expected volume fractions of the four phases, corresponding to their theoretical molecular masses, were used to further restrain the SA procedure. In independent DAMMIN calculations, the relevant curves were fitted separately to yield the low resolution models of the corresponding subcomplexes.

Molecular Modeling—Rigid body modeling was done using the atomic models of the C-terminal domain of Pex5p (residues; 280–602, Protein Data Bank (PDB) code 1FCH) (2), Pex5p(C)/SCP2 (residues; 335–639/22–143, PDB code 2C0L) (4), and Pex14p(N) (Personnel communication) as input for the programs BUNCH and SASREF (28). Note that the residue numbers of 1FCH correspond to those of the Protein Data Bank and differ from 2C0L. The scattering amplitudes of the individual subunit structures from the ternary Pex5p(F)·Pex14p(N)·SCP2 complex were calculated from their atomic coordinates by the program CRYSOL (29).

To model the structure of Pex5p(F) in solution, the unknown N-terminal segment was first represented by interconnected chains composed of dummy residues. An SA protocol implemented in the program BUNCH (28) was employed to find a native-like configuration of the N-terminal fragment without steric clashes, fitting the experimental scattering from the full-length protein. The results of multiple BUNCH runs were averaged to determine common structural features using the programs DAMAVER (30) and SUPCOMB (31). The latter program aligns two arbitrary low or high resolution models represented by ensembles of points by minimizing a dissimilarity measure called normalized spatial discrepancy (NSD). For every point (bead or atom) in the first model, the minimum value among the distances between this point and all points in the second model is found, and the same is done for the points in the second model. These distances are added and normalized against the average distances between the neighboring points for the two models. Generally, NSD values close to one indicate that the two models are similar. DAMAVER generates the average model of the set of superimposed structures and also specifies the most typical model (*i.e.* that having the lowest average normalized spatial discrepancy with all other models in the set).

The resulting low resolution structure of the N-terminal part of Pex5p(F) was further used to model the ternary Pex5p(F)·Pex14p(N)·SCP2 complex. The position of SCP2 with respect to Pex5p(C) was fixed as in the crystallographic model of truncated Pex5p(C)·SCP2 complex (as in 2C0L), and the portion of Pex5p(C) unresolved in the latter model was added (as in 1FCH). This model was treated as a single rigid body in all subsequent calculations. The six copies of Pex14p(N) were

added, and their positions were refined by rigid body modeling using the program SASREF (28). This program employs a SA protocol to generate an interconnected assembly of subunits without steric clashes fitting the scattering data. Intersubunit contacts were loosely defined for the Pex14p(N) interaction with the W*XXX*(F/Y) motifs (residues 28–32, 118–122, 140–144, 159–163, 243–247, and 308–312) of Pex5p(F), based on NMR data (10) (Fig. 1*A*).

To select the most typical *ab initio* model of the complex and estimate its possible conformational space, 10 MONSA reconstructions were pairwise aligned using SUPCOMB (31). For each model, the values of NSD against all other models were computed, and an average value ⟨NSD⟩ was calculated. The most typical reconstruction was the one with the lowest average NSD with respect to the other reconstructions. The same selection was made for the SASREF reconstruction to select the most typical rigid body model.

SAXS Titration Experiments—Pex5p(F) and Pex14p(N), at 3.7 and 6.3 mg/ml, respectively, were dialyzed in buffer III. Solutions containing appropriate molar ratios of Pex5p(F) and Pex14p(N) (1:1–1:8, 1:10, and 1:20) were incubated at room temperature for 1 min to form complexes, and their SAXS patterns were recorded at 15 °C after the addition of 20 mM TCEP-dissolved buffer III and pH-adjusted. The MMs and R_g values were determined as described above.

DLS—The R_h of Pex5p(F), Pex5p(F)·Pex14p(N), Pex5p(F)·SCP2, Pex5p(F)·Pex14p(N)·SCP2, and Pex5p(C)·Pex14p(N)·SCP2 were measured by DLS using DynaPro-MS (Protein Solutions) at 4 °C. The protein concentrations were 0.5 mg/ml in buffer II at 4 °C. For each sample, 0.1 ml was centrifuged at 14,000 rpm for 20 min and passed into a 12-μl chamber quartz cuvette. The data were analyzed using the Dynamics 5.0 software. A control size exclusion chromatography experiment showed no interaction between Pex14p(N) and Pex5p(C)-short (315–639) (data not shown).

Circular dichroism (CD)—CD spectra of Pex5p(F), Pex5p(N), and Pex5p(C) were measured on a Jasco J-810 spectropolarimeter equipped with a Peltier temperature control unit using a cylindrical 1-mm Hellma Quartz Suprasil cuvette in the wavelength range 190–240 nm. The protein concentrations were 0.6–1.5 μM in 10 mM sodium phosphate buffer (pH 7.4) at 10 °C. The spectra were averaged over three scans, corrected for buffer background, and scaled against the camphor-10-sulfonic acid ellipticity calibration.

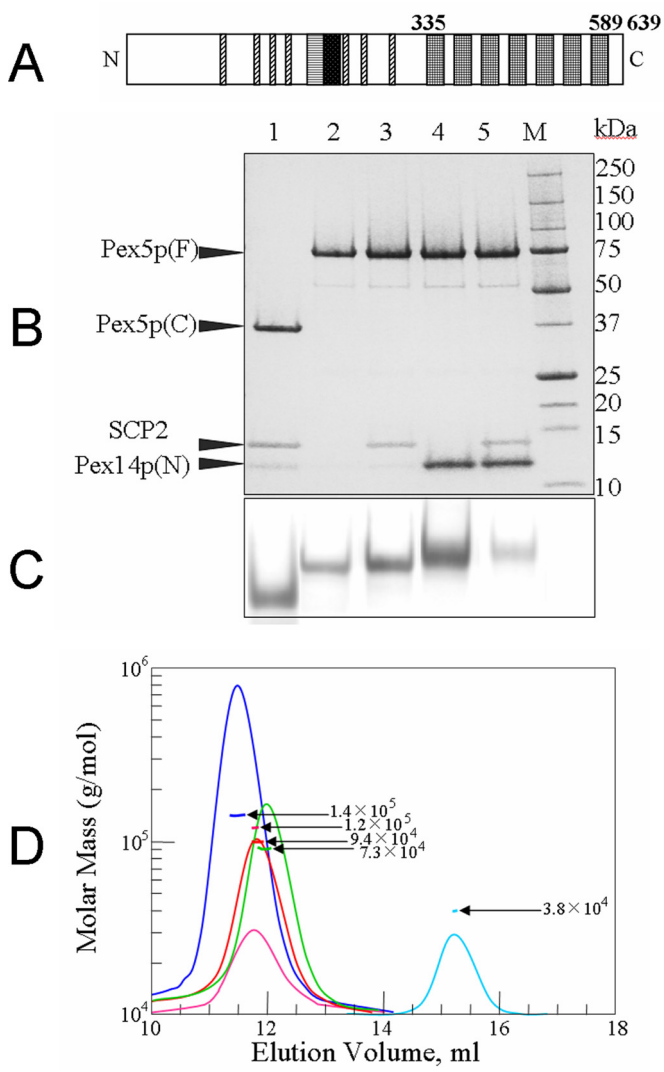

FIGURE 1. **Biophysical properties of the human Pex5p receptor.** *A*, schematic presentation of human Pex5p(F). It contains seven W*XXX*(F/Y) motifs (*hatched boxes*; residues 118–122, 140–144, 159–163, 184–188, 243–247, 257–261, 308–312), Pex7 binding sites (*dot*; residues 191–222), a Pex13 Src homology 3 domain binding site (*bar*; residues 184–192), and seven TPR domains (*gray boxes*; residues 335–589) (4). Shown are Coomassie Brilliant Blue-stained SDS-PAGE gel (*B*) and native gel of purified Pex5p(C)·Pex14p(N)·SCP2 (*lane 1*), Pex5p(F) (*lane 2*), Pex5p(F)·SCP2 (*lane 3*), Pex5p(F)·Pex14p(N) (*lane 4*), and Pex5p(F)·Pex14p(N)·SCP2 (*lane 5*) (*C*). *Lane M*, protein markers and their MMs. *D*, SLS measurement of Pex5p(F). The continuous line corresponds to the UV elution profile. The *numbers above the UV trace* correspond to the MM of each sample.

RESULTS

Determination of the Molar Mass and Stoichiometry of the Pex5p(F)·Pex14p(N) Complex by SLS and Native Gel Analysis—All Pex5p-containing constructs were first analyzed to assess their oligomeric states in solution. For this, recombinant human Pex5p(F), Pex5p(F)·Pex14p(N), Pex5p(F)·SCP2, Pex5p(F)·Pex14p(N)·SCP2, and Pex5p(C)·Pex14p(N)·SCP2 were studied by native gel analysis (Fig. 1*C*) and by SLS (Fig. 1*D*). The average MMs from SLS presented in Table 1 indicate that Pex5p(F) alone is monomeric and forms complexes with multiple Pex14p(N)s at the concentrations used for gel filtration. The complex formation was verified by SDS and native gels, and the purified constructs were found to be homogeneous in native gels (Fig. 1*C*) but displayed the expected bands of the individual components in SDS gels (Fig. 1*B*). The interaction between Pex5p(F) and Pex14p(N) was further assessed by SLS titration. The apparent MM of Pex5p(F)·Pex14p(N) increases as illustrated in Table 2 and Fig. 2 up to a molar ratio of 1:6. Above a 1:7 molar ratio, the value saturates, suggesting that Pex5p(F) binds six or seven Pex14p(N)s in solution. This result was further verified by a SAXS titration experiment (see below).

Molecular Parameters from SAXS—The structural parameters of Pex5p(F) determined from the experimental scattering pattern (Fig. 3*A*) are given in Table 1. The estimated MM and hydrated particle volume agree well with those predicted from the primary structure, indicating that Pex5p(F) is monomeric in solution. The R_g and D_{max} values are 5.2 ± 0.1 and 18 ± 1 nm,

TABLE 1

Overall parameters of Pex5p constructs from DLS, SLS, and SAXS

R_g, D_{max}, V_p, and R_h are, respectively, the experimental radius of gyration, maximum size, excluded volume, and hydrodynamic radius. MM_{exp}, MM_{sls}, and MM_{th} and molar masses were determined by SAXS and SLS and calculated from the appropriate primary sequences, respectively. C_h and C_{sls} are the concentration of samples used for DLS and SLS, respectively. χ_M is the overall discrepancy between the experimental data and computed scattering curves from the *ab initio* MONSA model, χ_{RB} denotes the discrepancy of a typical rigid body model by SASREF assuming 1:6:1 ternary complex formation of Pex5p(F)·Pex14p(N)·SCP2.

Sample	DLS R_h (C_h)	SLS MM_{sls} (C_{sls})	R_g	D_{max}	V_p	MM_{exp}	MM_{th}	χ_M	χ_{RB}
	nm ($mg \cdot ml^{-1}$)	*kDa* ($mg \cdot ml^{-1}$)	*nm*	*nm*	nm^3	*kDa*	*kDa*		
Ternary complex, Pex5p(F)·Pex14p(N)·SCP2 (1:6:1)	4.3 (0.7)	140 ± 3 (0.6)	6.0 ± 0.1	20.0 ± 1	267 ± 10	140 ± 4	135	3.09	2.68
Binary complex, Pex5p(F)·Pex14p(N) (1:6)	4.2 (0.7)	116 ± 3 (0.1)	5.8 ± 0.1	19.0 ± 1	234 ± 7	120 ± 4	121	2.15	2.04
Binary complex, Pex5p(F)·SCP2 (1:1)	4.1 (0.5)	94 ± 3 (0.3)	5.5 ± 0.1	20.0 ± 1	197 ± 4	87 ± 3	85	2.08	1.94
Pex5p(F) protein	3.9 (0.3)	73 ± 3 (0.1)	5.2 ± 0.1	18.0 ± 1	181 ± 4	75 ± 3	72	2.24	1.92
Ternary complex Pex5p(C)·Pex14p(N)·SCP2 (1:1:1)	2.8 (0.8)	38 ± 1 (0.2)	2.8 ± 0.1	9.0 ± 1	110 ± 5	54 ± 3	57	2.17	1.53

TABLE 2

SLS and SAXS titration experiments on the binary complex Pex5p(F)/Pex14p(N)

C_{exp} denotes the solute concentration of the SAXS experiments. R_g is the experimental radius of gyration, and C_{sls} is the concentration of samples used for SLS.

Molar ratio Pex5p(F):Pex14p(N)	SLS		SAXS			
	C_{sls}	MM_{sls}	C_{exp}	R_g	MM_{exp}	MM_{th}
	$mg \cdot ml^{-1}$	*kDa*	$mg \cdot ml^{-1}$	*nm*	*kDa*	*kDa*
1:0	0.03	67 ± 2.0	0.46	3.5 ± 0.1	73 ± 5	72
1:1	0.08	81 ± 1.5	0.49	4.4 ± 0.2	80 ± 5	79
1:2	0.07	91 ± 3.4	0.53	4.6 ± 0.2	93 ± 7	86
1:3	0.06	85 ± 4.3	0.57	4.8 ± 0.2	110 ± 10	94
1:4	0.12	96 ± 1.5	0.6	4.9 ± 0.2	118 ± 10	101
1:5	0.1	100 ± 2.0	0.64	5.3 ± 0.2	128 ± 10	108
1:6	0.14	108 ± 1.3	0.67	5.4 ± 0.2	134 ± 10	113
1:7	0.12	105 ± 4.2	0.71	5.1 ± 0.2	125 ± 10	123
1:8			0.74	4.9 ± 0.2	117 ± 10	120

respectively, confirming recent observations indicating that Pex5p(F) is elongated (15, 16). The distance distributions $p(r)$ computed from the experimental data of all complexes investigated are given in Fig. 3*B*. The bell-shaped $p(r)$ of the truncated Pex5p(C)·Pex14p(N)·SCP2 complex (*curve 5*) is typical for globular particles (32), pointing to a compact shape of the construct. The other $p(r)$ functions have similar skewed profiles characteristic of elongated particles with a maximum at around 5 nm, suggesting that little or no compaction occurs when complexes with Pex14p(N) and SCP2 are formed. The $p(r)$ functions for complexes in the presence of the scaffold ligand Pex14p, Pex5p(F)·Pex14p(N) (*curve 2*), and Pex5p(F)·Pex14p(N)·SCP2 (*curve 1*) go above that for the binary cargo Pex5p(F)·SCP2 complex (*curve 3*) in the distance range between 6 and 12 nm, suggesting that this range mostly corresponds to the distances between Pex14p molecules in the complex. Further evidence about the conformation of the N terminus of Pex5p(F) is provided by the comparison of the sizes of different constructs determined by DLS and SAXS. For the truncated Pex5p(C)·Pex14p(N)·SCP2 complex, the hydrodynamic radius R_h from DLS (Table 1, *last row*, *first data column*) coincides with the radius of gyration R_g from SAXS (Table 1, *last row*, *third data column*), as expected for globular particles. In contrast, for all remaining constructs involving the full-length Pex5p, R_g significantly exceeds R_h. Such an effect is observed for flexible macromolecules. For unfolded proteins, the ratio R_g/R_h is about 1.5 (33). Therefore, given that this type of effect has not been observed for complexes involving the C-terminal cargo binding part of the receptor only, the data demonstrate that the observed molecular flexibility must originate from the N-terminal part of the receptor regardless of whether is of the ligands investigated (SCP2, Pex14p(N)) are bound.

Determination of the Stoichiometry of the Complex between Human Pex5p(F) and Pex14p(N) by SAXS—For a quantitative assessment of the binding interaction between Pex5p(F) and Pex14p(N), a SAXS titration was performed. The MM of Pex14p(N) is 10 times smaller than that of Pex5p(F), and if Pex14p(N) forms a complex with Pex5p(F), the apparent R_g and $I(0)$ is expected to increase. In contrast, if a fraction of unbound Pex14p(N) exists in the solution with Pex5p(F)·Pex14p(N) complexes, the two values are expected to decrease. As illustrated in Fig. 2*A* (Table 2), the apparent R_g and MM_{exp} values do increase up to a 1:6 Pex5p(F)·Pex14p(N) molar ratio and start to decrease above a 1:7 molar ratio. This result is consistent with those of the SLS titration experiment (see above) and is further corroborated by the native gel analysis (Fig. 2*B*), also displaying saturation above a 1:6 molar ratio.

Ab Initio Modeling of Pex5p(F)—The structure of Pex5p(F) protein in the absence of ligands, restored *ab initio* from the scattering pattern (Fig. 3*A*), is shown in Fig. 4*A*. It was obtained by averaging 10 independent reconstructions yielding a good fit to the experimental data in the entire scattering range (the discrepancy is given in Table 1). The model depicts a compact globular domain, which may correspond to the C-terminal part of Pex5p and an elongated tail. Several independent runs yielded reproducible *ab initio* models of the Pex5p(F)·Pex14p(N)·SCP2 complex, which neatly fitted the experimental scattering profiles (Table 1). A typical model presented in Fig. 4*B* demonstrates that the Pex5p(F)·Pex14p(N)·SCP2 complex adopts an elongated conformation in solution. The overall appearance of the model is similar to the shape of Pex5p(F) alone, implying that there are no large structural rearrangements upon complex formation with Pex14p(N) and SCP2. The overlap of this model with the shape of apoPex5p(F) (Fig. 4*A*) indicates that both models have similar overall shapes. The model of the complex displays additional mass, suggesting possible binding sites of Pex14p(N) molecules. The above results of the *ab initio* modeling suggest that the overall structure of Pex5p(F) remains extended in complex with Pex14p(N) and SCP2. For a more detailed analysis, the program MONSA was employed to build a multiphase bead model of the complex by simultaneous fitting of the available data. Several independent runs yielded reproducible solutions, and the most typical one is shown in Fig. 5. The overall shape of the MONSA model is similar to the one from DAMMIN (Fig. 4*B*), but the former model highlights the shapes of individual parts of the ternary complex. The shape of N terminus of Pex5p(F) in the complex remains elongated when the Pex14p(N) molecules are bound to

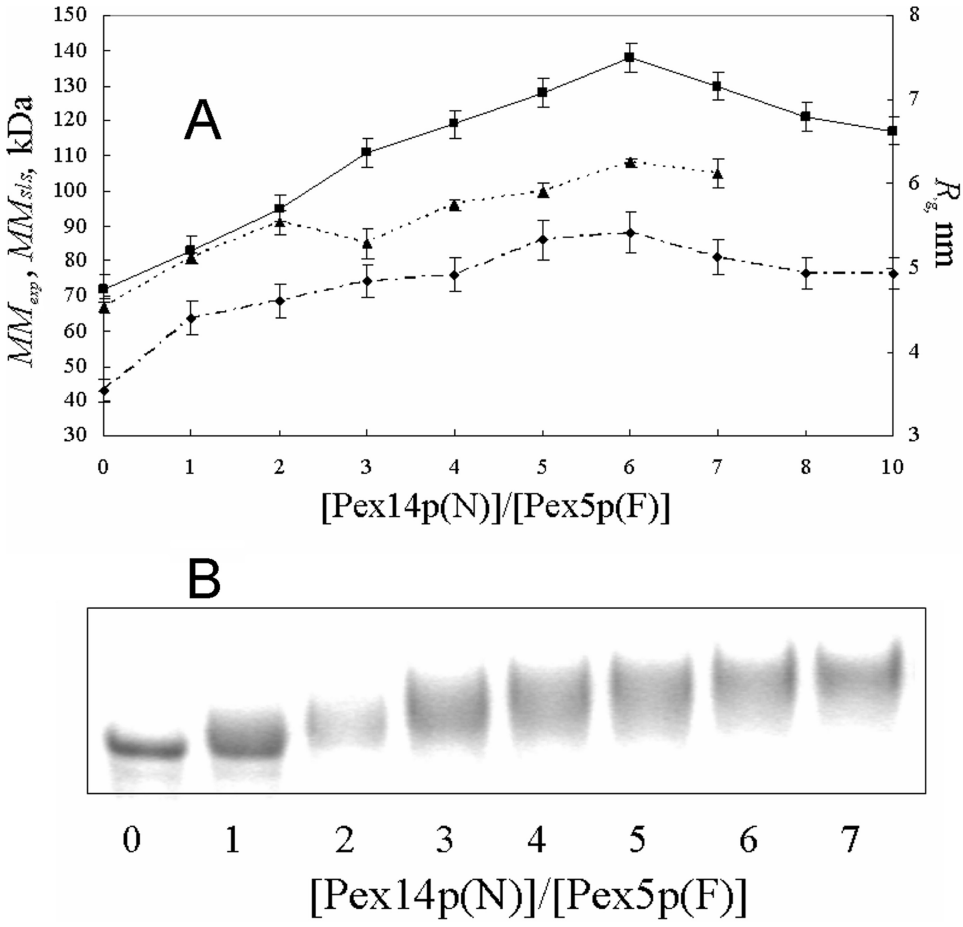

FIGURE 2. **Titration experiments on the Pex5p(F)·Pex14p(N) complex.** *A*, the experimental SAXS radius of gyration R_g (*dot-dashed line*) and the molecular masses MM_{exp} (*solid line*, SAXS) and MM_{sls} (*dashed line*, SLS) as a function of the molar ratio Pex14p(N):Pex5p(F). *B*, Coomassie Brilliant Blue-stained native gel of Pex5p(F)·Pex14p(N), for molar ratios from 0 to 7 from *left* to *right*.

the N-terminal region of Pex5p(F). The Pex14p(N) moiety depicted within the overall shape of the complex (*red beads* in Fig. 5) appears essentially intermingled with that of the N terminus of Pex5p(F). At the same time the C-terminal domain of Pex5p(F) displays a globular form, matching previous structural data (5, 34).

Molecular Modeling—In an independent approach, a model of the ternary complex was constructed making use of the high resolution models of the individual domains. These domains were treated as rigid bodies, whereas the fragments with unknown structure were represented as flexible chains of dummy residues. First, the model of Pex5p(F) was generated consisting of the compact C-terminal domain and the N terminus containing 322 dummy residues. Multiple runs of BUNCH (28) starting from random initial configurations yielded an extended configuration of N-terminal region of Pex5p(F), providing a good fit to the data (Fig. 4*A*). Twenty reconstructions were pairwise compared, and the most typical model, which had the smallest average normalized spatial discrepancy with respect to the other models, was selected. This model overlaps well with the *ab initio* model of Pex5p(F) (Fig. 4*A*).

The program SASREF (28) was then used for rigid body modeling of the ternary Pex5p(F)·Pex14p(N)·SCP2 complex by simultaneously fitting all available scattering profiles. The N-terminal region of Pex5p(F) was taken as rigid body from the BUNCH reconstruction, and SASREF was run for two stoichiometries of the Pex5p(F)·Pex14p(N)·SCP2 (1:6:1 and 1:5:1). The individual Pex14p(N) molecules were required to have contacts with the appropriate binding sites of Pex5p(F) (see "Experimental Procedures"). Multiple runs of SASREF yielded reproducible models whereby calculations with a 1:6:1 stoichiometry gave somewhat better overall fits (see Table 1). The most typical model of the ternary complex (with the lowest average NSD value) displays the overall shape similar to that provided by DAMMIN (Fig. 4*B*). A comparison with the most typical multiphase MONSA model in Fig. 5 reveals that the two independent modeling approaches predict a similar organization of the complex. The Pex14p(N) molecules (*red traces* in Fig. 4*B*), similar to the MONSA model in Fig. 5, do not build a compact cluster but, instead, are spread over the surface of the entire N terminus.

The variability and accessible conformational space of the Pex5p(F)·Pex14p(N)·SCP2 is illustrated in supplemental Fig. S1. It is evident that the ternary complex displays rather flexible organization of the individual components, but the overall shape of all the models remains similar. The average NSD values for the most typical *ab initio* and most typical rigid body models are 1.31 and 1.45, respectively, indicating a good reproducibility of the results.

CD—To independently verify the SAXS results suggesting a disordered conformation of the N-terminal part of Pex5p(F), the full-length Pex5p, Pex5p(N), and Pex5p(C) were analyzed by CD. The results (Fig. 6) indicate that Pex5p(C) is largely α-helical, whereas Pex5p(N) shows little secondary structure. The absence of secondary structure further corroborates the SAXS-based model displaying an extended Pex5p(F).

DISCUSSION

The SAXS model of the solution structure of human Pex5p(F) and the SLS results indicate that in solution human Pex5p(F) is an elongated monomer at the concentrations used (0.14–2.5 mg/ml), in agreement with recent studies (15, 16). Furthermore, the shape obtained from SAXS reveals that the unfolded N-terminal region and the globular C-terminal TPR domain are arranged in line. The Pex5p(F) tetramer was also observed by electron microscopy (EM) (35) together with filamentous aggregates.

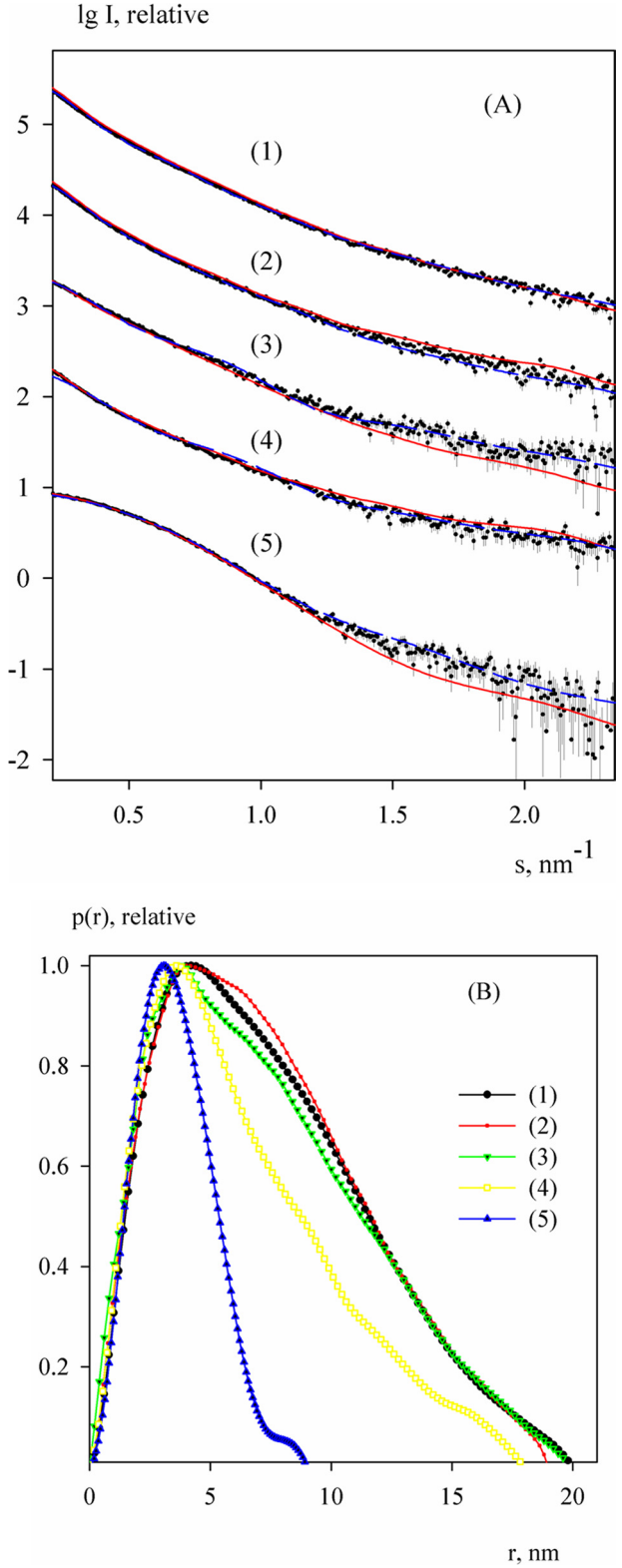

FIGURE 3. **Experimental x-ray scattering pattern and distance distribution functions of Pex5p(F).** *A*, scattering profiles of Pex5p(F)·Pex14p(N)·SCP2 complexes. *Curve 1*, ternary complex of Pex5p(F)·Pex14p(N)·SCP2 (1:6:1 ratio). *Curves 2* and *3* represent binary complexes of Pex5p(F)·Pex14p(N) (1:6 ratio) and Pex5p(F)·SCP2 (1:1 ratio), respectively. *Curve 4* corresponds to Pex5p(F) protein, and *curve 5* to the truncated Pex5p(C)·Pex14p(N)·SCP2 (1:1:1 ratio) complex. *Dots with error bars* denote the experimental scattering data. The fits obtained by MONSA and SASREF are displayed as *solid* (*red*) and *dashed* (*blue*) lines. The plot displays the logarithm of the scattering intensity as a function of momentum transfer. The curves are arbitrarily displaced along the vertical logarithmic axis for clarity. *B*, distance distribution functions of Pex5p(F)·Pex14p(N)·SCP2 constructs computed from the experimental x-ray scattering patterns using GNOM and normalized to a maximum value of unity. The notations of the curves are the same as in *A*.

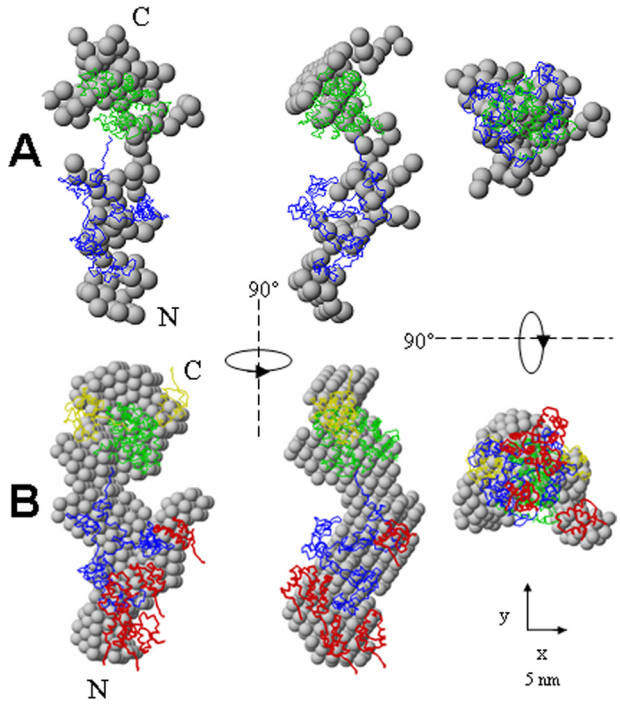

FIGURE 4. **Structural models of Pex5p constructs.** Pex5p(F) alone (*A*) and Pex5p(F)·Pex14p(N)·SCP2 complex (*B*) for 1:6:1 molar ratio obtained by DAMMIN (*gray spheres*) superimposed with the rigid body model of the complex by SASREF. The C-terminal part of Pex5p (taken from crystallographic structure 1FCH (2)) is displayed as *green* C traces. The N-terminal part of Pex5p restored from SASREF is shown as *blue* C traces. The five Pex14p(N) proteins bound to the Pex5p(F) N terminus are drawn as *red* C traces, the sixth Pex14p(N) and SCP2 proteins are indicated by *yellow* C traces. *Right* and *bottom views* are rotated counterclockwise by 90° around the *y* and *x* axes, respectively.

We performed titration experiments by SAXS and SLS, which independently pointed to a Pex5p(F)·Pex14p(N) stoichiometry of about (1:6). These results are in good agreement with previous surface plasmon resonance experiments and fluorescence titration analysis of the interaction of Pex14p (1–78) with synthetic peptides (W*XXX*(F/Y) motifs), which proposed 1:6–7 complexes (8, 9). Previous data indicated, however, that not all Pex14p binding motifs in Pex5p are required for its function as an PTS1-import receptor (9). Electrophoretic experiments of the related Pex5p(F)·Pex14p(F) complex, isolated from rat liver, also indicated that more than one Pex14p ligand is bound to the Pex5p receptor, with a probable stoichiometry of about 1:5 (17). Of course, the methods employed for titration in the present study (SAXS and SLS) are not site-specific, *i.e.* they give the overall stoichiometry only and do not provide evidence about the occupancy of the specific Pex14p(N) binding sites on Pex5p(F).

No significant structural changes are detected by SAXS upon complex formation of human Pex5p(F), *i.e.* Pex5p(F)·Pex14p(N)·SCP2 retains an elongated shape, and if any, only minor conformational changes occur. Dynamics or instability in the protein fold would be expected for a region involved in protein-protein

interaction (36). The natively disordered region and extended Stokes radius can lead to "Fly casting" (37), where proteins can "fish" for binding partners. This is a particularly effective means of recruiting protein assemblies at the membrane surfaces. For example, in mammalian endocytosis, the native disordered region of adapter proteins recruits clathrin at the membrane surface (38). Pex5p(F) containing a significant unfolded N-terminal domain could connect with other proteins or the membrane across a relatively large distance. It is, therefore, possible for Pex5p(F) to present several binding sites to a relatively wide area, as the SAXS model showed, and to improve the efficiency with which proteins are assembled or become recruited to the membrane. In addition, the long and unfolded N-terminal domain of Pex5p(F) on the membrane could allow the binding partners to remain close together, which facilitates the efficient exchange of binding partners after each binding interaction has fulfilled its purpose. By analogy with these observations, the unstructured N-terminal part of Pex5p(F) and its elongated shape might be necessary for the Pex5p(F)·Pex14p(N)·PTS1 complex to interact with other substrates, such as lipids and membrane proteins during the formation of functional multiprotein complexes (36, 39, 40).

The binding properties of human Pex5p differ significantly from those of its equivalent in lower organisms. In higher eukaryotes (plants, mammals), Pex5p and Pex7p interact with Pex13p and Pex14p at the peroxisomal membrane during the docking step of peroxisomal protein import, whereas in lower eukaryotes (yeast, fungi), they bind to Pex13p, Pex14p, and Pex17p (41). EM studies have revealed that *Hansenula polymorpha* Pex5p(F) (Hp Pex5p) is a globular tetramer near pH 7 which exists in two forms (open/closed) and undergoes a conformational change upon interaction with Hp Pex20p (42). A large conformational change was also observed by fluorescence spectroscopy in Hp Pex5p upon interaction with Hp Pex8p, an intraperoxisomal protein for which there is no equivalent in humans (18). In contrast, our data on the

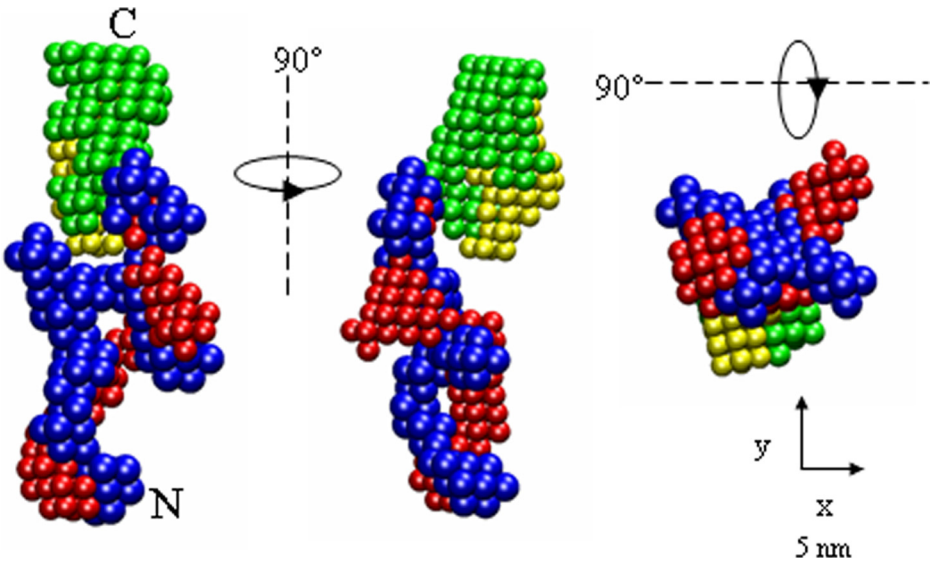

FIGURE 5. ***Ab initio* bead model of the Pex5p(F)·Pex14p(N)·SCP2 complex obtained by MONSA.** The colors of the components in the complex are the same as in Fig. 4. The *right* and *bottom views* are rotated counterclockwise by the 90° around the *y* and *x* axes, respectively.

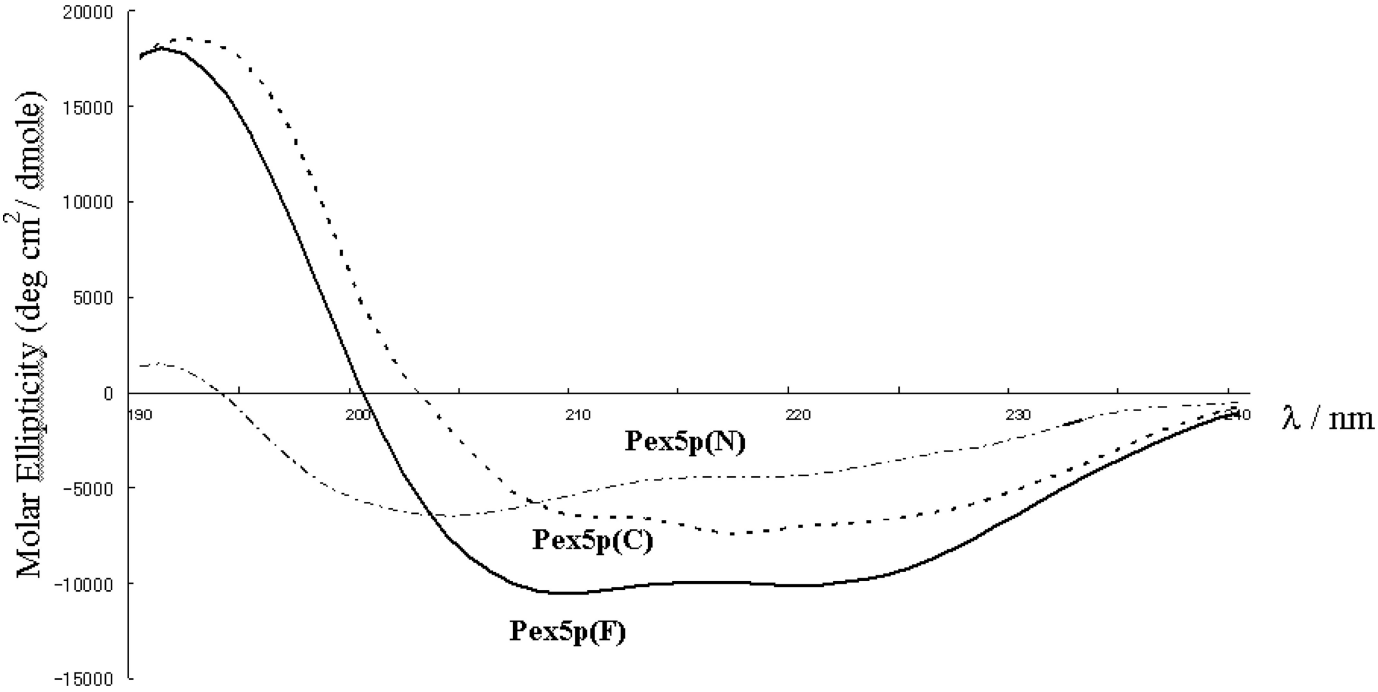

FIGURE 6. **Far-UV CD spectra of Pex5p(F) (*full line*), Pex5p(N) (*dot-dashed line*), and Pex5p(C) (*dashed line*).** Pex5p(N) has little secondary structure, whereas Pex5p(C) is α-helical.

human Pex5p receptor do not indicate either homo-oligomerization of any of the Pex5p fragments or major structural changes upon Pex14p(N)/SCP2 binding. It remains to be determined whether the different findings could be attributed to variations in experimental conditions or they reflect possible taxonomic differences in terms of architecture and function of the Pex5p receptor.

Acknowledgments—K. S. thanks Dr. M. H. J. Koch, Dr. W. A. Stanley, Dr. C. Williams, and Dr. K. Fodor for critical reading the manuscript and valuable comments. We thank Dr. H. Mertens (EMBL, Hamburg Outstation) for help in preparing the manuscript.

REFERENCES

1. Purdue, P. E., and Lazarow, P. B. (2001) *Annu. Rev. Cell Dev. Biol.* **17,** 701–752
2. Gatto, G. J., Jr., Geisbrecht, B. V., Gould, S. J., and Berg, J. M. (2000) *Nat. Struct. Biol.* **7,** 1091–1095
3. Kumar, A., Roach, C., Hirsh, I. S., Turley, S., deWalque, S., Michels, P. A., and Hol, W. G. (2001) *J. Mol. Biol.* **307,** 271–282
4. Stanley, W. A., Filipp, F. V., Kursula, P., Schüller, N., Erdmann, R., Schliebs, W., Sattler, M., and Wilmanns, M. (2006) *Mol. Cell* **24,** 653–663
5. Stanley, W. A., Fodor, K., Marti-Renom, M. A., Schliebs, W., and Wilmanns, M. (2007) *FEBS Lett.* **581,** 4795–4802
6. Azevedo, J. E., and Schliebs, W. (2006) *Biochim. Biophys. Acta* **1763,** 1574–1584
7. Nito, K., Hayashi, M., and Nishimura, M. (2002) *Plant. Cell Physiol.* **43,** 355–366
8. Saidowsky, J., Dodt, G., Kirchberg, K., Wegner, A., Nastainczyk, W., Kunau, W. H., and Schliebs, W. (2001) *J. Biol. Chem.* **276,** 34524–34529
9. Otera, H., Setoguchi, K., Hamasaki, M., Kumashiro, T., Shimizu, N., and Fujiki, Y. (2002) *Mol. Cell. Biol.* **22,** 1639–1655
10. Neufeld, C., Filipp, F. V., Simon, B., Neuhaus, A., Schüller, N., David, C., Kooshapur, H., Madl, T., Erdmann, R., Schliebs, W., Wilmanns, M., and Sattler, M. (2009) *EMBO J.* **28,** 745–754
11. Gouveia, A. M., Guimarães, C. P., Oliveira, M. E., Sá-Miranda, C., and Azevedo, J. E. (2003) *J. Biol. Chem.* **278,** 4389–4392
12. Holroyd, C., and Erdmann, R. (2001) *FEBS Lett.* **501,** 6–10
13. Madrid, K. P., De Crescenzo, G., Wang, S., and Jardim, A. (2004) *Mol. Cell. Biol.* **24,** 7331–7344
14. Schnell, D. J., and Hebert, D. N. (2003) *Cell* **112,** 491–505
15. Carvalho, A. F., Costa-Rodrigues, J., Correia, I., Costa Pessoa, J., Faria, T. Q., Martins, C. L., Fransen, M., Sá-Miranda, C., and Azevedo, J. E. (2006) *J. Mol. Biol.* **356,** 864–875
16. Costa-Rodrigues, J., Carvalho, A. F., Fransen, M., Hambruch, E., Schliebs, W., Sá-Miranda, C., and Azevedo, J. E. (2005) *J. Biol. Chem.* **280,** 24404–24411
17. Gouveia, A. M., Reguenga, C., Oliveira, M. E., Sa-Miranda, C., and Azevedo, J. E. (2000) *J. Biol. Chem.* **275,** 32444–32451
18. Wang, D., Visser, N. V., Veenhuis, M., and van der Klei, I. J. (2003) *J. Biol. Chem.* **278,** 43340–43345
19. Stanley, W. A., Versluis, K., Schultz, C., Heck, A. J., and Wilmanns, M. (2007) *Arch. Biochem. Biophys.* **461,** 50–58
20. Wen, J., Arakawa, T., and Philo, J. S. (1996) *Anal. Biochem.* **240,** 155–166
21. Koch, M. H., and Bordas, J. (1983) *Nucl. Instrum. Methods* **208,** 461–469
22. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) *J. Appl. Crystallogr.* **36,** 1277–1282
23. Guinier, A. (1939) *Ann. Phys.* **12,** 161–237
24. Svergun, D. I. (1992) *J. Appl. Crystallogr.* **25,** 495–503
25. Porod, G. (1982) in *General Theory. Small-angle X-ray Scattering* (Glatter, O., and Kratky, O., eds) pp. 17–51, Academic Press, Inc., London
26. Svergun, D. I. (1999) *Biophys. J.* **76,** 2879–2886
27. Svergun, D. I., and Nierhaus, K. H. (2000) *J. Biol. Chem.* **275,** 14432–14439
28. Petoukhov, M. V., and Svergun, D. I. (2005) *Biophys. J.* **89,** 1237–1250
29. Svergun, D. I., Barberato, C., and Koch, M. H. J. (1995) *J. Appl. Crystallogr.* **28,** 768–773
30. Volkov, V. V., and Svergun, D. I. (2003) *J. Appl. Crystallogr.* **36,** 860–864
31. Kozin, M. B., and Svergun, D. I. (2001) *J. Appl. Crystallogr.* **34,** 33–41
32. Feigin, L. A., and Svergun, D. I. (1987) *Structure Analysis by Small-angle X-ray and Neutron Scattering* (Taylor, G. W., ed) Plenum Press, New York
33. Rubinstein, M., and Colby, R. H. (2003) *Polymer Physics*, Oxford University Press, Oxford, UK
34. Stanley, W. A., Sokolova, A., Brown, A., Clarke, D. T., Wilmanns, M., and Svergun, D. I. (2004) *J. Synchrotron Radiat.* **11,** 490–496
35. Schliebs, W., Saidowsky, J., Agianian, B., Dodt, G., Herberg, F. W., and Kunau, W. H. (1999) *J. Biol. Chem.* **274,** 5666–5673
36. Dyson, H. J., and Wright, P. E. (2005) *Nat. Rev. Mol. Cell Biol.* **6,** 197–208
37. Shoemaker, B. A., Portman, J. J., and Wolynes, P. G. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97,** 8868–8873
38. Dafforn, T. R., and Smith, C. J. (2004) *EMBO Rep.* **5,** 1046–1052
39. Dodt, G., Warren, D., Becker, E., Rehling, P., and Gould, S.J. (2001) *J. Biol. Chem.* **276,** 41769–41781
40. Kerssen, D., Hambruch, E., Klaas, W., Platta, H. W., de Kruijff, B., Erdmann, R., Kunau, W. H., and Schliebs, W. (2006) *J. Biol. Chem.* **281,** 27003–27015
41. Brown, L. A., and Baker, A. (2003) *J. Cell. Mol. Med.* **7,** 388–400
42. Moscicka, K. B., Klompmaker, S. H., Wang, D., van der Klei, I. J., and Boekema, E. J. (2007) *FEBS Lett.* **581,** 1758–1762

Solution Structure of Human Pex5·Pex14·PTS1 Protein Complexes Obtained by Small Angle X-ray Scattering

Kumiko Shiozawa, Petr V. Konarev, Christian Neufeld, Matthias Wilmanns and Dmitri I. Svergun

J. Biol. Chem. 2009, 284:25334-25342.
doi: 10.1074/jbc.M109.002311 originally published online July 6, 2009
