Interaction Surface and Topology of Get3-Get4-Get5 Protein Complex, Involved in Targeting Tail-anchored Proteins to Endoplasmic Reticulum*‡§

Yi-Wei Chang‡1, Tai-Wen Lin‡1, Yi-Chuan Li§†, Yu-Shan Huang§, Yuh-Ju Sun§, and Chwan-Deng Hsiao‡2

From the ‡1Institute of Molecular Biology, Academia Sinica, Taipei 115 and the §2Institute of Bioinformatics and Structural Biology, National Tsing Hua University and the ‡3National Synchrotron Radiation Research Center, Hsinchu 300, Taiwan

Background: The GET system mediates the insertion of tail-anchored (TA) proteins into the endoplasmic reticulum membrane.
Results: The protein complex comprising most of the cytosolic portions of the GET system shows an extended conformation in solution.
Conclusion: The ternary complex Get3-Get4-Get5 forms an elongated structure with 2:2:2 stoichiometry for conducting TA protein delivery.
Significance: The structure provides a framework for TA protein insertion into the membrane.

Recent work has uncovered the “GET system,” which is responsible for endoplasmic reticulum targeting of tail-anchored proteins. Although structural information and the individual roles of most components of this system have been defined, the interactions and interplay between them remain to be elucidated. Here, we investigated the interactions between Get3 and the Get4-Get5 complex from Saccharomyces cerevisiae. We show that Get3 interacts with Get4-Get5 via an interface dominated by electrostatic forces. Using isothermal titration calorimetry and small-angle x-ray scattering, we further demonstrate that the Get3 homodimer interacts with two copies of the Get4-Get5 complex to form an extended conformation in solution.

The tail-anchored (TA) proteins are a distinct class of membrane proteins characterized by a single C-terminal transmembrane domain (TMD) that is anchored into the phospholipid bilayer that surrounds cellular organelles. The N-terminal cytosolic portion is thereby correctly orientated for its particular intracellular function (1). TA proteins are found mostly in the nuclear envelope, peroxisomes, mitochondria, endoplasmic reticulum, and Golgi, and they have many diverse functions, such as regulating apoptosis (e.g. the Bcl-2 family) (2), operating vesicular transport via intracellular membrane fusion (e.g. SNAREs) (3), translocating proteins into mitochondria (e.g. Tom5, Tom6, Tom7, and Tom22) (4) or across the ER membrane (e.g. Sec61β and Sec61γ) (5), and assisting in folding or degradation of membrane proteins (1, 6). Biogenesis of TA proteins requires a post-translational targeting mechanism for correct localization and membrane insertion (6). During targeting, the composition of the C-terminal and flanking regions of the TMD helps to determine the destination of a TA protein (7): variations in length and hydrophobicity in the TMD can be distinguished by different protein trafficking systems, resulting in localization into the membrane of specific cellular organelles.

In Saccharomyces cerevisiae, the biomolecular machinery involved in recognizing and delivering the TMD of TA proteins bound for the ER has been identified and is composed of the proteins Get1, Get2, Get3, Get4, Get5, and Sgt2 (8–11). Briefly, the pathway begins when Sgt2 recognizes the TMD of a nascent TA protein by its Met-rich C-terminal domain and then binds to a ubiquitin-like domain of Get5 via its N-terminal dimerization domain (11, 12). The N-terminal region of Get5 (referred to as Get5N) then binds to the C-terminal region of Get4 to form a stable complex, predominantly through hydrophobic interactions (13). Mediated by interactions between the N-terminal regions of Get4 and Get3 (13, 14), Sgt2 becomes close enough to Get3 to enable the TMD of the TA protein to be loaded into the hydrophobic groove on the Get3 dimer (11, 15). The membrane protein Get2 then acts as receptor for recruiting the Get3-TA protein complex onto the ER surface, and another membrane protein, Get1, triggers the release of the TMD of the TA protein from Get3, thereby achieving membrane insertion (8, 16–18).

Several studies have examined the structural interactions, motions, and interplay between the proteins in this system, including crystallographic structures of the Get3 dimer with and without bound nucleotide (15, 19–22), the structure of Get4 complexed with the N-terminal domain of Get5 (referred to as Get4/5N) (13, 14), and structures of the Get3 dimer complexed to the cytosolic domains of Get1 and Get2 (16). Detailed structural information of other interactions involving the GET proteins still needs to be elucidated, e.g. the Get3 interaction with Get4 and Get5 and the Get5 interaction with Sgt2. Here,
we investigated the detailed interaction between dimeric Get3 and Get4/5N using computational docking and isothermal titration calorimetry (ITC) together with protein mutagenesis. Using small-angle x-ray scattering (SAXS), we also determined the solution structure of the multiprotein complex consisting of Get3, Get4, and the Get5 N-terminal domain (Get3-Get4/5N). Our results provide a better understanding of the composition, arrangement, and stoichiometry of the Get3-Get4/5N complex.

EXPERIMENTAL PROCEDURES

Construction of Plasmids for Protein Expression—Plasmids were constructed using standard molecular cloning and recombinant DNA techniques. Expression plasmids containing GET3 were constructed using PCR amplification of its coding sequence from genomic DNA of S. cerevisiae. During amplification, an NdeI restriction endonuclease site was introduced at the 5'-end of the coding sequence, and an Xhol site was introduced after the stop codon. The NdeI-Xhol fragment thus obtained was cloned into the pET-21b(+) vector (Novagen) using the same two restriction sites. The resulting plasmid, pET-21b(+)/Get3, was used to express recombinant Get3 in Escherichia coli. To express Get4 and Get5N simultaneously, the coding sequences of full-length GET4 and the first 59 amino acid residues of GET5 were inserted into the multiple cloning sites MCS2 and MCS1, respectively, of vector pRSFDuet-1 (Novagen). Plasmid pRSFDuet-1/(Get4/5N) was used to simultaneously express His6-tagged Get4 and untagged Get5N. The Get4/5N mutant constructs used for the ITC assays were produced using a QuikChange site-directed mutagenesis kit (Stratagene) using pRSFDuet-1/(Get4/5N) as the template.

Protein Expression and Purification—Expression of recombinant Get3 was induced by the addition of isopropyl β-D-thiogalactopyranoside (to give a final concentration of 1 mM) to an E. coli BL21(DE3) culture harboring plasmid pET-21b(+)/Get3. Cells were collected by centrifugation and then resuspended in homogenization buffer (20 mM Tris-HCl, 500 mM NaCl, and 5 mM imidazole, pH 7.9) with protease inhibitors (cOmplete ULTRA protease inhibitor mixture tablets, EDTA-free, Roche Applied Science). The cells were homogenized using a Microfluidizer (Microfluidics) under high pressure, and the lysate was centrifuged at 40,000 × g for 1 h at 4°C. The supernatant was then loaded onto a nickel-Sepharose affinity column (HisTrap HP column, 5 ml, GE Healthcare) and eluted with a linear imidazole gradient (5–300 mM imidazole in 20 mM Tris-HCl and 500 mM NaCl, pH 7.9). Fractions containing Get3 were pooled and further purified by size-exclusion chromatography using a HiLoad 16/60 Superdex 200 column (GE Healthcare) equilibrated with buffer containing 20 mM Tris-HCl, pH 7.9, and 300 mM NaCl. The same procedure was used to purify the recombinant Get4/5N complexes. For SAXS data collection, purified Get3 dimer and Get4/5N were mixed at a 1:1 molar ratio and dialyzed against dialysis buffer (10 mM Tris-HCl, 50 mM NaCl, and 2 mM MgCl2, pH 7.9). The Get3-Get4/5N complex was further purified by size-exclusion chromatography using the Superdex 200 column and buffer conditions described above.

Molecular Docking—Crystal structures of the S. cerevisiae nucleotide-free Get3 dimer (Protein Data Bank code 3H84) (20) and the Get4/5N complex (code 2WPV) (13) were docked using the ClusPro v1.0 protein-protein docking server (23, 24) with the DOT algorithm. According to our previous yeast two-hybrid results (13), a Protein Data Bank mask was applied to residues 149–292 of Get4 and the entire sequence of Get5N to specify repulsion of these residues during the Get3 binding calculation. Using the recommended default procedure, 20,000 models were obtained in total and ranked based on their electrostatic and desolvation energies. The 1500 solutions with the lowest values for electrostatic energy and the 500 solutions with the lowest values for desolvation free energy were then clustered and ranked using the ClusPro algorithm (24). After clustering, van der Waals minimization was carried out using CHARMM (25) to remove potential side chain clashes. The top-scoring model in the electrostatically favored category was used for all other studies in this work.

ITC—Experiments were carried out using the MicroCal iTC200 system (GE Healthcare). For all isotherms, the sample cell was filled with 35 μM Get3 in 10 mM sodium phosphate, pH 7.4, and 2 mM MgCl2, and the syringe was loaded with 350 μM Get4/5N or the same concentration of the various mutants in the same buffer. Each binding isotherm was obtained from 20 injections of Get4/5N into Get3 at 25°C. For the first titration, an injection volume of 1 μl was used; for subsequent titrations, 2 μl of Get4/5N was injected with an interval of 240 s. The stirring speed and reference power were 1000 rpm and 5 μcal/s, respectively. Binding isotherms were integrated and analyzed using Origin v7.0 software (MicroCal).

SAXS—SAXS data for the Get3-Get4/5N complex were collected at 20°C using the 14-keV beamline BL23A1 at the National Synchrotron Radiation Research Center. Data were collected using a 1 mg/ml protein solution in 10 mM Tris-HCl, 50 mM NaCl, and 2 mM MgCl2, pH 7.9. Data were recorded for a total q range from 0.0072 to 0.3357 Å⁻¹. Background scattering from the buffer was subtracted, and data were scaled using the software PRIMUS (26). Scattering data were analyzed using the program GNOM (27). Individual ab initio conformations of the Get3-Get4/5N complex, each composed of 1508 dummy residues, were calculated using GASBOR (28) and then averaged using DAMAVER (29). The superposition of the Get3-Get4/5N atomic model onto the averaged SAXS model was carried out using SUPCOMB (30). The pair-distance distribution function (P(r)) was calculated using GROMOS. Radii of gyration (Rg) were computed from the Guinier plot using PRIMUS and GNOM. The calculated scattering curve and goodness of fit (χ) based on the atomic model of Get3-Get4/5N were obtained using CRYSOL (31).

RESULTS

Computational Docking of Get4/5N with Nucleotide-free Get3 Dimer—Based on protein pulldown, yeast two-hybrid, and chromatographic assays, it has been reported that the N-terminal portion of Get4 (residues 1–148) interacts directly with Get3 (13, 14). Here, we used a computational docking method to investigate this interaction in detail. Hu et al. (20) proposed that TA proteins first bind to an “open” nucleotide-free form of Get3, and upon ATP binding, Get3 closes around the TMD in a protective fashion. Wang et al. (11) reported that
the Get4-Get5-Sgt2 complex conducts the “handoff” of the TA protein to Get3. This model predicts that Get4-Get5 needs to first recognize the open form of Get3 to take delivery of the TA protein. Even though Chartron et al. (14) demonstrated that the binding affinity of Get3 toward the immobilized Get4-Get5 complex (Protein Data Bank code 2WPV), A, close-up of the Get3-Get4/5N interface. Residues on Get4 with strong interactions with Get3 are shown in blue, and those on Get3 with strong and weak interactions with Get4 are shown in red and magenta, respectively. The relative strength of the interactions was determined individually from ITC experiments on site-directed mutants of Get3 and Get4. C, view of the Get3-Get4/5N complex orthogonal to that shown in B. There was no significant difference in binding (as determined by ITC) between wild-type and mutant proteins for single-point mutations of those residues shown in gray.

**TABLE 1**

**ITC data for Get4/5N binding to Get3**

<table>
<thead>
<tr>
<th>Protein in cell</th>
<th>Protein in syringe</th>
<th>$K_d \times 10^{-7}$ M</th>
<th>Stoichiometry of binding$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Get3</td>
<td>Get4/5N</td>
<td>6.2</td>
<td>0.90</td>
</tr>
<tr>
<td>Get4(K12E)/5N</td>
<td>Get4/5N</td>
<td>357.1</td>
<td>0.35</td>
</tr>
<tr>
<td>Get4(K15E)/5N</td>
<td>Get4/5N</td>
<td>107.4</td>
<td>0.97</td>
</tr>
<tr>
<td>Get4(R19E)/5N</td>
<td>Get4/5N</td>
<td>ND$^b$</td>
<td>ND</td>
</tr>
<tr>
<td>Get4(K23E)/5N</td>
<td>Get4/5N</td>
<td>45.4</td>
<td>0.35</td>
</tr>
<tr>
<td>Get4(Y30A)/5N</td>
<td>Get4/5N</td>
<td>8.7</td>
<td>0.83</td>
</tr>
<tr>
<td>Get4(R42E)/5N</td>
<td>Get4/5N</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Get4(R45E)/5N</td>
<td>Get4/5N</td>
<td>74.1</td>
<td>0.94</td>
</tr>
<tr>
<td>Get3(Y250A)</td>
<td>Get4/5N</td>
<td>7.8</td>
<td>1.14</td>
</tr>
<tr>
<td>Get3(K253R)</td>
<td>Get4/5N</td>
<td>161.3</td>
<td>0.67</td>
</tr>
<tr>
<td>Get3(K256R)</td>
<td>Get4/5N</td>
<td>9.3</td>
<td>0.88</td>
</tr>
<tr>
<td>Get3(K259E)</td>
<td>Get4/5N</td>
<td>13.2</td>
<td>0.90</td>
</tr>
<tr>
<td>Get3(K297E)</td>
<td>Get4/5N</td>
<td>18.1</td>
<td>1.08</td>
</tr>
<tr>
<td>Get3(D300R)</td>
<td>Get4/5N</td>
<td>180.2</td>
<td>0.53</td>
</tr>
<tr>
<td>Get3(D303R/E304R)</td>
<td>Get4/5N</td>
<td>67.1</td>
<td>0.80</td>
</tr>
<tr>
<td>Get3(E307R/D308R)</td>
<td>Get4/5N</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

$^a$ Determined from the fit to the data (see “Experimental Procedures”) and representing [protein in the syringe]/[protein in the cell] at the point of saturation of the protein in the cell.

$^b$ ND, not determined.

the Get4-Get5-Sgt2 complex conducts the “handoff” of the TA protein to Get3. This model predicts that Get4-Get5 needs to first recognize the open form of Get3 to take delivery of the TA protein. Even though Chartron et al. (14) demonstrated that the binding affinity of Get3 toward the immobilized Get4-Get5 complex in pulldown assays could be enhanced by using adenine nucleotides, the reciprocal binding assay using immobilized Get3 against Get4-Get5 showed no nucleotide dependence. Moreover, the crystal structures of the nucleotide-free and ADP-bound forms of Get3 from *S. cerevisiae* were both shown to adopt the open form (19, 20, 22), which suggests that the reported nucleotide dependence of the interaction between Get3 and Get4-Get5 may not be directly relevant to the open/closed conformation of the Get3 dimer for Get4-Get5 recognition. We therefore examined the crystal structures of the nucleotide-free open form of Get3 from *S. cerevisiae* (Protein Data Bank code 3H84) (20) and Get4/5N (code 2WPV) (13) and submitted these in the protein-protein docking server ClusPro (23, 24). Based on published yeast two-hybrid data from our laboratory (13), residues 149–292 of Get4 and the entire sequence of Get5N were designated as repulsive regions to Get3 binding during the calculation. Based on gel filtration data (supplemental Fig. S1), purified Get3-Get4/5N disassociates into Get3 and Get4/5N at relatively high salt concentrations, indicating that interactions between Get3 and Get4/5N are predominantly electrostatic. The results of docking calculations were therefore
ranked using this criterion. The docking algorithm was run twice, and 30 solutions were generated from the input structures by each run. The top-scoring structures were identical. The resulting computational docking model of the Get3-Get4/5N complex had an elongated shape (Fig. 1A), revealing a complementary electrostatic potential on the interface between the open dimer of Get3 and Get4/5N. Interactions are made between the charged residues Lys-12, Lys-15, Arg-19, Lys-23, Arg-42, and Arg-45 of N-terminal α-helices A and B of Get4 and the charged residues Glu-253, Lys-293, Lys-297, Asp-300, Asp-303, Glu-304, Glu-307, and Asp-308 of α-helices α11 and α12 of Get3 (Fig. 1B).

Binding Affinities Determined by ITC for Get3 and Get4 Mutants in Get3-Get4/5N Complex—To investigate the interactions between Get3 and Get4/5N predicted by the computational docking, charged residues at the interface (Fig. 1A) were individually mutated to oppositely charged amino acids. ITC experiments were then performed with mutant Get3 and wild-type Get4/5N and with wild-type Get3 and mutant Get4/5N (Table 1 and supplemental Fig. S2). For Get3, the point mutations E253R and D300R and the double mutant D303R/E304R caused a significant reduction in binding affinity toward Get4/5N, and the E307R/D308R double mutant abolished the interaction with Get4/5N. Lys-293 and Lys-297 of Get3 are located at the edge of the interface, and as predicted, mutants K293E and K297E showed only a small reduction in binding affinity. Get3(E258R), used as a negative control, showed no change in binding affinity. For Get4, the K12E, K15E, K23E, and R42E mutants exhibited significantly lower binding affinities, presumably indicating significant disruptions in the interaction between Get3 and Get4/5N, and the mutations R19E and R42E of Get4 appeared to abolish binding to Get3. Although Get3(Y250A) and Get4(Y30A) have previously shown a significant decrease in partner protein capture (14), a similar effect was not observed in our ITC assay. Nonetheless, our ITC results are fully consistent with our computationally generated docking model. The ITC results also revealed a 1:1 stoichiometry between Get3 and Get4/5N, and the functional form of the Get3-Get4-Get5 complex must therefore be composed of two molecules each of Get3, Get4, and Get5. Using this molecular composition, the following SAXS experiment was performed.

Structure of Get3-Get4/5N in Solution Determined by SAXS—To obtain further support for the proposed binding model between Get3 and Get4/5N, we determined the con-
formation of the complex in solution using SAXS. The curve-fitted SAXS data are shown in Fig. 2A. Curve fitting was repeated 10 times, and the resulting 10 output structures, each consisting of two copies of Get3, Get4, and Get5N (i.e., 1508 dummy residues in total), were averaged and generated as a molecular envelope. A protein model of the Get3-Get4/5N complex (i.e., two Get4/5N complexes bound one per side to a Get3 open dimer) was generated by applying 2-fold symmetry to the computational docking model and then superimposed into the envelope (Fig. 2B). Significantly, the envelope of Get3-Get4/5N had an elongated shape and aligned well with the 2-fold symmetry model. The observed SAXS curve was also in good agreement with the theoretical curve calculated from the model with 2-fold symmetry (with goodness of fit ($\chi^2$) of 5.18) (Fig. 2C). The distance distribution ($P(r)$) functions calculated from primary scattering data and from the model are compared in Fig. 2D. The slight dissimilarity may be due to structural flexibility of the protein complex in solution. Very similar $R_g$ values between Get3-Get4/5N in solution (46.52 Å calculated by GNOM and 46.4 Å calculated by PRIMUS) and the 2-fold symmetry model (46.38 Å calculated by GNOM) strongly support the correctness of the model, however. These results indicate that, in solution, the Get3 dimer binds to the termini of two Get4/5N molecules, resulting in an elongated overall structure of the Get3-Get4/5N complex. This is consistent with the binding model proposed by the molecular docking calculations and the stoichiometry of the complex determined by ITC.

**DISCUSSION**

We recently reported the crystal structure of *S. cerevisiae* Get4/5N and predicted a potential Get3-binding site located at the N-terminal region of Get4 based on conserved surface residues (13). Confirmation of this location was obtained with a yeast two-hybrid assay using Get3 with different bisections of Get4 (13). Using mutagenesis and immunoprecipitation, Chartron et al. (14) subsequently identified several key residues at the Get3-Get4 interaction surface. Using size-exclusion chromatography and domain-swapping experiments, they also determined that the Get4-Get5 complex is dimeric in solution, and dimerization is mediated via the C terminus of Get5. Here, we used several different methods to further elucidate the interaction between Get3 and Get4/5N. First, size-exclusion chromatography revealed a salt-sensitive interaction between Get3 and Get4/5N, suggesting that the interaction is predominantly electrostatic. Second, computational docking of Get4/5N with the open form of Get3 predicted a highly charged interface, the location of which correlates well with the N-terminal region of Get4 in our previous prediction. The charged residues distributed on this surface are also consistent with those reported by Chartron et al. Third, we used ITC to analyze the involvement of individual charged residues at the interface of Get3 and Get4. By mutating one residue at a time, our results went beyond those of Chartron et al., where neighboring charged residues were simultaneously mutated in pairs to examine the effect on binding. Our ITC experiments revealed a 1:1 stoichiometry between Get3 and Get4/5N, indicating that the interface is composed of an equal number of both Get3 and Get4/5N.
Moreover, even though Chartron et al. demonstrated a nucleotide dependence of the interaction between Get3 and Get4-Get5, Get3 and Get4/5N showed a remarkable affinity in the absence of nucleotide in our ITC experiment. Accordingly, we believe that nucleotides may facilitate the interaction but are not essential. Chartron et al. reported that nucleotide binding to Get3 might cause regional conformational change at certain residues, such as Tyr-250, which was shown to participate in the interaction between Get3 and Get4-Get5 and further enhance the interaction. However, this phenomenon was not detected in our nucleotide-free ITC assay. Fourth, solution SAXS experiments clearly revealed an elongated molecular shape for the Get3-Get4/5/N complex. Because Get4-Get5 has been identified as dimeric in solution (14, 15), the functional Get3-Get4-Get5 complex is assumed to comprise two Get3 molecules and two Get4-Get5 complexes. A molecular model consisting of one Get3 dimer bound to one Get4/5N complex on either side fits well into the SAXS-determined molecular envelope. In our model, binding of dimeric Get4-Get5 to the Get3 dimer would then cause the ubiquitin-like domains of Get5 to become spatially positioned to enable binding of Sgt2 to Get3, thereby facilitating handoff of the TA protein.

Using size-exclusion chromatography together with multangle laser light scattering, Chartron et al. (32) recently demonstrated that the tetraticopeptide repeat region of a dimer of the N-terminal (dimerization) domain of Sgt2 interacts with the Get4-Get5 dimer and is the dominant species in solution. The stoichiometry between Get5 and Sgt2 was also confirmed by SAXS using the purified Sgt2 N-terminal dimerization domain coupled to the Get5 ubiquitin-like domain. Thus, based on our results and those of Chartron et al., the architecture of the cytosolic portion of the GET system can be visualized: the Get3 dimer is tethered by the two Get4 moieties from the Get4-Get5 dimer and awaits delivery of the TA protein. Within the Get4-Get5 dimer, only one Get5 molecule interacts with the Sgt2 dimer to guide the delivery of the TA protein from Sgt2 to the Get3 dimer.

Recent studies have shown that Get1 and Get2 interact competitively with Get3 on a conserved peptide motif, DELYED (16, 17). Upon binding to the closed form of the Get3 dimer, Get2 could be competitively displaced by Get1, which triggers opening of the Get3 dimer and release of the TA protein. In this study, we have shown that the Get4-Get5-binding site on Get3 also contains the DELYED motif. This suggests that Get4-Get5 may also use a similar mechanism to regulate Get3 conformation for uptake of the TA protein from Sgt2, although the details of this mechanism remain to be elucidated. A proposed working framework of the TA protein membrane insertion cycle by the GET system is summarized in Fig. 3.

Acknowledgment—We are grateful for access to synchrotron radiation beamline BL23A1 at the National Synchrotron Radiation Research Center.

REFERENCES


