Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the Sgt2 dimerization domain complexed with the Get5 UBL domain involved in the targeting of tail-anchored membrane proteins to the endoplasmic reticulum

Jung-Yu Tung,[a]‡ Yi-Chuan Li,[a,b]‡ Tai-Wen Lin[a,c] and Chwan-Deng Hsiao[a]*

[a]Institute of Molecular Biology, Academia Sinica, Taipei 115, Taiwan, [b]Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu 300, Taiwan, and [c]Molecular Cell Biology, Taiwan International Graduate Program, Graduate Institute of Life Sciences, National Defense Medical Center and Academia Sinica, Taipei 115, Taiwan

‡ These authors contributed equally to this work.

Correspondence e-mail: hsiao@gate.sinica.edu.tw

Received 16 May 2013
Accepted 13 July 2013

PDB Reference:
Sgt2N–Get5Ubl, 3zdm

The insertion of tail-anchored membrane (TA) proteins into the appropriate membrane is a post-translational event that requires stabilization of the transmembrane domain and targeting to the proper destination. Sgt2, a small glutamine-rich tetratricopeptide-repeat protein, is a heat-shock protein cognate (HSC) co-chaperone that preferentially binds endoplasmic reticulum-destined TA proteins and directs them to the GET pathway *via* Get4 and Get5. The N-terminal domain of Sgt2 seems to exert dual functions. It mediates Get5 interaction and allows substrate delivery to Get3. Following the N-terminus of Get5 is a ubiquitin-like (Ubl) domain that interacts with the N-terminus of Sgt2. Here, the crystal structure of the Sgt2 dimerization domain complexed with the Get5 Ubl domain (Sgt2N–Get5Ubl) is reported. This complex reveals an intimate interaction between one Sgt2 dimer and one Get5 monomer. This research further demonstrates that hydrophobic residues from both Sgt2 and Get5 play an important role in cell survival under heat stress. This study provides detailed molecular insights into the specific binding of this GET-pathway complex.

1. Introduction

Targeted delivery of membrane proteins is a critical process. The targeting pathway for tail-anchored membrane (TA) proteins has only recently begun to be understood (Borgese *et al.*, 2003, 2007; Stefer *et al.*, 2011). These proteins comprise a large and diverse class of integral membrane proteins that are found in all organisms (Leznicki *et al.*, 2011; Rabu *et al.*, 2009; Borgese *et al.*, 2003, 2007; Kutay *et al.*, 1993). Characteristically, TA proteins have a single transmembrane helix (TM) at their extreme C-terminus (Kutay *et al.*, 1993). Owing to this topological constraint, these proteins cannot insert into membranes following the signal recognition particle (SRP)-dependent cotranslational pathway that is typically used by most integral membrane proteins. Instead, the newly characterized GET (guided entry of tail-anchored proteins) pathway is the major vehicle for localizing TA proteins to the endoplasmic reticulum membrane in yeast (Hegde & Keenan, 2011). There are at least six proteins (Get1, Get2, Get3, Get4, Get5 and Sgt2) in this pathway (Denic, 2012; Chartron, Clemons *et al.*, 2012; Mateja *et al.*, 2009; Schuldiner *et al.*, 2008; Hegde & Keenan, 2011). Sgt2, Get4 and Get5 form a sorting complex that accepts the newly synthesized TA proteins from the ribosome (Bozkurt *et al.*, 2010; Chang *et al.*, 2010, 2012). The TA proteins are then loaded onto a targeting chaperone, Get3. Get3 is an ATPase (Bozkurt *et al.*, 2009; Schuldiner *et al.*, 2008; Jonikas *et al.*, 2009; Stefanovic & Hegde, 2007). Upon ATP hydrolysis and conformational changes, Get3 transfers the TA protein to the Get1–Get2 receptor complex

(Mariappan *et al.*, 2011; Stefer *et al.*, 2011; Hegde & Keenan, 2011), which facilitates the entry of the TA protein into the membrane (Kubota *et al.*, 2012).

Sgt2 participates in the sorting of TA proteins to the ER with the aid of the Get4–Get5 complex (Liou *et al.*, 2007). Sgt2 is a small glutamine-rich tetratricopeptide-repeat (TPR)-containing co-chaperone protein that is highly conserved across eukaryotes (Liou *et al.*, 2007; Wang *et al.*, 2010; Schantl *et al.*, 2003). Sgt2 consists of three domains: (i) an N-terminal dimerization domain (Sgt2N) that binds to the ubiquitin-like domain of Get5 (Get5Ubl), (ii) a central TPR domain composed of three TPR motifs that mediates protein–protein interaction and possibly acts as an HSC co-chaperone (Wang *et al.*, 2010; Chartron *et al.*, 2011; D'Andrea & Regan, 2003) and (iii) a C-terminal glutamine- and methionine-rich domain that binds hydrophobic substrates such as the transmembrane domains of ER-destined TA proteins (Liou & Wang, 2005). Sgt2 serves as a scaffold for binding Get4/5 and other proteins required to mediate the sorting of TA proteins to the ER membrane (Chang *et al.*, 2010; Wang *et al.*, 2010; Chartron *et al.*, 2011; Kohl *et al.*, 2011). The *in vivo* activity of Sgt2 relies on its N- and C-terminal domains (Kohl *et al.*, 2011), whereas the central TPR domain may be involved in chaperone interactions, but its exact function is not clear (Liou & Wang, 2005).

Get5 was originally identified with a decreased mating phenotype (Kohl *et al.*, 2011). Get5 is required for cell survival under heat stress and for efficient mating (Hu *et al.*, 2006; Cohnen *et al.*, 2010). Structurally, Get5 is a multi-domain protein (Simpson *et al.*, 2010; Chang *et al.*, 2012). The N-terminal domain of Get5 forms a tight interaction with the C-terminus of Get4 (Chang *et al.*, 2010; Chartron *et al.*, 2010). Penultimate to the N-terminal domain is a ubiquitin-like domain (Get5Ubl) that binds the N-terminal domain of Sgt2 (Chang *et al.*, 2010; Chartron *et al.*, 2010, 2011). The C-terminal 35 residues of Get5 form a homodimerization domain necessary for biological function (Chartron *et al.*, 2010). The crystal and solution structures of the Get5 dimerization domain have been elucidated by Chartron, VanderVelde, Rao *et al.* (2012). Unlike other yeast proteins that contain a ubiquitin-like domain (Hicke *et al.*, 2005), Get5 does not interact with polyubiquinated proteins nor does it bind to the 26S proteasome (Hu *et al.*, 2006). Recent reports have indicated that nuclear import of Get5 is necessary for the heat stress-induced response and for recruitment to cytoplasmic stress granules (Arhzaouy & Ramezani-Rad, 2012; Cohnen *et al.*, 2010).

Biochemical and genetic studies have linked Get5 to both Sgt2 and Get4. Get5 and Get4 act as an adaptor complex linking Sgt2 to Get3, but the interactions among these proteins in the complex are unclear. In an effort to understand the structural basis for the interaction of Sgt2 and Get5, several related structures have been determined, including the Sgt2N NMR structure (PDB entries 2lxb and 4asv), the Get5Ubl crystal structure (PDB entries 4goc and 4a20) and the Sgt2N–Get5Ubl complex NMR structure (PDB entries 2lxc and 4asw) (Chartron, VanderVelde & Clemons, 2012; Simon *et al.*, 2013). However, the high-resolution structure of the Sgt2–Get5 complex remains to be elucidated. Here, we determined the crystal structure of the N-terminal domain of Sgt2 (Sgt2N) complexed with the ubiquitin-like domain of Get5 (Get5Ubl) at 1.8 Å resolution. Our data reveal an interplay that is strongly influenced by both electrostatic forces and a central hydrophobic interaction. Through a combination of crystal structure analysis and *in vivo* complementation assays, we elucidated a crucial interaction of the Sgt2–Get5 complex in the GET pathway.

2. Materials and methods

2.1. Construction of plasmids

Plasmids were constructed by standard molecular-cloning and recombinant DNA technologies. Briefly, the *Saccharomyces cerevisiae* S228C gene encoding Sgt2 was PCR-amplified using the forward primer 5′-TACTGTCATATGTC-AGCATCAAAAGAAGAAATTG-3′ and the reverse primer 5′-TACTGTCTCGAGCTATTGCTTGTTCTCATTGTCTG-GTG-3′. These products were inserted between the *Nde*I and *Xho*I restriction sites of a pET15b vector (Novagen) for Sgt2 protein expression. Sgt2N (residues 1–72) was derived from full-length Sgt2 by introducing a stop codon after nucleotide 216 by site-directed mutagenesis. Similarly, the fragment encoding Get5 was amplified using the forward primer 5′-AGTACTCATATGAACGCCGCCGTCCACTT-3′ and the reverse primer 5′-AGTATGCTCGAGTTATTTGGCCAG-AGACCAGCC-3′. The ubiquitin-like domain of Get5 (Get5Ubl; residues 71–151) was derived from residues 71–212 of Get5 by introducing a stop codon after nucleotide 453 by site-directed mutagenesis.

2.2. Expression and purification of recombinant proteins

The gene encoding Sgt2N or Get5Ubl was expressed in *Escherichia coli* BL21 (DE3) cells cultured in Luria–Bertani (LB) broth at 310 K. Expression of recombinant Sgt2N (residues 1–72) was induced by adding isopropyl β-D-1-thio-galactopyranoside (to a final concentration of 1 m*M*) to bacterial cultures harbouring the plasmid Sgt2N/pET15b. Cells were collected by centrifugation and were suspended in homogenization buffer (20 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl, 5 m*M* imidazole). The cells were homogenized using a microfluidizer (Microfluidics) and the lysate was centrifuged at 40 000*g* for 1 h at 277 K. The supernatants were then loaded onto a Ni^{2+}–NTA affinity column (HisTrap HP column; 5 ml; GE Healthcare) and eluted with a linear imidazole gradient (5–300 m*M* imidazole in 20 m*M* Tris–HCl, 500 m*M* NaCl pH 8.0). Fractions containing Sgt2N were pooled and were further purified by size-exclusion chromatography on a 16/60 Superdex 75 column (GE Healthcare) equilibrated with a buffer consisting of 20 m*M* Tris–HCl pH 7.5, 500 m*M* NaCl. The recombinant Get5Ubl (residues 71–151) was isolated similarly. The Sgt2N–Get5Ubl protein complex was obtained by mixing purified Sgt2N and Get5Ubl in a 1:1 molar ratio and was dialyzed against 20 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl. The Sgt2N–Get5 complex was further purified by size-exclusion

Table 1
X-ray diffraction data and structure-refinement statistics.

Values in parentheses are for the highest resolution shell.

	Sgt2N–Get5Ubl–$C_2H_7HgO_4P$		Sgt2N–Get5Ubl
	High remote	Inflection	Native
Source	BL13B1, NSRRC	BL13B1, NSRRC	BL44XU, SPring-8
Detector	Q315r	Q315r	MX-225HE
Oscillation interval (°)	1.0	1.0	1.0
Wavelength (Å)	1.0050	1.0085	1.00000
Space group	*C*2	*C*2	*C*2
Unit-cell parameters			
a (Å)	138.62	138.65	137.12
b (Å)	31.24	31.25	34.13
c (Å)	92.26	92.28	92.08
β (°)	105.85	105.85	105.84
Resolution range (Å)	30.0–2.4	30.0–2.4	30.0–1.8
Multiplicity	6.8 (6.2)	6.8 (6.2)	6.9 (6.0)
Completeness (%)	96.5 (86.4)	96.2 (83.0)	96.9 (97.3)
$\langle I/\sigma(I)\rangle$	31.7 (3.9)	32.2 (4.4)	34.1 (3.2)
$R_{\rm merge}$† (%)	7.9 (27.4)	7.4 (26.3)	7.0 (51.8)
Wilson *B* value (Å^2)			25.31
Mosaicity (°)	1.1	1.1	1.0
Figure of merit	0.45	0.45	
Refinement			
Resolution range (Å)			30.0–1.8
No. of reflections			35615
R factor‡/$R_{\rm free}$§ (%)			18.67/22.61
R.m.s.d.			
Bonds (Å)			0.014
Angles (°)			1.343
No. of atoms			
Overall			3492
Protein			3036
Water			456
B factors (Å^2)			
Overall			31.8
Protein			30.4
Water			41.5

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity from multiple observations of symmetry-related reflections. ‡ $R = \sum_{hkl}||F_{\rm obs}| - |F_{\rm calc}||/\sum_{hkl}|F_{\rm obs}|$, where $|F_{\rm obs}|$ and $|F_{\rm calc}|$ are the observed and calculated structure-factor amplitudes. § $R_{\rm free}$ was calculated on the basis of 10% of the total number of reflections that were randomly omitted from the refinement.

chromatography on a 16/60 Superdex 75 column (GE Healthcare).

2.3. Crystallization, data collection and structural determination of the Sgt2N–Get5Ubl complex

For crystallization trials, the Sgt2N–Get5Ubl complex was first concentrated to 13.5 mg ml^{-1} in a buffer consisting of 20 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl using a Centricon concentrator (Millipore). The hanging-drop vapour-diffusion method was then performed at 298 K by mixing 1 µl Sgt2N–Get5Ubl with an equal volume of crystal screening solutions. Rod-shaped crystals appeared after 3–7 d in 100 m*M* MIB buffer (sodium malonate, imidazole and boric acid in a 2:3:3 molar ratio) and 25%(*v*/*v*) polyethylene glycol 1500 pH 7.0. Diffraction data were collected from a single crystal of Sgt2N–Get5Ubl complex on beamline BL44XU of SPring-8 in Japan. Data integration and scaling were performed using the *HKL*-2000 package (Otwinowski & Minor, 1997). The crystals belonged to space group *C*2, with unit-cell parameters $a = 137.12$, $b = 34.13$, $c = 92.08$ Å, $\beta = 105.84°$, and diffracted to 1.8 Å resolution.

Because no structure similar to Sgt2N could be used for molecular-replacement phasing, heavy-atom derivatives were used to solve the phase problem using the multi-wavelength anomalous dispersion method. After an extensive search, one useful mercury derivative was obtained by cocrystallizing 1.0 m*M* ethylmercury phosphate ($C_2H_7HgO_4P$) in a modified reservoir consisting of 100 m*M* MIB buffer, 25%(*v*/*v*) polyethylene glycol 1500. Cryogenic multiwavelength anomalous dispersion data were collected on a Quantum 315 charge-coupled device detector (Area Detector Systems Corp.) using a synchrotron-radiation X-ray source: beamline BL13B1 at the National Synchrotron Radiation Research Center (NSRRC) in Taiwan. Two energies were chosen near the absorption high remote and the Hg edge for Sgt2N–Get5Ubl: 1.0050 and 1.0085 Å. A remote energy was selected as a reference wavelength at 0.8550 Å. *SOLVE* (Terwilliger & Berendzen, 1999) was used to locate the mercury site and to generate the initial multi-wavelength anomalous dispersion phases at 2.4 Å resolution. Extension of the initial phases to 1.8 Å and preliminary automated model building were carried out by *RESOLVE* (Terwilliger, 2003, 2004). *Coot* (Emsley *et al.*, 2010) was used to examine the electron-density maps and for manual model building. Further refinement was performed using *PHENIX* (Adams *et al.*, 2010). After refinement, the *R* factor of the final model was 18.67% for all reflections above 2σ between 25.59 and 1.8 Å resolution and the $R_{\rm free}$ value was determined to be 22.61% using a randomly distributed 4.8% of the reflections. The Ramachandran plot showed that 93.8% of the residues lie within the most favoured regions and 6.2% lie in additional allowed regions; no residues were found in the generously allowed or disallowed regions (Laskowski *et al.*, 1993). Statistics of data collection and structure refinement are listed in Table 1. The final model contained four chains of Sgt2N molecules and two chains of Get5Ubl molecules. The four Sgt2N molecules are comprised of protein residues 2–70 (chains *A* and *D*) and residues 2–51 (chains *B* and *E*). The two Get5Ubl molecules are comprised of protein residues 72–150 (chains *C* and *F*). The figures were produced using *PyMOL* (http://www.pymol.org). The atomic coordinates for the Sgt2N–Get5Ubl complex have been deposited in the Protein Data Bank under accession code 3zdm.

2.4. Yeast strains and *in vivo* complementation assay

Residues that participate in the Get5–Sgt2 interaction were chosen for point-mutation studies. Seven residues in Get5 (Ile81, Lys85, Lys118, Leu120, Lys122, Lys124 and His127) and six residues in Sgt2 (Asp28, Asp31, Val35, Asp38, Cys39 and Glu42) were individually replaced with alanines using the QuikChange site-directed mutagenesis kit (Stratagene). Four residues (Lys110 and Lys136 of Get5, and Glu24 and Glu47 of Sgt2) that do not participate in Get5–Sgt2 interactions were also mutated to alanine as controls. *S. cerevisiae* strain BY4741 with an S288C background was used as host cells (Liou *et al.*, 2007). Empty YEp33 plasmids were transformed into the

BY4741, BY4741Δ*sgt2* or BY4741Δ*get5* strains. YEp33 plasmids carrying Sgt2 or Get5 mutants were transformed into the BY4741Δ*sgt2* or BY4741Δ*get5* strains, respectively. Transformants were selected at 303 K on SC-ura plates. Each colony was suspended in YPD medium and grown to mid-log phase. The A_{600} of each cell culture was adjusted to 0.5 by adding YPD. Serial dilutions (fivefold) of these cultures were spotted onto YPD plates and incubated at 303 or 312 K to evaluate growth behaviour.

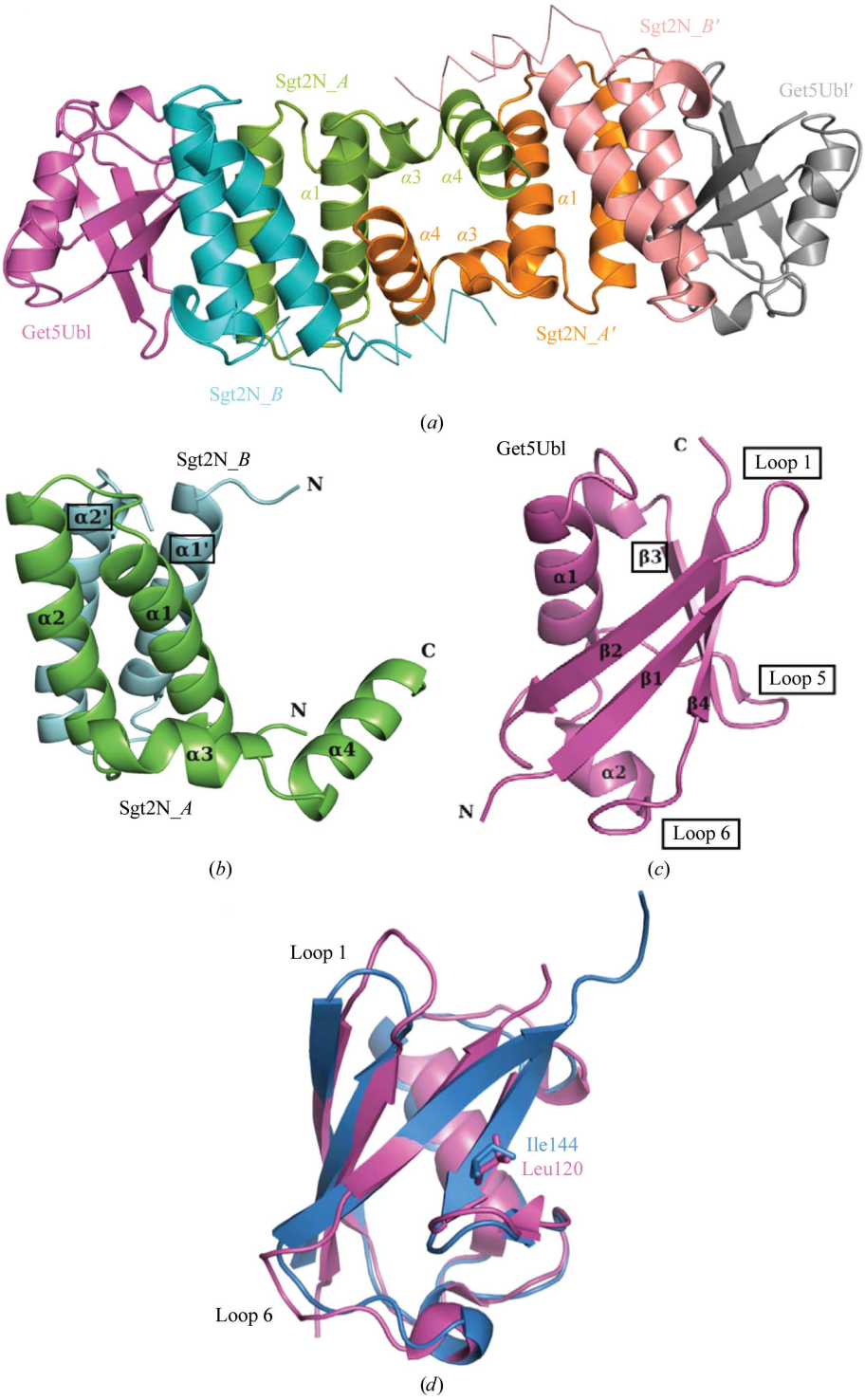

Figure 1
Crystal structure of the Sgt2N–Get5Ubl complex. (*a*) A ribbon diagram of the *S. cerevisiae* Sgt2N–Get5Ubl complex showing Get5Ubl (magenta, Get5Ubl; grey, Get5Ubl′) and the Sgt2N homodimer (green, Sgt2N_*A*; cyan, Sgt2N_*B;* orange; Sgt2N_*A*′; salmon, Sgt2N_*B*′). The individual overall structures of (*b*) the Sgt2N homodimer and (*c*) Get5Ubl are shown. Secondary structures participating in complex interactions are boxed. (*d*) Comparison of *S. cerevisiae* Get5Ubl (blue) and *Homo sapiens* erythrocyte ubiquitin (magenta; PDB entry 1ubq). The highly conserved residues Leu120 from *S. cerevisiae* and Ile44 from *H. sapiens* ubiquitin are indicated by sticks. The highly divergent loops 1 and 6 are also labelled.

3. Results and discussion

3.1. Crystal structure of the Sgt2N–Get5Ubl complex

To elucidate the molecular basis of the interaction of Sgt2 and Get5, we determined the crystal structure of the Sgt2N–Get5Ubl complex from *S. cerevisiae*. The overall crystal structure of the Sgt2N–Get5Ubl complex is shown in Fig. 1(*a*). The final atomic model comprises two almost identical Sgt2N–Get5Ubl complexes per asymmetric unit, with a root-mean-square deviation of 0.06 Å for the C^{α} coordinates. Each complex contains one Get5Ubl molecule and an Sgt2N dimer. We labelled the two Sgt2N monomers in the dimer Sgt2N_*A* and Sgt2N_*B*.

Sgt2N has four helices and does not contain an obvious internal consensus or any common helical repeat features. The first two helices (α1 and α2) are of identical length and the two Sgt2N monomers form a four-helix bundle (Fig. 1*b*). The interface of the Sgt2N dimer consists of large hydrophobic residues from α1 and α2 of the two Sgt2N monomers. The two short α3 and α4 helices (residues 52–72) are packed against the four-helix bundle away from the dimer interface. We noted that the α3 and α4 helices were missing in Sgt2N_*B* and Sgt2N_*B*′ owing to a lack of molecular contact. In the asymmetric unit, the α4 helix of Sgt2N_*A* stacks against the α1 helices of Sgt2N_*A*′ and Sgt2N_*B*′. The α4 helix of Sgt2N_*A*′ also interacts with the α1 helices of Sgt2N_*A* and Sgt2N_*B* in the same manner. However, there is no interaction between Sgt2N_*B* and Sgt2N_*B*′ (Fig. 1*a*). Since our crystal structure indicated that Sgt2N might form a tetramer, we performed analysis ultracentrifugation (AUC) experiments to

further confirm this observation. The AUC results showed only dimer formation in solution (data not shown). Thus, the tetramer may result from crystallographic packing.

The structure of Get5Ubl (residues 71–151 of Get5) is a typical β-grasp fold with a mixture of α-helices and β-strands (Fig. 1*c*). The main architecture consists of a one-sided β-sheet of four antiparallel β-strands (β1 to β4) grasping the α1 helix (residues 98–108). Compared with the structure of ubiquitin (PDB entry 1ubq; Vijay-Kumar *et al.*, 1987), loop 1 and loop 6 of Get5Ubl appear to be distinctly longer than those of ubiquitin (Fig. 1*d*). The root-mean-square deviation between Get5Ubl and ubiquitin is 1.3 Å for the C^{α} atoms. These two

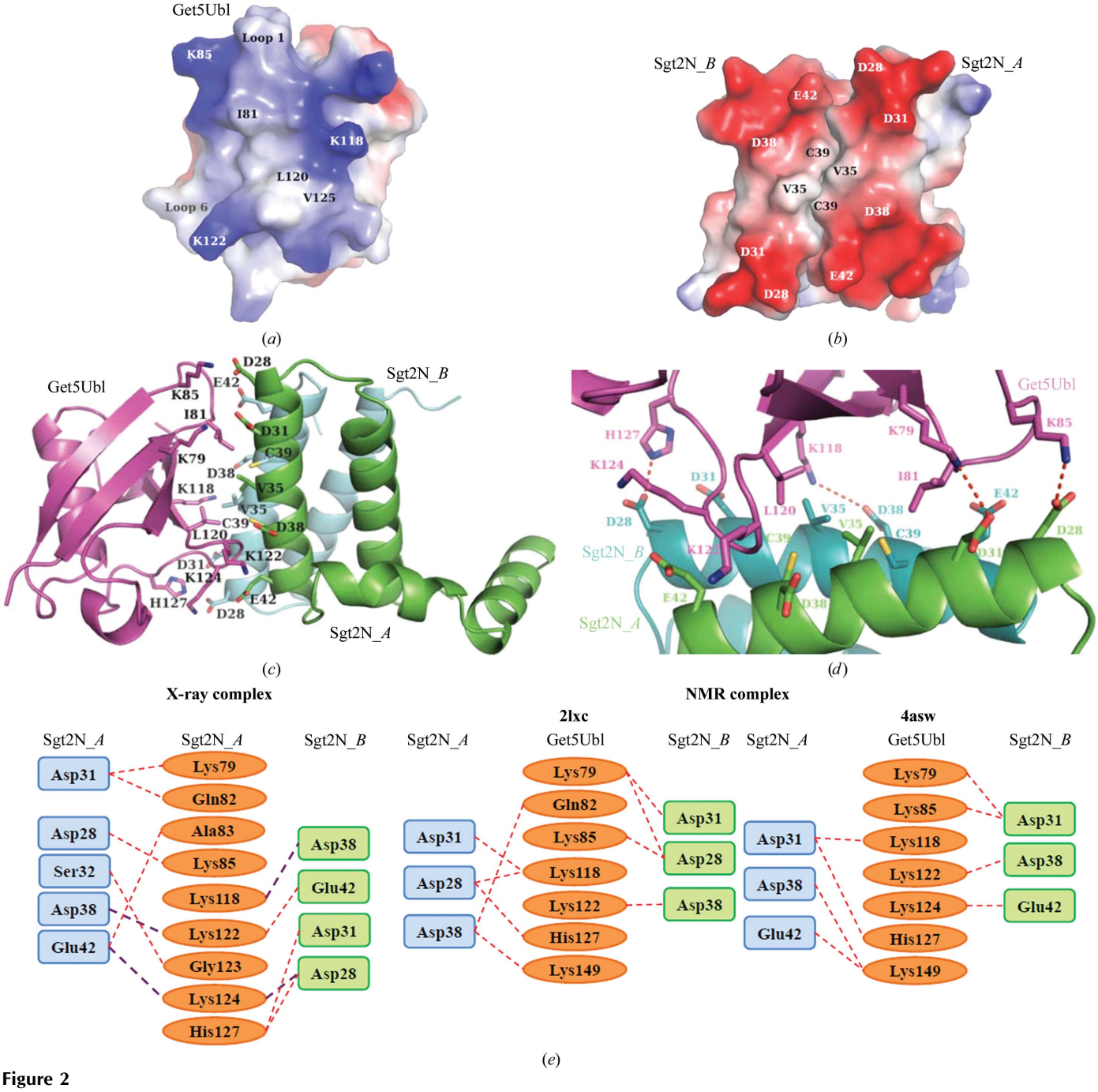

Figure 2
The interaction surfaces of the Sgt2N–Get5Ubl complex. (*a*) The Get5Ubl interaction surface is composed of a hydrophobic area at the centre which elongates to loop 6. The other surrounding residues are mostly positively charged residues (coloured blue). The residues and loops participating in interactions are labelled. (*b*) In contrast, the interaction surface of the Sgt2N homodimer has a high electrostatic potential and exhibits a hydrophobic area at the core surrounded by negatively charged residues (coloured red). (*c*) Interacting residues in the Sgt2N–Get5Ubl complex; (*d*) a magnification of the interaction surface. The colours are conserved from the previous figures and hydrophilic interactions are indicated by red dots. (*e*) Detailed hydrophilic interactions between Sgt2N and Get5Ubl from the X-ray study and two NMR studies. The red dashed lines represent direct hydrogen bonds and the purple dashed lines represent interactions mediated by water molecules.

loops in both Get5Ubl (Ile81, Ala82 and Pro84 form loop 1; Val137, Thr138, Pro139 and Ala140 form loop 6) and ubiquitin are composed mainly of hydrophobic residues and create an extended hydrophobic path from the central hydrophobic patch composed of Leu120, Gly123 and Val125 (Fig. 2*a*).

3.2. Composite binding site and interactions between Sgt2N and Get5Ubl

In the Sgt2N–Get5Ubl complex, the two α2 helices of the Sgt2N dimer are arranged in an antiparallel conformation and interact with Get5Ubl. The binding surface of Get5Ubl is created by loop 1, β3 and loop 5. Leu120 of Get5Ubl is located on the β3 strand and corresponds to the conserved Ile44 in ubiquitin and ubiquitin-like proteins. In ubiquitin, Ile44 is part of a hydrophobic core (named the Ile44 patch) and interacts with its binding partners. In the Sgt2N–Get5Ubl complex Get5Ubl binds to the Sgt2N dimer *via* this hydrophobic patch. The interaction buries a surface area of ~620 Å^2. The binding-surface area between Sgt2N_*A* and Get5Ubl is 300.9 Å^2 and the shape-complementarity (shape correlation statistic) score (Lawrence & Colman, 1993; Zhang *et al.*, 2009) is 0.66. The binding-surface area between Sgt2N_*B* and Get5Ubl is 316.5 Å^2 and yields a high shape-complementarity score of 0.79. These analyses suggest that the two monomers of the Sgt2N dimer bind asymmetrically to Get5Ubl.

Sgt2 contains 18% glutamic and aspartic acids in its sequence and this is reflected by the electrostatic surface (Fig. 2*b*). The top and bottom sides of the Sgt2 dimer distribute strong negative charges created by Glu27, Asp28, Asp31, Glu42 and Glu47. Interestingly, the central part of the Sgt2 α2 helix has a distinctive hydrophobic core created by the hydrophobic residues Val35 and Cys39 that are exposed on the protein surface. In contrast, Ile, Leu and Lys are predominantly found in the Get5 sequence. The distinctive hydrophobic Ile44 patch of the ubiquitin superfamily is composed of Ile81, Leu120 and Val125 in Get5Ubl. This hydrophobic patch is surrounded by basic residues including Lys85, Lys118, Lys122, Lys125 and Lys136 (Fig. 2*a*).

The hydrophobic area of Sgt2N involving Val35 and Cys39 interacts with Ile81, Ala83, Leu120 and Gly123 of Get5Ubl (Fig. 2*c*) and is completely buried at the binding interface. The distance between the two Cys39 residues of the Sgt2N dimer is 5.9 Å and they do not form a disulfide bond. It should be noted that free cysteines display stronger hydrophobicity than disulfide-bonded cysteines (Nagano *et al.*, 1999). Hence, Cys39 of Sgt2N contributes significantly to the binding of Get5, and replacing Cys39 with alanine reduced the binding affinity of Sgt2N for Get5 ~127-fold (Chartron, VanderVelde & Clemons, 2012).

Leu120 of Get5Ubl interacts with Cys39 of Sgt2N_*A* and Val35 of Sgt2N_*B* (Fig. 2*d*). Similarly, Ile81 of Get5Ubl interacts with Cys39 of Sgt2N_*B* and Val35 of Sgt2N_*A*. The Asp28 residues of Sgt2N_*A* and Sgt2N_*B* form hydrogen bonds to Lys85 and His127 of Get5, respectively. In addition, Ser32 and Glu42 of Sgt2N_*B* form hydrogen bonds to the main-chain O atom of Gly123 and the main-chain N atom of Ala83 of Get5Ubl, respectively (Fig. 2*e*). The acidic residues Asp31, Asp38 and Glu42 of both Sgt2N monomers form a cluster of hydrogen bonds to the basic residues Lys79, Lys85, Lys118 and Lys122 of Get5Ubl. The interaction with Lys122 is mediated *via* water molecules (Fig. 2*e*). However, minor differences in the interactions of the two Sgt2N molecules with Get5Ubl in addition to the strengths of these interactions can be observed.

In general, Asp28, Asp31, Val35, Csy39 and Glu42 of each Sgt2N monomer and Ser32 of Sgt2N_*B* are in direct contact with Get5Ubl in the Sgt2N–Get5Ubl complex. Lys79–Lys85 (β1 and loop 1) and Lys118–His127 (β3 and loop 5) of Get5Ubl are responsible for the primary contacts with Sgt2N (Fig. 2*c*). These interactions consti-

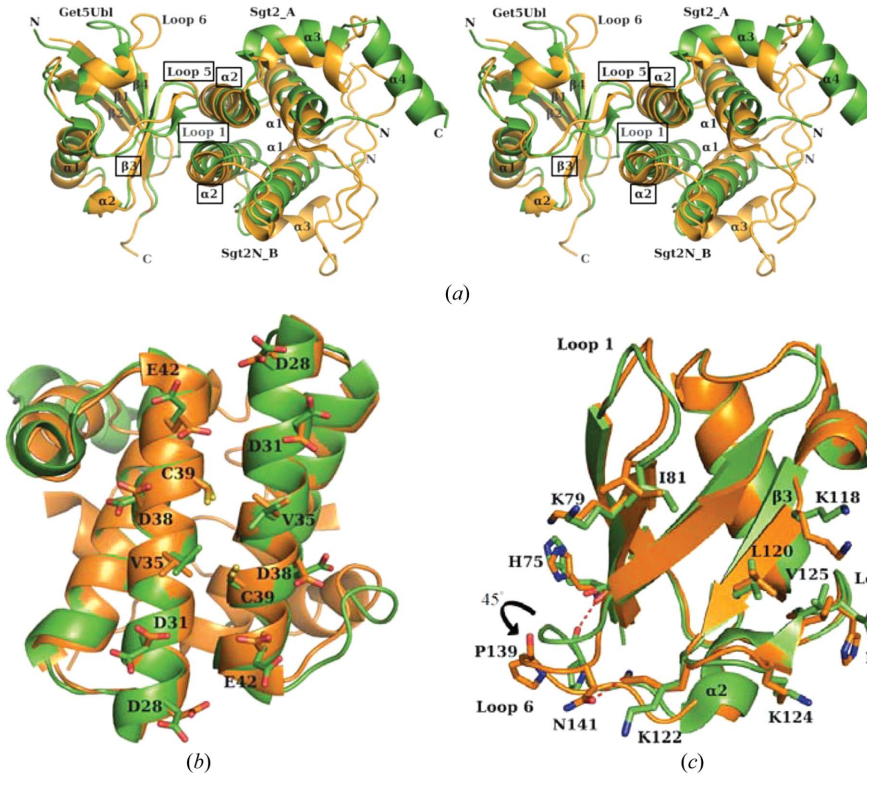

Figure 3
Comparison of the crystal structure and the NMR structure of the Sgt2N–Get5Ubl complex. (*a*) Stereoview of the complexes revealing orientation differences. The superimposed complexes are based on Get5Ubl. The crystal structure (green) has a different orientation compared with the NMR structure (orange; PDB entry 2lxc). Secondary structures participating in complex interactions are boxed. (*b*) The superimposed crystal and NMR renderings of the Sgt2N homodimer are highly similar. Differences in the Get5-interacting residues are indicated as sticks. (*c*) Superimposed crystal and NMR renderings of the Get5Ubl structures. Loops 1 and 6 are internally flexible. The Sgt2-interacting residues are indicated by sticks.

tute a substantial hydrophobic core supplemented by electrostatic interactions.

3.3. Comparison with other Sgt2N and Get5Ubl structures

Recently, Chartron, VanderVelde & Clemons (2012) and Simon *et al.* (2013) characterized the complex of Sgt2N (residues 1–72) and Get5Ubl (residues 73–153) from *S. cerevisiae* using solution NMR and experimental restraints. In these two studies, the authors used reciprocal chemical shift perturbation experiments to determine the interaction interface between Sgt2N and Get5Ubl. Although the X-ray crystal structure and NMR structures (PDB entries 2lxc and 4asw; Chartron, VanderVelde & Clemons, 2012; Simon *et al.*, 2013) of the Sgt2N–Get5Ubl complex have similar overall arrangements (Fig. 3*a*), there are several structural disparities. The overall root-mean-square deviation between the crystal and the NMR complex structures is approximately 5.8 Å for the C^{α} atoms and the greatest structural differences are at loops 1, 5, 6 and β3 of Get5Ubl. These regions are involved in the interactions with the Sgt2N dimer.

Close inspection of the crystal structure and the two NMR structures reveals several differences in hydrogen-bonding patterns on the binding surface. A comparison of the detailed interactions between the Sgt2N dimer and Get5Ubl is shown in Fig. 2(*e*). Most acidic residues of Sgt2N form hydrogen bonds to the basic residues of Get5Ubl on the binding surface in both the crystal structure and the NMR structure. The interacting residues of Sgt2N are mainly Asp28, Asp31, Asp38 and Glu42 in the crystal structure (Fig. 2*e*). In contrast to the crystal structure, the interacting residues differ in both NMR structures (PDB entries 2lxc and 4asw). For instance, Glu42 in the 2lxc structure and Asp28 in the 4asw structure were not involved in protein–protein interactions. In the crystal structure, Glu42 from the two Sgt2N molecules contributes three further interactions with the main-chain N atoms of Ala83, Lys122 and Lys124 of Get5Ubl. Asp28 of Sgt2N_*A* contributes a hydrogen bond to Lys85 of Get5Ubl, and Asp28 of Sgt2N_*B* interacts with Lys124 and His127 of Get5Ubl. In addition, several interactions between the hydrophilic residues of two molecules are mediated by water molecules in the complexed crystal structure, as mentioned in the previous section and shown in Fig. 2(*e*) (purple dashed lines). For example, Asp38 of Sgt2N_*A* interacts with Lys122 of Get5Ubl through a water molecule. In conclusion, the crystal structure displays more hydrogen bonds either by direct interaction or through water molecules than the NMR structures.

We then compared the structures of the Sgt2N molecule in the complex and free forms. Superimposition of the free form (PDB entry 2lxb; Chartron, VanderVelde & Clemons, 2012) and the complexed form (this work) from the crystal structure (Fig. 3*b*) revealed an overall root-mean-square deviation for C^{α} coordination of approximately 4.22 Å. Although the four-helix bundle composed of the α1 and α2 helices (Get5 binding surface) does not display a striking conformational change, the side chains of the binding residues display obvious differences between the free-form and complex-form structures. Asp28, Asp31, Asp38 and Glu42 of the α2 helix of Sgt2N are flipped into a different orientation in complexed Sgt2N (Fig. 3*b*). However, the α3 and α4 helices of Sgt2N are different as they exhibited no secondary structure in the solution structure of the free-form Sgt2N. The α3 and α4 helices of Sgt2N are likely to be considerably disordered, and the lack of additional physical contacts results in the absence of these residues from one of the Sgt2N dimers (Sgt2N_*B*). The side chains of many of the binding residues, especially the acidic residues Asp28, Asp31, Asp38 and Glu42 in the α2 helix of Sgt2N, are flipped into a different orientation in complexed Sgt2N (Fig. 3*b*).

To further investigate the conformational change of Get5Ubl when binding to the Sgt2N dimer, we superimposed the crystal structures of free (PDB entry 4goc) and complexed Get5Ubl (Fig. 3*c*). The overall C^{α} root-mean-square deviation is approximately 1.5 Å. The largest structural difference was found in loop 1, β3, loop 5, α2 and loop 6, where loop 1, β3 and loop 5 of Get5Ubl are associated with the Sgt2N-binding region. In the free-form Get5Ubl structure, Lys122 forms a hydrogen bond to Asn141 and drags loop 6 outward. Interestingly, Lys122 is located near the surface, allowing binding to Asp38 of Sgt2N. This interaction results in an approximately 45° backwards curvature shift of the β1 strand and the formation of a hydrogen bond between the main-chain O atom of Pro139 and the main-chain N atom of His75 of Get5 (Fig. 3*c*). In addition, there are several basic residues (Lys79, Lys118, Lys122, Lys124 and His127) that exhibit different orientations (Fig. 3*c*).

3.4. *In vivo* complementation assay of the Sgt2N–Get5Ubl complex by site-directed mutagenesis

It has been shown that deletion of components in the GET pathway leads to a wide variety of phenotypes, such as increased sensitivity towards higher temperatures (312 K; Schuldiner *et al.*, 2008). To further confirm and understand whether the residues involved in protein–protein interactions are crucial for cell survival, we generated several point mutants from full-length Sgt2 and Get5 for yeast-growth experiments. As shown in Fig. 4, the Sgt2 and Get5 knockouts were sensitive to elevated temperature. Additionally, the Get5 knockout exhibited more severe defects in TA protein sorting than the Sgt2 knockout (Fig. 4). This phenomenon can be rescued by expressing wild-type Sgt2 or Get5 under the control of the native promoter on a plasmid (Fig. 4). Furthermore, we derived several variants with single amino-acid mutations within the region responsible for the interaction between Sgt2 and Get5. Most Sgt2 mutants were able to restore growth and displayed only minor defects compared with the wild type at 312 K (Fig. 4*a*). However, the V35A Sgt2 mutant displayed significant defects in growth at elevated temperature.

Similar results were obtained for the Get5 mutant variants. The most significant defect was observed for the Leu120 mutant (Fig. 4*b*), a highly conserved residue in the ubiquitin superfamily. The level of impairment observed for the L120A mutant was close to that observed for the ΔGet5 mutant. The

result indicates that Leu120 is crucial for the function of Get5. Previous binding-affinity studies using isothermal titration calorimetry (ITC) showed that the V35A and C39A Sgt2N mutants have a 133-fold lower binding affinity, while the L120A mutant of Get5Ubl exhibits a 1000-fold lower binding affinity, compared with the wild type (Chartron, VanderVelde & Clemons, 2012). However, the cell-survival analyses of the Val35 and Cys39 Sgt2 mutants were not consistent with the biochemical results. The V35A mutant of Sgt2N showed severe impairment, but the C39A mutant displayed only a minor defect (Fig. 4*a*). These results suggested that although both Val35 and Cys39 of Sgt2 are directly involved in the binding of the Sgt2–Get5 complex, Val35 is functionally more important. Interestingly, the L120A mutant of Get5 showed a similar defect to the Sgt2 knockout. This suggests that Leu120 of Get5 is the most important residue involved in the binding of Sgt2. This also implies that blocking the binding of Sgt2 to Get5 results in loss of Sgt2 from the GET pathway.

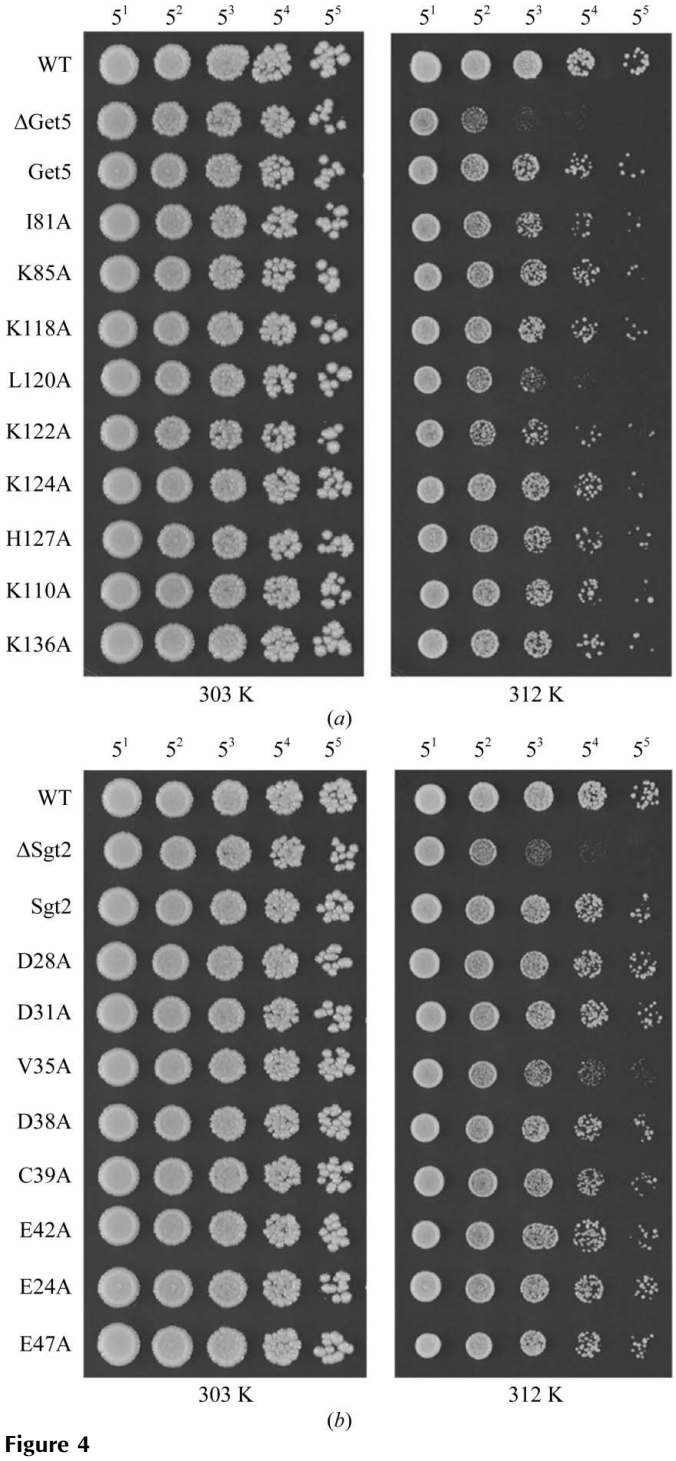

Figure 4
Complementation assays for Get5 and Sgt2. Serial dilutions of fresh yeast cultures spotted onto YPD plates and incubated at 303 or 312 K for 3 d. Empty and full-length vectors of (*a*) Sgt2 and (*b*) Get5 were used as negative and positive controls, respectively. The dilution factors for the yeast cultures are labelled above the panels.

In contrast to mutations of hydrophobic residues, replacements of hydrophilic residues caused only minor defects and the cells were viable at elevated temperature. These findings are consistent with recent ITC and SPR biophysical experiments (Chartron, VanderVelde & Clemons, 2012). Taken together, the data suggests that hydrophobic interactions are indispensable for Sgt2–Get5 complex formation and cell survival. However, the role of the hydrophilic residues is of significant interest as they are also important participants in formation of the Sgt2–Get5 complex.

4. Conclusions

Sgt2 plays an important role in the yeast GET pathway in sorting TA proteins with the aid of the Get4–Get5 complex. Interrupting the formation of this complex affects cell viability at elevated temperatures (Fig. 4). Previous biochemical and genetic studies have linked Sgt2 to Get4 and Get5 (Liou *et al.*, 2007). Subsequent reports determined that the N-terminal domain of Sgt2 binds the ubiquitin-like domain of Get5 (Chartron *et al.*, 2011; Chang *et al.*, 2010, 2012; Wang *et al.*, 2010), and both the N-terminal and the C-terminal domains of Sgt2 have been implicated in TA-protein delivery (Kohl *et al.*, 2011). Deletion of either domain of Sgt2 results in cell-viability defects at elevated temperature or under stress (Kohl *et al.*, 2011).

Get5 is required for cell survival under heat stress and for efficient mating (Hu *et al.*, 2006). The role of Get5 in the GET pathway is unclear. A recent report showed that Get5 can co-localize with Pab1, a poly(A)-binding protein, and that the translocation of Get5 correlated with the accumulation of cytoplasmic stress granules under elevated temperature (Arhzaouy & Ramezani-Rad, 2012). The main components of the stress granules are molecular chaperones that sequester, protect and possibly repair proteins that were unfolded or misfolded during heat or other stresses (Buchan *et al.*, 2011; Grousl *et al.*, 2009). More recently, it has been demonstrated that all of the cytosolic components of the GET system, including Get3, Get4, Get5 and Sgt2, together with Hsp104, Hsp42 and Ssa2, were translocated into foci containing TA proteins after short-term glucose starvation (Powis *et al.*, 2012). Furthermore, Get3 can prevent the aggregation of denatured proteins *in vitro*. Hence, the GET complex might have functional role(s) other than TA-protein delivery (Powis *et al.*, 2012). Interestingly, the translocation of Get3 into stress

granules requires the presence of Get5 but not of Sgt2 (Powis *et al.*, 2012). This is not consistent with the notion that Sgt2 is the most upstream factor for delivering newly synthesized TA proteins. The data suggest that Get5 might be able to recruit cytosolic TA proteins away from the ribosomes. Our complementation assays indicating that Get5 deletion has more severe defects than Sgt2 deletion are consistent with previous studies. It is possible that the Get5 deletion strain has also lost the ability to translocate Get3 and Sgt2 to Get-positive stress granules under stress.

Since the factors required for TA-protein delivery were systematically identified in 2009 (Jonikas *et al.*, 2009), the structures of all of the cytosolic members and the cytosolic portions of the membrane-bound members (Get1 and Get2) have been resolved at least in domain structures. Interestingly, all of the cytosolic factors are dimeric and highly conserved among fungi and mammals even though additional factors have been acquired in mammals during evolution. Get4 itself does not form a dimer, but two Get4 molecules can interact with one Get5 dimer *via* strong hydrophobic interactions and thereby could be considered as a dimeric Get4–Get5 complex. Furthermore, structural evidence has demonstrated that the stoichiometry of interaction between Get3, Get4 or Get5 with Sgt2 is 2:2 (Chang *et al.*, 2010, 2012; Bozkurt *et al.*, 2009; Nagano *et al.*, 1999; Chartron *et al.*, 2011). Here, we examined the stoichiometry of the Sgt2N–Get5Ubl complex. Our crystal structure revealed that Get5Ubl interacts with the interface of an Sgt2N dimer and results in a 2:1 stoichiometry (one Sgt2N dimer to one Get5Ubl molecule). Based on this model, the stoichiometry of a Get4–Get5–Sgt2 complex would be 2:2:4. To investigate whether the C-terminal dimerization domain of Get5 affects the binding ratio of Sgt2N, we performed equilibrium ultracentrifugation of Sgt2N and Get5 containing the ubiquitin and C-terminal domains (Get5Ubl-C; residues 71–212). Results (data not shown) indicated that two Sgt2N and two Get5Ubl-C molecules (2:2) can form a complex. Therefore, dimerization of Get5 might lead to rearrangement of the Get5 ubiquitin-like domain and thereby block one potent Sgt2-binding site. Similarly, a previous study has shown that the Get4–Get5–Sgt2 complex exhibited a molecular weight corresponding to a 2:2:2 stoichiometry by multi-angle light-scattering measurement and it was hypothesized that the association of Get4 and Get5 could lead to a stereo effect and exclude the binding of the second Sgt2 dimer (Chartron *et al.*, 2011). More recently, it was reported that the interaction of Get5 and Sgt2 exhibits rapid association and dissociation rates (Chartron, VanderVelde & Clemons, 2012). Thus, Sgt2 could rapidly dissociate from one of the two potent binding sites on one Get5 dimer, resulting in the observation of the binding of a single Sgt2 molecule. However, more structural evidence will be required in order to understand the exclusive effect in molecular detail. Furthermore, the biological function of dimerization remain to be addressed.

In addition to the responses of single amino-acid mutations under heat stress, our crystal structure provided a detailed picture of the interactions in the Sgt2N–Get5Ubl complex. The hydrophobic interactions between Sgt2 and Get5 are the most important forces, while the hydrophilic interactions account for the rapid binding and dissociation. Get5 appears to be essential for a rapid cellular response. The complete mechanism and function of Sgt2 or Get5 in the process remains unclear, but these structural and cellular response data contribute a clearer understanding of the GET pathway and of the targeting of TA proteins to the ER membrane.

We thank Dr M. F. Tam for critical reading of the manuscript and for useful comments. We are grateful for access to the synchrotron-radiation beamlines BL13B1 at the National Synchrotron Radiation Research Center in Taiwan and BL44XU at SPring-8 in Japan. This work was supported by research grants from Academia Sinica (to C-DH) and the National Science Council (NSC101-2311-B-001-022-MY3 to C-DH), Taiwan, Republic of China.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Arhzaouy, K. & Ramezani-Rad, M. (2012). *PLoS One*, **7**, e52956.
Borgese, N., Brambillasca, S. & Colombo, S. (2007). *Curr. Opin. Cell Biol.* **19**, 368–375.
Borgese, N., Colombo, S. & Pedrazzini, E. (2003). *J. Cell Biol.* **161**, 1013–1019.
Bozkurt, G., Stjepanovic, G., Vilardi, F., Amlacher, S., Wild, K., Bange, G., Favaloro, V., Rippe, K., Hurt, E., Dobberstein, B. & Sinning, I. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 21131–21136.
Bozkurt, G., Wild, K., Amlacher, S., Hurt, E., Dobberstein, B. & Sinning, I. (2010). *FEBS Lett.* **584**, 1509–1514.
Buchan, J. R., Yoon, J.-H. & Parker, R. (2011). *J. Cell Sci.* **124**, 228–239.
Chang, Y.-W., Chuang, Y.-C., Ho, Y.-C., Cheng, M.-Y., Sun, Y.-J., Hsiao, C.-D. & Wang, C. (2010). *J. Biol. Chem.* **285**, 9962–9970.
Chang, Y.-W., Lin, T.-W., Li, Y.-C., Huang, Y.-S., Sun, Y.-J. & Hsiao, C.-D. (2012). *J. Biol. Chem.* **287**, 4783–4789.
Chartron, J. W., Clemons, W. M. & Suloway, C. J. (2012). *Curr. Opin. Struct. Biol.* **22**, 217–224.
Chartron, J. W., Gonzalez, G. M. & Clemons, W. M. (2011). *J. Biol. Chem.* **286**, 34325–34334.
Chartron, J. W., Suloway, C. J., Zaslaver, M. & Clemons, W. M. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 12127–12132.
Chartron, J. W., VanderVelde, D. G. & Clemons, W. M. (2012). *Cell Rep.* **2**, 1620–1632.
Chartron, J. W., VanderVelde, D. G., Rao, M. & Clemons, W. M. (2012). *J. Biol. Chem.* **287**, 8310–8317.
Cohnen, A., Bielig, H., Hollenberg, C. P., Hu, Z. & Ramezani-Rad, M. (2010). *Cytoskeleton*, **67**, 635–649.
D'Andrea, L. D. & Regan, L. (2003). *Trends Biochem. Sci.* **28**, 655–662.
Denic, V. (2012). *Trends Biochem. Sci.* **37**, 411–417.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Grousl, T., Ivanov, P., Frýdlová, I., Vasicová, P., Janda, F., Vojtová, J., Malínská, K., Malcová, I., Nováková, L., Janosková, D., Valásek, L. & Hasek, J. (2009). *J. Cell Sci.* **122**, 2078–2088.
Hegde, R. S. & Keenan, R. J. (2011). *Nature Rev. Mol. Cell Biol.* **12**, 787–798.
Hicke, L., Schubert, H. L. & Hill, C. P. (2005). *Nature Rev. Mol. Cell Biol.* **6**, 610–621.
Hu, Z., Potthoff, B., Hollenberg, C. P. & Ramezani-Rad, M. (2006). *J. Cell Sci.* **119**, 326–338.
Jonikas, M. C., Collins, S. R., Denic, V., Oh, E., Quan, E. M., Schmid, V., Weibezahn, J., Schwappach, B., Walter, P., Weissman, J. S. & Schuldiner, M. (2009). *Science*, **323**, 1693–1697.

Kohl, C., Tessarz, P., von der Malsburg, K., Zahn, R., Bukau, B. & Mogk, A. (2011). *Biol. Chem.* **392**, 601–608.
Kubota, K., Yamagata, A., Sato, Y., Goto-Ito, S. & Fukai, S. (2012). *J. Mol. Biol.* **422**, 366–375.
Kutay, U., Hartmann, E. & Rapoport, T. A. (1993). *Trends Cell Biol.* **3**, 72–75.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lawrence, M. C. & Colman, P. M. (1993). *J. Mol. Biol.* **234**, 946–950.
Leznicki, P., Warwicker, J. & High, S. (2011). *Biochem. J.* **436**, 719–727.
Liou, S.-T., Cheng, M.-Y. & Wang, C. (2007). *Cell Stress Chaperones*, **12**, 59–70.
Liou, S.-T. & Wang, C. (2005). *Arch. Biochem. Biophys.* **435**, 253–263.
Mariappan, M., Mateja, A., Dobosz, M., Bove, E., Hegde, R. S. & Keenan, R. J. (2011). *Nature (London)*, **477**, 61–66.
Mateja, A., Szlachcic, A., Downing, M. E., Dobosz, M., Mariappan, M., Hegde, R. S. & Keenan, R. J. (2009). *Nature (London)*, **461**, 361–366.
Nagano, N., Ota, M. & Nishikawa, K. (1999). *FEBS Lett.* **458**, 69–71.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Powis, K., Schrul, B., Tienson, H., Gostimskaya, I., Breker, M., High, S., Schuldiner, M., Jakob, U. & Schwappach, B. (2012). *J. Cell Sci.* **126**, 473–483.
Rabu, C., Schmid, V., Schwappach, B. & High, S. (2009). *J. Cell Sci.* **122**, 3605–3612.
Schantl, J. A., Roza, M., De Jong, A. P. & Strous, G. J. (2003). *Biochem. J.* **373**, 855–863.
Schuldiner, M., Metz, J., Schmid, V., Denic, V., Rakwalska, M., Schmitt, H. D., Schwappach, B. & Weissman, J. S. (2008). *Cell*, **134**, 634–645.
Simon, A. C., Simpson, P. J., Goldstone, R. M., Krysztofinska, E. M., Murray, J. W., High, S. & Isaacson, R. L. (2013). *Proc. Natl Acad. Sci. USA*, **110**, 1327–1332.
Simpson, P. J., Schwappach, B., Dohlman, H. G. & Isaacson, R. L. (2010). *Structure*, **18**, 897–902.
Stefanovic, S. & Hegde, R. S. (2007). *Cell*, **128**, 1147–1159.
Stefer, S., Reitz, S., Wang, F., Wild, K., Pang, Y.-Y., Schwarz, D., Bomke, J., Hein, C., Löhr, F., Bernhard, F., Denic, V., Dötsch, V. & Sinning, I. (2011). *Science*, **333**, 758–762.
Terwilliger, T. C. (2003). *Acta Cryst.* D**59**, 38–44.
Terwilliger, T. (2004). *J. Synchrotron Rad.* **11**, 49–52.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Vijay-Kumar, S., Bugg, C. E. & Cook, W. J. (1987). *J. Mol. Biol.* **194**, 531–544.
Wang, F., Brown, E. C., Mak, G., Zhuang, J. & Denic, V. (2010). *Mol. Cell*, **40**, 159–171.
Zhang, Q., Sanner, M. & Olson, A. J. (2009). *Proteins*, **75**, 453–467.