Biosci. Rep. (2013) / **33** / art:e00033 / doi 10.1042/BSR20120107

Original Paper

OPEN ACCESS

Bioscience Reports

www.bioscirep.org

Bioscience Reports

Low-resolution structure of the soluble domain GPAA1 (yGPAA1$^{70–247}$) of the glycosylphosphatidylinositol transamidase subunit GPAA1 from *Saccharomyces cerevisiae*

Wuan Geok SAW*, Birgit EISENHABER†, Frank EISENHABER†‡§ and Gerhard GRÜBER*†[1]

*Nanyang Technological University, School of Biological Sciences, 60 Nanyang Drive, Singapore 637551, Republic of Singapore, †Bioinformatics Institute, Agency for Science, Technology and Research (A*STAR), 30 Biopolis Street, 07-01 Matrix, Singapore 138671, Republic of Singapore, ‡School of Computer Engineering, Nanyang Technological University (NTU), 50 Nanyang Drive, Singapore 637553, Republic of Singapore, and §Department of Biological Sciences (DBS), National University of Singapore (NUS), 8 Medical Drive, Singapore 117597, Republic of Singapore

Synopsis

The GPI (glycosylphosphatidylinositol) transamidase complex catalyses the attachment of GPI anchors to eukaryotic proteins in the lumen of ER (endoplasmic reticulum). The *Saccharomyces cerevisiae* GPI transamidase complex consists of the subunits yPIG-K (Gpi8p), yPIG-S (Gpi17p), yPIG-T (Gpi16p), yPIG-U (CDC91/GAB1) and yGPAA1. We present the production of the two recombinant proteins yGPAA1$^{70–247}$ and yGPAA1$^{70–339}$ of the luminal domain of *S. cerevisiae* GPAA1, covering the amino acids 70–247 and 70–339 respectively. The secondary structural content of the stable and monodisperse yGPAA1$^{70–247}$ has been determined to be 28% α-helix and 27% β-sheet. SAXS (small-angle X-ray scattering) data showed that yGPAA1$^{70–247}$ has an R_g (radius of gyration) of 2.72 ± 0.025 nm and D_{max} (maximum dimension) of 9.14 nm. These data enabled the determination of the two domain low-resolution solution structure of yGPAA1$^{70–247}$. The large elliptical shape of yGPAA1$^{70–247}$ is connected via a short stalk to the smaller hook-like domain of 0.8 nm in length and 3.5 nm in width. The topological arrangement of yGPAA1$^{70–247}$ will be discussed together with the recently determined low-resolution structures of yPIG-K$^{24–337}$ and yPIG-S$^{38–467}$ from *S. cerevisiae* in the GPI transamidase complex.

Key words: glycosylphosphatidylinositol lipid anchor, glycosylphosphatidylinositol transamidase, GPAA1, PIG-K, PIG-S, post-translational modification.

Cite this article as: Saw, W.G., Eisenhaber, B., Eisenhaber, F. and Grüber, G. (2013) Low-resolution structure of the soluble domain GPAA1 (yGPAA1$^{70–247}$) of the glycosylphosphatidylinositol transamidase subunit GPAA1 from *Saccharomyces cerevisiae*. Biosci. Rep. **33**(2), art:e00033.doi:10.1042/BSR20120107

INTRODUCTION

GPI (glycosylphosphatidylinositol) lipid anchoring is an alternative post-translational modification for cellular systems to anchor proteins to the outer leaflet of the cell membrane [1,2]. Approximately 0.5% of all proteins in higher eukaryotes are capable of being GPI lipid-anchored [3], including cell surface receptors like urokinase receptor, CD14, CD16, enzymes [ALP (alkaline phosphatase) and 5′-nucleotidase], adhesion proteins and antigens [Thy-1, DAF (decay-accelerating factor) and MACIF (membrane attack complex inhibition factor)]. GPI lipid-anchoring of proteins has medical implications in a variety of conditions, including paroxysmal nocturnal haemoglobinuria, prion disease pathogenesis, AD (Alzheimer's disease) and parasitic diseases such as malaria [1,4].

The GPI lipid anchor is biosynthesized stepwise by a series of enzymes located in the ER (endoplasmic reticulum) membrane [1]. Typically, the nascent substrate proprotein is exported into the ER after ribosomal synthesis via the signal peptide pathway. The so-called transamidase complex in the ER lumen recognizes a specific GPI lipid anchor attachment signal located at the C-terminus of protein substrates [5–7]. This GPI lipid anchor attachment signal consists of a sequence motif from residue $\omega - 11$ to the very C-terminus of the protein substrate, with a region of polar residues between $\omega - 11$ and $\omega - 1$, a short

Abbreviations used: BV, bed volume; D_{max}, maximum dimension; DTT, dithiothreitol; ER, endoplasmic reticulum; GPI, glycosylphosphatidylinositol; GST, glutathione transferase; IPTG, isopropyl β-D-thiogalactopyranoside; MM, molecular mass; NSD, normalized spatial discrepancy; R_g, radius of gyration; SAXS, small-angle X-ray scattering; TM, transmembrane.

[1] To whom correspondence should be addressed (email ggrueber@ntu.edu.sg).

stretch of small polar residues from $\omega-1$ to $\omega+2$, including the ω-site for cleavage and the GPI attachment, a linker region that contains moderately polar residues, followed by a hydrophobic tail from $\omega+9$ to the C-terminus of the proprotein (Figure 1) [1]. During the modification process the C-terminal propeptide after the ω-site is removed and replaced with a pre-synthesized GPI moiety by the GPI transamidase complex [1,8]. The GPI lipid anchor is a complex organic structure made up of a lipid-modified phosphatidylinositol, a tetrasaccharide with variable elaborations and phosphoethanolamine subunits [2,9]. Although the general outline of the GPI lipid anchor biosynthesis pathway is common among eukaryotes, there are distinctive taxon-specific differences that might be useful for species-specific inhibitor design as described for suppression of unicellular parasites such as *Trypanosoma* [10,11].

The GPI transamidase complex is made up of at least five subunits named PIG-K, PIG-T, PIG-U, PIG-S and GPAA1 in humans [12–16]. PIG-K is the catalytic subunit in this complex and proposed to be responsible for the C-terminal proteolytic processing of the substrate proteins [17–19]. GPAA1 is predicted to consist of an N-terminal TM (transmembrane) region, a globular luminal domain and six C-terminal TM segments in humans, yeast and trypanosomes [1,20]. GPAA1 is discussed to play a role in structural function and/or in supplying the mature GPI lipid anchor to the protein substrate [1].

The molecular mechanisms of the GPI transamidase complex are the most non-understood aspect of the GPI lipid anchor biology. Therefore, biochemical and structural studies are needed to delineate the biological functions of the transamidase subunits. Structural studies of the transamidase subunits have proved difficult because of their solubility issue. Recently, the low-resolution solution structures of yPIG-K (yPIG-K$^{24–337}$) and yPIG-S (yPIG-S$^{38–467}$) from *Saccharomyces cerevisiae* have been solved [21]. yPIG-K$^{24–337}$ consists of an egg-like domain and a small globular segment that are linked by a short stalk, while yPIG-S$^{38–467}$ appears like an open-hand domain attached to a wrist made up by the smaller domain. Since GAA1/GPAA1 is described to interact with GPI transamidase subunits and in particular with Gpi8p/PIG-K [13,22], we wanted to extend the structural insight into the subunit ensemble of the GPI transamidase. Based on sequence analytic considerations, we generated the two constructs yGPAA1$^{70–247}$ and yGPAA1$^{70–339}$ of the luminal domain of *S. cerevisiae* GPAA1. In comparison with yGPAA1$^{70–339}$ the recombinant protein yGPAA1$^{70–247}$ showed a higher degree of stability as well as monodispersity and enabled us to observe the first low-resolution solution structure of this protein by SAXS (small-angle X-ray scattering).

EXPERIMENTAL

Biochemicals

Restriction enzymes were purchased from Fermentas. Chemicals for gel electrophoresis were received from Serva. All other chemicals were at least of analytical grade and received from BIOMOL, Merck, Roth, Sigma or Serva.

Cloning of luminal domain of yGPAA1

To amplify the luminal domain of *yGPAA1* genes encoding residues 70–339 and 70–247 (UniProt P39012, EMBL U53880.1), three oligonucleotide primers 5′-CTTTGGATCCTCTGAATGGAACATTTTGAGGGGCTATC-3′ (forward primer), 5′-AATACTCGAGCTACGATTGGTGAAATTTTTCCAAAAGG-3′ (yGPAA1$^{70–339}$ reverse primer) and 5′-CTTTCTCGAGCTAATGTTCCGTAATGGATATAGCGATGTTG-3′ (yGPAA1$^{70–247}$ reverse primer), incorporating BamHI and XhoI restriction sites (underlined), were designed. Following digestion with BamHI and XhoI, the PCR products were ligated into the pGEX-6P-1 vector. The vectors containing *yGPAA1*$^{70–339}$ and *yGPAA1*$^{70–247}$ genes were then transformed into *Escherichia coli* cells [strain BL21(DE3)]. The cells were grown on 100 μg/ml ampicillin-containing LB (Luria–Bertani) liquid cultures until an attenuance (D_{600}) of 0.6 was reached. To induce the production of GST (glutathione transferase)-tagged yGPAA1$^{70–339}$ and yGPAA1$^{70–247}$, the cultures were supplemented with IPTG (isopropyl β-D-thiogalactopyranoside) to a final concentration of 1 mM, and incubated for 16 h at 20 °C.

Purification of recombinant yGPAA1$^{70–247}$

Cells were lysed on ice by sonication for 3 × 1 min in buffer A [25 mM sodium phosphate, pH 7.0, 150 mM NaCl, 2 mM DTT (dithiothreitol) and 2 mM Pefabloc^{SC} (BIOMOL)]. The lysate was cleared by centrifugation at 12 500 ***g*** for 25 min and then the supernatant was filtered (0.45 μm; Millipore). For affinity purification of GST–yGPAA1$^{70–247}$, the filtered supernatant was incubated with the glutathione–Sepharose 4B (GE Healthcare) for 1 h, equilibrated with buffer B (25 mM sodium phosphate, pH 7.0, 150 mM NaCl and 2 mM DTT). Subsequently, the medium was washed with 5 BV (bed volume) of buffer B twice and followed by 10 BV of buffer C (50 mM Tris/HCl, pH 7.5, 150 mM NaCl, 2 mM DTT and 1 mM EDTA). The on-column cleavage of GST tag from GST–yGPAA1$^{70–247}$ was performed by incubating the medium with bound proteins together with PreScission Protease (GE Healthcare) for 16 h at 4 °C. Following incubation, the cleaved yGPAA1$^{70–247}$ protein was collected and applied to gel filtration column (Superdex 75 HR 10/30 column, GE Healthcare) in buffer D (50 mM Tris/HCl, pH 7.5, and 150 mM NaCl). Fractions containing yGPAA1$^{70–247}$ were identified by SDS/PAGE [23], pooled and concentrated using Amicon Ultra-4 Centrifugal Unit [10 kDa MM (molecular mass) cut-off] (Millipore).

CD spectroscopy

Steady-state CD spectra were measured in far-UV light (188–260 nm) using a CHIRASCAN spectropolarimeter (Applied Photophysics). Spectra were collected in a 60 μl quartz cell (Hellma) with a path length of 0.1 mm at 20 °C and a step resolution of 1 nm. The readings were for an average of 2 s at each

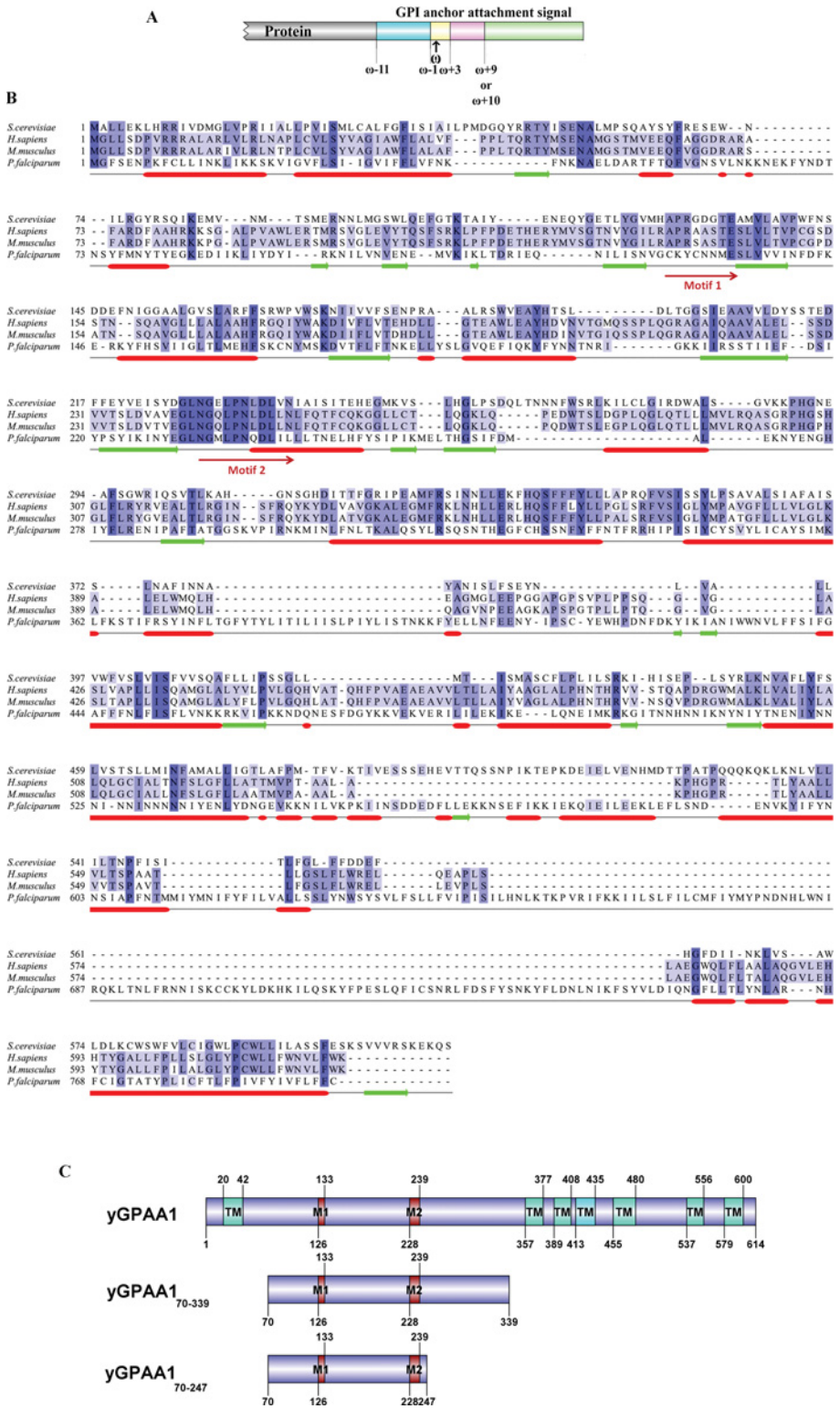

Figure 1 **The GPI lipid anchor motif in substrate proteins and the sequence architecture of GPAA1**
(**A**) Schematic diagram of the GPI lipid anchor attachment signal of a GPI-anchored proprotein. The attachment signal is mainly classified into four regions: the polar region from $\omega-11$ to $\omega-1$, a short segment with small polar residues from $\omega-1$ to $\omega+2$, a moderately polar linker and a stretch of hydrophobic residues from $\omega+9$ to $\omega+10$ to the C-terminus of the proprotein. The cleavage site and GPI anchor attachment site is at position ω. (**B**) Amino acid sequence alignment of *S. cerevisiae* GPAA1 (UniProt P39012) with the GPAA1 from *Homo sapiens* (O43292), *M. musculus* (Q9WTK3), *P. falciparum*

wavelength and the recorded ellipticity values were the average of three determinations for each sample. CD spectroscopy of yeast GPAA1$^{70–247}$ (1.1 mg/ml) was performed in buffer of 50 mM Tris/HCl, pH 7.5 and 150 mM NaCl. The spectrum for the buffer was subtracted from the spectrum of the protein. CD values were converted into mean residue molar ellipticity (Θ) in units of degree · cm^2/dmol using the software Chirascan Version 1.2, Applied Photophysics. This baseline-corrected spectrum was used as input for computer methods to obtain prediction of secondary structure. The CD spectrum was analysed using the K2D3 web server [24].

X-ray scattering experiments and data analysis

The synchrotron radiation X-ray scattering data for yGPAA1$^{70–247}$ were collected following the standard procedures on the X33 SAXS camera [25,26] of EMBL Hamburg located on a bending magnet (sector D) on the storage ring DORIS III of the DESY (Deutsches Elektronen Synchrotron). A photon counting Pilatus 1 M pixel detector (67 mm × 420 mm) was used at a sample–detector distance of 2.4 m covering the range of momentum transfer $0.1<s<4.5$/nm [$s=4p\ \sin(q)/l$, where q is the scattering angle and $l=0.15$ nm is the X-ray wavelength]. The s-axis was calibrated by the scattering pattern of silver-behenate salt (d-spacing 5.84 nm). The scattering from the buffer alone was measured before and after each sample measurement and the average of the scattering before and after each sample was used for background subtraction. The scattering pattern from yGPAA1$^{70–247}$ was measured at protein concentrations of 2.3 and 5.8 mg/ml respectively in a buffer composed of 50 mM Tris/HCl, pH 7.5 and 150 mM NaCl. The protein as well as the buffer samples were injected automatically using the sample-changing robot for solution scattering experiments at the SAXS station X33 [27]. All data processing steps were performed automatically using the program package PRIMUS [28]. The forward scattering $I(0)$ and the R_g (radius of gyration) were evaluated using the Guinier approximation [29] assuming that for spherical particles at very small angles ($s<1.3/R_g$) the intensity is represented by $I(s)=I(0)\ \exp[-(sR_g)^2/3]$. These parameters were also computed from the entire scattering patterns using the indirect transform package GNOM [30], which also provide the distance distribution function $\rho(r)$ of the particle as defined by:

$$\rho(r) = 2\pi \int I(s) sr \sin(sr) ds$$

The MM of both proteins was calculated by comparison with the forward scattering from the reference solution of BSA. From this procedure a relative calibration factor for the MM can be calculated using the known MM of BSA (66.4 kDa) and the concentration of the reference solution by applying

$$MM_p = I(0)_p/c_p \times \frac{MM_{st}}{I(0)_{st}/c_{st}}$$

where $I(0)_p, I(0)_{st}$ are the scattering intensities at zero angle of the studied and the BSA standard protein respectively, MM_p, MM_{st} are the corresponding MMs and c_p, c_{st} are the concentrations. Errors have been calculated from the upper and the lower $I(0)$ error limit estimated by the Guinier approximation. Low-resolution models of yeast GPAA1$^{70–247}$ were built by the program DAMMIN [31] as described in [32].

RESULTS AND DISCUSSION

Sequence analytic finding for yeast GPAA1

The sequence architecture of yeast GPAA1 (UniProt P39012 [33]) as predicted by sequence-analytic tools [1,34,35] consists of an N-terminal TM helical segment (20–42), a large luminal supposedly globular domain (42–356), followed by six TM helical regions (357–377, 389–408, 413–435, 455–480, 537–556 and 579–600) (Figures 1B and 1C). The yeast GPAA1 has a sequence identity to the human, *Mus musculus* and *Plasmodium falciparum* proteins of 26 %, 26 % and 19.4 % respectively. Within the predicted luminal domain located between the first TM segment and the second TM region, two conserved motifs are found at residues 126–133 ($RXPRX_3TE$) and 228–239 (NGX_2PNXDX_2N) (Figure 1C) which had been suggested to have a role in supplying the GPI lipid anchor to the transamidase complex [1]. These two motifs are proposed to be located at the membrane–lumen interface of the ER. Furthermore, the luminal domain of GPAA1 is essential for the assembly of GPI transamidase complex [22]. Here, two constructs were designed, in which one construct (yGPAA1$^{70–339}$) covers the predicted luminal part of yGPAA1 and one shorter luminal construct, yGPAA1$^{70–247}$, that still include the conserved motifs (Figure 1C).

Expression and purification of the luminal domain of yGPAA1

The *S. cerevisiae* GPAA1 genes, encoding the residues 70–339 and 70–247, have been amplified and cloned, and the GST-tagged proteins yGPAA1$^{70–339}$ and yGPAA1$^{70–247}$ have been produced in significant amounts in BL21(DE3) cells (Figure 2A). However, degradation of GST–yGPAA1$^{70–339}$ was observed in the

(C0H5L4), *Trypanosoma brucei* (Q7YTW5) and *Trypanosoma cruzi* (Q4DAV8) using Clustal Omega [39] in Jalview 2.8 [40]. The secondary structures presented were predicted based on yeast GPAA1 using JNet Secondary Structure Prediction server [41,42], with red cylinders (α-helices) and green arrows (β-strands). Two conserved motifs are labelled with red arrows. (**C**) Schematic diagram of the yGPAA1 protein sequence encoding 614 residues, showing that yGPAA1 contains seven potential TM segments, which are situated at residues 20–42, 357–377, 389–408, 413–435, 455–480, 537–556 and 579–600. The two conserved motifs ($RXPRX_3TE$ and NGX_2PNXDX_2N) within the globular domain of yGPAA1 [1] are found at residues 126–133 (M1) and 228–239 (M2). The constructs of residues 70–339 and 70–247 were created, which cover the luminal domain of yGPAA1 and two conserved motifs. The Figure has been generated using the program described in [43].

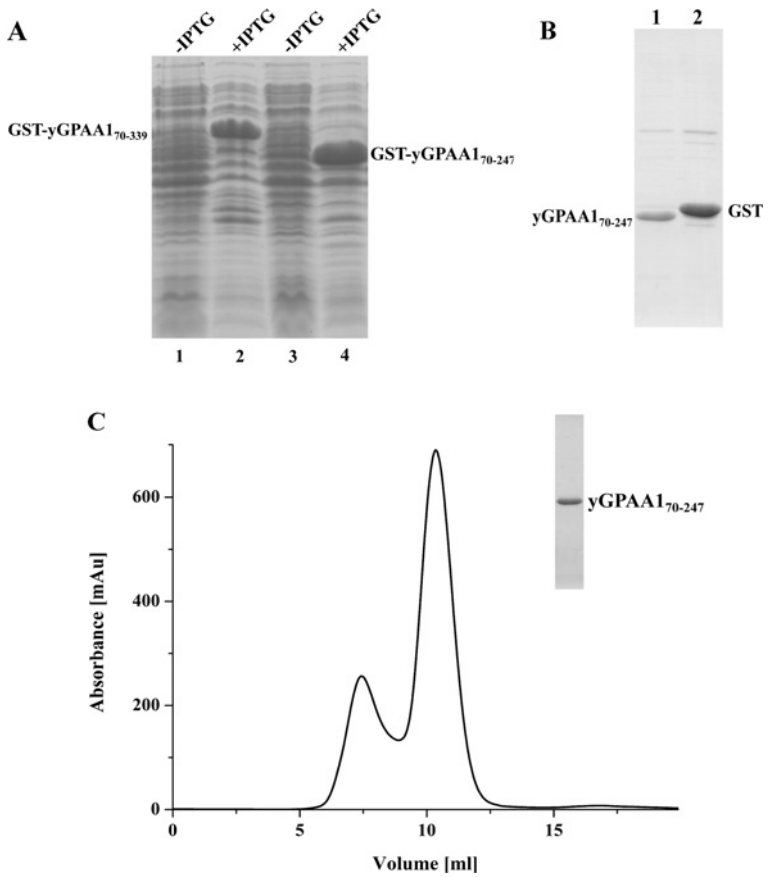

Figure 2 Production and purification of recombinant yGPAA1$^{70–247}$
(**A**) SDS/PAGE (17 % total acrylamide and 0.4 % cross-linked acrylamide) of produced GST–yGPAA1$^{70–339}$ (lanes 1 and 2) and GST–yGPAA1$^{70–247}$ (lane 3 and 4) in the absence (−ITPG) and presence (+IPTG) of 1 mM IPTG. Degradation was observed in GST–yGPAA1$^{70–339}$. (**B**) The eluted yGPAA1$^{70–247}$ (lane 1) and GST (lane 2) after GST tag cleavage by PreScission Protease at 4 °C, 16 h. (**C**) The eluted yGPAA1$^{70–247}$ was subsequently applied on to a Superdex 75 HR 10/30 column at a flow rate of 0.5 ml/min. The insert shows SDS/PAGE of the pooled protein fractions (grey area in the chromatogram).

corresponding SDS/PAGE (Figure 2A, lane 2). The same phenomenon was also described previously, where the recombinant full-length protein of the luminal loop of human GPAA1 underwent proteolysis in cells [22]. Therefore, only the recombinant GST–yGPAA1$^{70–247}$ was selected for further experiments.

GST–yGPAA1$^{70–247}$ was affinity-bound to the glutathione–Sepharose 4B medium from the crude lysate, before subjected to the on-column cleavage of GST tag by PreScission protease (Figure 2B). The eluted fractions were further purified via size-exclusion chromatography (Superdex 75 HR 10/30 column) as demonstrated by SDS/PAGE, revealing the high purity of the protein (Figure 2C). A Superdex 75 gel filtration column was calibrated by determining the K_{av} values for a set of standard proteins of known MM. Comparison of the K_{av} for yGPAA1$^{70–247}$ with the standard proteins suggests a native MM of approximately 24 kDa.

Secondary structure content of yGPAA1$^{70–247}$

The secondary structure of recombinant yGPAA1$^{70–247}$ was determined from CD spectra, measured between 188 and 260 nm (Figure 3). The average secondary structure content was 28 % α-helix and 27 % β-sheet. This result is consistent with secondary structure predictions based on yGPAA1$^{70–247}$ amino acid sequence using the web server PredictProtein [36,37] (36 % α-helix, 25 % β-sheet and 38 % random coil). The molar ellipticity values at 208 nm and at 222 nm are in a ratio of 1.26.

Shape determination of yGPAA1$^{70–247}$ in solution

The high protein purity allowed SAXS experiments to be performed, with the aim to determine the first low-resolution structures of yeast GPAA1$^{70–247}$ in solution. The final composite scattering curve of yGPAA1$^{70–247}$ is shown in Figure 4(A). Inspection of the Guinier plot of the protein at low angles showed good data quality and no protein aggregation (Figure 4B). The R_g of yGPAA1$^{70–247}$ is 2.72 ± 0.025 nm and the D_{max} (maximum dimension) of the protein is 9.14 nm (Figure 4C). Comparison of forward scattering $I(0)$ between yGPAA1$^{70–247}$ and reference protein, BSA, yields a MM of 28 kDa,

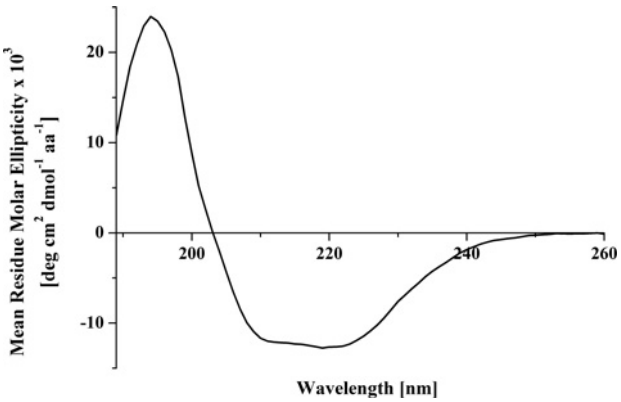

Figure 3 Far-UV CD spectrum of recombinant yGPAA1$^{70–247}$
The protein was measured in 50 mM Tris/HCl, pH 7.5, and 150 mM NaCl.

indicating that yGPAA1$^{70–247}$ is monomeric at the concentration used. Qualitative analysis of the distance distribution function $\rho(r)$ suggests that yGPAA1$^{70–247}$ consists of a major portion, yielding a principal maximum in the $\rho(r)$ at around 3.0 nm (Figure 4C), whereas the separated protuberance domain giving rise to a shoulder from 6.5 to 9.14 nm.

The gross structure of yGPAA1$^{70–247}$ was restored *ab initio* from the scattering patterns in Figure 4(A), using the program DAMMIN. The obtained shape for yGPAA1$^{70–247}$ yields a good fit to the experimental data in the entire scattering range. The corresponding fit, shown in Figure 4(A), has a discrepancy of $\chi^2 = 1.3$. All 10 independent reconstructions yielded a reproducible shape (Figure 4D). The NSD (normalized spatial discrepancy), which is a measure of similarity between sets of three-dimensional points [29], was computed between all 10 reconstructions, with a range of NSD from 0.396 to 0.476. The reconstruction with least NSD was selected, which has NSD = 0.396. The protein appears as a two-domain molecule with a large elliptical shape, followed by a hook-like domain, and connected via a narrowed 0.8 nm short stalk (average diameter 0.8 nm). The major domain has dimensions of approximately 7.1 nm × 4.8 nm, whereby the smaller domain is approximately 0.8 nm × 3.5 nm.

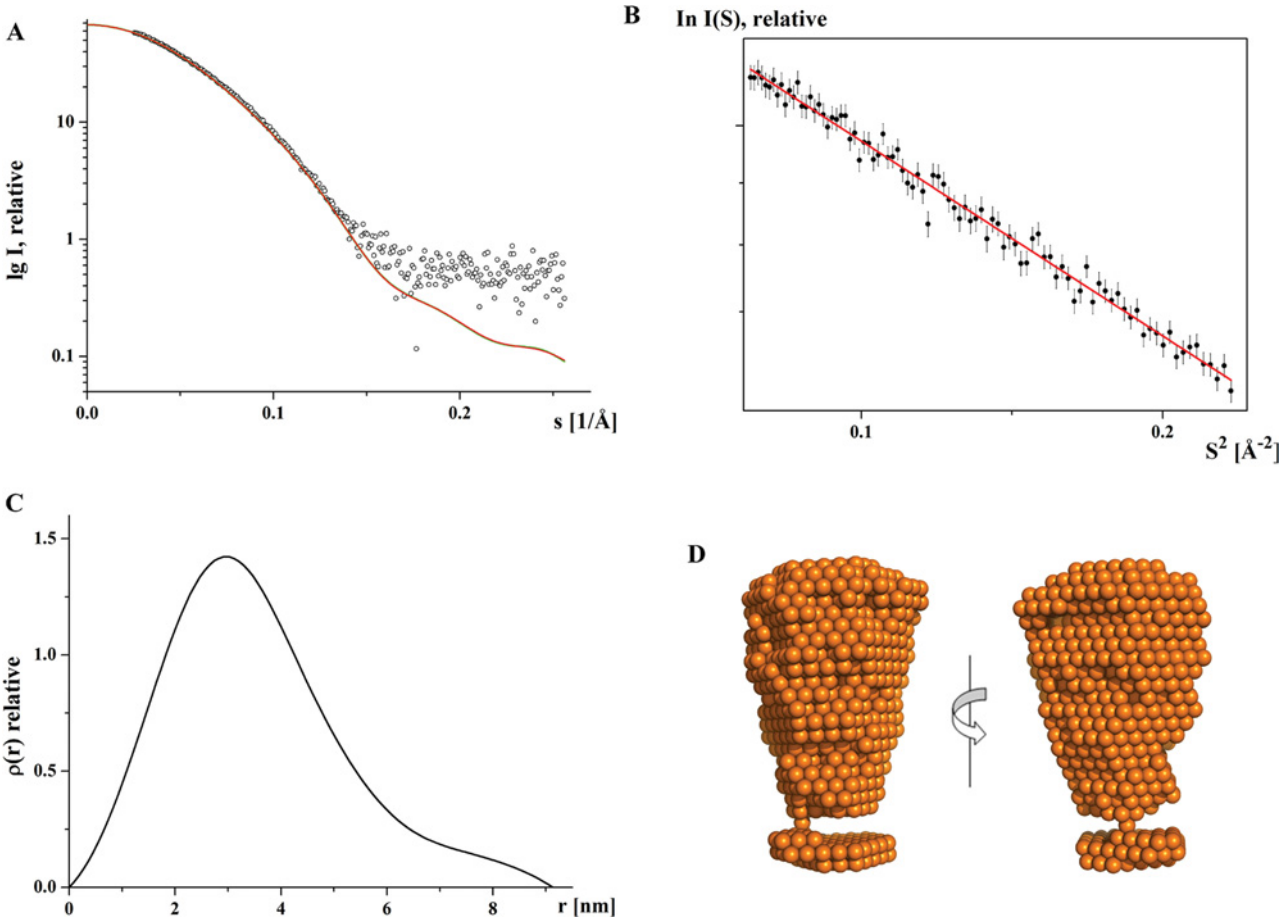

Figure 4 SAXS data of yGPAA1$^{70–247}$
(**A**) Experimental scattering data (o) and the fitting curve (–; green: experimental, red: calculated from *ab initio* model) for yGPAA1$^{70–247}$. (**B**) Guinier plot with the linear fit (red line). (**C**) The distance distribution function of the same protein. (**D**) Low-resolution solution structure of yGPAA1$^{70–247}$ derived from SAXS data. The right-hand model is rotated by 180° around the *y*-axis.

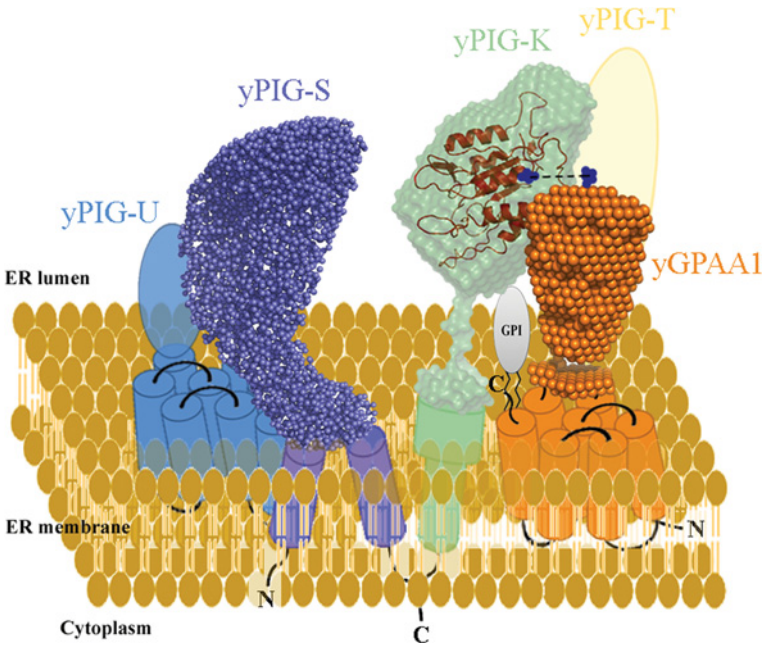

Figure 5 Proposed topological model of the subunits yPIG-K, yPIG-S, yPIG-T, yPIG-U and yGPAA1 inside the *S. cerevisiae* GPI transamidase complex

The low-resolution solution structure of yGPAA1$^{70–247}$ is shown together with the low-resolution structures of yPIG-K$^{28–337}$ and yPIG-S$^{38–467}$ described in Toh et al. [21]. A calculated three-dimensional model of yPIG-K, including the residues 28–292, is superimposed on the low-resolution solution structure yPIG-K$^{24–337}$ as described recently [21]. The broken line indicates the predicted cross-link between subunit yPIG-K and yPIG-T via a disulfide bond between amino acid residue Cys^{85} of PIG-K (blue spheres) and a conserved cysteine residue in PIG-T (blue spheres).

Insight into the structural topological arrangement of GPI transamidase

The eukaryotic transamidase complex is composed of five major subunits: PIG-K, PIG-S, PIG-T, PIG-U and GPAA1. The substrate protein interacts as folded protein with this complex [38]. The recently determined first solution structures of yPIG-K$^{24–337}$ and yPIG-S$^{38–467}$ from yeast provided the basis for a better understanding of the structural shape and arrangement of these subunits on the ER lumen side. The hand-wrist shaped yPIG-S$^{38–467}$ (Figure 5) with its concave surfaces allows this protein to interact with other subunits and to fulfil the function of stabilizing a substrate protein–transamidase complex. yPIG-K$^{24–377}$ has been described as an elongated particle consisting of an egg-like portion and a small globular segment linked together by an 1.9 nm long stalk. The egg-like portion involves the active site machinery (Figure 5) [21], which belongs to the C13 cysteine peptide family. The conserved residue Cys^{85} of yPIG-K$^{24–377}$ is exposed to the solvent [21] and proposed to form a disulfide bridge with a cysteine residue in subunit PIG-T of the yeast transamidase, bringing both subunits in close proximity (Figure 5) [21].

PIG-K co-precipitates with GPAA1 [13]. Both subunits are described to be in close proximity. PIG-K is known to bind the ω-site of the GPI anchor attachment signal in the substrate protein sequence [1]. The removal of the C-terminal TM domains in GPAA1 yields a non-functional GPI transamidase, while the assembly of the whole complex is not affected, suggesting the luminal domain of GPAA1 is important in the interaction with PIG-K, PIG-S and PIG-T [22]. The elongated shape of yGPAA1$^{70–247}$ with its elliptical domain, which is connected via a short stalk to the smaller hook-like domain, has a total length of 9.14 nm and would provide sufficient space to interact with yPIG-K$^{24–377}$, having a D_{max} of 10.3 nm (Figure 5) [21]. yGPAA1$^{70–247}$ includes the two conserved stretches of polar residues, i.e. RXPRX$_3$TE between residues 126–133 and NGX$_2$PNXDX$_2$N in the amino acid sequence 228–239. These two motifs have been suggested to play a role in the binding of the free GPI lipid anchor and to supply it to PIG-K [1]. The arrangement of yGPAA1$^{70–247}$ and yPIG-K$^{24–377}$ in the model presented brings both subunits in close proximity, which allows the entrance of the GPI lipid anchor for the subsequent catalytic event. The hydrophobic tail of the substrate would then supposedly be stabilized by residues of the TM helices at the N- and C-termini of yGPAA1$^{70–247}$, predicted to be involved in GPI interaction [1].

In summary, the cloning and production of the recombinant yGPAA1$^{70–247}$ provided a pure and monodisperse protein with a proper secondary structural content. The protein has been demonstrated as an elongated protein with a major elliptical- and a smaller hook-like domain that connected via a short stalk. Together with the recently determined low-resolution structures of yPIG-K$^{24–337}$ and yPIG-S$^{38–467}$ [21], the presented low-resolution structure of yGPAA1$^{70–247}$ in solution shines new light into the concerted interaction of the transamidase complex

subunits together with its substrate protein and provides the basis for a better understanding of the structural shape and arrangement of these subunits on the ER lumen side.

AUTHOR CONTRIBUTION

Wuan Geok Saw designed and conducted experiments, and performed data analysis. Birgit Eisenhaber and Frank Eisenhaber contributed to the experimental design and discussions. Gerhard Grüber designed the study and experiments, and supervised the research. All authors participated in writing the paper.

ACKNOWLEDGEMENTS

W.G.S. thanks the authority of Nanyang Technological University for awarding a research scholarship. We acknowledge the EMBL-Outstation, Hamburg, Germany, for provision of synchrotron radiation facility, and we thank M. Roessle for his great help in collecting SAXS data. We are grateful to M. S. S. Manimekalai (SBS, NTU) for her support in SAXS data analysis.

FUNDING

This work was supported by the New Initiative Fund FY2010, NTU.

REFERENCES

1 Eisenhaber, B., Maurer-Stroh, S., Novatchkova, M., Schneider, G. and Eisenhaber, F. (2003) Enzymes and auxiliary factors for GPI lipid anchor biosynthesis and post-translational transfer to proteins. BioEssays **25**, 367–385
2 Paulick, M. G. and Bertozzi, C. R. (2008) The glycosylphosphatidylinositol anchor: a complex membrane-anchoring structure for proteins. Biochemistry **47**, 6991–7000
3 Eisenhaber, B., Bork, P. and Eisenhaber, F. (2001) Post-translational GPI lipid anchor modification of proteins in kingdoms of life: analysis of protein sequence data from complete genomes. Protein Eng. **14**, 17–25
4 Taylor, D. and Hooper, N. (2011) GPI-anchored proteins in health and disease. In Post-Translational Modifications in Health and Disease (Vidal, C. J., ed.), pp. 39–55, Springer, New York
5 Eisenhaber, B., Bork, P. and Eisenhaber, F. (1998) Sequence properties of GPI-anchored proteins near the omega-site: constraints for the polypeptide binding site of the putative transamidase. Protein Eng. **11**, 1155–1161
6 Eisenhaber, B. and Eisenhaber, F. (2010) Prediction of post-translational modification of proteins from their amino acid sequences. In Data Mining Techniques for the Life Sciences (Carugo, O. and Eisenhaber, F., eds), pp. 365–384, Humana Press, New York
7 Eisenhaber, B., Schneider, G., Wildpaner, M. and Eisenhaber, F. (2004) A sensitive predictor for potential GPI lipid modification sites in fungal protein sequences and its application to genome-wide studies for *Aspergillus nidulans*, *Candida albicans*, *Neurospora crassa*, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. J. Mol. Biol. **337**, 243–253
8 Orlean, P. and Menon, A. K. (2007) Thematic review series: lipid posttranslational modifications. GPI anchoring of protein in yeast and mammalian cells, or: how we learned to stop worrying and love glycophospholipids. J. Lipid Res. **48**, 993–1011
9 Eisenhaber, B. and Eisenhaber, F. (2007) Posttranslational modifications and subcellular localization signals: indicators of sequence regions without inherent 3D structure? Curr. Protein Pept. Sci. **8**, 197–203
10 Izquierdo, L., Nakanishi, M., Mehlert, A., Machray, G., Barton, G. J. and Ferguson, M. A. J. (2009) Identification of a glycosylphosphatidylinositol anchor-modifying β1–3 *N*-acetylglucosaminyl transferase in *Trypanosoma brucei*. Mol. Microbiol. **71**, 478–491
11 Urbaniak, M. D., Yashunsky, D. V., Crossman, A., Nikolaev, A. V. and Ferguson, M. A. J. (2008) Probing enzymes late in the trypanosomal glycosylphosphatidylinositol biosynthetic pathway with synthetic glycosylphosphatidylinositol analogues. ACS Chem. Biol. **3**, 625–634
12 Hong, Y., Ohishi, K., Kang, J. Y., Tanaka, S., Inoue, N., Nishimura, J., Maeda, Y. and Kinoshita, T. (2003) Human PIG-U and yeast Cdc91p are the fifth subunit of GPI transamidase that attaches GPI-anchors to proteins. Mol. Biol. Cell **14**, 1780–1789
13 Ohishi, K., Inoue, N., Maeda, Y., Takeda, J., Riezman, H. and Kinoshita, T. (2000) Gaa1p and Gpi8p are components of a glycosylphosphatidylinositol (GPI) transamidase that mediates attachment of GPI to proteins. Mol. Biol. Cell **11**, 1523–1533
14 Ohishi, K., Inoue, N. and Kinoshita, T. (2001) PIG-S and PIG-T, essential for GPI anchor attachment to proteins, form a complex with GAA1 and GPI8. EMBO J. **20**, 4088–4098
15 Ohishi, K., Nagamune, K., Maeda, Y. and Kinoshita, T. (2003) Two subunits of glycosylphosphatidylinositol transamidase, GPI8 and PIG-T, form a functionally important intermolecular disulfide bridge. J. Biol. Chem. **278**, 13959–13967
16 Fraering, P., Imhof, I., Meyer, U., Strub, J.-M., van Dorsselaer, A., Vionnet, C. and Conzelmann, A. (2001) The GPI transamidase complex of *Saccharomyces cerevisiae* contains Gaa1p, Gpi8p, and Gpi16p. Mol. Biol. Cell **12**, 3295–3306
17 Maxwell, S. E., Ramalingam, S., Gerber, L. D., Brink, L. and Udenfriend, S. (1995) An active carbonyl formed during glycosylphosphatidylinositol addition to a protein is evidence of catalysis by a transamidase. J. Biol. Chem. **270**, 19576–19582
18 Ramalingam, S., Maxwell, S. E., Medof, M. E., Chen, R., Gerber, L. D. and Udenfriend, S. (1996) COOH-terminal processing of nascent polypeptides by the glycosylphosphatidylinositol transamidase in the presence of hydrazine is governed by the same parameters as glycosylphosphatidylinositol addition. Proc. Natl. Acad. Sci. U.S.A. **93**, 7528–7533
19 Sharma, D. K., Hilley, J. D., Bangs, J. D., Coombs, G. H., Mottram, J. C. and Menon, A. K. (2000) Soluble GPI8 restores glycosylphosphatidylinositol anchoring in a trypanosome cell-free system depleted of lumenal endoplasmic reticulum proteins. Biochem. J. **351**, 717–722
20 Nagamune, K., Ohishi, K., Ashida, H., Hong, Y., Hino, J., Kangawa, K., Inoue, N., Maeda, Y. and Kinoshita, T. (2003) GPI transamidase of *Trypanosoma brucei* has two previously uncharacterized (trypanosomatid transamidase 1 and 2) and three common subunits. Proc. Natl. Acad. Sci. U.S.A. **100**, 10682–10687
21 Toh, Y. K., Kamariah, N., Maurer-Stroh, S., Roessle, M., Eisenhaber, F., Adhikari, S., Eisenhaber, B. and Grüber, G. (2011) Structural insight into the glycosylphosphatidylinositol transamidase subunits PIG-K and PIG-S from yeast. J. Struct. Biol. **173**, 271–281
22 Vainauskas, S., Maeda, Y., Kurniawan, H., Kinoshita, T. and Menon, A. K. (2002) Structural requirements for the recruitment of Gaa1 into a functional glycosylphosphatidylinositol transamidase complex. J. Biol. Chem. **277**, 30535–30542
23 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**, 680–685

24 Louis-Jeune, C., Andrade-Navarro, M. A. and Perez-Iratxeta, C. (2012) Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. Proteins: Struct. Funct. Bioinform. **80**, 374–381

25 Roessle, M. W., Klaering, R., Ristau, U., Robrahn, B., Jahn, D., Gehrmann, T., Konarev, P., Round, A., Fiedler, S., Hermes, C. et al. (2007) Upgrade of the small-angle X-ray scattering beamline X33 at the European Molecular Biology Laboratory, Hamburg. J. Appl. Crystallogr. **40**, s190–s194

26 Round, A. R., Franke, D., Moritz, S., Huchler, R., Fritsche, M., Malthan, D., Klaering, R., Svergun, D. I. and Roessle, M. (2008) Automated sample-changing robot for solution scattering experiments at the EMBL Hamburg SAXS station X33. J. Appl. Crystallogr. **41**, 913–917

27 Svergun, D. I. (1993) A direct indirect method of small-angle scattering data treatment. J. Appl. Crystallogr. **26**, 258–267

28 Gunier, A. and Fournet, G. (1955) Small Angle Scattering of X-Rays, Wiley, New York

29 Svergun, D. I., Petoukhov, M. V. and Koch, M. H. J. (2001) Determination of domain structure of proteins from X-ray solution scattering. Biophys. J. **80**, 2946–2953

30 Petoukhov, M. V. and Svergun, D. I. (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. Biophys. J. **89**, 1237–1250

31 Svergun, D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys. J. **76**, 2879–2886

32 Armbrüster, A., Svergun, D. I., Coskun, Ü., Juliano, S., Bailer, S. M. and Grüber, G. (2004) Structural analysis of the stalk subunit Vma5p of the yeast V-ATPase in solution. FEBS Lett. **570**, 119–125

33 Consortium, T. U. (2012) Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic Acids Res. **40**, D71–D75

34 Ooi, H. S., Kwo, C. Y., Wildpaner, M., Sirota, F. L., Eisenhaber, B., Maurer-Stroh, S., Wong, W. C., Schleiffer, A., Eisenhaber, F. and Schneider, G. (2009) ANNIE: integrated *de novo* protein sequence annotation. Nucleic Acids Res. **37**, W435–W440

35 Schneider, G., Sherman, W., Kuchibhatla, D., Ooi, H. S., Sirota, F. L., Maurer-Stroh, S., Eisenhaber, B. and Eisenhaber, F. (2012) Protein sequence-structure-function-network links discovered with the ANNOTATOR software suite: application to ELYS/Mel-28. In Computational Medicine (Trajanoski, Z., ed.), pp. 111–143, Springer, Vienna

36 Rost, B., Yachdav, G. and Liu, J. (2003) The PredictProtein server. Nucleic Acids Res. **32**, W321–W326

37 Rost, B. (1996) PHD: predicting one-dimensional protein structure by profile-based neural networks. Methods Enzymol. **266**, 525–539

38 Spurway, T. D., Dalley, J. A., High, S. and Bulleid, N. J. (2001) Early events in glycosylphosphatidylinositol anchor addition. J. Biol. Chem. **276**, 15975–15982

39 Sievers, F., Wilm, A., Dineen, D. G., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J. et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. **7**, 539

40 Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. and Barton, G. J. (2009) Jalview version 2: a multiple sequence alignment editor and analysis workbench. Bioinformatics **25**, 1189–1191

41 Cole, C., Barber, J. D. and Barton, G. J. (2008) The Jpred 3 secondary structure prediction server. Nucleic Acids Res. **36**, W197–W201

42 Cuff, J. A. and Barton, G. J. (2000) Application of multiple sequence alignment profiles to improve protein secondary structure prediction. Proteins **40**, 502–511

43 Ren, J., Wen, L., Gao, X., Jin, C., Xue, Y. and Yao, X. (2009) DOG 1.0: illustrator of protein domain structures. Cell Res. **19**, 271–273

Received 25 October 2012/2 January 2013; accepted 23 January 2013

Published as Immediate Publication 4 March 2013, doi 10.1042/BSR20120107