SUPPLEMENTARY METHODS

Peptides

All peptides were chemically synthesized and purified by Peptide Protein Research Ltd, Fareham, UK.

Cloning, expression and purification

PALB2-C

The vector pOZC-PALB2 was used as a template for PCR, as it contained a full-length coding sequence for human PALB2. Primers were designed to amplify the regions encompassing the C-terminal WD40-repeat (PALB2-C; residues 835-1186), with additional flanking restriction sites (*Nde*I and *Eco*RI) to facilitate cloning into the donor vector pTWO-B - a in-house modified pFastBac1 vector (Invitrogen, Paisley, UK) that encodes an N-terminal hexahistidine affinity tag followed by a rhinovirus 3C-protease site. A recombinant baculovirus was generated from this donor vector, as per the manufacturer's recommended protocol.

Insect cells (*Spodoptera frugiperda*, *Sf*9) were cultured in shaker flasks, containing SF900II media (supplemented with penicillin and streptomycin), with growth conditions of 27 ºC and 150 rpm. Recombinant virus, encoding PALB2-C, was amplified from cells infected at a cell density of 2.0 x 10^6 cells/ml, a multiplicity of infection (MOI) of 0.1, and a 3-day incubation period. For expression, *Sf*9 cells were infected at the same cell density, but with an MOI of 2, and a 3-day incubation period. Cells were harvested by centrifugation at 1000 x *g* for 5 minutes, and the resulting cell pellet stored at -80 ºC until required.

The cell pellet from 5 litres of expression culture was resuspended, on ice, in 45 ml of Buffer A: 50 mM HEPES-NaOH pH 7.5, 250 mM NaCl, 10 mM imidazole, supplemented with protease inhibitors (Roche, Burgess Hill, UK). Cells were lysed through a combination of the thawing process, hand-homogenisation and a brief sonication step. Cell debris and precipitated material were then removed by a high-speed centrifugation step at 40 000 x *g* for 30 minutes, at 10 ºC.

The supernatant arising from this step was filtered through a 5 μm filter (Sartorius Stedim, Epsom, UK) then applied to a batch/gravity column containing 10 ml of Talon resin (TaKaRa Bio, Saint-Germain-en-Laye, France) equilibrated in Buffer A. The column containing the cell extract and

resin was rotated/rolled at 4 ºC for a period of 1 hour to facilitate protein binding. The resin was allowed to pack under gravity, and then washed with successive applications of Buffer A (approximately 250 ml in total). Bound protein was eluted from the column with the application of Buffer B: 50 mM HEPES-NaOH pH 7.5, 250 mM NaCl, 300 mM Imidazole. Fractions containing PALB2-C were identified by SDS-PAGE and then pooled and concentrated to a final volume of 10 ml (Vivaspin 20, 10 kD MWCO; Sartorius Stedim). The affinity tag was cleaved from the protein (when required) by incubation with rhinovirus 3C-protease overnight at 4 ºC. The sample was then applied to a Superdex S200 (GE Healthcare, Little Chalfont, UK) size exclusion column equilibrated in Buffer C: 10 mM HEPES-NaOH pH 7.5, 250 mM NaCl, 5 mM DTT, 1 mM EDTA. Again, fractions containing PALB2-C were identified by SDS-PAGE, then pooled and concentrated. The purified protein was flash-frozen on dry-ice and stored at -80 ºC until required.

EMSY-N

A codon-optimised sequence (for expression in E.coli) corresponding to the N-terminal region of EMSY (EMSY-N; amino acids 1-139) was generated by gene synthesis (MWG Biotech, Ebersberg, Germany), subcloned into the expression vector pTWO-E (A W Oliver, ICR), then transformed into the E.coli strain Rosetta2(DE3)pLysS (Merck Chemicals, Nottingham, UK). A 250 ml flask containing 50 ml of L-broth (1% w/v tryptone, 0.5% w/v NaCl, 0.5% w/v yeast extract), supplemented with 50 µg/ml carbenicillin and 34 µg/ml chloramphenicol, was inoculated with a single transformed bacterial colony. This was grown, in an orbital incubator set at 37 °C and 225 rpm, until the absorbance of the culture (at 600 nm) reached 0.6. This ‘starter culture’ was then stored overnight at 4 °C. From the starter culture, 10 ml was used to inoculate a 2 l flask containing 1 l of L-broth supplemented, as before, with antibiotics. The culture was grown in an orbital incubator set at 37 °C and 225 rpm. Once the A_{600} of the cell culture had reached 0.6, expression of EMSY-N was induced by the addition of IPTG to a final concentration of 1 mM. The culture was grown for a further 3 hours, when the cells were harvested by centrifugation (3100 g, 10 minutes, 10 °C), and the pellet stored at -80 °C until required.

The cell pellet from 2 l of cell culture was resuspended, on ice, in 5 ml of Buffer A: 50 mM HEPES-NaOH pH 7.5, 250 mM NaCl, 10 mM imidazole, supplemented with protease inhibitors (Roche, Burgess Hill, UK). Cells were lysed by sonication (15 x 5 second bursts, on ice, at 35 %

amplitude, stepped microprobe, Jencons Ultrasonic Processor), after which cell debris and precipitated material were removed by a centrifugation step at 16 000 x *g* for 10 minutes at 10 °C.

EMSY-N was purified by the same chromatographic steps as those used for PALB2-C, with the exception of using a Sephadex S75 (GE Healthcare) column at the gel filtration stage.

Crystallisation, Phasing and Data Refinement

PALB2-C

Initial crystallisation trials were carried out by the vapour diffusion method in hanging drops, at a protein concentration of 10 mg/ml and a temperature of 20 ˚C. Crystals were observed after 24 hours in condition 34 of 'The PEGs Suite' (Qiagen, Crawley, UK). This condition was optimised, again in vapour diffusion experiments at 20 ˚C, to mixing 1 µl of protein (10 mg/ml in 10 mM HEPES-NaOH pH 7.5, 250 mM NaCl, 5 mM DTT, 1 mM EDTA) with 1 µl of precipitant containing 100 mM MES pH 6.0, 50 mM KH_2PO_4 and PEG 8000 at a concentrations between 12 and 20 % (w/v). Crystals were cryoprotected for data collection by step-wise soaking in buffers containing increasing amounts of glycerol, to a final concentration of 30% (v/v).

Heavy atom soaks

Crystals were transferred into a stabilisation solution of 100 mM HEPES-NaOH pH 7.5, 200 mM NaCl, 18% (w/v) PEG 8000, 30 % (v/v) Ethylene Glycol. $KAu(CN)_2$ was then added (from a 100 mM stock solution in 100 mM HEPES-NaOH pH 7.5) to achieve a final concentration of 10 mM. Crystals were soaked overnight (~20 hours), before being frozen for data collection.

Co-crystallisation with BRCA2 peptide

Buffer exchange of PALB2-C was carried out using a Micro Biospin 6 column (Bio-Rad, Hemel Hempstead, UK). PALB2-C (72 µl at 9 mg/ml in 20 mM HEPES-NaOH pH 7.5, 100 mM NaCl, 1 mM DTT, 1 mM EDTA) was mixed with BRCA2(21-39) peptide (18 µl at 4 mg/ml in 50 mM HEPES-NaOH pH 7.5, 100 mM NaCl), to give a 1:2 molar ratio of protein:peptide. Crystallographic screening in hanging drops at 20˚C produced two related conditions, 67 and 84,

from the 'Classics' screen (Qiagen). These conditions were optimised, again in vapour diffusion experiments at 20˚C, to mixing 2 μl of the protein-peptide complex with 1.6 μl of 0.1 M Na-Cacodylate pH 6.5, 0.2 M magnesium acetate, 10% (w/v) PEG 8000 plus an additional 0.4 μl of BRCA2(21-39) peptide. Crystals were cryoprotected for data collection by step-wise soaking in buffers containing increasing amounts of glycerol, to a final concentration of 30% (v/v).

Data Collection and Refinement

All diffraction data were collected on station ID14.1 at ESRF, Grenoble, France. Data were process sed and scaled using the software packages Mosflm (Leslie, 1995) and Scala (CCP4, 1994).

The PALB2-C structure was phased using single wavelength anomalous dispersion using the KAu $(CN)_2$ soaked crystals. The positions of 4 Au atoms were determined using SHELX (Sheldrick, 2008) then refined using SHARP (Vonrhein et al, 2007) to give an interpretable electron density map at 2.1 Å. Automated model building was then carried out using the programs of the PHENIX suite (Adams et al, 2002). Manual model building and crystallographic refinement used Coot (Emsley & Cowtan, 2004) and the PHENIX suite to refine the model against the native dataset at 1.9 Å. This model was then used to solve the PALB2-C / BRCA2(21-39) co-crystal dataset to 2.1 Å using PHASER (McCoy, 2007).

Biochemical and Biophysical Experiments

Peptide Binding / Competition Assays

Experiments were carried out with biotinylated peptides corresponding to either residues 21-39 of human BRCA2 (Biotin-GG-KALDGPISLNWFEELSSEA) or to the same sequence containing a single point mutant W31C (Biotin-GG-KALDGPISLNCFEELSSEA). An excess of peptide was incubated for 1 hour at room temperature with 50 μl of NeutrAvidin beads (Thermo Scientific, Rockford, USA) pre-equilbrated in Binding Buffer: 20 mM HEPES.NaOH pH 7.5, 100 mM NaCl, 1 mM DTT, 1 mM EDTA, 0.01 % (v/v) IGEPAL CA-630. Beads were equilibrated with purified PALB2-C or EMSY-N (0.3 mg/ml in Binding Buffer) for a period of 1 hour at 4 °C, then extensively washed, before being resuspended in SDS loading buffer and analysed by SDS-PAGE.

For the protein competition experiments EMSY-N was prebound to the NeutrAvidin beads, then equilibrated with increasing concentrations of PALB2-C in Binding Buffer (0.06, 0.15, 0.3 and 0.6 mg/ml) for a period of 1 hour at 4˚C, before washing and analysis as before.

For the peptide competition experiments, protein was mixed with increasing concentrations of a non-biotinylated BRCA2(21-39) peptide (KALDGPISLNWFEELSSEA) then incubated with NeutrAvidin beads for a period of 1 hour, before washing and analysis as before.

Fluorescence Polarisation

A fluorescein-labelled peptide (Fluorescein-GG-KALDGPISLNWFEELSSEA) at a concentration of 10 nM was incubated, for a period of 30 minutes at 4 °C, with increasing concentrations of PALB2-C in FP Buffer (20 mM HEPES/NaOH pH 7.5, 100 mM NaCl, 1mM EDTA, 1mM DTT, 0.01% (v/v) IGEPAL CA-630). The samples (typically a volume of 70 μl) were then transferred to a black 96-well polypropylene plate (U96 MicroWell Plate, Thermo Fisher Scientific, UK) for measurement of fluorescence polarisation in a Victor X 2030 Multilabel Reader (Perkin Elmer LAS Ltd, UK). Two 0.1 second measurements were collected for each well with excitation / emission wavelengths of 485 and 535 nm respectively, with either parallel or perpendicular polarisers in-line. Background fluorescence in wells containing only FP-buffer was subtracted from all values obtained for the samples. Polarisation, in millipolarisation units (mP) was calculated using the following equation:

$$\text{Fluorescence Polarisation} = (I_{parallel} - I_{perpendicular}) / (I_{parallel} + I_{perpendicular}) \times 1000$$

where $I_{parallel}$ = emission measured parallel to the excitation plane and $I_{perpendicular}$ = emission erpendicular to the excitation plane.

Polarisation data were analysed using GraphPad Prism 5.0 by non-linear fitting with a one-site specific binding model.

Mammalian Expression and M2-Agarose Immunoprecipitation

Single point mutants of full-length PALB2, in the mammalian expression vector pOZ-FH-C, were generated using a QuickChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) following the manufacturer's recommended protocol. HEK293T cells were obtained from ATCC and

maintained in DMEM containing 10 % (v/v) fetal calf serum. Cells were transfected in 6 well plates using FUGENE6 (Roche, Burgess Hill, UK) according to the manufacturer's instructions. Whole-cell extracts were prepared in 1 ml of NETN420 buffer (20 mM Tris-HCl pH 7.5, 420 mM NaCl, 1 mM EDTA, 0.5 % (v/v) IGEPAL CA-630) as described in Xia *et al.,* 2006, forty-eight hours after transfection. Flag-tagged PALB2 was precipitated from these extracts with anti-FLAG M2 Agarose beads (Sigma-Aldrich, Gillingham, UK), for a period of 20 hours at 4˚C.

SUPPLEMENTARY FIGURE LEGENDS

Fig S1: Electron density for BRCA2(21-39) peptide bound to PALB2-C

The mesh (coloured blue) corresponds to the electron density from an Fo-Fc omit map contoured at 2.0 σ. The fitted BRCA2(21-39) peptide is shown by the 'stick' representation (magenta) and PALB2-C by the cartoon representation (yellow/green).

Fig S2: Amino acid sequence conservation of the WD40 domain across PALB2 homologues

Molecular surface representations of PALB2-C coloured according to amino acid conservation scores: from red (absolutely conserved) to white (not conserved).

Two separate regions of high amino acid conservation can be visualised on (A) the 'Bottom' face - corresponding to the BRCA2 binding site determined in this study, and (B) the 'Top' face – suggesting the WD40 domain of PALB2 may bind / recruit other proteins in addition to BRCA2.

Conservation scores were calculated from multiple sequence alignments of vertebrate PALB2 amino acid sequences (see Figure S3) by the ConSurf web-server (http://consurf.tau.ac.il).

Fig S3: Multiple sequence alignment of the C-terminal WD40 region of PALB2 homologues

Amino acid sequences corresponding to homologues of human PALB2 were obtained either from the UniProt database (http://www.uniprot.org) or from a BLAST search (http://blast.ncbi.nlm.nih.gov) using the human sequence as a query.

The multiple sequence alignment was generated using MultAlin (http://bioinfo.genotoul.fr/multalin) and formatted for presentation with ESPript (http://espript.ibcp.fr/ESPript/ESPript).

Amino acid sequences are numbered according to the human protein. Secondary structure elements corresponding to that of the PALB2-C structure presented in this manuscript are also shown (PDB: 2W18) and labelled using standard 'WD40' nomenclature (see manuscript, Fig 1B).

Fig S4. Competitive binding of PALB2 and EMSY to BRCA2-N

(A) Overlap of binding sites on the BRCA2 N-terminus sequence - for PALB2, as determined from the crystal structure presented here, and for EMSY, inferred from a yeast-two-hybrid alanine scan (Hughes-Davies et al, 2003).

(B) Co-precipitation of an N-terminal EMSY construct (1-139) by the biotinylated BRCA2(21-39) peptide. Gels are Coomassie stained (left) and a western blot against the His6 affinity-tag appended to EMSY-N (right). Bands marked with an asterisk arise from the NeutrAvidin resin (see Experimental Procedures).

(C) Co-precipitation of EMSY-N by biotinylated BRCA2(21-39) is progressively eliminated by the presence of increasing amounts of a non-biotinylated version of the same peptide. However, incubation with increasing amounts of an unrelated XRCC1 peptide, does not eliminate binding, indicating that the EMSY-N / BRCA(21-39) interaction is specific.

(D) Co-precipitation of EMSY-N by biotinylated BRCA2(21-39) is progressively eliminated by the presence of increasing amounts of PALB2-C in the reaction, showing direct competition of the two proteins for BRCA2 binding.

REFERENCES FOR SUPPLEMENTARY INFORMATION

Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr 58: 1948-1954

CCP4 (1994) Programs for protein crystallography. Acta Crystallographica D50: 760-763

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60: 2126-2132

Hughes-Davies L, Huntsman D, Ruas M, Fuks F, Bye J, Chin SF, Milner J, Brown LA, Hsu F, Gilks B, Nielsen T, Schulzer M, Chia S, Ragaz J, Cahn A, Linger L, Ozdag H, Cattaneo E, Jordanova ES, Schuuring E, Yu DS, Venkitaraman A, Ponder B, Doherty A, Aparicio S, Bentley D, Theillet C, Ponting CP, Caldas C, Kouzarides T (2003) EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. *Cell* **115**: 523-535.

Leslie AGW. (1995) MOSFLM Users Guide. MRC Laboratory of Molecular Biology, Cambridge, U.K.

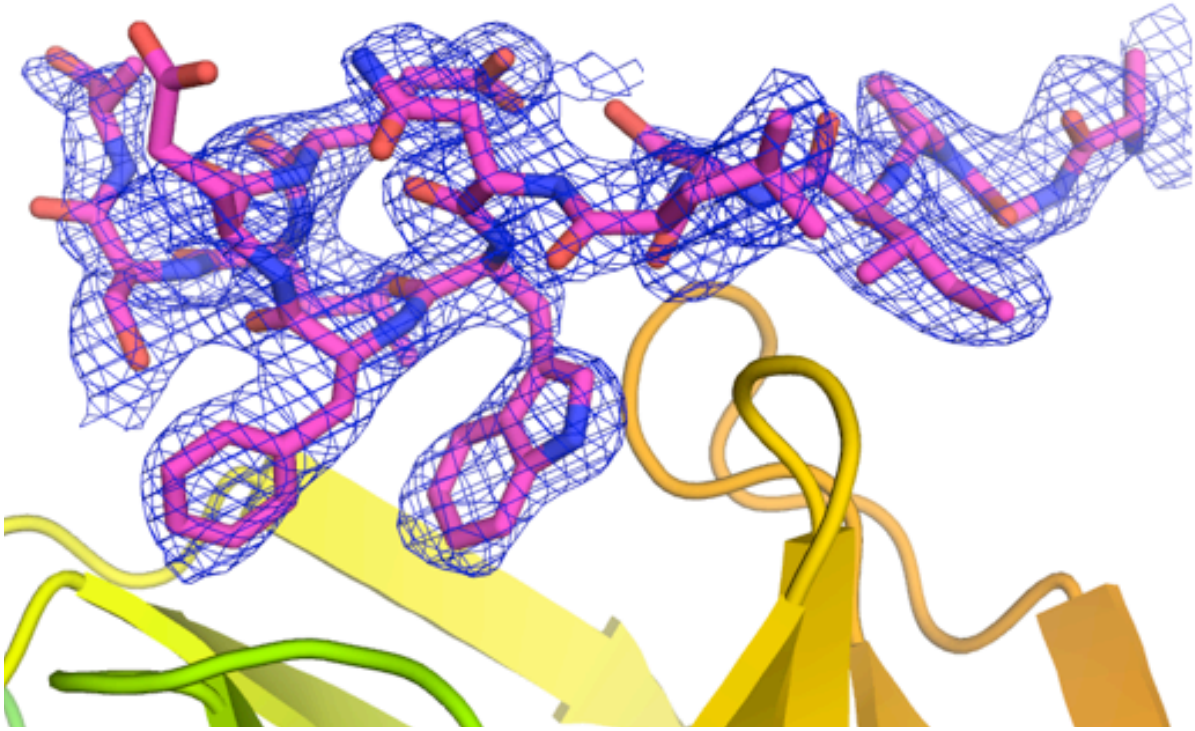
McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr D Biol Crystallogr 63: 32-41

Sheldrick GM (2008) A short history of SHELX. Acta Crystallogr A 64: 112-122

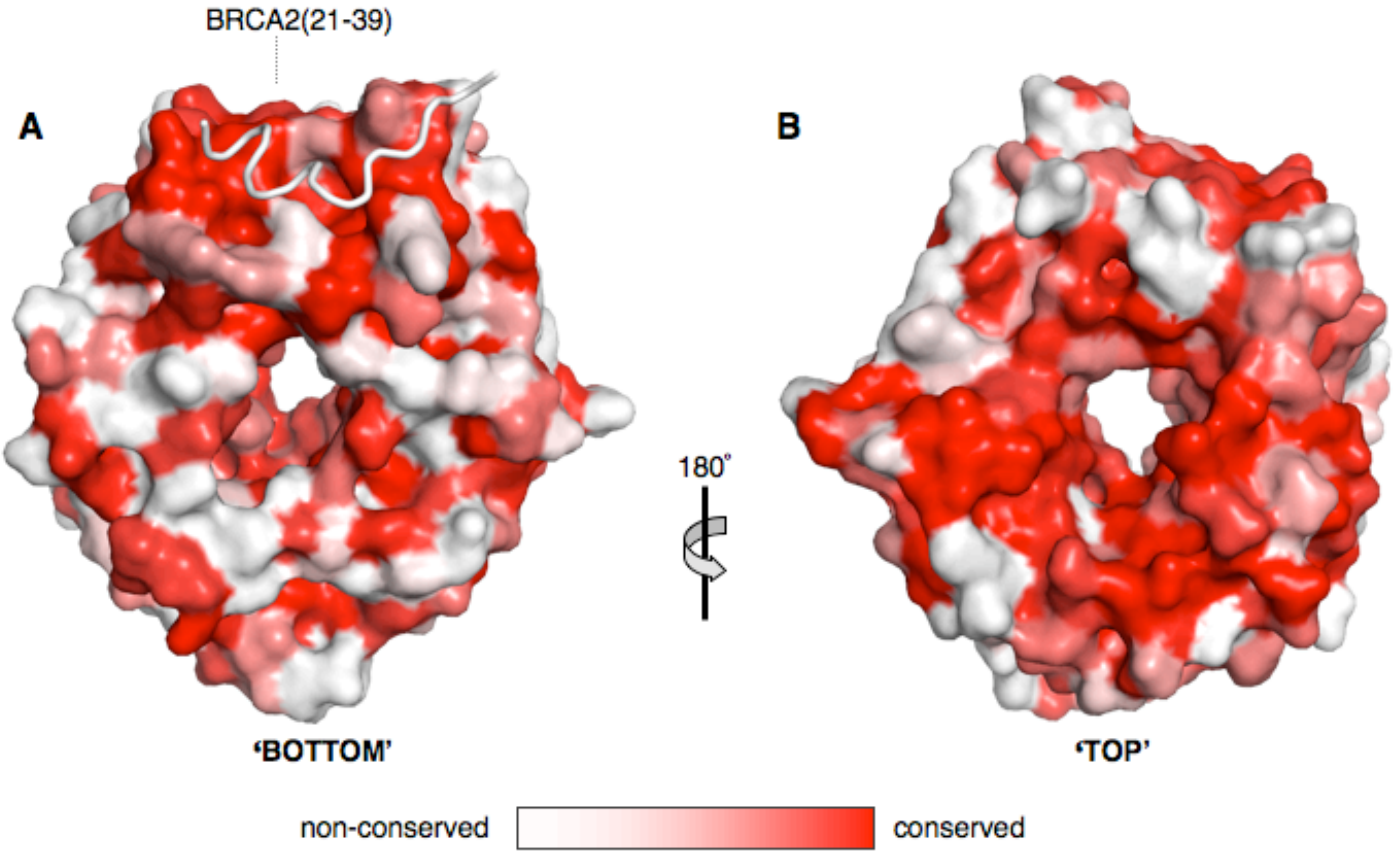
Vonrhein C, Blanc E, Roversi P, Bricogne G (2007) Automated structure solution with autoSHARP. Methods Mol Biol 364: 215-230

Xia B, Sheng Q, Nakanishi K, Ohashi A, Wu J, Christ N, Liu X, Jasin M, Couch FJ, Livingston DM (2006) Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Mol Cell* 22: 719-729

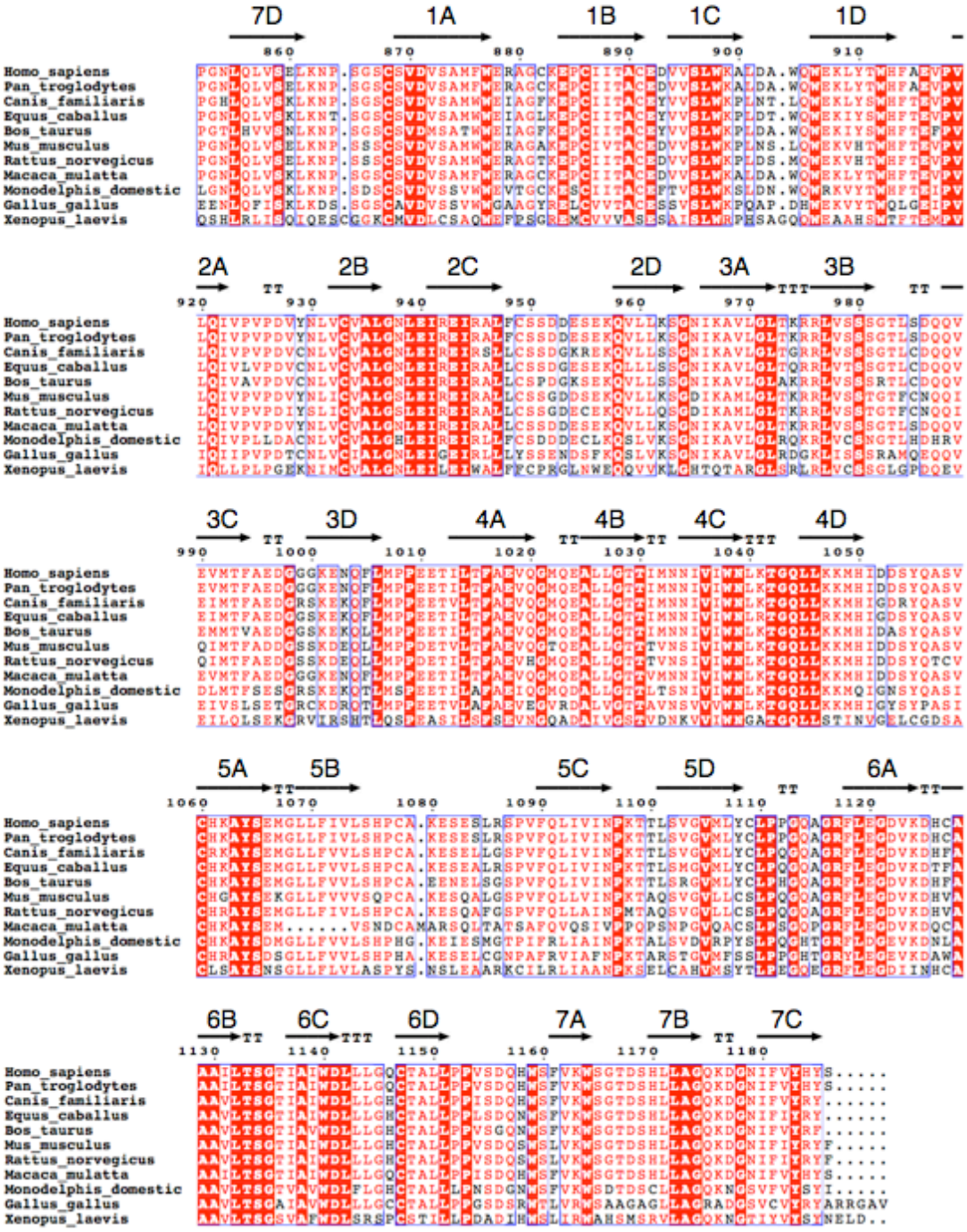
SUPPLEMENTARY FIG S1



SUPPLEMENTARY FIG S2



SUPPLEMENTARY FIG S3



SUPPLEMENTARY FIG S4