Crystal structure of human WBSCR16, an RCC1-like protein in mitochondria

Masako Koyama [1,†], **Taeko Sasaki** [1], **Narie Sasaki** [1], **and Yoshiyuki Matsuura** [1,2,*]

[1]*Division of Biological Science, and* [2]*Structural Biology Research Center, Graduate School of Science, Nagoya University, Japan*.

† Present address: Laboratory of Structural Biology, Research Institute for Science and Engineering, Waseda University, Japan.

* Correspondence to: Yoshiyuki Matsuura, Division of Biological Science, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan.
E-mail: matsuura.yoshiyuki@d.mbox.nagoya-u.ac.jp

Running title: Crystal structure of human WBSCR16

Keywords: WBSCR16; RCC1-like protein; Williams-Beuren syndrome; mitochondrial translation; mitoribosome

Supplementary Material:

Supporting Information (Figures S1-S2 and Table S1) is supplied as a PDF file.

Filename: koyama_2017ProteinSci_supporting_information.pdf

Abstract

WBSCR16 (Williams-Beuren Syndrome Chromosomal Region 16) gene is located in a large deletion region of Williams-Beuren syndrome (WBS), which is a neurodevelopmental disorder. Although the relationship between WBSCR16 and WBS remains unclear, it has been reported that WBSCR16 is a member of a functional module that regulates mitochondrial 16S rRNA abundance and intra-mitochondrial translation. WBSCR16 has RCC1 (Regulator of Chromosome Condensation 1)-like amino acid sequence repeats but the function of WBSCR16 appears to be different from that of other RCC1 superfamily members. Here we demonstrate that WBSCR16 localizes to mitochondria in HeLa cells, and report the crystal structure of WBSCR16 determined to 2.0 Å resolution using multi-wavelength anomalous diffraction. WBSCR16 adopts the seven-bladed β-propeller fold characteristic of RCC1-like proteins. A comparison of the WBSCR16 structure with that of RCC1 and other RCC1-like proteins reveals that, although many of the residues buried in the core of the β-propeller are highly conserved, the surface residues are poorly conserved and conformationally divergent.

Introduction

Mitochondrial ribosomes (mitoribosomes) are specialized for the synthesis and insertion of membrane proteins that are essential for the oxidative phosphorylation system.[1] Although recent high-resolution cryo-electron microscopy has provided rich insights into the structure and function of mitoribosomes,[2] the current understanding of the mitoribosome biogenesis pathway and the factors involved in regulation of the intra-mitochondrial translation machinery still remains far from complete.[3]

Interestingly, a recent study using a CRISPR-based screen (as a new approach to systematically identify human genes essential for oxidative phosphorylation) identified a previously unrecognized protein-RNA functional module consisting of NGRN, WBSCR16, RPUSD4, TRUB2, and FASTKD2 required for mitochondrial 16S rRNA abundance and intra-mitochondrial translation.[4] WBSCR16 (Williams-Beuren syndrome chromosomal region 16), one of the components of the newly identified functional module, is encoded in the deletion region of Williams-Beuren syndrome (WBS).[5] WBS is a neurodevelopmental disorder characterized by cardiovascular malformations, mental retardation, and a specific facial dysmorphism. WBS is caused by a heterozygous deletion of many genes on chromosome 7q11.23. Although WBSCR16 gene is not located in the critical region for WBS, a large deletion including the entire WBSCR16 gene demonstrates a severe phenotype of WBS.[6] There are no known point mutation(s) or partial deletion(s) of the WBSCR16 gene associated with WBS.

WBSCR16 has RCC1 (Regulator of Chromosome Condensation 1)-like amino acid sequence repeats.[7] RCC1 is the only known guanine-nucleotide exchange factor for the Ran GTPase,[8] and is localized to chromosomes throughout the cell cycle.[9,10]

Chromosome-anchored RCC1 generates Ran-GTP in the vicinity of chromosomes and thus provides a spatial signal that controls nucleo-cytoplasmic transport, nuclear envelope formation, and mitotic spindle assembly.[11] Here we describe the X-ray crystal structure of human WBSCR16 to be used as a starting point to understand its structure-function relationship.

Results and Discussion

Structure determination

Previous immunoprecipitation study by Arroyo *et al*. detected co-precipitation of WBSCR16 with mitochondrial proteins and mitochondrial 16S rRNA,[4] indicating that WBSCR16 is a mitochondrial protein. We analyzed subcellular localization of C-terminally FLAG-tagged WBSCR16 expressed in HeLa cells by immunostaining using anti-FLAG antibody. As expected, WBSCR16 localized specifically to mitochondria [Fig. 1(A)]. Consistently, the TargetP web server[12] predicted the N-terminal 31 amino acids of WBSCR16 as a mitochondrial targeting peptide (mTP) [Fig. 1(B)]. Because the N-terminal mitochondrial targeting sequence would be cleaved off upon import into mitochondria, we initially attempted to crystallize WBSCR16 residues 32-464 (a construct devoid of the N-terminal mitochondrial targeting sequence). Although we obtained crystals of this construct, the crystals diffracted X-rays only to 7.0 Å resolution. To obtain crystals suitable for structure determination, we employed a strategy to optimize the construct by shortening one of the predicted surface loops. We made a deletion mutant by replacing residues 146-153 (in one of the predicted surface loops of WBSCR16)[7] with two glycine residues. This mutant formed well-ordered crystals that diffracted X-rays to 2.0 Å resolution. The crystals belonged to the space group $I2$, with unit-cell parameters a = 111.0 Å, b = 54.6 Å, c = 140.5 Å, and β = 95.8°. The structure was determined using SeMet MAD (multi-wavelength anomalous diffraction) phasing method, and the final model was refined to free and working R-factor values of 21.8% and 18.0%, respectively (Table 1). There were two molecules of WBSCR16 per asymmetric unit, with a solvent content of 46.5%. The refined model consists of residues 59-462 for one molecule of WBSCR16, residues 57-82, 84-145, and 155-464 for another molecule of WBSCR16, and 598 water molecules. The two

WBSCR16 molecules in the asymmetric unit were essentially identical [a root mean square deviation (r.m.s.d.) of Cα atoms of 0.73 Å], indicating that crystal packing interactions had not significantly altered their conformation. A representative portion of the final $2F_o–F_c$ map is shown in Supporting Information Figure S1. The final model had excellent stereochemistry (Table 1). Although Y291 was a Ramachandran outlier, the electron density unambiguously showed the location of its carbonyl oxygen (Supporting Information Figure S1).

Overall structure of human WBSCR16

In the crystal structure, WBSCR16 adopts a seven-bladed β-propeller structure, as expected from the RCC1-like sequence repeats. Fig. 1(C) shows the top view of the overall structure of WBSCR16, whereas Fig. 1(D) shows its right-side view. In the top view [Fig. 1(C)], the N-terminus of the RCC1-like domain (D59) is located at the bottom surface. Each blade is composed of four antiparallel β-strands with loops between each strand. The innermost strands (strands A) of each blade run roughly parallel to the central axis of the propeller. The outer strands (strands B, C, and D) tilt as they radiate away from the center of the propeller. The inner strands (strands A, B, and C) are roughly seven residues long, whereas the outermost strands (strands D) are subdivided into two much shorter strands (strands D1 and D2). Similar to RCC1,[13] ring closure in WBSCR16 is accomplished by 2 + 2 arrangement of the β-strands of the first blade (B1): the blade B1 is composed of the strands C and D from the N-terminus of the β-propeller domain, and the strands A and B from the C-terminus of WBSCR16. The propeller has pseudo seven-fold symmetry, as is evident in the overlay of the structures of the seven blades [Fig. 1(E)] and also in the similarity of the amino acid sequences of the seven blades [Fig. 1(F)].

Comparison of the human WBSCR16 structure with that of other proteins

A similarity search using the DALI web server[14] revealed that the structure of WBSCR16 is more closely related to that of RCC1 and RCC1-like proteins than to any other β-propeller structures such as that of WD40-repeat[15] proteins or FG-GAP (phenylalanyl-glycyl-glycyl-alanyl-prolyl)-repeat[16] proteins (Supporting Information Table S1). We used the ENDscript web server[17] to analyze both sequences and crystal structures of RCC1 and RCC1-like proteins (with the same 2 + 2 arrangement of β-strands in the first blade) in detail. Fig. 2(A) shows the comparison of human WBSCR16 with human RCC1 (PDB code, 1I2M),[18] Prp20p (yeast RCC1; PDB code, 3OF7),[19] human RPGR (PDB code, 4JHP),[20] and the third RCC1-like domain of human HERC1 (PDB code, 4O2W; unpublished). As shown in Fig. 2(A), many of the residues invariant or highly conserved among the WBSCR16 repeats [boxed residues in Fig. 1(F)], are also highly conserved in RCC1 and other RCC1-like proteins. Almost all of these highly conserved residues are buried in the core of the β-propeller [Fig. 2(A, B)], and so are probably important for structural integrity of the β-propeller. On the other hand, almost all of the residues exposed on the surface of the β-propeller, especially in the C-D loops and D-A loops, are poorly conserved and conformationally divergent [Fig. 2(A, B)], suggesting that many of the surface residues of WBSCR16 are potentially involved in WBSCR16-specific function.

Comparison of human WBSCR16 with its homologs

To try to identify the surface residues that are potentially important for specific interactions of the RCC1-like domain of WBSCR16 homologs, we analyzed the sequence conservation of WBSCR16 using the ConSurf web server.[21] For this analysis, we built a model of the loop (deleted for crystallization) in the blade B2 using the

program Modeller[22] [as marked by a dashed circle in Supporting Information Figure S2(A)], and the surface conservation profile of WBSCR16 (together with surface electrostatic potential of human WBSCR16) is shown in Supporting Information Figure S2(B, C). There are a number of highly conserved residues on the surface of WBSCR16 that might be involved in interactions with its binding partners. Notably, there is a patch of highly conserved residues on the top surface of the β-propeller that has predominantly negative potential. The bottom surface of the β-propeller, on the other hand, has predominantly positive electrostatic potential. Also noteworthy are the basic loops of the blades B1 and B2 [as marked by dashed circles in Supporting Information Figure S2(B)] protruding from the β-propeller. Because WBSCR16 is involved in the regulation of mitochondrial 16S rRNA abundance and intra-mitochondrial translation,[4] the basic residues exposed on the surface are intriguing as potential binding sites for nucleic acids such as rRNA. However, the sequence conservation of the surface basic patches of WBSCR16 is not quite high. The functional significance of the exposed residues is unclear at present. Identification of binding partners of WBSCR16 and extensive functional analyses will be required to elucidate the structure-function relationship of WBSCR16.

In summary, we have used X-ray crystallography and established the structure of human WBSCR16, which demonstrates that WBSCR16 has seven-bladed β-propeller fold (the RCC1 fold) with unique surface features. Although the function of WBSCR16 is only beginning to be understood, the availability of this structure will likely have significance for understanding the poorly characterized molecular mechanism underlying intra-mitochondrial translation, which is paramount to mitochondrial respiration and thus to cell viability, growth, and differentiation.

Materials and Methods

Cell culture

The HeLa cell line expressing mitochondria-targeted DsRed (HeLa-Su9)[23] was cultured in Dulbecco's Modified Eagle Medium (D-MEM; Wako, Osaka, Japan) containing 10% fetal bovine serum (FBS) and 1 μg/ml puromycin.

Analysis of WBSCR16 localization

To analyze the localization of WBSCR16, a human *WBSCR16* fragment was amplified by PCR from human cDNA using the primers 5'-AAGCTTCACCATGGCGCTGGTGGTGGCGTTGGTG-3' and 5'-GGTACCGTGATGAATGACTTGGCCAGGGTCACC-3', and cloned into p3xFLAG-CMV-14 expression vector (SIGMA). For immunostaining, HeLa-Su9 cells were plated onto an eight-well cover glass chamber (AGC Techno Glass Co., LTD., 2×10^4 cells per well) and cultured in D-MEM containing 10% FBS. After 1 day, the cells were transfected with the expression plasmid for WBSCR16-FLAG (200 ng per well) using jetPRIME (Polyplus-transfection), in accordance with the manufacturer's protocol. Two days after transfection, fixation of the cells was performed using 8% (wt/vol) paraformaldehyde in PBS (137 mM NaCl, 3 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4, pH 7) at 37 °C for 3 min. Then, the cells were permeabilized with 0.5% (vol/vol) Triton X-100 in PBS at room temperature for 10 min. After fixation and permeabilization, cells were incubated with rabbit antibodies against FLAG (F7425; Sigma) in PBS with 1% (wt/vol) skimmed milk powder. The primary antibody was detected with goat Alexa 488-conjugated anti-rabbit (A-11008, Thermo Fisher Scientific). After immunostaining, cells were incubated with 10 μg/ml Hoechst 33342 in

PBS at room temperature for 10 min. The fluorescent images of immunostained cells were acquired using a spinning disk confocal system (CellVoyager CV1000; Yokogawa Electric). To create maximum-intensity projection images and adjust the brightness and contrast, we used CV1000 software (Yokogawa Electric) and Fiji (http://fiji.sc/).

Protein expression and purification

To construct a plasmid for expressing human WBSCR16 (UniProt code, Q96I51) with N-terminal GST tag in *E. coli* cells, the cDNA encoding WBSCR16 (residues 32-464, in which residues 146-153 were replaced with two glycine residues) was subcloned into pGEX-TEV[24] bacterial expression vector. The construct was verified by DNA sequencing. GST-WBSCR16 was expressed in *E. coli* strain BL21-CodonPlus(DE3)RIL (Stratagene) at 18 °C in 2× TY medium. After harvesting, the pellet was frozen in liquid nitrogen and stored at −20 °C until needed.

For purification, the frozen pellet of the cells expressing GST-WBSCR16 was thawed at room temperature and resuspended in buffer A [30 mM Tris-HCl (pH7.5), 0.5 M NaCl, 7 mM 2-mercaptoethanol, 1 mM PMSF] and lysed by sonication on ice. All subsequent purification steps were performed at 4 °C. Tween20 was added to the clarified lysate to a final concentration of 0.05%. After incubating the clarified lysate with Glutathione-sepharose 4B resin (GE Healthcare) for 3 h, the resin was washed with buffer B [30 mM Tris-HCl (pH7.5), 0.3 M NaCl, 2 mM 2-mercaptoethanol, 0.05% Tween20]. The GST-tag was removed with His-TEV protease (0.07 mg/ml) overnight in buffer B containing 0.2 mM AEBSF. WBSCR16 released from the resin was finally purified by gel filtration over Superdex200 (GE Healthcare) in buffer C [10 mM Tris-HCl (pH7.5), 0.25 M NaCl, 2 mM 2-mercaptoethanol]. Fractions containing >95% pure WBSCR16, as assessed by SDS-PAGE, were pooled and concentrated using a

Millipore concentrator (Mol. wt. cutoff 10,000).

The selenomethionine (SeMet)-substituted human WBSCR16 (residues 32-464, in which residues 146-153 were replaced with two glycine residues) was expressed in the same strain, BL21-CodonPlus(DE3)RIL (Stratagene), which is not auxotrophic for methionine. Methionine biosynthesis was inhibited by growth conditions as described.[25] A preculture, grown in 2× TY medium at 37 °C was inoculated into minimal medium containing 1× M9 supplemented with 20 μg/ml thiamine, 20 μg/ml biotin and 50 μg/ml ampicillin and grown overnight at 28 °C to an OD_{600} of 0.7. A mixture of L-amino acids was added as solids (per liter of culture: 50 mg of SeMet (Wako), 50 mg of leucine, isoleucine, valine, and 100 mg of lysine, threonine, phenylalanine). After 15 min, protein expression was induced by the addition of 0.8 mM IPTG and the culture was grown overnight at 18 °C. SeMet-substituted WBSCR16 was purified as described for the non-substituted WBSCR16.

Crystallization, data collection, and structure determination

Crystals of WBSCR16 were obtained using hanging drop vapor diffusion method at 20 °C. Both native and SeMet-substituted WBSCR16 crystals were grown by mixing equal volumes (1.5 μl each) of the protein solution (10 mg/ml WBSCR16 in buffer C) with the precipitant solution consisting of 0.1 M HEPES-NaOH (pH 7.0), 15% PEG20000. Rod-shaped crystals grew to a maximum dimension of 0.02 × 0.02 × 1.0 mm in one month. Crystals were serially transferred to 0.1 M HEPES-NaOH (pH 7.0), 15% PEG20000, 20% glycerol in four steps and flash-cooled in liquid nitrogen. Diffraction datasets for the native WBSCR16 crystals were collected at SPring-8 beamline BL26B2 at 100 K. Two-wavelength MAD diffraction datasets at the Se absorption edge for the SeMet-substituted WBSCR16 crystals were collected at Photon

Factory beamline BL-1A at 100 K. The SeMet-substituted crystals were isomorphous with the native crystals.

Diffraction data were indexed and integrated using MOSFLM and further processed using CCP4 programs.[26] The structure was determined by SeMet MAD phasing using CRANK.[27] Twelve Se sites (of twelve SeMet residues in the two molecules of human WBSCR16 in the asymmetric unit) were located by program AFRO/CRUNCH2 and refined using BP3.[27] Following density modification using SOLOMON,[28] a large fraction of WBSCR16 was automatically built using BUCCANEER.[29] Iterative cycles of manual rebuilding using COOT[30] and refinement using PHENIX[31] against the native dataset to 2.0 Å resolution yielded a final model with R_{work} = 18.0% (R_{free} = 21.8%). The final model geometry was validated by MolProbity.[32] Data collection and refinement statistics are summarized in Table 1. Structural figures were produced using CCP4MG[33] and PyMOL (DeLano Scientific).

ConSurf analysis of evolutionary conservation

The evolutionary conservation profile of WBSCR16 was estimated using ConSurf.[21] A CSI-BLAST search for homologs of the human WBSCR16 sequence was performed against the UNIREF90 database with an E-value cutoff of 0.00001, minimal % ID of 35% for homologs and maximal % ID of 95% between sequences. A total of 150 homologous sequences were retrieved and multiply aligned using MAFFT. Calculation of position-specific conservation scores was performed using the Bayesian method.

Accession number

The coordinates and structure factors of human WBSCR16 have been deposited in the

PDB with accession code 5XGS.

Supplementary Material

Additional Supporting Information (Figures S1-S2 and Table S1) may be found in the online version of this article.
Filename: koyama_2017ProteinSci_supporting_information.pdf

Acknowledgements

We thank Hidemi Hirano for technical assistance and discussion. We also thank the staff of SPring-8 and Photon Factory for assistance during data collection and Prof. N. Ishihara in Kurume University for providing the HeLa-Su9 cell line.

Conflict of Interest Statement

The authors have no conflict of interest to declare.

References

1. Ott M, Amunts A, Brown A (2016) Organization and Regulation of Mitochondrial Protein Synthesis. Annu Rev Biochem 85:77-101.
2. Greber BJ, Ban N (2016) Structure and Function of the Mitochondrial Ribosome. Annu Rev Biochem 85:103-132.
3. De Silva D, Tu YT, Amunts A, Fontanesi F, Barrientos A (2015) Mitochondrial ribosome assembly in health and disease. Cell Cycle 14:2226-2250.
4. Arroyo JD, Jourdain AA, Calvo SE, Ballarano CA, Doench JG, Root DE, Mootha VK (2016) A Genome-wide CRISPR Death Screen Identifies Genes Essential for Oxidative Phosphorylation. Cell Metab 24:875-885.
5. Merla G, Ucla C, Guipponi M, Reymond A (2002) Identification of additional transcripts in the Williams-Beuren syndrome critical region. Human Genetics 110:429-438.
6. Schubert C, Laccone F (2006) Williams-Beuren syndrome: determination of deletion size using quantitative real-time PCR. Int J Mol Med 18:799-806.
7. Hadjebi O, Casas-Terradellas E, Garcia-Gonzalo FR, Rosa JL (2008) The RCC1 superfamily: from genes, to function, to disease. Biochim Biophys Acta 1783:1467-1479.
8. Bischoff FR, Ponstingl H (1991) Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. Nature 354:80-82.
9. Ohtsubo M, Okazaki H, Nishimoto T (1989) The RCC1 protein, a regulator for the onset of chromosome condensation locates in the nucleus and binds to DNA. J Cell Biol 109:1389-1397.
10. Moore W, Zhang C, Clarke PR (2002) Targeting of RCC1 to chromosomes is required for proper mitotic spindle assembly in human cells. Curr Biol 12:1442-1447.
11. Dasso M (2002) The Ran GTPase: theme and variations. Curr Biol 12:R502-508.
12. Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. Nature Protocols 2:953-971.
13. Renault L, Nassar N, Vetter I, Becker J, Klebe C, Roth M, Wittinghofer A (1998) The 1.7 A crystal structure of the regulator of chromosome condensation (RCC1) reveals a seven-bladed propeller. Nature 392:97-101.

14. Holm L, Rosenstrom P (2010) Dali server: conservation mapping in 3D. Nucleic Acids Res 38:W545-549.
15. Xu C, Min J (2011) Structure and function of WD40 domain proteins. Protein Cell 2:202-214.
16. Springer TA (1997) Folding of the N-terminal, ligand-binding region of integrin alpha-subunits into a beta-propeller domain. Proc Natl Acad Sci U S A 94:65-72.
17. Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res 42:W320-324.
18. Renault L, Kuhlmann J, Henkel A, Wittinghofer A (2001) Structural basis for guanine nucleotide exchange on Ran by the regulator of chromosome condensation (RCC1). Cell 105:245-255.
19. Wu F, Liu Y, Zhu Z, Huang H, Ding B, Wu J, Shi Y (2011) The 1.9A crystal structure of Prp20p from Saccharomyces cerevisiae and its binding properties to Gsp1p and histones. J Struct Biol 174:213-222.
20. Watzlich D, Vetter I, Gotthardt K, Miertzschke M, Chen YX, Wittinghofer A, Ismail S (2013) The interplay between RPGR, PDEdelta and Arl2/3 regulate the ciliary targeting of farnesylated cargo. EMBO Rep 14:465-472.
21. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res 38:W529-533.
22. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779-815.
23. Taguchi N, Ishihara N, Jofuku A, Oka T, Mihara K (2007) Mitotic phosphorylation of dynamin-related GTPase Drp1 participates in mitochondrial fission. J Biol Chem 282:11521-11529.
24. Matsuura Y, Stewart M (2004) Structural basis for the assembly of a nuclear export complex. Nature 432:872-877.
25. Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J Mol Biol 229:105-124.
26. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS (2011) Overview of the

CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr 67:235-242.

27. Pannu NS, Waterreus WJ, Skubak P, Sikharulidze I, Abrahams JP, de Graaff RA (2011) Recent advances in the CRANK software suite for experimental phasing. Acta Crystallogr D Biol Crystallogr 67:331-337.
28. Abrahams JP, Leslie AG (1996) Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallogr D Biol Crystallogr 52:30-42.
29. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr D Biol Crystallogr 62:1002-1011.
30. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60:2126-2132.
31. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66:213-221.
32. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66:12-21.
33. McNicholas S, Potterton E, Wilson KS, Noble ME (2011) Presenting your structures: the CCP4mg molecular-graphics software. Acta Crystallogr D Biol Crystallogr 67:386-394.

Figure legends

Figure 1. Subcellular localization and crystal structure of human WBSCR16. (A) Localization of WBSCR16 to mitochondria in HeLa-Su9 cells. WBSCR16-FLAG was transiently expressed in HeLa-Su9 cells expressing mitochondria-targeted DsRed (DsRed-Mito). Cells were fixed and stained with anti-FLAG antibody (green) and Hoechst 33342 (blue). Scale bar, 10 μm. (B) Schematic representation of full-length WBSCR16. mTP, mitochondrial targeting peptide. (C), (D) Stereoview of a ribbon drawing of the crystal structure of human WBSCR16 (PDB code, 5XGS) in (C) top view and (D) a side view. The seven blades are numbered B1-B7 along the amino acid sequence. The Cα of D59, the N-terminus of the RCC1-like domain of WBSCR16, is shown as a black sphere. (E) Superposition of the seven blades. Each blade is colored as in (C) and (D). (F) Structure-based sequence alignment of the seven blades. Highly conserved regions are boxed. The names of basic residues in C-D loops of the blades B1 and B2 are written in blue. The names of residues in β-strands are written in red. The positions of β-strands are indicated by red arrows. Sequence logo, generated by the program WebLogo (http://weblogo.berkeley.edu/), is shown below the sequence alignment. The x-axis of the sequence logo shows the amino acid position and the y-axis represents the information content measured in bits. The overall height of each stack of letters indicates the sequence conservation at that position, and the height of a letter within the stack indicates the relative frequency of the amino acid.

Figure 2. Comparison of human WBSCR16 with RCC1 and other RCC1-like proteins. (A) Multiple sequence alignment of human WBSCR16, human RCC1, yeast Prp20p, human RPGR, and the third RCC1-like domain of human HERC1. Similarity scores were calculated using the ENDscript web server, according to a matrix based on physicochemical properties. Residue names are written in black if score is below 0.7

(low similarity); they are in red on a yellow background if score is in the range 0.7-1 (high similarity); they are in white on a red background in case of strict identity. The positions of β-strands of WBSCR16 are indicated by arrows above the sequence. Solvent accessibility of WBSCR16 residues is rendered by a bar below the sequence (blue is accessible, cyan is intermediate, white is buried) (B) Stereoview of a sausage drawing of WBSCR16. For this drawing, the structures of human RCC1 (PDB code, 1I2M, chain B), yeast Prp20p (PDB code, 3OF7, chain A), human RPGR (PDB code, 4JHP, chain C), and the third RCC1-like domain of human HERC1 (PDB code, 4O2W, chain A) were superposed onto the structure of human WBSCR16, and the radius of the sausage is proportional to the mean r.m.s.d. per residue between Cα pairs. To visualize sequence conservation, the sausage is color ramped from white [similarity score calculated in (A) below 0.7] to red (strict identity).

Table 1. Crystallographic statistics.

Crystal	Native	SeMet	
Data collection			
Space group	*I*2	*I*2	
Unit cell dimensions			
a, *b*, *c* (Å)	111.0, 54.6, 140.5	111.4, 54.2, 140.6	
α, *β*, *γ* (degree)	90, 95.8, 90	90, 95.9, 90	
X-ray source	SPring-8 BL26B2	Photon Factory BL-1A	
		Peak	*Inflection*
Wavelength (Å)	1.00000	0.97850	0.97900
Resolution range (Å)[a]	34.9-2.0 (2.05-2.00)	26.0-2.4 (2.49-2.40)	27.2-2.4 (2.49-2.40)
Total observations[a]	214273	208700	208754
Unique reflections[a]	56417	32982	33033
Completeness (%)[a]	99.1 (98.3)	99.3 (97.7)	99.4 (98.7)
R_{merge}[a]	0.17 (0.84)	0.09 (0.27)	0.07 (0.18)
Mean $I/\sigma(I)$[a]	6.5 (1.8)	15.4 (6.9)	18.1 (9.0)
Mean *I* half-set correlation CC(1/2)[a]	0.987 (0.531)	0.996 (0.963)	0.997 (0.982)
Multiplicity[a]	3.8 (3.8)	3.8 (3.1)	3.8 (3.1)
Refinement			
Resolution range (Å)[a]	34.9-2.0		
No. of reflections	56397		
R_{work} (%)[a]	18.0 (27.7)		
R_{free} (%)[a]	21.8 (31.6)		
No. of atoms			
Protein	6016		
Water	598		
No. of amino acids	796		
Mean B-factor (Å^2)			
Protein	19.6		
Water	24.3		
RMSD from ideality			
Bond lengths (Å)	0.002		
Bond angles (degree)	0.531		
Protein geometry[b]			
Rotamer outliers (%)	0		
Ramachandran favored (%)	98.73		
Ramachandran outliers (%)	0.25		
Cβ deviations > 0.25 Å (%)	0		
MolProbity score (percentile)	1.01 (100)		
PDB code	5XGS		

[a] Values in parentheses are for the highest-resolution shell.

[b] MolProbity was used to analyze the structures.

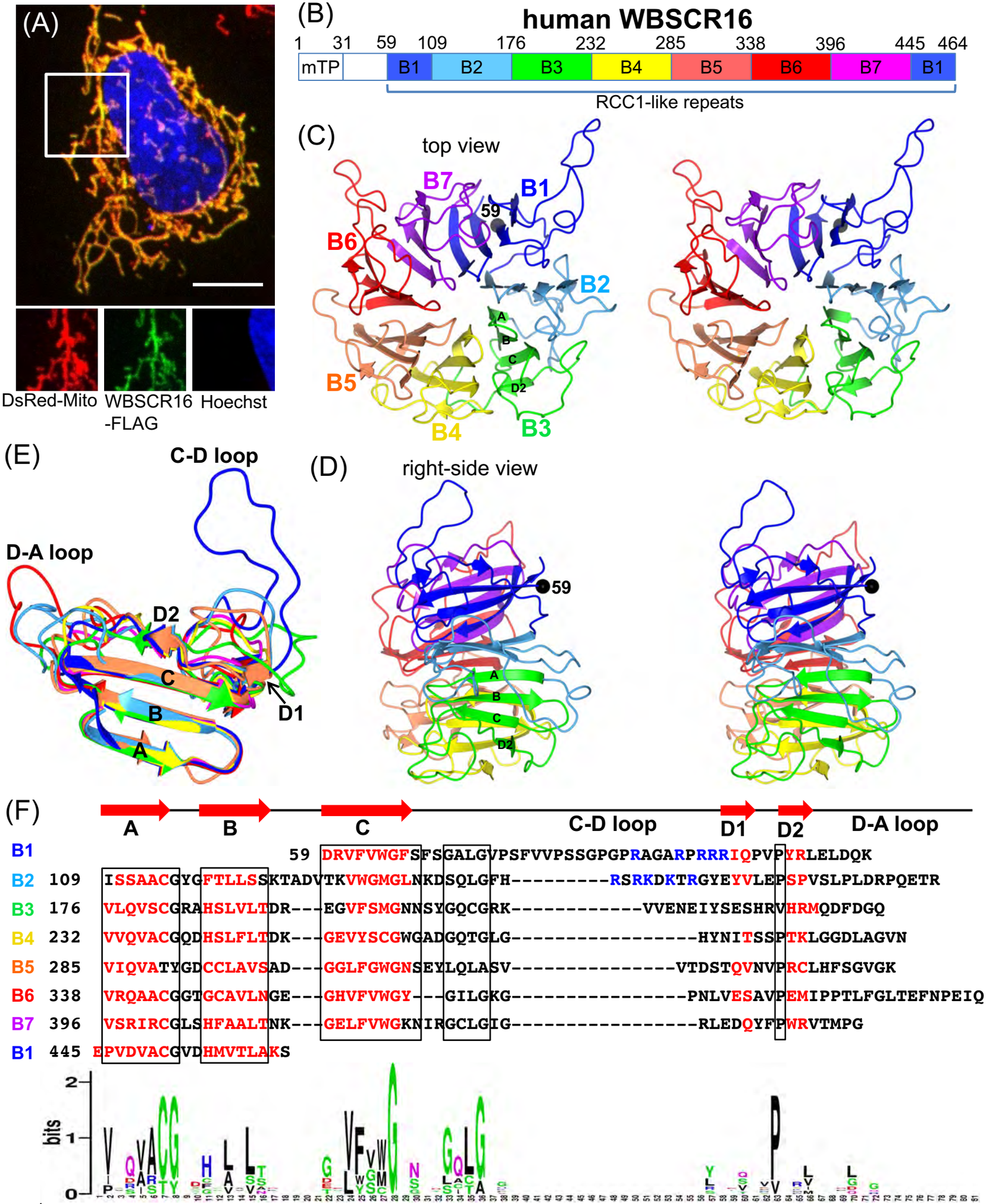
(A)
DsRed-Mito
WBSCR16 -FLAG
Hoechst
(B)
human WBSCR16
1 31 59 109 176 232 285 338 396 445 464
mTP
B1
B2
B3
B4
B5
B6
B7
B1
RCC1-like repeats
(C)
top view
B7
B1
59
B6
B2
A
B
C
D2
B5
B4
B3
(E)
C-D loop
D-A loop
D2
C
B
A
D1
(D)
right-side view
59
A
B
C
D2
(F)
A
B
C
C-D loop
D1
D2
D-A loop
B1
B2 109 ISSAACGYGFTLLSSKTADVTKVWGMGLNKDSQLGFH---------RSRKDKTRGYEYVLEPSPVSLPLDRPQETR
B3 176 VLQVSCGRAHSLVLTDR---EGVFSMGNNSYGQCGRK-----------VVENEIYSESHRVHRMQDFDGQ
B4 232 VVQVACGQDHSLFLTDK---GEVYSCGWGADGQTGLG----------------HYNITSSPTKLGGDLAGVN
B5 285 VIQVATYGDCCLAVSAD---GGLFGWGNSEYLQLASV--------------VTDSTQVNVPRCLHFSGVGK
B6 338 VRQAACGGTGCAVLNGE---GHVFVWGY---GILGKG---------------PNLVESAVPEMIPPTLFGLTEFNPEIQ
B7 396 VSRIRCGLSHFAALTNK---GELFVWGKNIRGCLGIG----------------RLEDQYFPWRVTMPG
B1 445 EPVDVACGVDHMVTLAKS
bits
2
1
0

Figure 1

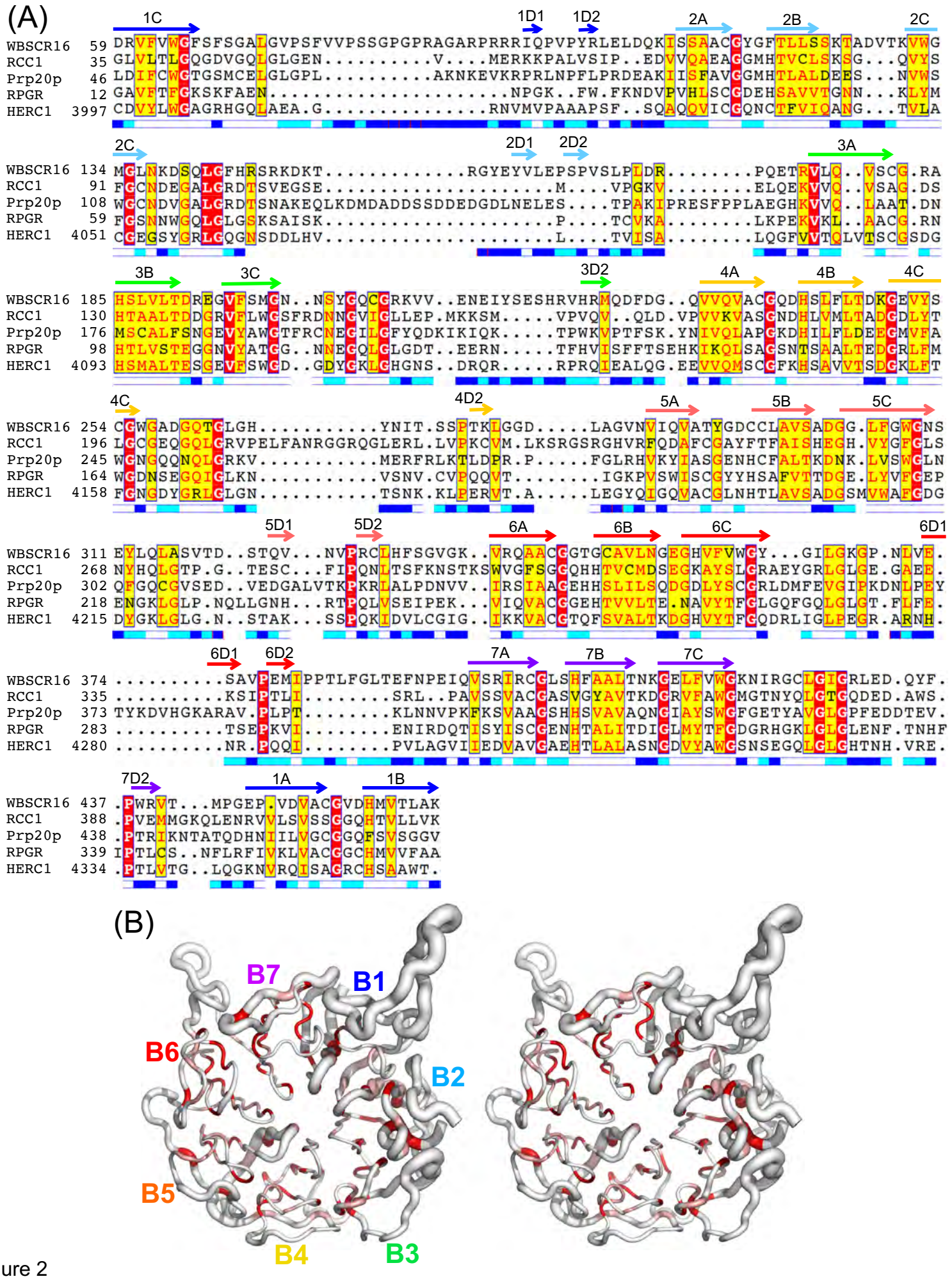
(A)
1C
1D1
1D2
2A
2B
2C
WBSCR16 59 DRVFVWGFSFSGALGVPSFVVPSSGPGPRAGARPRRRIQPVPYRLELDQKISSAACGYGFTLLSSKTADVTKVWG
RCC1 35 GLVLTLGQGDVGQLGLGEN.....V....MERKKPALVSIP..EDVVQAEAGGMHTVCLSKSG...QVYS
Prp20p 46 LDIFCWGTGSMCELGLGPL.....AKNKEVKRPRLNPFLPRDEAKIISFAVGGMHTLALDEES...NVWS
RPGR 12 GAVFTFGKSKFAEN.....NPGK..FW.FKNDVPVHLSCGDEHSAVVTGNN...KLYM
HERC1 3997 CDVYLWGAGRHGQLAEA.G.....RNVMVPAAAPSF..SQAQQVICGQNCTFVIQANG...TVLA
2C
2D1
2D2
3A
WBSCR16 134 MGLNKDSQLGFHRSRKDKT.....RGYEYVLEPSPVSLPLDR.....PQETRVLQ..VSCG.RA
RCC1 91 FGCNDEGALGRDTSVEGSE.....M....VPGKV.....ELQEKVVQ..VSAG.DS
Prp20p 108 WGCNDVGALGRDTSNAKEQLKDMDADDSSDDEDGDLNELES....TPAKIPRESFPPLAEGHKVVQ..LAAT.DN
RPGR 59 FGSNNWGQLGLGSKSAISK.....P....TCVKA.....LKPEKVKL..AACG.RN
HERC1 4051 CGEGSYGRLGQGNSDDLHV.....L....TVISA.....LQGFVVTQLVTSCGSDG
3B
3C
3D2
4A
4B
4C
WBSCR16 185 HSLVLTDREGVFSMGN..NSYGQCGRKVV..ENEIYSESHRVHRMQDFDG..QVVQVACGQDHSLFLTDKGEVYS
RCC1 130 HTAALTDDGRVFLWGSFRDNNGVIGLLEP.MKKSM.....VPVQV..QLD.VPVVKVASGNDHLVMLTADGDLYT
Prp20p 176 MSCALFSNGEVYAWGTFRCNEGILGFYQDKIKIQK.....TPWKVPTFSK.YNIVQLAPGKDHILFLDEEGMVFA
RPGR 98 HTLVSTEGGNVYATGG..NNEGQLGLGDT..EERN.....TFHVISFFTSEHKIKQLSAGSNTSAALTEDGRLFM
HERC1 4093 HSMALTESGEVFSWGD..GDYGKLGHGNS..DRQR.....RPRQIEALQG.EEVVQMSCGFKHSAVVTSDGKLFT
4C
4D2
5A
5B
5C
WBSCR16 254 CGWGADGQTGLGH.....YNIT.SSPTKLGGD.....LAGVNVIQVATYGDCCLAVSADGG.LFGWGNS
RCC1 196 LGCGEQGQLGRVPELFANRGGRQGLERL.LVPKCVM.LKSRGSRGHVRFQDAFCGAYFTFAISHEGH.VYGFGLS
Prp20p 245 WGNGQQNQLGRKV.....MERFRLKTLDPR.P.....FGLRHVKYIASGENHCFALTKDNK.LVSWGLN
RPGR 164 WGDNSEGQIGLKN.....VSNV.CVPQQVT.....IGKPVSWISCGYYHSAFVTTDGE.LYVFGEP
HERC1 4158 FGNGDYGRLGLGN.....TSNK.KLPERVT.A.....LEGYQIGQVACGLNHTLAVSADGSMVWAFGDG
5D1
5D2
6A
6B
6C
6D1
WBSCR16 311 EYLQLASVTD..STQV...NVPRCLHFSGVGK..VRQAACGGTGCAVLNGEGHVFVWGY...GILGKGP.NLVE.
RCC1 268 NYHQLGTP.G..TESC...FIPQNLTSFKNSTKSWVGFSGGQHHTVCMDSEGKAYSLGRAEYGRLGLGE.GAEE.
Prp20p 302 QFGQCGVSED..VEDGALVTKPKRLALPDNVV..IRSIAAGEHHSLILSQDGDLYSCGRLDMFEVGIPKDNLPEY
RPGR 218 ENGKLGLP.NQLLGNH...RTPQLVSEIPEK...VIQVACGGEHTVVLTE.NAVYTFGLGQFGQLGLGT.FLFE.
HERC1 4215 DYGKLGLG.N..STAK...SSPQKIDVLCGIG..IKKVACGTQFSVALTKDGHVYTFGQDRLIGLPEGR.ARNH.
6D1
6D2
7A
7B
7C
WBSCR16 374 .....SAVPEMIPPTLFGLTEFNPEIQVSRIRCGLSHFAALTNKGELFVWGKNIRGCLGIGRLED.QYF.
RCC1 335 .....KSIPTLI.....SRL..PAVSSVACGASVGYAVTKDGRVFAWGMGTNYQLGTGQDED.AWS.
Prp20p 373 TYKDVHGKARAV.PLPT.....KLNNVPKFKSVAAGSHHSVAVAQNGIAYSWGFGETYAVGLGPFEDDTEV.
RPGR 283 .....TSEPKVI.....ENIRDQTISYISCGENHTALITDIGLMYTFGDGRHGKLGLGLENF.TNHF
HERC1 4280 .....NR.PQQI.....PVLAGVIIEDVAVGAEHTLALASNGDVYAWGSNSEGQLGLGHTNH.VRE.
7D2
1A
1B
WBSCR16 437 .PWRVT...MPGEP.VDVACGVDHMVTLAK
RCC1 388 .PVEMMGKQLENRVLSVSSGGQHTVLLVK
Prp20p 438 .PTRIKNTATQDHNIILVGCGGQFSVSGGV
RPGR 339 IPTLCS..NFLRFIVKLVACGGCHMVVFAA
HERC1 4334 .PTLVTG..LQGKNVRQISAGRCHSAAWT.
(B)
B7
B1
B6
B2
B5
B4
B3

Figure 2

SUPPORTING INFORMATION

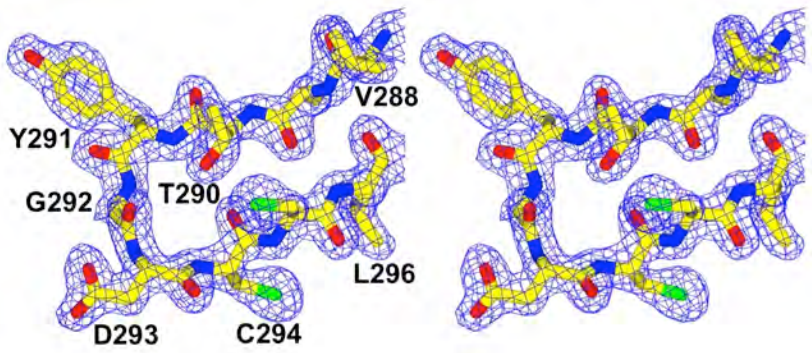

Figure S1. Stereoview of a piece of $2F_O-F_C$ electron density map. The map was contoured at 1.5σ and corresponds to the electron densities for residues 288-296.

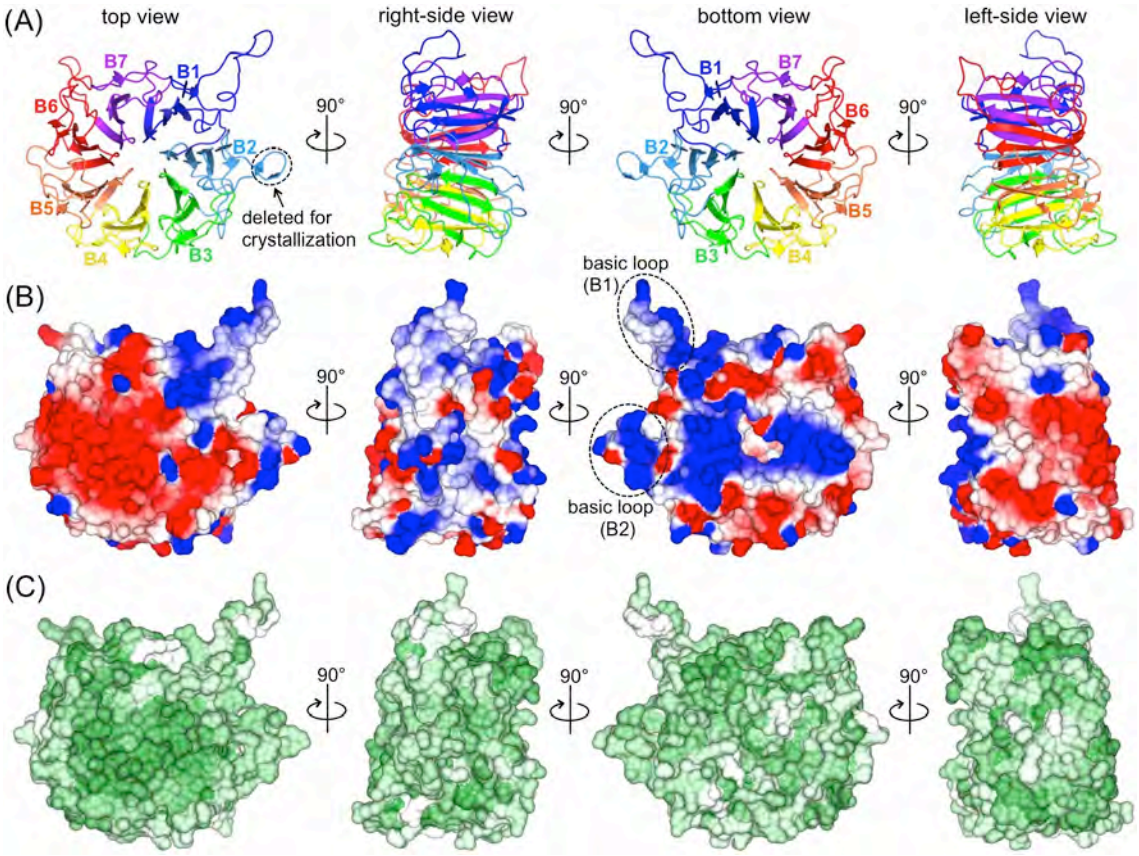

Figure S2. Surface properties of human WBSCR16. For this drawing, a model of a surface loop of the blade B2 (deleted for crystallization; marked by a dashed circle in (A)) was built using the program Modeller. (A) Ribbon representation of human WBSCR16. Each blade is colored as in Fig. 1(C). (B) Electrostatic potential of human WBSCR16. The surface potential was calculated using CCP4MG. Molecular surface is colored according to electrostatic potential, shaded from blue (potential +150 mV) through white (potential 0 mV) to red (potential −150 mV). Each image is in the same orientation as in (A). The protruding surface loops (enriched in basic residues) in the blades B1 and B2 are marked by dashed circles. (C) Sequence conservation of WBSCR16. Conservation scores were calculated by the Bayesian method using the ConSurf web server (http://consurf.tau.ac.il/). The surface representation is colored according to the conservation scores, shaded from white (the least conserved residues) to green (the most conserved residues).

Table S1. Structural homologs of WBSCR16.

PDB ID	*Z*-score	r.m.s.d (Å)	Seq. ID to WBSCR16 (%)	Protein Name	Category
1I2M	46.5	2.1	28	RCC1	RCC1
4JHN	45.9	2.1	28	RPGR	RCC1-like
4O2W	45.5	2.3	27	HERC1	RCC1-like
3KCI	45.0	2.3	26	HERC2	RCC1-like
4DNV	42.6	2.3	28	UVR8	RCC1-like
3OF7	40.4	2.5	23	Prp20p	RCC1
4D4P	34.6	2.4	18	Ats1p	RCC1-like
3QI0	29.0	2.1	27	BLIP-II	RCC1-like
2EBS	17.9	3.1	11	OXG-RCBH	WD40
5E6S	17.4	4.2	8	Integrin alpha-L	FG-GAP
5JPQ	17.3	4.0	12	Rrp9	WD40
2QC5	17.1	3.8	6	Streptogramin B lactonase	WD40
2Z2O	17.1	3.8	5	Virginiamycin B lyase	WD40
3A0F	17.0	3.2	7	Xyloglucanase	WD40
5ES4	16.9	4.0	10	Integrin alpha-X	FG-GAP
4H5J	16.7	3.7	6	Sec12p	WD40
3V4V	16.5	4.1	10	Integrin alpha-4	FG-GAP
2GOP	16.4	3.1	7	Trilobed protease	WD40
3OW8	16.4	3.9	12	WD repeat-containing protein 61	WD40
4BL0	16.3	4.2	11	Bub3p	WD40
3FM0	16.3	3.8	11	Protein CIAO1	WD40
3MMY	16.2	4.0	12	mRNA export factor	WD40
5FKS	16.2	3.5	10	Endo-1,4-beta-glucanase/ xyloglucanase, putative, gly74A	WD40
2CN2	16.2	3.4	8	Beta-1,4-xyloglucan hydrolase	WD40
1B9X	16.2	4.0	9	Protein (transducin)	WD40

Structural homologs of WBSCR16 were identified by search with the DALI web server (http://ekhidna.biocenter.helsinki.fi/dali_server/), and those with *Z*-score values above 16.2 are summarized here. The structural homologs can be classified into three categories: RCC1, RCC1-like proteins, and other β-propeller proteins (such as WD40 repeat proteins and FG-GAP repeat proteins), as noted in the Category column.