

Molecular Biology

Entry Clone Accession: IMAGE:2823112

Entry Clone Source: MGC (S611G Mutation introduced by PCR)

SGC Construct ID: RECQLA-c533

Protein Region: C49-Q616

Vector: pNIC-CTHF

Tag: C-terminal His6 and Flag tag, cleavable with TEV protea

Host: BL21(DE3)-R3-pRARE2

Protein Sequence (with tag underlined)

MCLEDSDAGASNEYDSSPAWNKEDFPWSGKVKDILQNVFKLEKFRPLQLETINVTMAG
KEVFLVMPTGGGKSLCYQLPALCSDGFTLVCPLISLMEDQLMVLKQLGISATMLNASSS
KEHVKWVHAEMVNKNSELKLIYVTPEKIAKSKMFMRSLEKAYEARRFTRIAVDEVHCCS
QWGHDFRPDYKALGILKRQFPNASLIGLTATATNHVLTDAQKILCIEKCFTFTASFNRPNL
YYEVRQKPSNTEDFIEDIVKLINGRYKGQSGIYCFSQKDSEQVTVSLQNLGIHAGAYHAN
LEPEDKTTVHRKWSANEIQVVVATVAFGMGIDKPDVRFVIHHSMSKSMENYYQESGRAG
RDDMKADCILYYGFGDIFRISSMVVMENVGQQKLYEMVSYCQNISKCRRVLMAQHFDE
VWNSEACNKMCDNCCKDSAERKNITEYCRDLIKLKQAEELNEKLTPLKLIDSWMGKG
AAKLRVAGVVAPTLPREDEKIIAHFLIQQYLKEDYSFTAYATISYKIGPKANLLNEAHA
ITMQVTKSTQNSFRAESSQTCHSEQSDKKMEAENLYFQSHHHHHHDYKDDDDK

DNA coding sequence:

ATGTGTTAGAGGATTCTGATGCCGGGGCAAGCAATGAATATGATTCTCACCTGCCGC
TTGGAATAAAGAAGATTTCCATGGTCTGGTAAAGTAAAGATATTCTGAAAATGTCT
TTAAACTGGAAAAGTTCAGACCACTTCAGCTGAAACTATTAACGTAACAATGGCTGG
AAAGGAGGTATTCAGCTTATGCCTACAGGAGGTGGAAAGAGCTTATGTTACCAAGTTA
CCAGCATTATGTTAGATGGTTTACACTCGTCATTGCCATTGATCTCTTTATGGAA
GACCAATTAAATGGTTTAAAACAATTAGGAATTTCAGCAACCATTAAATGCTTCTAG
TTCTAAGGAGCATGTTAAATGGGTTCATGCTGAAATGGTAAATAAAACTCCGAGTTA
AAGCTGATTATGTGACTCCAGAGAAAATTGCAAAAAGCAAAATGTTATGTCAAGAC
TAGAGAAAGCCTATGAAGCAAGGGAGATTACTCGAATTGCTGTGGATGAAGTTCACTG
CTGTAGTCAGTGGGGACATGATTTCAGACCTGATTATAAGGCACCTGGTATCTAAAGC
GGCAGTCCCTAACGCATCACTAATTGGCTGACTGCAACTGCAACAAATCACGTTT
GACGGATGCTCAGAAAATTGTCATTGAAAAGTGTAACTTTACAGCTTCTTTA
ATAGGCCAAATCTATATTATGAGGTTGGCAGAAGCCCTAACACTGAAGATTTATT
GAGGATATTGTAAGCTCATTAATGGGAGATAAAAGGCAATCAGGAATCATATATTG
TTTCTCAGAAAGACTCTGAACAAAGTTACGGTAGTTGCAGAATCTGGATTCA
GCAGGTGCTTACCATGCCAATTGGAGCCAGAAGATAAGACCACAGTTCATAGAAAAT
GGTCAGCCAATGAAATTCAAGTAGTAGTGGCAACTGTTGCATTGGTATGGATTGA
TAAGCCAGATGTGAGGTTGTTATCCATCATTCAATGAGTAAATCCATGGAAAATTATT
ACCAAGAGAGTGGACGTGCAGGTCGAGATGACATGAAAGCAGACTGTATTTGTACT
ACGGCTTGGAGATATTCAAGATAAGTCATGGTGGTAGGGAAAATGTGGGACA
GCAGAAGCTTATGAGATGGTATCATACTGTCAAAACATAAGCAAATGTCGTGTTG
TTGATGGCTAACATTGATGAAGTATGGAACCTCAGAACAGCATGTAACAAAATGTGCG
ATAACTGCTGTAAGACAGTCATTGAAAGAAAAGAACATAACAGAGTACTGCAGAG
ATCTAATCAAGATCCTGAAGCAGGCAGAGGAACCTGAATGAAAAACTCACTCCATTGA
AACTGATTGATTCTGGATGGAAAGGGTGAGCAAAACTGAGAGTAGCAGGTGTTG
TGGCTCCACACTCCTCGTGAAGATCTGGAGAAGATTATTGCACACTTCTAATACA
GCAGTATCTAAAGAAGACTACAGTTACAGCTTACAGCTTATGCTACCTTCGTATTGAAAA
TAGGACCTAAAGCTAACCTCTGAACAATGAGGCACATGCTATTACTATGCAAGTGAC
AAAGTCCACGCAGAACTCTTCAGGGCTGAATCGTCTCAAACATTGTCATTCTGAACAA

AGTGATAAAAAGATGGAGGCAGAGAACCTCTACTTCCAATCGCACCATCATCACCACC
ATGATTACAAGGATGACGACGATAAGTGA

Protein Expression

Medium: Terrific Broth

Antibiotics: Kanamycin, 50 μ g/ml

Procedure: The expression plasmid was transformed into the host strain and plated on LB-agar containing 50 μ g/ml kanamycin and 35 μ g/ml chloramphenicol. Several colonies were combined to inoculate a 1-ml culture in TB (+ 50 μ g/ml kanamycin, 35 μ g/ml chloramphenicol). The culture was grown overnight, glycerol was added to 15% v/v (from a 60% stock), and the resulting glycerol stock was frozen at -80°C in 100 μ l aliquots. A loopful of cells from the glycerol stock was inoculated into 100-ml of TB medium containing 50 μ g/ml kanamycin and 35 μ g/ml chloramphenicol and grown overnight at 37°C.

Six cultures of 1-L each of TB medium (+ 50 μ g/ml kanamycin) in 2.5L UltraYield baffled flasks were inoculated with 10 ml each of the overnight culture. The cultures were grown at 37°C until OD₆₀₀ of 2-3 and then shifted to 25°C. After 30 minutes, IPTG was added to 0.1 mM, and growth continued overnight. The cells were collected by centrifugation, the pellets were scraped out and transferred to 50-ml Falcon tubes and frozen at -80°C.

Protein Purification

Buffers: Lysis buffer: 50 mM HEPES buffer, pH 7.5, 500 mM NaCl, 10 mM imidazole, 0.5 mM TCEP, 1x Protease Inhibitors Cocktail Set VII (Calbiochem, 1/1000 dilution), and 15 units/ml Benzonase. 2x Lysis buffer contains the same components at double concentration.

Affinity buffer: 50 mM HEPES buffer, pH 7.5, 500 mM NaCl, 10 mM imidazole, 0.5 mM TCEP;

Wash buffer: 50 mM HEPES buffer, pH 7.5, 500 mM NaCl, 30 mM imidazole, 0.5 mM TCEP;

Elution buffer: 50 mM HEPES buffer, pH 7.5, 500 mM NaCl, 250 mM imidazole, 0.5 mM TCEP.

GF buffer: 10 mM HEPES, pH 7.5, 500 mM NaCl, 5% glycerol, 0.5 mM TCEP.

Procedure: Frozen cell pellets (86 g) were thawed briefly in a bath of warm water (20 - 37°C) then transferred to ice. One volume (i.e. 1 ml for every gram of cells) of 2x lysis buffer was added, followed by 1x lysis buffer to a total volume of 300-ml. The cells were resuspended by agitating and disrupted by sonication. Nucleic acids and cell debris were removed by adding 0.15% PEI (polyethyleneimine) from a 5% (w/v, pH 7.5) stock, stirring for 15 minutes, then centrifugation for 20 minutes at 25,000 x g.

Column 1: The cell extract was loaded onto a Histrap FF 5 ml column (GE Healthcare) at 5 ml/minute on an AKTA-express system. The column was then washed with 10 volumes of loading buffer, 10 volumes of wash buffer, then eluted with elution buffer at 4 ml/min. The eluted peak of A₂₈₀ was automatically collected.

Column 2: The eluted fraction was loaded and fractionated on a HiLoad 16/60 Superdex 200 gel filtration column in GF buffer at 1.2 ml/min. 2-ml fractions were collected at the A₂₈₀ peaks. The fractions were analyzed by SDS-PAGE and relevant fractions were pooled. The protein was concentrated to 10 mg/ml using a centricon centrifugal device (30 kDa MWCO).

Mass-spec Verification: ESI-TOF was used to verify the intact mass of the protein. Observed mass 67222, calculated mass 67294; (unexplained mass difference of -72). This may partly arise from cleavage of the N-terminal methionine.

Structure Determination

Crystallization: RECQ1 protein was buffer exchanged into 10 mM HEPES (pH 7.5), 10 mM NaCl, 5% glycerol, 1 mM DTT, diluted to 15 μ M, and combined with 20 μ M of a single-stranded oligonucleotide (5'- GGATCTCGACGCTCTCCCTT-3'). The complex was then concentrated by ultrafiltration (Centricon) to 150 μ M protein (10 mg/ml). Crystals were grown at 4°C by vapor diffusion from nanoliter sitting drops against a reservoir solution containing 17% PEG3350, 5% ethylene glycol, 0.18M potassium citrate, 0.1M Bis-tris-propane pH 6.5. Crystals were soaked in reservoir solution supplemented with an additional 20% ethylene glycol prior to flash-cooling in liquid nitrogen.

Data Collection: Diffraction data were collected on beamline I04 at the Diamond Light Source (Harwell, UK). Data were processed with XDS (4) (tailed duplex), MOSFLM (5) (ssDNA complex) and SCALA (6)

The apparent symmetry of the ssDNA complex was orthorhombic but both intensity moments and L-test indicated that the data were twinned. Data were processed assuming monoclinic symmetry with a beta angle close to 90° with MOSFLM and SCALA. Free flags were generated by PHENIX taking into consideration the known pseudo-symmetry. The structure was solved using the coordinates of DNA-free enzyme (PDB code: 2V1X) as a search model in PHASER. This procedure correctly placed 3 of the 4 protein chains present in the asymmetric unit. The fourth protein chain was placed manually using knowledge of the RECQ1's known dimeric arrangement followed by rigid-body refinement. Initial maps indicated the presence of two pseudo tailed DNA duplexes that is sandwiched between the winged-helix domain's beta-hairpins of adjacent molecules in the crystal. The single stranded oligonucleotides were initially modeled as poly-T and the sequence was then tentatively assigned based on comparison with the dsDNA tailed duplex complex and the visible start point of the longer single strand that extends between the RecA domains (assigned as the first nucleotide of sequence as no additional electron density was visible in the 5' direction). Twinned refinement was carried out with PHENIX with additional torsional restraints using the high resolution apo enzyme coordinates (PDB code: 2V1X) as a reference molecule, NCS restraints, grouped B-factor and TLS refinement. The final ssDNA model comprises four RECQ1 chains and four single-stranded oligonucleotides. The pseudo-duplex (assigned chains P/Q) associated with protein chains A & B is the better resolved than that associated with the other two chains.