Entry Clone Source: MGC

Entry Clone Accession: BC030775

SGC Construct ID: ERAP1A-c200

GenBank GI number: gi|94818901

Vector: pFB-CT10HF. Details [PDF]; Sequence [FASTA] or [GenBank]

Amplified construct sequence:

CTTAAGAAGGAGATATACTATGGTGTTTCT
GCCCCTCAAATGGTCCCTTGCAACCATGTC
ATTTCTACTTTCCTCACTGTTGGCTCTCTT
AACTGTGTCCACTCCTTCATGGTGTCAGAG
CACTGAAGCATCTCCAAAACGTAGTGATGG
GACACCATTTCCTTGGAATAAAATACGACT
TCCTGAGTACGTCATCCCAGTTCATTATGA
TCTCTTGATCCATGCAAACCTTACCACGCT
GACCTTCTGGGGAACCACGAAAGTAGAAAT
CACAGCCAGTCAGCCCACCAGCACCATCAT
CCTGCATAGTCACCACCTGCAGATATCTAG
GGCCACCCTCAGGAAGGGAGCTGGAGAGAG
GCTATCGGAAGAACCCCTGCAGGTCCTGGA
ACACCCCCGTCAGGAGCAAATTGCACTGCT
GGCTCCCGAGCCCCTCCTTGTCGGGCTCCC
GTACACAGTTGTCATTCACTATGCTGGCAA
TCTTTCGGAGACTTTCCACGGATTTTACAA
AAGCACCTACAGAACCAAGGAAGGGGAACT
GAGGATACTAGCATCAACACAATTTGAACC
CACTGCAGCTAGAATGGCCTTTCCCTGCTT
TGATGAACCTGCCTTCAAAGCAAGTTTCTC
AATCAAAATTAGAAGAGAGCCAAGGCACCT
AGCCATCTCCAATATGCCATTGGTGAAATC
TGTGACTGTTGCTGAAGGACTCATAGAAGA
CCATTTTGATGTCACTGTGAAGATGAGCAC
CTATCTGGTGGCCTTCATCATTTCAGATTT
TGAGTCTGTCAGCAAGATAACCAAGAGTGG
AGTCAAGGTTTCTGTTTATGCTGTGCCAGA
CAAGATAAATCAAGCAGATTATGCACTGGA
TGCTGCGGTGACTCTTCTAGAATTTTATGA
GGATTATTTCAGCATACCGTATCCCCTACC
CAAACAAGATCTTGCTGCTATTCCCGACTT
TCAGTCTGGTGCTATGGAAAACTGGGGACT
GACAACATATAGAGAATCTGCTCTGTTGTT
TGATGCAGAAAAGTCTTCTGCATCAAGTAA
GCTTGGCATCACAATGACTGTGGCCCATGA
ACTGGCTCACCAGTGGTTTGGGAACCTGGT
CACTATGGAATGGTGGAATGATCTTTGGCT
AAATGAAGGATTTGCCAAATTTATGGAGTT
TGTGTCTGTCAGTGTGACCCATCCTGAACT
GAAAGTTGGAGATTATTTCTTTGGCAAATG
TTTTGACGCAATGGAGGTAGATGCTTTAAA
TTCCTCACACCCTGTGTCTACACCTGTGGA

AAATCCTGCTCAGATCCGGGAGATGTTTGA
TGATGTTTCTTATGATAAGGGAGCTTGTAT
TCTGAATATGCTAAGGGAGTATCTTAGTGC
TGACGCATTTAAAAGTGGTATTGTACAGTA
TCTCCAGAAGCATAGCTATAAAAATACAAA
AAACGAGGACCTGTGGGATAGTATGGCAAG
TATTTGCCCTACAGATGGTGTAAAAGGGAT
GGATGGCTTTTGCTCTAGAAGTCAACATTC
ATCTTCATCCTCACATTGGCATCAGGAAGG
GGTGGATGTGAAAACCATGATGAACACTTG
GACACTGCAGAAGGGTTTTCCCCTAATAAC
CATCACAGTGAGGGGGAGGAATGTACACAT
GAAGCAAGAGCACTACATGAAGGGCTCTGA
CGGCGCCCCGGACACTGGGTACCTGTGGCA
TGTTCCATTGACATTCATCACCAGCAAATC
CGACATGGTCCATCGATTTTTGCTAAAAAC
AAAAACAGATGTGCTCATCCTCCCAGAAGA
GGTGGAATGGATCAAATTTAATGTGGGCAT
GAATGGCTATTACATTGTGCATTACGAGGA
TGATGGATGGGACTCTTTGACTGGCCTTTT
AAAAGGAACACACACAGCAGTCAGCAGTAA
TGATCGGGCGAGTCTCATTAACAATGCATT
TCAGCTCGTCAGCATTGGGAAGCTGTCCAT
TGAAAAGGCCTTGGATTTATCCCTGTACTT
GAAACATGAAACTGAAATTATGCCCGTGTT
TCAAGGTTTGAATGAGCTGATTCCTATGTA
TAAGTTAATGGAGAAAAGAGATATGAATGA
AGTGGAAACTCAATTCAAGGCCTTCCTCAT
CAGGCTGCTAAGGGACCTCATTGATAAGCA
GACATGGACAGACGAGGGCTCAGTCTCAGA
GCGAATGCTGCGGAGTCAACTACTACTCCT
CGCCTGTGTGCACAACTATCAGCCGTGCGT
ACAGAGGGCAGAAGGCTATTTCAGAAAGTG
GAAGGAATCCAATGGAAACTTGAGCCTGCC
TGTCGACGTGACCTTGGCAGTGTTTGCTGT
GGGGGCCCAGAGCACAGAAGGCTGGGATTT
TCTTTATAGTAAATATCAGTTTTCTTTGTC
CAGTACTGAGAAAAGCCAAATTGAATTTGC
CCTCTGCAGAACCCAAAATAAGGAAAAGCT
TCAATGGCTACTAGATGAAAGCTTTAAGGG
AGATAAAATAAAAACTCAGGAGTTTCCACA
AATTCTTACACTCATTGGCAGGAACCCAGT
AGGATACCCACTGGCCTGGCAATTTCTGAG
GAAAAACTGGAACAAACTTGTACAAAAGTT
TGAACTTGGCTCATCTTCCATAGCCCACAT
GGTAATGGGTACAACAAATCAATTCTCCAC
AAGAACACGGCTTGAAGAGGTAAAAGGATT
CTTCAGCTCTTTGAAAGAAAATGGTTCTCA
GCTCCGTTGTGTCCAACAGACAATTGAAAC
CATTGAAGAAAACATCGGTTGGATGGATAA
GAATTTTGATAAAATCAGAGTGTGGCTGCA
AAGTGAAAAGCTTGAACGTATGGCAGAGAA
CCTCTACTTCCAATCGCACCATCATCACCA

TCACCATCACCACCATGATTACAAGGATGA
CGACGATAAGTGAGGATCC

Final protein sequence (tag sequence in lowercase)
MVFLPLKWSLATMSFLLSSLLALLTVSTPS
WCQSTEASPKRSDGTPFPWNKIRLPEYVIP
VHYDLLIHANLTTLTFWGTTKVEITASQPT
STIILHSHHLQISRATLRKGAGERLSEEPL
QVLEHPRQEQIALLAPEPLLVGLPYTVVIH
YAGNLSETFHGFYKSTYRTKEGELRILAST
QFEPTAARMAFPCFDEPAFKASFSIKIRRE
PRHLAISNMPLVKSVTVAEGLIEDHFDVTV
KMSTYLVAFIISDFESVSKITKSGVKVSVY
AVPDKINQADYALDAAVTLLEFYEDYFSIP
YPLPKQDLAAIPDFQSGAMENWGLTTYRES
ALLFDAEKSSASSKLGITMTVAHELAHQWF
GNLVTMEWWNDLWLNEGFAKFMEFVSVSVT
HPELKVGDYFFGKCFDAMEVDALNSSHPVS
TPVENPAQIREMFDDVSYDKGACILNMLRE
YLSADAFKSGIVQYLQKHSYKNTKNEDLWD
SMASICPTDGVKGMDGFCSRSQHSSSSSHW
HQEGVDVKTMMNTWTLQKGFPLITITVRGR
NVHMKQEHYMKGSDGAPDTGYLWHVPLTFI
TSKSDMVHRFLLKTKTDVLILPEEVEWIKF
NVGMNGYYIVHYEDDGWDSLTGLLKGTHTA
VSSNDRASLINNAFQLVSIGKLSIEKALDL
SLYLKHETEIMPVFQGLNELIPMYKLMEKR
DMNEVETQFKAFLIRLLRDLIDKQTWTDEG
SVSERMLRSQLLLLACVHNYQPCVQRAEGY
FRKWKESNGNLSLPVDVTLAVFAVGAQSTE
GWDFLYSKYQFSLSSTEKSQIEFALCRTQN
KEKLQWLLDESFKGDKIKTQEFPQILTLIG
RNPVGYPLAWQFLRKNWNKLVQKFELGSSS
IAHMVMGTTNQFSTRTRLEEVKGFFSSLKE
NGSQLRCVQQTIETIEENIGWMDKNFDKIR
VWLQSEKLERMaenlyfqshhhhhhhhhhd
ykddddk

Tags and additions: C-terminal, TEV cleavable decahistidine tag and Flag tag. Tag sequence: aenlyfq*shhhhhhhhhhdykddddk

Host: *Trichoplusia Ni* (High five)

Expression:
High five cells were grown in Insect Express medium at 27°C. Cells were infected at a density of $2x10^6$/ml with recombinant baculovirus (virus stock P2; 1ml of virus stock/1l of cell culture) Culture was supplemented with FCS to final concentration 1%. 120 hours post-infection the cultures were collected and centrifuged for 30min at 2000rpm. Cellular pellets were discarded and supernatants were used as a source of protein.

Purification:
Step 1: Ni-affinity, Ni-Sepharose - (GE Healthcare) purification in batch
Step 2: Superdex 200 Column , HiPrep 16/60 (Amersham)
Buffers:

Washing buffer: 50mM HEPES pH 7.5, 500mM NaCl, 10mM Imidazole, 5% glycerol, 1mM PMSF, 0.5mM TCEP
Elution buffer: 50mM HEPES pH 7.5, 500mM NaCl, 5% glycerol, 250mM Imidazole, 0.5mM TCEP
GF buffer: 10mM HEPES pH 7.5, 500mM NaCl, 5% glycerol, 0.5mM TCEP
Procedure: Supernatant was supplemented with Tris buffer pH 8.0 to final concentration 50mM, NaCl to final concentration 300mM, and $NiSO_4$ to final concentration 1mM. Solution was supplemented with PMSF to final concentration 1mM and 1tablet of protease inhibitors/ per 1l of solution. Ni resin was added and incubation was performed in room temprature for 4h. Suspension was loaded on gravity column and after washing with 20 volumes of washing buffer, protein was eluted in 4 elution fractions (10ml each). Protein fractions were analysed by SDS-PAGE. Target protein containing fractions were concentrated using Amicon Ultra-15 concentrators with 30kDa cut-off, and purified on a gel filtration column (Superdex 200) on an Akta Express system. Fractions containing protein were analysed by SDS-PAGE.

Concentration and buffer exchange:
Using Amicon Ultra-15 concentrators with 30kDa cutoff, the sample was concentrated to 17.1mg/ml. Concentrations were determined from the absorbance at 280 nm using a NanoDrop spectrophotometer.

Mass spectrometry characterization: The calculated mass of the construct was 110554.5Da, and the observed mass (ESI-MS) was 111563Da, suggesting glycosylation of the protein.

Crystallisation: Crystals were grown by vapor diffusion at 20°C in 150nl sitting drops. The drops were prepared by mixing 100nl of protein solution and 50nl of precipitant consisting of 1.4M KCitrate; 0.1M Cacodylate pH 5.7 and 0.17mMn-Dodecyl-Beta-D-maltoside. Crystals were flash-cooled in liquid nitrogen with 25 % glycerol as cryoprotectant. A mercury derivative was prepared by adding 1µl of reservoir solution supplemented with 10mM Thiomersal to a crystallization drop. After 10 minutes, a crystal was transferred to a cryo solution containing reservoir solution supplemented with 25% ethylene glycol and the crystal was immediately flash frozen in liquid nitrogen. Data on native and derivative crystals were collected on beamline I03 at the Diamond Light Source. One mercury site was immediately identified with the program HYSS and subsequent phase refinement with SHARP and solvent flattening with SOLOMON resulted in an electron density map of excellent quality

Data collection: Resolution: 2.7 Å native protein 3.4 Å (Hg-derivative). **X-ray source:** Diamond Light Source beamline IO3.