ZMPSTE24A (2YPT) Materials & Methods

Entry Clone Source: MGC

Entry Clone Accession: IMAGE: 5269064

SGC Construct ID: ZMPSTE24A-c000 / ZMPSTE24A-c011

GI Number: 10269

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

DNA sequence:

c000:

```
CTTAAGAAGGAGATATACTATGGGGA
TGTGGGCATCGCTGGACGCTTTGTGG
GAGATGCCGGCCGAGAAGCGTATCTT
CGGGGCCGTGCTGCTCTTTTCCTGGA
CAGTGTATCTTTGGGAGACCTTCCTA
GCACAGCGGCAGAGAAGGATATATAA
AACAACAACTCATGTACCACCGGAGT
TAGGACAGATCATGGATTCTGAAACA
TTTGAGAAATCTCGACTCTATCAACT
GGATAAAAGCACTTTCAGCTTCTGGT
CAGGACTCTATTCAGAGACTGAAGGC
ACTCTTATTCTTCTCTTTGGAGGAAT
ACCTTATCTCTGGAGACTTTCTGGAC
GGTTCTGTGGTTATGCTGGCTTTGGA
CCAGAATATGAGATCACTCAGTCCCT
GGTGTTTCTGCTGTTGGCTACACTTT
TCAGTGCATTGGCTGGTTTGCCATGG
AGTCTTTATAATACTTTTGTGATAGA
AGAAAAACATGGCTTCAATCAACAGA
CTTTGGGGTTCTTCATGAAAGATGCA
ATCAAGAAATTTGTTGTGACTCAGTG
CATTTTGTTGCCTGTGTCTTCACTTC
TACTTTACATTATTAAAATTGGGGGT
GACTATTTTTTTATTTATGCCTGGCT
GTTCACATTAGTTGTGTCTCTGGTTC
TTGTCACAATCTATGCTGATTATATT
GCCCCTTTATTTGACAAATTCACACC
TCTGCCTGAGGGAAAGCTTAAAGAAG
AAATTGAAGTAATGGCAAAGAGTATT
GACTTTCCTTTGACGAAGGTGTATGT
TGTGGAAGGATCTAAACGCTCTTCCC
ACAGCAATGCTTATTTTTATGGCTTC
TTCAAGAACAAGCGAATAGTTTTGTT
TGACACTCTACTAGAAGAGTACTCTG
TACTAAACAAAGACATCCAGGAGGAT
TCTGGCATGGAACCCCGCAATGAGGA
AGAAGGGAACAGTGAAGAAATAAAAG
```

DNA sequence:

c011:

```
CTTAAGAAGGAGATATACTATGGGGA
TGTGGGCATCGCTGGACGCTTTGTGG
GAGATGCCGGCCGAGAAGCGTATCTT
CGGGGCCGTGCTGCTCTTTTCCTGGA
CAGTGTATCTTTGGGAGACCTTCCTA
GCACAGCGGCAGAGAAGGATATATAA
AACAACAACTCATGTACCACCGGAGT
TAGGACAGATCATGGATTCTGAAACA
TTTGAGAAATCTCGACTCTATCAACT
GGATAAAAGCACTTTCAGCTTCTGGT
CAGGACTCTATTCAGAGACTGAAGGC
ACTCTTATTCTTCTCTTTGGAGGAAT
ACCTTATCTCTGGAGACTTTCTGGAC
GGTTCTGTGGTTATGCTGGCTTTGGA
CCAGAATATGAGATCACTCAGTCCCT
GGTGTTTCTGCTGTTGGCTACACTTT
TCAGTGCATTGGCTGGTTTGCCATGG
AGTCTTTATAATACTTTTGTGATAGA
AGAAAAACATGGCTTCAATCAACAGA
CTTTGGGGTTCTTCATGAAAGATGCA
ATCAAGAAATTTGTTGTGACTCAGTG
CATTTTGTTGCCTGTGTCTTCACTTC
TACTTTACATTATTAAAATTGGGGGT
GACTATTTTTTTATTTATGCCTGGCT
GTTCACATTAGTTGTGTCTCTGGTTC
TTGTCACAATCTATGCTGATTATATT
GCCCCTTTATTTGACAAATTCACACC
TCTGCCTGAGGGAAAGCTTAAAGAAG
AAATTGAAGTAATGGCAAAGAGTATT
GACTTTCCTTTGACGAAGGTGTATGT
TGTGGAAGGATCTAAACGCTCTTCCC
ACAGCAATGCTTATTTTTATGGCTTC
TTCAAGAACAAGCGAATAGTTTTGTT
TGACACTCTACTAGAAGAGTACTCTG
TACTAAACAAAGACATCCAGGAGGAT
TCTGGCATGGAACCCCGCAATGCTGC
TGCTGGGAACAGTGAAGAAATAAAAG
```

CTAAAGTTAAAAATAAGAAACAAGGA TGTAAAAATGAGGAGGTACTCGCTGT ACTAGGCCATGAACTGGGGCACTGGA AGTTGGGACATACAGTCAAAAATATC ATTATTAGCCAGATGAATTCTTTCCT GTGTTTTTTTTTATTTGCTGTATTAA TTGGTCGAAAGGAGCTTTTTGCTGCA TTTGGTTTTTATGATAGCCAACCCAC TCTTATTGGACTATTGATCATCTTCC AGTTTATTTTTTCACCTTACAATGAG GTTCTTTCTTTTTGCCTAACAGTCCT AAGCCGCAGATTTGAGTTTCAAGCTG ATGCATTTGCCAAGAAACTTGGGAAG GCTAAAGACTTATATTCTGCTTTAAT CAAACTTAACAAAGATAACTTGGGAT TCCCTGTTTCTGACTGGTTGTTCTCA ATGTGGCATTATTCTCATCCTCCACT GCTAGAGAGACTTCAAGCTTTGAAAA CTATGAAGCAACACGCAGAGAACCTC TACTTCCAATCGCACCATCATCACCA TCACCATCACCACCATGATTACAAGG ATGACGACGATAAGTGAGGATCC ww	CTAAAGTTAAAAATAAGAAACAAGGA TGTAAAAATGAGGAGGTACTCGCTGT ACTAGGCCATGAACTGGGGCACTGGA AGTTGGGACATACAGTCAAAAATATC ATTATTAGCCAGATGAATTCTTTCCT GTGTTTTTTTTTATTTGCTGTATTAA TTGGTCGAAAGGAGCTTTTTGCTGCA TTTGGTTTTTATGATAGCCAACCCAC TCTTATTGGACTATTGATCATCTTCC AGTTTATTTTTTCACCTTACAATGAG GTTCTTTCTTTTTGCCTAACAGTCCT AAGCCGCAGATTTGAGTTTCAAGCTG ATGCATTTGCCAAGAAACTTGGGAAG GCTAAAGACTTATATTCTGCTTTAAT CAAACTTAACAAAGATAACTTGGGAT TCCCTGTTTCTGACTGGTTGTTCTCA ATGTGGCATTATTCTCATCCTCCACT GCTAGAGAGACTTCAAGCTTTGAAAA CTATGAAGCAACACGCAGAGAACCTC TACTTCCAATCGCACCATCATCACCA TCACCATCACCACCATGATTACAAGG ATGACGACGATAAGTGAGGATCC
Expressed sequence (small letters refer to tag sequence): c000 MGMWASLDALWEMPAEKRIFGAVLLF SWTVYLWETFLAQRQRRIYKTTTHVP PELGQIMDSETFEKSRLYQLDKSTFS FWSGLYSETEGTLILLFGGIPYLWRL SGRFCGYAGFGPEYEITQSLVFLLLA TLFSALAGLPWSLYNTFVIEEKHGFN QQTLGFFMKDAIKKFVVTQCILLPVS SLLLYIIKIGGDYFFIYAWLFTLVVS LVLVTIYADYIAPLFDKFTPLPEGKL KEEIEVMAKSIDFPLTKVYVVEGSKR SSHSNAYFYGFFKNKRIVLFDTLLEE YSVLNKDIQEDSGMEPRNEEEGNSEE IKAKVKNKKQGCKNEEVLAVLGHELG HWKLGHTVKNIIISQMNSFLCFFLFA VLIGRKELFAAFGFYDSQPTLIGLLI IFQFIFSPYNEVLSFCLTVLSRRFEF QADAFAKKLGKAKDLYSALIKLNKDN LGFPVSDWLFSMWHYSHPPLLERLQA LKTMKQaenlyfq^shhhhhhhhhhd ykddddk	**Expressed sequence (small letters refer to tag sequence):** c011 MGMWASLDALWEMPAEKRIFGAVLLF SWTVYLWETFLAQRQRRIYKTTTHVP PELGQIMDSETFEKSRLYQLDKSTFS FWSGLYSETEGTLILLFGGIPYLWRL SGRFCGYAGFGPEYEITQSLVFLLLA TLFSALAGLPWSLYNTFVIEEKHGFN QQTLGFFMKDAIKKFVVTQCILLPVS SLLLYIIKIGGDYFFIYAWLFTLVVS LVLVTIYADYIAPLFDKFTPLPEGKL KEEIEVMAKSIDFPLTKVYVVEGSKR SSHSNAYFYGFFKNKRIVLFDTLLEE YSVLNKDIQEDSGMEPRNAAAGNSEE IKAKVKNKKQGCKNEEVLAVLGHELG HWKLGHTVKNIIISQMNSFLCFFLFA VLIGRKELFAAFGFYDSQPTLIGLLI IFQFIFSPYNEVLSFCLTVLSRRFEF QADAFAKKLGKAKDLYSALIKLNKDN LGFPVSDWLFSMWHYSHPPLLERLQA LKTMKQHAENLYFQaenlyfq^shhh hhhhhhhdykddddk
Tags and additions: C-terminal, TEV cleavable decahistidine / FLAG tag.	

Host: *Spodoptera frugiperda* (SF9) insect cells

Growth Medium & Induction Protocol: Insect cells with a density of $2x10^6$ per litre of cell culture in SF900 medium (Invitrogen) were infected with recombinant baculovirus (5ml P2 virus per litre cell culture) and incubated for 72 hours. Cells were harvested by centrifugation and the pellet was flash frozen in liquid N_2.

Extraction buffer, extraction method: Frozen pellets were thawed and re-suspended in extraction buffer supplemented with 1% DDM or 1% OGNG/ 0.1% CHS and incubated at 4°C on a tube rotator. The extracted protein was separated from insoluble membranes by centrifugation and collected for purification.

Extraction/wash buffer: 50 mM HEPES, pH 7.5; 200 mM NaCl

Column 1: Co-affinity. Cobalt Talon (Clontech), 0.75 ml of 50 % slurry/ 1L of cells in 1.5 x 10 cm column, washed with extraction/ wash buffer.

Column 1 Buffers:
Extraction/wash buffer: 50 mM HEPES, pH 7.5; 200 mM NaCl; 20 mM imidazole, detergent at 3x CMC.
Elution buffer: 50 mM HEPES, pH 7.5; 200 mM NaCl; 250 mM imidazole, detergent at 3xCMC

Column 1 Procedure: The solubilised membrane protein was batch bound to Cobalt Talon resin equilibrated with extraction buffer at 4°C for one hour and subsequently poured into a gravity flow glass column. The column was then washed with 30CV of wash buffer followed by elution in 1CV fractions until all the protein was eluted.

Column 2: Size Exclusion Chromatography. Superdex200 (GE healthcare)

Column 2 Buffer:
Gel Filtration buffer: 20 mM HEPES pH 7.5, 200 mM NaCl,detergent at 1.2xCMC

Column 2 Procedure: The protein was concentrated and applied to a Superdex 200 gel filtration column equilibrated with gel filtration buffer using an AKTA purifier system.

Enzymatic Treatment: TEV protease (1:4, TEV:protein) was added overnight at 4°C to purified protein. The protease was removed and tag cleaved protein was collected by binding to Cobalt talon and collecting the flow-through.

Mass spec characterization:
The purified protein was homogeneous and had an experimental mass of 55517Da (c000) or 55345Da (c011), as expected from the primary sequence including loss of the N-terminal methionine. Masses were determined by LC-MS, using an Agilent LC/MSD TOF system with reversed-phase HPLC coupled to electrospray ionisation and an orthogonal time-of-flight mass analyser. Proteins were desalted prior to mass spectrometry by rapid elution off a C3 column with a gradient of 5-95% methanol in water with 0.1% formic acid.

Protein Concentration: Protein was concentrated to ~20mg/ml using a 100kDa cut-off concentrator and back diluted to 9-11mg/ml.

Crystallization: Crystals were grown at 20°C from sitting drops (150-300nl). Two crystal forms were observed depending on the purification detergent used. Monoclinic (P21) crystals were obtained from protein (c011) purified in 0.015% DDM and crystallized with a reservoir solution containing 30% (v/v) PEG 400, 0.1M lithium sulfate, 0.1M sodium chloride, 0.1M MES, pH 6.5. Triclinic (P1) crystals were obtained from protein (c000) purified in 0.12%

OGNG/0.012% CHS, and crystallised with a reservoir solution containing 30-42%(v/v) PEG 400, 0.1M HEPES pH 7.5, 0.1M calcium chloride. Crystals were cryocooled by transfer into liquid nitrogen without the addition of a cryoprotectant. Diffraction of the P1 crystal form was improved by transferring crystallization plates to 6°C and equilibration for at least 24 hours prior to direct flash cooling of crystals.

Data Collection:
Resolution: 3.4Å
All data were collected at 100K on beamlines I04 (P21 form) and I24 (P1 form) at Diamond Light Source using helical / straight line scans with a beamsize of 10µm x 10-30µm (wxh).

Structure Solution: Phases were obtained using SIRAS with data from P1 crystals derivatised with 1mM EMTS. A total of 12 mercury sites were located with SHELXD (4 ZMPSTE24 molecules per asymmetric unit) and phase refinement was carried out by SOLVE/RESOLVE to 3.8Å resolution. Initial electron density maps, calculated after four-fold averaging / phase extension to 3.4Å, allowed an initial backbone trace to be built. Subsequent cross-crystal averaging between the P1 (4 mols/AU) and $P2_1$ (4 mols/AU) forms gave a substantial improvement in the quality of the experimental phases and allowed tracing and sequence assignment for the transmembrane and protease domains. The structure was refined with BUSTER using all data to 3.4Å with appropriate NCS and TLS restraints to final R/Rfree values of 24.6 and 26.4% respectively. The final P1 model comprises 4 copies of ZMPSTE24 with each copy encompassing residues 10-107, 116-285, 322-472, a single zinc ion and two phospholipid alkyl chains.