SI19: ORF8

Table 1: General Information

1	**Protein Name (according to NCBI Reference Sequence NC_045512.2)**
	ORF8
2	**Region/Name/Further Specification**
3	**Sequence of fl protein (according to NCBI Reference Sequence NC_045512.2)**
	MKFLVFLGIITTVAAFHQECSLQSCTQHQPYVVDDPCPIHFYSKWYIRVGARKSAPLIELCVDEA GSKSPIQYIDIGNYTVSCSPFTINCQEPKLGSLVVRCSFYEDFLEYHDVRVVLDFI
4	**Protein boundaries of expressed construct (according to NCBI Reference Sequence NC_045512.2)**
ORF8	aa 1-121 (fl ORF8 = ORF8)
ORF8 m	aa 1-121 (fl ORF8) with L84S mutation (~ isolate 2019-nCoV_HKU-SZ-002a_2020).
ΔORF8	aa 16-121 (without signal peptide = ΔORF8)
5	**Ratio for construct design (detailed and comprehensible)**
ORF8	fl protein
ΔORF8	Protein after the hypothetical cleavage of the N-terminal Signal Peptide
6	**Sequence homology (to SCoV)**
ORF8	Identity: 31.7%; similarity: 70.7%
ΔORF8	Identity: 40.5%; similarity: 66.7%
7	**Published structures (SCoV2 or homologue variants)**
	SCoV2: 7JTL, 7JX6
8	**(Published) assignment (SCoV2 or homologue variants)**
	-

Bacterial

Table 2: Protein Expression

1	**Expression vector**
ORF8 m	pPK1154 (GenScript)
ΔORF8	pET22b (+) (Merck/Novagen)
2	**Purification-/Solubility-Tag**
ORF8 m	N-terminal His_6-SUMO

ΔORF8	N-terminal His_6-GST
3	**Cleavage Site**
ORF8m	Ulp1
ΔORF8	TEV
4	**Molecular weight / Extinction coefficient / pI - of cleaved protein**
ORF8m	13.80 kDa / 15,930 M^{-1} cm^{-1} / 5.42
ΔORF8	12.54 kDa / 15,930 M^{-1} cm^{-1} / 5.15
5	**Comments on sequence of expressed construct**
ORF8m	No artificial residues due to Ulp1-cleavage and construct design.
ΔORF8	N-terminal “GAMG” three artificial residues due to TEV-cleavage and construct design.
6	**Used expression strain**
ORF8m	*E. coli* BL21 (DE3)
ΔORF8	*E. coli* BL21 (DE3) pLysS
7	**Cultivation medium**
ORF8m	LB / M9 (uniformly ^{15}N-labelled)
ΔORF8	LB
8	**Induction system**
	IPTG inducible T7 promoter
9	**Induction of protein expression**
ORF8m	0.5 mM IPTG at OD_{600} 0.6
ΔORF8	0.5 mM IPTG at OD_{600} 0.6-0.7
10	**Cultivation temperature and time**
ORF8m	16-20°C for 16-18 h
ΔORF8	18°C for 16-18 h

Table 3a: Protein Purification (ORF8)

1	**Buffer List**
A	10 mM NaPi (pH 8.0), 300 mM NaCl, 10 mM imidazole, 0.5 mM DTT (Cell disruption).
B	10 mM NaPi (pH 8.0), 300 mM NaCl, 10 mM imidazole, 0.5 mM DTT (Solubilization of pellet).

C	10 mM NaPi (pH 8.0), 300 mM NaCl, 0.5 mM DTT (IMAC).
D	50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM DTT, 0.2% (w/v) NP40.
E	50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM DTT.
2	**Purification steps (with corresponding buffer(s) and incubation times)**
A	Cell disruption in buffer **1A** (plus one tablet of EDTA free protease inhibitor cocktail (Merck)) by French-press.
B	Solubilization of pellet after lysis **1B** (plus one tablet of EDTA free protease inhibibitor cocktail (Merck)).
C	IMAC (Nickel-NTA-Agarose, QIAGEN) by hand, elution with 250 mM imidazole in buffer **1C**.
D	Ulp1-cleavage (Protein/Ulp1 ratio 10:1) o.n. at 21°C in buffer **1D**.
E	Rebuffer in buffer **1E**.

Table 3b: Protein Purification (ΔORF8)

1	**Buffer List**
A	50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 5% (v/v) glycerol, 50 mM imidazole (cell disruption/IMAC).
B	50 mM Tris-HCl (pH 8.0), 150 mM NaCl (TEV-cleavage).
C	20 mM NaPi (pH 7.4), 150 mM NaCl, 1 mM EDTA (SEC final buffer).
2	**Purification steps (with corresponding buffer(s) and incubation times)**
A	Cell disruption in buffer **1A** (supplemented with 0.5 mg/ml Lysozyme, 10 µg/ml DNaseI, 5 mM MgCl2, cOmplete™ EDTA-free protease inhibitors) by incubation for 30 min at rt followed by sonication at 43% amplitude for 2 minutes (1 s on, 1 s off). Extraction of the periplasmatic fraction: added 0.1% (v/v) Triton to the total sample after sonication, and incubated 15 min at 4°C. Centrifugation at 24.700 g for 40 min at 4°C. Recovering of the soluble fraction and filtration using 0.45 µm syringe filters.
B	IMAC (HisTrap FF Crude (GE Healthcare), ÄKTA Pure 25 M1 (GE Healthcare)), binding with buffer **1A** supplemented with 50 mM imidazole, elution with imidazole gradient up to 500 mM in buffer **1A**.
C	TEV-cleavage (Protein/TEV ratio 1:10) at 4°C, o.n. in buffer **1B**.
D	Inv. IMAC (HisTrap FF Crude (GE Healthcare), ÄKTA Pure 25 M1 (GE Healthcare)), binding with buffer **1A** supplemented with 50 mM imidazole, elution with imidazole gradient up to 500 mM in buffer **1A**.
E	SEC on Increase 10/300 S75 (GE Healthcare) at 4°C in buffer **1C**.

Table 4: Final sample

1	**Yield**
ORF8 m	<0.5 mg/L LB mg/mL ^{15}N-M9 medium
ΔORF8	0.5 mg/L LB medium
2	**Stability**
ORF8	Not determined.

 doi:10.3389/fmolb.2021.653148

m	
ΔORF8	No significant precipitation or degradation observed after storage at 4°C for 1 week.
3	**Comment on applicability**
ORF8 m	Weak expression into solube fraction, 30%/70% soluble/inclusion bodies. After purification extremely low yield for NMR studies.
ΔORF8	Very low yield. It would be very expensive to prepare a labelled sample for NMR studies.

Additional information (bacterial expression)

Constructs	**Conditions**	**Comments**
ORF8 (with L84S mutation); His_6 (pPK1151 (Genscript)), TEV-cleavage site, N-terminal "GS" two artificial residues.	As above for ORF8m, only LB medium.	No expression.

Cell-free

Table 2: Cell-free Protein Synthesis

1	**Expression vector**
	pEU-E01-MCS (Cell-Free Sciences)
2	**Purification-/Solubility-Tag**
ORF8	C-terminal Strep tag II (WSHPQFEK)
ΔORF8	N-terminal Strep tag II (WSHPQFEK)
3	**Cleavage Site**
	-
4	**Molecular weight / Extinction coefficient / pI - of cleaved protein**
ORF8	15.00 kDa / 21,805 M^{-1} cm^{-1} / 5.64
ΔORF8	13.53 Da / 21,805 M^{-1} cm^{-1} / 5.39
5	**Comments on sequence of expressed construct**
ORF8	C-terminal "SAWSHPQFEK" ten artificial residues due to construct design.
ΔORF8	N-terminal "M" and C-terminal "SAWSHPQFEK" eleven artificial residues due to construct design.
6	**Feeding buffer**
	30 mM HEPES-KOH (pH 7.6), 100 mM potassium acetate, 2.7 mM magnesium acetate, 16 mM creatine phosphate, 0.4 mM spermidine, 1.2 mM ATP, 0.25 mM GTP, 4 mM DTT and 6 mM (average concentration) amino acid mix and 0.05% (w/v) Brij-58.
7	**Translation mix**
	50% (v/v) mRNA, 50% (v/v) home-made WGE, 40 μg/mL creatine kinase, and 6 mM (average concentration) amino acid mix 0.05% (w/v) Brij-58.

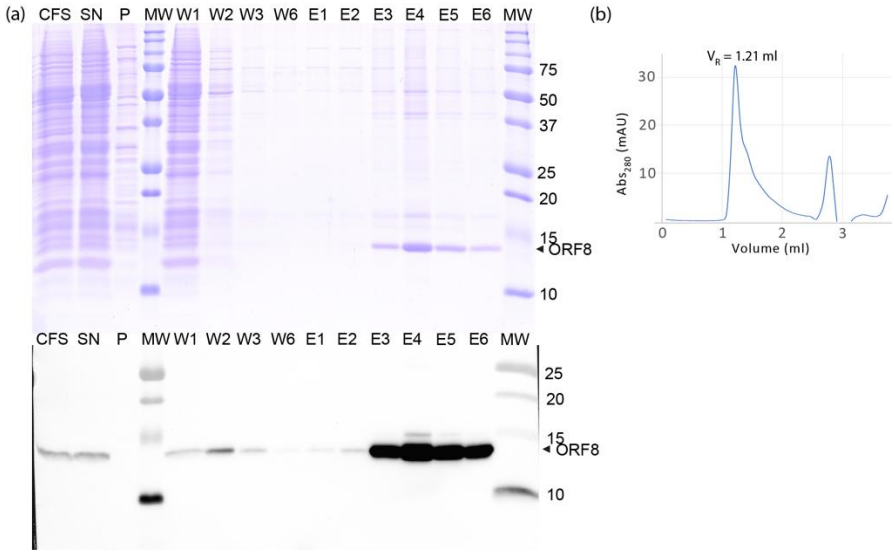
8	**Protein synthesis temperature and time**
	22°C for 16 h without agitation (bilayer method).

Table 3: Protein Purification (ORF8 and ΔORF8

1	**Buffer List**
A	100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA, 0.1% (w/v) DDM (wash buffer).
B	100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin, and 0.1% (w/v) DDM (elution buffer).
2	**Purification steps (with corresponding buffer(s) and incubation times)**
A	Harvest total CFS.
B	Incubate with benzonase for 30 min on a wheel, at rt.
C	Centrifuge for 30 min at 20,000 g, 4°C.
D	Harvest the soluble fraction (SN).
E	Equilibrate the Strep-Tactin column (IBA Lifesciences) with 2 CV of **1A** (all steps performed on the bench by gravity).
F	Load SN onto the column.
G	Wash the column with 5 CV of **1A**.
H	Elute the protein of interest with **1B**.

Table 4: Final sample

1	**Yield**
	ORF8: 0.62 mg/ml WGE after purification. Total of 683 µg for the NMR samples ΔORF8: small scale production; 0,28 ug/ml WGE
1b	**A260/280 ratio**
	0.7
2	**Stability**
	Stable at 4ºC for weeks.
3	**Comment on applicability**
	Protein very sensitive to dilution-concentration steps. Purity is sufficient for NMR as other cell-free proteins are not labelled.



(a) WG-CFPS in presence of detergent and Strep-tag purification of ORF8. SDS-PAGE (upper panel) and WB (lower panel). **(b) SEC profile of ORF8.**