SI4: Nsp5

Table 1: General Information

1	**Protein Name (according to NCBI Reference Sequence NC_045512.2)**
	ORF1a and ORF1ab; nsp5
2	**Region/Name/Further Specification**
	3C-like protease ($3CL^{pro}$) / main protease (M^{pro})
3	**Sequence of fl protein (according to NCBI Reference Sequence NC_045512.2)**
	SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSN HNFLVQAGNVQLRVIGHSMQNCVLKLKVDTANPKTPKYKFVRIQPGQTFSVLACYNGSPSGVY QCAMRPNFTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGNFYGPFVDRQ TAQAAGTDTTITVNVLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILG PLSAQTGIAVLDMCASLKELLQNGMNGRTILGSALLEDEFTPFDVVRQCSGVTFQ
4	**Protein boundaries of expressed construct (according to NCBI Reference Sequence NC_045512.2)**
	aa 1-306 (fl nsp5)
5	**Ratio for construct design**
	fl protein
6	**Sequence homology (to SCoV)**
	Identity: 96%; similarity: 99.7%
7	**Published structures (SCoV2 or homologue variants)**
	SCoV: PDB 1P9U, 6LU7 SCoV2: PDB 6Y2E, 5R7Y, 6Y84, 7K3T
8	**(Published) assignment (SCoV2 or homologue variants)**
	SCoV: BMRB 17251

Table 2: Protein Expression

1	**Expression vector**
	pE-SUMO (LifeSensors)
2	**Purification-/Solubility-Tag**
	N-terminal His_6-SUMO
3	**Cleavage Site**
	Ulp1
4	**Molecular weight / Extinction coefficient / pI - of cleaved protein**
	33.80 kDa / 32,890 M^{-1} cm^{-1} / 5.95
5	**Comments on sequence of expressed construct**
	No artificial residues due to TEV-cleavage and construct design.

6	**Used expression strain**
	E. coli BL21 (DE3)
7	**Cultivation medium**
	LB / M9 (uniformly ^{15}N-labelled)
8	**Induction system**
	IPTG inducible T7 promoter
9	**Induction of protein expression**
	0.2 mM IPTG at OD_{600} 0.6-0.7
10	**Cultivation temperature and time**
	18-20°C for 16-18 h

Table 3: Protein Purification

1	**Buffer List**
A	50 mM NaPi (pH 7.5), 300 mM NaCl, 5 mM imidazole, 5% (v/v) glycerol, 10 mM bME (cell disruption / IMAC).
B	50 mM NaPi (pH 7.0), 300 mM NaCl, 10 mM bME, 5% (v/v) glycerol (dialysis after IMAC / Ulp1-cleavage).
C	25 mM NaPi (pH 7.5), 150 mM NaCl, 2 mM TCEP-HCl (SEC buffer).
D	10 mM NaPi (pH 7.0), 0.5 mM TCEP-HCl (final NMR buffer).
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 microfluidization.
B	IMAC (HisTrap HP (GE Healthcare), ÄKTA start (GE Healthcare)), elution with imidazole gradient up to 500 mM in buffer **1A**.
C	Ulp1-cleavage (1 mg TEV protease per 50 mL protein solution) o.n. in buffer **1B**.
D	Inv. IMAC (HisTrap HP (GE Healthcare), ÄKTA start (GE Healthcare)), elution with 500 mM imidazole in buffer **1A**.
E	SEC (HiLoad 26/600 SD 75 pg (GE Healthcare), ÄKTApurifier (GE Healthcare)) in buffer **1C** (elution volume 170-210 mL).
F	NMR sample preparation in buffer **1D**.

Table 4: Final sample

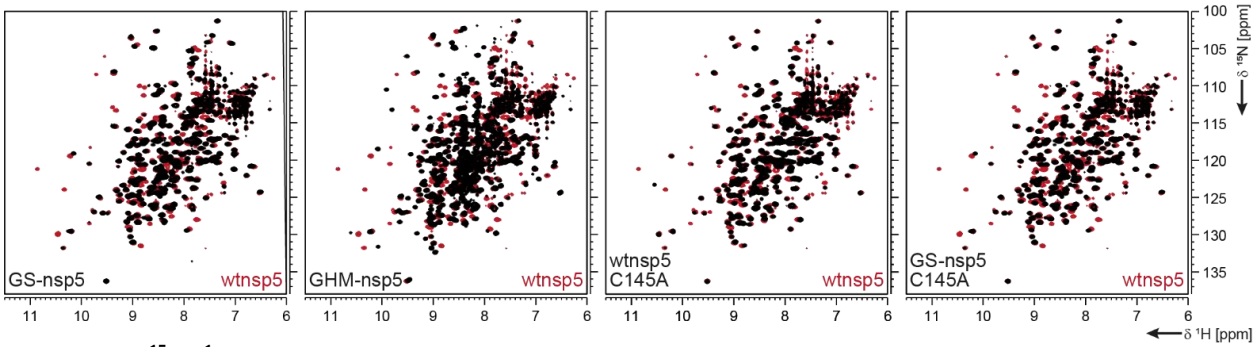
1	**Yield**
	55 mg/L ^{15}N-M9 medium
2	**Stability**
	No significant precipitation or degradation observed after storage at -80°C for a month.

3	**Comment on applicability**
	Suitable for NMR structure determination, fragment screening, interaction studies.

Additional information

	Constructs	**Conditions**	**Comments**
A	aa 1-306 (fl nsp5) C145A mutation; His_6-SUMO (pE-SUMO (LifeSensors)), Ulp1-cleavage site, no N-terminal artificial residues.	Native (as above)	Comparable to fl nsp5 expression and purification, similar yield (80 mg/L ^{15}N-M9 medium).
B	aa 1-306 (fl nsp5); His_6-SUMO (pE-SUMO (LifeSensors)), Ulp1-cleavage site, N-terminal "GS" two artificial residues.	Native (as above)	Comparable to fl nsp5 expression and purification, similar yield (55 mg/L ^{15}N-M9 medium, 36 mg/L ^{13}C,^{15}N-M9 medium, 20 mg/L ^{2}H,^{13}C,^{15}N E. coli-OD2 CDN medium (Silantes)).
C	aa 1-306 (fl nsp5) C145A mutation; His_6-SUMO (pE-SUMO (LifeSensors)), Ulp1-cleavage site, N-terminal "GS" two artificial residues.	Native (as above)	Comparable to fl nsp5 expression and purification, similar yield (55 mg/L ^{15}N-M9 medium).
D	aa 1-306 (fl nsp5); His_6 (pet28a(+) (GenScript)), TEV-cleavage site; N-terminal "GHM" three artificial residues.	Native (as above) **IMAC buffer** (**1A**): 25 mM Tris-HCl (pH 8.0), 300 mM NaCl, 5 mM imidazole, 5% (v/v) glycerol, 10 mM bME	Comparable to fl nsp5 purification, however, less expression/yield (35 mg/L ^{15}N-M9 medium, 10 mg/L ^{13}C,^{15}N-M9 medium).
E	aa 1-306 (fl nsp5); GST and His_6-tag (pET-28a(+) (GenScript)), TEV and auto cleavage site for M^{pro}, N-terminal „GS" and C-terminal "GPHHHHHH" ten artificial residues.	**IMAC buffer:** 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 20 mM Imidazole. **SEC-buffer:** 50 mM NaPi (pH 7.6), 50 mM NaCl, 0.02% (w/v) NaN_3. **NMR-buffer:** 50 mM NaPi (pH 7.6), 50 mM NaCl, 0.02% (w/v) NaN_3, 5 mM bME.	Yields 20 mg/L ^{15}N-M9 medium. The protein is stable up to 350 µM in NMR buffer at 25°C for at least 7 days. At 50 µM and at 4°C, the protein is stable for ~15 days. The protein is not suitable for freeze/thaw and results in precipitation.
F	aa 1-306 (fl nsp5); C-terminal His_6-tag (pET21b(+) (GenScript)), human rhinovirus 3-C protease cleavage site, N-terminal "M" additional aa, however our mass spectrum results suggest that M1 was removed by *E. coli* methionine aminopeptidase.	**IMAC buffer:** 20 mM Tris-HCl (pH 7.33), 150 mM NaCl, 20 mM imidazole. **Storage buffer:** 20 mM Tris-HCl (pH 7.33), 150 mM NaCl.	Yields 5 mg/L ^{15}N-M9 medium. Stable for 2-3 weeks at 4°C at low micromolar concentration.
G	aa 1-306 (fl nsp5) C145A mutation; His_6-GB1 (pET24a(+) (GenScript)), TEV-cleavage site, no artificial residues.	**IMAC buffer:** 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 20 mM imidazole, 0.5 mM TCEP-HCl. **SEC/NMR buffer:** 10 mM NaPi (pH 7.0), 0.5 mM TCEP-HCl.	Yields ≥ 70 mg/L ^{15}N, ^{2}H,^{15}N-M9, and ^{2}H^{13}C^{15}N-M9 medium. 1-2 mM sample stable for several weeks at 25°C. Negligible precipitation on freeze-thaw. Samples stable for ≥ 3 months at 80°C. Sample precipitation in buffer: 10 mM NaPi (pH 7.0), 0.4 M GdnHCl.
H	aa 1-306 (fl nsp5); His_6-GB1 (pET24a(+) (GenScript)), TEV-	As above (**G**).	Negligible expression when induced in ^{15}N-M9 medium at

	cleavage site, no artificial residues.		25˚C, 30˚C, and 37˚C, with 0.5-1 mM IPTG.
I	aa 1-306 (fl nsp5); His_6-GST (pGEX-6p-1 (Genewiz)), autolytic and HRV 3C cleavage site, no artificial residues.	**IMAC buffer:** 25 mM Tris-HCl (pH 7.8), 150 mM NaCl, 5 mM imidazole, 1 mM bME. **Cleavage buffer:** 25 mM Tris-HCl (pH 7.8), 150 mM NaCl, 1mM DTT. **SEC buffer:** 25 mM Tris-HCl (pH 7.8), 150 mM NaCl, 1 mM DTT, 1 mM EDTA.	40-60 mg/mL autoinduction Media ZYM-5052. Stored at 1 mg/mL at -20°C with 30% glycerol in SEC buffer. Also stored at 25 mg/mL at -80 ˚C in SEC buffer. Flash frozen. Neither show loss of activity compared to non-frozen samples.



Overlays of [^{15}N, ^{1}H]-BEST TROSY spectra of wtnsp5 (red) with the other constructs (black). From left to right: N-terminally GS added nsp5 (GS-nsp5), GHM added (GHM-nsp5), the active site mutants C145A with native N-terminus (wtnsp5 C145A), and GS added mutant (GS-nsp5 C145A).

 doi:10.3389/fmolb.2021.653148