SI2: Nsp2

Table 1: General Information

1	**Protein Name (according to NCBI Reference Sequence NC_045512.2)**
	ORF1a and ORF1ab; nsp2
2	**Region/Name/Further Specification**
	C-terminal IDR (CtDR)
3	**Sequence of fl protein (according to NCBI Reference Sequence NC_045512.2)**
	AYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQLDFIDTKRGVYCCREHEHEIAWYTE RSEKSYELQTPFEIKLAKKFDTFNGECPNFVFPLNSIIKTIQPRVEKKKLDGFMGRIRSVYPVASP NECNQMCLSTLMKCDHCGETSWQTGDFVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYCP ACHNSEVGPEHSLAEYHNESGLKTILRKGGRTIAFGGCVFSYVGCHNKCAYWVPRASANIGCN HTGVVGEGSEGLNDNLLEILQKEKVNINIVGDFKLNEEIAIILASFSASTSAFVETVKGLDYKAFK QIVESCGNFKVTKGKAKKGAWNIGEQKSILSPLYAFASEAARVVRSIFSRTLETAQNSVRVLQK AAITILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKLKPVL DWLEEKFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKEIKESVQTFFKLVNKFLALCADSII IGGAKLKALNLGETFVTHSKGLYRKCVKSREETGLLMPLKAPKEIIFLEGETLPTEVLTEEVVLK TGDLQPLEQPTSEAVEAPLVGTPVCINGLMLLEIKDTEKYCALAPNMMVTNNTFTLKGG
4	**Protein boundaries of expressed construct (according to NCBI Reference Sequence NC_045512.2)**
	aa 557-601 of complete nsp2 (Ct-DR)
5	**Ratio for construct design**
	Based on disorder predictions (PrDOS (Ishida and Kinoshita, 2007))
6	**Sequence homology (to SCoV)**
	Identity: 55%; similarity: 68%
7	**Published structures (SCoV2 or homologue variants)**
	-
8	**(Published) assignment (SCoV2 or homologue variants)**
	SCoV: 50687

Table 2: Protein Expression

1	**Expression vector**
	Home made plasmid derived from pET28b(+) (EMD Biosciences) containing the codifying sequence for thioredoxin A from *E. coli* and TEV protease cleavage site instead of thrombin.
2	**Purification-/Solubility-Tag**
	N-terminal His_6-Trx
3	**Cleavage Site**
	TEV
4	**Molecular weight / Extinction coefficient / pI - of cleaved protein**

	4.92 kDa / - / 3.9
5	**Comments on sequence of expressed construct**
	N-terminal „G“, one artificial residue due to TEV-cleavage.
6	**Used expression strain**
	E. coli BL21 star (DE3)
7	**Cultivation medium**
	LB / M9 (uniformly ^{15}N or ^{13}C,^{15}N-labelled)
8	**Induction system**
	IPTG inducible T7 promoter
9	**Induction of protein expression**
	0.5 mM IPTG at OD_{600} 0.6
10	**Cultivation temperature and time**
	37ºC until induction. Following induction, incubation at 25ºC for 17 h

Table 3: Protein Purification

1	**Buffer List**
A	50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 10 mM imidazole (cell lysis, IMAC1 and 2).
B	5 mM Tris-HCl (pH 8.0), 20 mM NaCl (dialysis after IMAC1/TEV cleavage).
C	5 mM histidine (pH 5.4), 5 mM NaCl (dialysis after IMAC2 and anionic IEC).
D	10 mM acetic acid (pH 4.3), 5 mM NaCl (dialysis after cationic IEC).
2	**Purification steps (with corresponding buffer(s) and incubation times)**
A	Cell lysis in **1A** (plus 5 µl Halt protease inhibitor (Thermo) and lysozyme 20 µg/ml).
B	IMAC1 (HisTrap crude 5 ml (Cytiva). Elution 10-500 mM imidazole in buffer **1A**.
C	Dialysis in buffer **1B** and TEV cleavage (4ºC, 17 h).
D	IMAC2 (after TEV cleavage) (HisTrap crude 5 ml (Cytiva). Elution 10-500 mM imidazole in buffer **1A** (protein expected in flow-through).
E	Dialysis in buffer **1C** (4ºC, 17 h).
F	Anionic IEC, elution 10-1,000 mM NaCl in buffer **1C**.
G	Dialysis in buffer **1C** (4ºC, 48 h).
H	Cationic IEC. Elution 10-1,000 mM NaCl in buffer **1D** (protein expected in flow-through).

 doi:10.3389/fmolb.2021.653148

Table 4: Final sample

1	**Yield**
	1.5 mg/L LB medium, 0.7-1.5 mg/L ^{13}C, ^{15}N-M9 medium
2	**Stability**
	No visible precipitation after two weeks at 4ºC.
3	**Comment on applicability**
	Suitable for NMR structure determination, fragment screening, interaction studies.

 doi:10.3389/fmolb.2021.653148