



# $^1\text{H}$ , $^{13}\text{C}$ and $^{15}\text{N}$ backbone chemical shift assignments of SARS-CoV-2 nsp3a

Nicola Salvi<sup>1</sup> · Luiza Mamigonian Bessa<sup>1</sup> · Serafima Guseva<sup>1</sup> · Aldo Camacho-Zarco<sup>1</sup> · Damien Maurin<sup>1</sup> · Laura Marino Perez<sup>1</sup> · Anas Malki<sup>1</sup> · Martin Hengesbach<sup>2,4</sup> · Sophie Marianne Korn<sup>3,4</sup> · Andreas Schlundt<sup>3,4</sup> · Harald Schwalbe<sup>2,4</sup> · Martin Blackledge<sup>1</sup>

Received: 19 October 2020 / Accepted: 24 December 2020 / Published online: 21 January 2021  
© The Author(s), under exclusive licence to Springer Nature B.V. part of Springer Nature 2021

## Abstract

The non-structural protein nsp3 from SARS-CoV-2 plays an essential role in the viral replication transcription complex. Nsp3a constitutes the N-terminal domain of nsp3, comprising a ubiquitin-like folded domain and a disordered acidic chain. This region of nsp3a has been linked to interactions with the viral nucleoprotein and the structure of double membrane vesicles. Here, we report the backbone resonance assignment of both domains of nsp3a. The study is carried out in the context of the international *covid19-nmr* consortium, which aims to characterize SARS-CoV-2 proteins and RNAs, providing for example NMR chemical shift assignments of the different viral components. Our assignment will provide the basis for the identification of inhibitors and further functional and interaction studies of this essential protein.

**Keywords** SARS-CoV-2 · Intrinsically disordered protein · Covid-19 · Viral replication

## Biological context

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) represents a significant threat to human health and to the stability of contemporary societies. The development of vaccines or potential inhibitors is essential if the disease is to be eradicated. One important step in the development of effective therapeutic strategies is the characterization of potentially druggable targets that constitute the functioning virus. The international *covid19-nmr* research consortium is dedicated to the measurement and rapid dissemination of

NMR-based parameters that will further deepen our understanding of the viral life cycle and in particular facilitate the identification of potential drug binding sites.

With over 1900 amino acids the non-structural protein nsp3 is the largest of the coronavirus proteins (Snijder et al. 2003). It comprises a large number of structured domains and disordered linker regions in SARS-CoV (Neuman 2016), and is an essential component of the replication-transcription complex (Lei et al. 2018). As such this protein represents a promising target for inhibitory strategies. Nsp3a constitutes the N-terminal section of nsp3, comprising an N-terminal ubiquitin-like domain (Ub11) and a hypervariable acidic intrinsically disordered region (IDR) (Neuman et al. 2014).

Nsp3a has been shown to interact with the nucleoprotein (N) in the related beta-coronavirus mouse hepatitis virus (MHV) (Keane and Giedroc 2013). The acidic IDR domain has a high Glu/Asp content and is predicted to be 23 amino acids longer in SARS-CoV-2 compared to SARS-CoV. In SARS-CoV, pulldown experiments show nsp3a to associate with viral proteins nsp8, nsp9, components of the replication/transcription complex, as well as different domains of nsp3 (Imbert et al. 2008). The ubiquitin-like domain colocalizes with the nucleoprotein in viral liquid droplets in SARS-CoV-2 (Carlson et al. 2020) and has been shown to bind single stranded RNA in SARS-CoV (Serrano et al.

✉ Martin Blackledge  
martin.blackledge@ibs.fr

<sup>1</sup> Univ. Grenoble Alpes, CNRS, CEA, IBS, 38000 Grenoble, France

<sup>2</sup> Institute for Organic Chemistry and Chemical Biology, Johann Wolfgang Goethe-University Frankfurt, Max-von-Laue-Str. 7, 60438 Frankfurt, Germany

<sup>3</sup> Institute for Molecular Biosciences, Johann Wolfgang Goethe-University Frankfurt, Max-von-Laue-Str. 9, 60438 Frankfurt, Germany

<sup>4</sup> Center for Biomolecular Magnetic Resonance (BMRZ), Johann Wolfgang Goethe-University Frankfurt, 60438 Frankfurt, Germany

2007). Recently nsp3 was shown to play a structural role in the formation of molecular pores in double membrane vesicles associated with the endoplasmic reticulum (Stertz et al. 2007; Wolff et al. 2020).

Here, we present the near-complete backbone resonance assignment of the two domains of nsp3a, providing the basic data required for screening of interaction partners and detailed mapping of the interaction using the homologous structures of nsp3a from SARS-CoV (Serrano et al. 2007) and MHV (Keane and Giedroc 2013).

## Methods and experiments

### Construct design

The primary sequence of nsp3a (amino acids 1–206 of nsp3) from SARS-CoV-2 was extracted from NCBI genome entry NC\_045512.2 [GenBank entry MN908947.3]. A commercially synthesized gene (GenScript Biotech) was codon-optimized for expression in *Escherichia coli* and subcloned in a pET21b(+) vector. Hexa-histidine and TEV-cleavage tags were included at the N-terminus to facilitate protein purification. After protease cleavage, the proteins contain N-terminal GAM- extensions.

### Sample preparation

The nsp3a plasmid was transformed into BL21 (DE3) *E. coli* cells and the protein expressed heterologously with an N-terminal His<sub>6</sub> tag. Cells were grown at 37 °C until OD<sub>600</sub> of 0.6–0.8, at which point protein expression was induced with IPTG and incubated for 5 h at 37 °C. Bacteria were harvested by centrifugation and the cell pellet resuspended in buffer A (50 mM Tris–HCl pH 8.0 and 250 mM NaCl) with protease inhibitors (complete, Roche). Cell lysis was performed by sonication, followed by centrifugation (45 min, 18,000 rpm at 5 °C).

The protein was purified by affinity chromatography on Ni–NTA agarose (ThermoFisher), washed with buffer A

supplemented with 20 mM imidazole and eluted with buffer A supplemented with 500 mM imidazole. TEV cleavage was achieved by incubation with TEV protease at 4 °C coupled with dialysis into buffer A supplemented with 2 mM DTT, and the protein concentrated and subjected to size exclusion chromatography on a HiLoad 16/600 Superdex 75 column (GE Healthcare) in NMR buffer (50 mM Na-phosphate, pH 6.5, 150 mM NaCl). For <sup>15</sup>N and <sup>13</sup>C isotope labelling, cells were grown in M9-minimal medium supplemented with <sup>15</sup>N-NH<sub>4</sub>-Cl and <sup>13</sup>C<sub>6</sub>-D-glucose (1 g/L each).

### NMR experiments

A suite of BEST- and BEST-TROSY (BT) double and triple resonance assignment experiments, including BEST-HNCA, BEST HN(CO)CA, BT-HNCO, BT-HN(CO)CACB and intraresidue BT-iHNCACB (Lescop et al. 2007; Solyom et al. 2013), were recorded on <sup>15</sup>N,<sup>13</sup>C-labeled samples (632 μM) at 298 K using a Bruker Avance III spectrometer equipped with a cryoprobe at a <sup>1</sup>H frequency of 850 MHz. The experimental parameters for these acquisitions are summarized in Table 1. All spectra were processed using NMRFX Analyst (Norris et al. 2016) and analysed using CCPNMR Analysis Assign (Skinner et al. 2016) and NMR-FAM-SPARKY (Lee et al. 2015). Manual assignment of residues was assisted using I-PINE (Lee et al. 2019).

### Assignments and data deposition

The <sup>15</sup>N,<sup>1</sup>H-HSQC of Nsp3a(1–206) is typical of a two-domain protein comprising both folded and unfolded domains (Fig. 1). The folded domain is well-resolved while the unfolded domain has a more restricted chemical shift dispersion in the <sup>1</sup>H dimension. Nevertheless, a high percentage of resonances could be assigned in both domains (92% <sup>1</sup>H<sup>N</sup>, 92% <sup>15</sup>N<sup>H</sup>, 93% <sup>13</sup>C<sub>α</sub>, 38% <sup>13</sup>C<sub>β</sub> and 83% <sup>13</sup>C<sup>γ</sup>). <sup>13</sup>C<sub>β</sub> assignments are low because of difficulty in detecting transfer to these nuclei in the 3D experiments. This assignment has been deposited in the biological magnetic resonance data-bank (BMRB ID: 50446).

**Table 1** Acquisition parameters used in the assignment experiments

| Experiments                          | Time domain data size (points) |                |                | Spectral width (ppm)    |                        |                         | ns | Delay time (s) |
|--------------------------------------|--------------------------------|----------------|----------------|-------------------------|------------------------|-------------------------|----|----------------|
|                                      | t <sub>1</sub>                 | t <sub>2</sub> | t <sub>3</sub> | F <sub>1</sub>          | F <sub>2</sub>         | F <sub>3</sub>          |    |                |
| <sup>1</sup> H- <sup>15</sup> N-HSQC | 300                            | 2048           |                | 28.0 ( <sup>15</sup> N) | 12.0 ( <sup>1</sup> H) |                         | 16 | 1              |
| BT-HNCO                              | 80                             | 1450           | 90             | 28.0 ( <sup>15</sup> N) | 12.0 ( <sup>1</sup> H) | 13.0 ( <sup>13</sup> C) | 8  | 0.2            |
| BT-HN(CA)CO                          | 80                             | 2048           | 90             | 28.0 ( <sup>15</sup> N) | 12.0 ( <sup>1</sup> H) | 13.0 ( <sup>13</sup> C) | 16 | 0.2            |
| BEST-HNCA                            | 128                            | 1452           | 74             | 28.0 ( <sup>15</sup> N) | 12.0 ( <sup>1</sup> H) | 30.0 ( <sup>13</sup> C) | 24 | 0.2            |
| BEST-HN(CO)CA                        | 128                            | 1452           | 74             | 28.0 ( <sup>15</sup> N) | 12.0 ( <sup>1</sup> H) | 30.0 ( <sup>13</sup> C) | 24 | 0.2            |
| BT-iHNCACB                           | 128                            | 2048           | 70             | 28.0 ( <sup>15</sup> N) | 12.0 ( <sup>1</sup> H) | 62.0 ( <sup>13</sup> C) | 32 | 0.2            |
| BT-HN(CO)CACB                        | 128                            | 2048           | 70             | 28.0 ( <sup>15</sup> N) | 12.0 ( <sup>1</sup> H) | 62.0 ( <sup>13</sup> C) | 32 | 0.2            |



- Lei J, Kusov Y, Hilgenfeld R (2018) Nsp3 of coronaviruses: structures and functions of a large multi-domain protein. *Antivir Res* 149:58–74. <https://doi.org/10.1016/j.antiviral.2017.11.001>
- Lescop E, Schanda P, Brutscher B (2007) A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J MagnReson* 187:163–169. <https://doi.org/10.1016/j.jmr.2007.04.002>
- Neuman BW (2016) Bioinformatics and functional analyses of coronavirus nonstructural proteins involved in the formation of replicative organelles. *Antivir Res* 135:97–107. <https://doi.org/10.1016/j.antiviral.2016.10.005>
- Neuman BW, Chamberlain P, Bowden F, Joseph J (2014) Atlas of coronavirus replicase structure. *Virus Res* 194:49–66. <https://doi.org/10.1016/j.virusres.2013.12.004>
- Norris M, Fetler B, Marchant J, Johnson BA (2016) NMRFX processor: a cross-platform NMR data processing program. *J Biomol NMR* 65:205–216. <https://doi.org/10.1007/s10858-016-0049-6>
- Serrano P, Johnson MA, Almeida MS et al (2007) Nuclear magnetic resonance structure of the N-terminal domain of nonstructural protein 3 from the severe acute respiratory syndrome coronavirus. *J Virol* 81:12049–12060. <https://doi.org/10.1128/JVI.00969-07>
- Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* 56:227–241. <https://doi.org/10.1007/s10858-013-9741-y>
- Skinner SP, Fogh RH, Boucher W et al (2016) CcpNmr analysis assign: a flexible platform for integrated NMR analysis. *J Biomol NMR* 66:111–124. <https://doi.org/10.1007/s10858-016-0060-y>
- Snijder EJ et al (2003) Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J Mol Biol* 331:991–1004. [https://doi.org/10.1016/s0022-2836\(03\)00865-9](https://doi.org/10.1016/s0022-2836(03)00865-9)
- Solyom Z, Schwarten M, Geist L et al (2013) BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *J Biomol NMR* 55:311–321. <https://doi.org/10.1007/s10858-013-9715-0>
- Stertz S, Reichelt M, Spiegel M et al (2007) The intracellular sites of early replication and budding of SARS-coronavirus. *Virology* 361:304–315. <https://doi.org/10.1016/j.virol.2006.11.027>
- Wolff G, Melia CE, Snijder EJ, Bárcena M (2020) Double-membrane vesicles as platforms for viral replication. *Trends Microbiol*. <https://doi.org/10.1016/j.tim.2020.05.009>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.