

^1H , ^{13}C , and ^{15}N backbone chemical-shift assignments of SARS-CoV-2 non-structural protein 1 (leader protein)

Wang, Ying; Kirkpatrick, John; Lage, Susanne zur; Korn, Sophie M.; Neißner, Konstantin; Schwalbe, Harald; Schlundt, Andreas; Carlomagno, Teresa

DOI:

[10.1007/s12104-021-10019-6](https://doi.org/10.1007/s12104-021-10019-6)

License:

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Wang, Y., Kirkpatrick, J., Lage, SZ, Korn, SM, Neißner, K, Schwalbe, H, Schlundt, A & Carlomagno, T 2021, ' ^1H , ^{13}C , and ^{15}N backbone chemical-shift assignments of SARS-CoV-2 non-structural protein 1 (leader protein)', *Biomolecular NMR Assignments*, vol. 15, no. 2, pp. 287–295. <https://doi.org/10.1007/s12104-021-10019-6>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.



^1H , ^{13}C , and ^{15}N backbone chemical-shift assignments of SARS-CoV-2 non-structural protein 1 (leader protein)

Ying Wang¹ · John Kirkpatrick^{1,2} · Susanne zur Lage² · Sophie M. Korn^{3,5} · Konstantin Neißner^{3,5} · Harald Schwalbe^{4,5} · Andreas Schlundt^{3,5} · Teresa Carlomagno^{1,2}

Received: 29 January 2021 / Accepted: 13 March 2021 / Published online: 26 March 2021
© The Author(s) 2021

Abstract

The current COVID-19 pandemic caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) has become a worldwide health crisis, necessitating coordinated scientific research and urgent identification of new drug targets for treatment of COVID-19 lung disease. The *covid19-nmr* consortium seeks to support drug development by providing publicly accessible NMR data on the viral RNA elements and proteins. The SARS-CoV-2 genome comprises a single RNA of about 30 kb in length, in which 14 open reading frames (ORFs) have been annotated, and encodes approximately 30 proteins. The first two-thirds of the SARS-CoV-2 genome is made up of two large overlapping open-reading-frames (ORF1a and ORF1b) encoding a replicase polyprotein, which is subsequently cleaved to yield 16 so-called non-structural proteins. The non-structural protein 1 (Nsp1), which is considered to be a major virulence factor, suppresses host immune functions by associating with host ribosomal complexes at the very end of its C-terminus. Furthermore, Nsp1 facilitates initiation of viral RNA translation via an interaction of its N-terminal domain with the 5' untranslated region (UTR) of the viral RNA. Here, we report the near-complete backbone chemical-shift assignments of full-length SARS-CoV-2 Nsp1 (19.8 kDa), which reveal the domain organization, secondary structure and backbone dynamics of Nsp1, and which will be of value to further NMR-based investigations of both the biochemical and physiological functions of Nsp1.

Keywords SARS-CoV-2 · Nsp1 · Non-structural proteins · New drug targets · 5' untranslated region · NMR spectroscopy

Biological context

The ongoing COVID-19 pandemic has initiated intense scientific research into the causative pathogen, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Previous studies have demonstrated that SARS-CoV-2 shows close sequence homology to SARS-CoV-1, which was responsible for the SARS epidemic in 2003, and also to MERS-CoV, which causes Middle-Eastern Respiratory Syndrome (Benedetti et al. 2020; A. Wu et al. 2020a, b). With a high transmissibility and potential to induce life-threatening acute respiratory distress syndrome, SARS-CoV-2 represents a severe threat to human health worldwide. Rapid identification and characterization of druggable protein targets within the SARS-CoV-2 genome is the first step towards development of targeted therapeutic treatments for COVID-19, which together with vaccination approaches, will be critically important in mitigating the impact of the pandemic.

SARS-CoV-2 belongs to the family of beta-coronaviruses, with an enveloped, positive-sense single-stranded

✉ Teresa Carlomagno
teresa.carlomagno@oci.uni-hannover.de;
covid19-nmr@dlist.server.uni-frankfurt.de

¹ Centre of Biomolecular Drug Research (BMWZ), Leibniz University Hannover, Schneiderberg 38, 30167 Hannover, Germany

² Group of NMR-Based Structural Chemistry, Helmholtz Centre for Infection Research, Inhoffenstrasse 7, 38124 Braunschweig, Germany

³ Institute for Molecular Biosciences, St Lucia, QLD 4072, Australia

⁴ Institute for Organic Chemistry and Chemical Biology, 60438 Frankfurt, Germany

⁵ Center for Biomolecular Magnetic Resonance (BMRZ), Johann Wolfgang Goethe-University Frankfurt, Max-von-Laue-Str. 7, 60438 Frankfurt, Germany

RNA genome (Gorbalenya et al. 2020), which encodes two large overlapping open reading frames (ORF1a and ORF1b) at the 5'-end, as well as four structural proteins and eight accessory proteins at the 3'-end (Lim et al. 2016; Zhou et al. 2020). After cell entry, the virus exploits the host translational machinery to produce the polypeptides corresponding to ORF1a and ORF1b. The polypeptides are then proteolytically cleaved into 16 functional non-structural proteins (Khailany et al. 2020), most of which are involved in assembling the host–viral replication/transcription complex (Masters 2006). Among these proteins, non-structural protein 1 (Nsp1), also known as the leader protein, plays a key role in hampering host gene expression.

SARS-CoV-2 Nsp1, the first N-terminal cleavage protein from the replicase polyprotein, is 180 residues in length, and shares ~85% amino-acid sequence identity with the homologous protein from SARS-CoV-1. SARS-CoV-1 Nsp1 has been structurally characterized using solution-state NMR methods, leading to an atomic-resolution structure of the globular domain comprising residues 13–128 and revealing short N-terminal and long C-terminal tails (residues 1–12 and 129–179, respectively) that are flexibly disordered (Almeida et al. 2007). Recent cryo-electron microscopy (cryo-EM) studies on Nsp1 from SARS-CoV-2 have demonstrated that host protein translation inhibition by Nsp1 is mediated by insertion of its C-terminus into the entrance of the mRNA tunnel in the small ribosomal subunit (Schubert et al. 2020; Thoms et al. 2020). In addition, it was shown biochemically that the 5' UTR of SARS-CoV-2 is capable of promoting translation initiation of viral mRNAs through binding to the N-terminal domain (NTD) of Nsp1 (Schubert et al. 2020). However, the precise mechanism by which SARS-CoV-2 escapes from the Nsp1-mediated translation inhibition—thereby switching the host translation machinery from host to viral protein synthesis—and which specific region of SARS-CoV-2 5' UTR binds to the NTD of Nsp1 are still unclear. Due to its dual role in inhibition of host-protein translation and stimulation of viral-protein translation, Nsp1 has been proposed as an attractive drug target for the treatment of COVID-19.

The research consortium *covid19-nmr* seeks to rapidly and publicly support the search for anti-viral COVID-19 drugs using an NMR-based screening approach that in the initial stage involves the production of all druggable proteins and RNAs from SARS-CoV-2 and their subsequent NMR chemical-shift assignment, followed by a second stage of solution-structure determination, drug-library screening and rational structure-based drug-design. Here we report the near-complete backbone assignment of full-length SARS-CoV-2 Nsp1 that represents the first step towards its structural characterization and provides a basis for residue-resolved drug-screening and protein–RNA interaction studies.

Methods and experiments

Construct design

The Nsp1 protein sequence studied here derives from the SARS-CoV-2 genomic sequence corresponding to NCBI GenBank entry NC_045512.2, which is identical to GenBank entry MN908947.3 (F. Wu et al. 2020a, b). A codon-optimized expression construct of SARS-CoV-2 full-length Nsp1 was inserted into the vector pETM-11 (EMBL bacterial expression vector database), containing an N-terminal His₆-tag, and a tobacco etch virus (TEV) cleavage site. Due to the nature of the TEV cleavage site, two artificial N-terminal residues (Gly[–1] and Ala[0]) are included in the cleaved protein, before the native Nsp1 sequence starts at Met[1] (corresponding to NCBI GenPept entry YP_009725297.1).

For the NTD of SARS-CoV-2 Nsp1, domain boundaries were defined in analogy to the available NMR structure (PDB codes 2hsx/2gdt (Almeida et al. 2007)) of the SARS-CoV-1 homologue Nsp1 (85% sequence identity). The expression construct contained residues 13–127 of the Nsp1 sequence. The construct was amplified from the full-length Nsp1 gene, obtained as an *E. coli*-codon-optimized DNA construct from GenScript Biotech (Netherlands). It was inserted into the pET3b-based vector pKM263, containing an N-terminal His₆-tag, a GST-tag and a tobacco etch virus (TEV) cleavage site. Four artificial N-terminal residues (Gly[–3], Ala[–2], Met[–1] and Gly[0]) remained after proteolytic TEV cleavage, resulting in a 12.9 kDa protein.

Protein expression and purification

Large-scale expressions of uniformly ¹³C,¹⁵N-labelled full-length Nsp1 in *E. coli* BL21 (DE3) cells were carried out in M9 minimal medium, containing 1 g/L ¹⁵NH₄Cl (Cambridge Isotope Laboratories), 2.5 g/L ¹³C₆-D-glucose (Cambridge Isotope Laboratories) and 50 µg/mL kanamycin. The cells were cultured at 37 °C until the OD₆₀₀ reached 0.7, at which point protein expression was induced with 0.6 mM isopropyl-beta-thiogalactopyranoside (IPTG). The temperature was reduced to 16 °C and expression continued for 22 h. After harvesting (4 °C, 4000 rpm for 40 min), cell pellets were stored at –20 °C. For protein purification, the cells were resuspended in buffer A (50 mM Tris–HCl, 500 mM sodium chloride, 100 mM sodium sulfate, 5% v/v glycerol, 5 mM imidazole, 1 mM Tris-(2-carboxyethyl)-phosphine (TCEP), pH 7.5) with one tablet of EDTA-free protease inhibitor cocktail (Roche), 100 µg of lysozyme (Carl Roth), and 50 µg of

deoxyribonuclease (DNase) (New England Biolabs) per 1 L of original culture volume. Cell lysis was performed by sonication (50% power; 5 s pulse/10 s rest duty-cycle, for 20 min). The cell lysate was cleared by centrifugation (4 °C, 18000 rpm for 1 h). The supernatant was filtered and loaded onto a HisTrap HP column (GE Healthcare), washed first with buffer A and then with buffer A containing additional 2 M LiCl, before elution with a linear gradient of buffer A to buffer B (composition same as buffer A but with 300 mM imidazole). Nsp1-containing fractions of the eluate were exchanged back into buffer A using a desalting column, and then incubated with TEV protease (0.5 mg per 1 L of original culture volume) overnight at 4 °C. The cleaved Nsp1 was separated from the TEV protease and residual uncleaved protein by passing the dialysis mixture over the HisTrap column and washing with buffer A. The flow-through was collected, concentrated and loaded onto a HiLoad Superdex 75 16/600 size-exclusion column (GE Healthcare), previously equilibrated in buffer C (50 mM Tris-HCl, 500 mM sodium chloride, 100 mM sodium sulfate, 1 mM EDTA, 1 mM TCEP, pH 7.5). Pure Nsp1-containing fractions were identified by SDS-PAGE, pooled and exchanged into NMR buffer (50 mM sodium phosphate (pH 6.5), 200 mM sodium chloride, 2 mM dithiothreitol, 2 mM ethylene diamine tetra-acetic acid, 0.01% w/v sodium azide, 0.001% w/v 3-(trimethylsilyl)propane-1-sulfonate) by means of repeated dilution/concentration using Amicon centrifugal concentrators (10-kDa molecular-weight cutoff).

Uniformly ^{15}N -labelled Nsp1-NTD (residues 13–127) was expressed in *E. coli* strain BL21 (DE3) in M9 minimal medium containing 1 g/L $^{15}\text{NH}_4\text{Cl}$ (Cambridge Isotope Laboratories) and 100 $\mu\text{g}/\text{mL}$ ampicillin. Expression was induced at an OD_{600} of 0.7 with 1 mM IPTG for 18 h at 16 °C. Cell pellets were resuspended in Buffer D (50 mM Tris-HCl, 300 mM NaCl, 10 mM imidazole, 4 mM DTT and 100 μL protease inhibitor mix (SERVA), pH 8.0). The supernatant was cleared by centrifugation (40 min, 10000 g, 4 °C) and subsequently loaded onto a Ni^{2+} -NTA gravity-flow column (Sigma Aldrich). Cleavage of the His₆-GST-tag was achieved overnight at 4 °C with 0.5 mg of TEV protease per 1L of culture while dialyzing into Buffer D. The TEV protease and the cleaved tag were removed via a second Ni^{2+} -NTA gravity-flow column. Further purification of Nsp1-NTD by size exclusion chromatography (HiLoad 16/600 SD 75, GE Healthcare) was carried out in Buffer E (25 mM NaPi, 250 mM NaCl, 2 mM TCEP, 0.02% w/v NaN_3 , pH 7.0). Fractions containing pure Nsp1-NTD were determined by SDS-PAGE, pooled and concentrated using Amicon centrifugal concentrators (3-kDa cutoff). Final NMR samples were prepared in Buffer E, containing 5% v/v D_2O at Nsp1-NTD concentrations of 15 μM .

NMR experiments

NMR samples (~550 μL total volume in 5-mm-diameter NMR tubes) were prepared with Nsp1 at a concentration of 300–400 μM , dissolved in NMR buffer. The protein appeared relatively stable over a period of several days, except for the gradual appearance of a set of small, sharp peaks characteristic of the dipeptides resulting from proteolysis. Appearance of these peaks was also accompanied by small shift-changes in the main set of peaks.

NMR experiments on full-length Nsp1 were recorded at 298 K on two Bruker Avance III-HD spectrometers running Topspin 3.2 software, with ^1H field-strengths of 850 MHz and 600 MHz, and equipped with inverse HCN CP-TCI (helium-cooled) and CPP-TCI (nitrogen-cooled) cryogenic probeheads, respectively.

2D ^{15}N -HSQC spectra were recorded using States-TPPI for frequency discrimination, with water suppression achieved via a combination of WATERGATE and water flip-back pulses to preserve the water magnetization (Bodenhausen & Ruben 1980; Piotto et al. 1992). Backbone resonance assignments were obtained from a standard-suite of 3D triple-resonance out-and-back-type spectra, comprising HNCO (Ikura et al. 1990; Kay et al. 1990), HN(CA)CO (Clubb et al. 1992; Kay et al. 1994), HNCA, HNCACB (Grzesiek & Bax 1992; Wittekind & Mueller 1993), HN(CO)CA and HN(CO)CACB (Bax & Ikura 1991; Yamazaki et al. 1994) spectra. With the exception of the HNCO experiment, which was recorded at both field-strengths, the through-carbonyl and the HN(CA)CO spectra were recorded at 600 MHz, while the remaining spectra were recorded at 850 MHz. Water suppression and frequency discrimination in the triple-resonance spectra were achieved as for the ^{15}N -HSQC spectra.

For the Nsp1 NTD, 2D ^{15}N -HSQC-TROSY experiments were measured at 950 MHz with acceleration of longitudinal ^1H relaxation between scans via the Band-Selective Excitation Short-Transient (BEST) approach (Favier & Brutscher 2011; Lescop et al. 2007; Solyom et al. 2013) using exclusively shaped-pulses on ^1H (bandwidth and offset of 4.5 and 8.5 ppm, respectively) and an inter-scan delay of 0.3 s.

NMR data were processed with NMRPipe v10.1 (Delaglio et al. 1995) and analysed in CcpNmr Analysis v2.4 (Vranken et al. 2005).

Assignments and data deposition

The ^{15}N -HSQC spectrum of SARS-CoV-2 full-length Nsp1 has the appearance expected for a protein comprising a well-folded globular domain together with an extended and largely disordered tail (Fig. 1). The amide peaks from the residues in the globular domain are widely dispersed with a ^1H chemical-shift dispersion of ~4 ppm, while the peaks

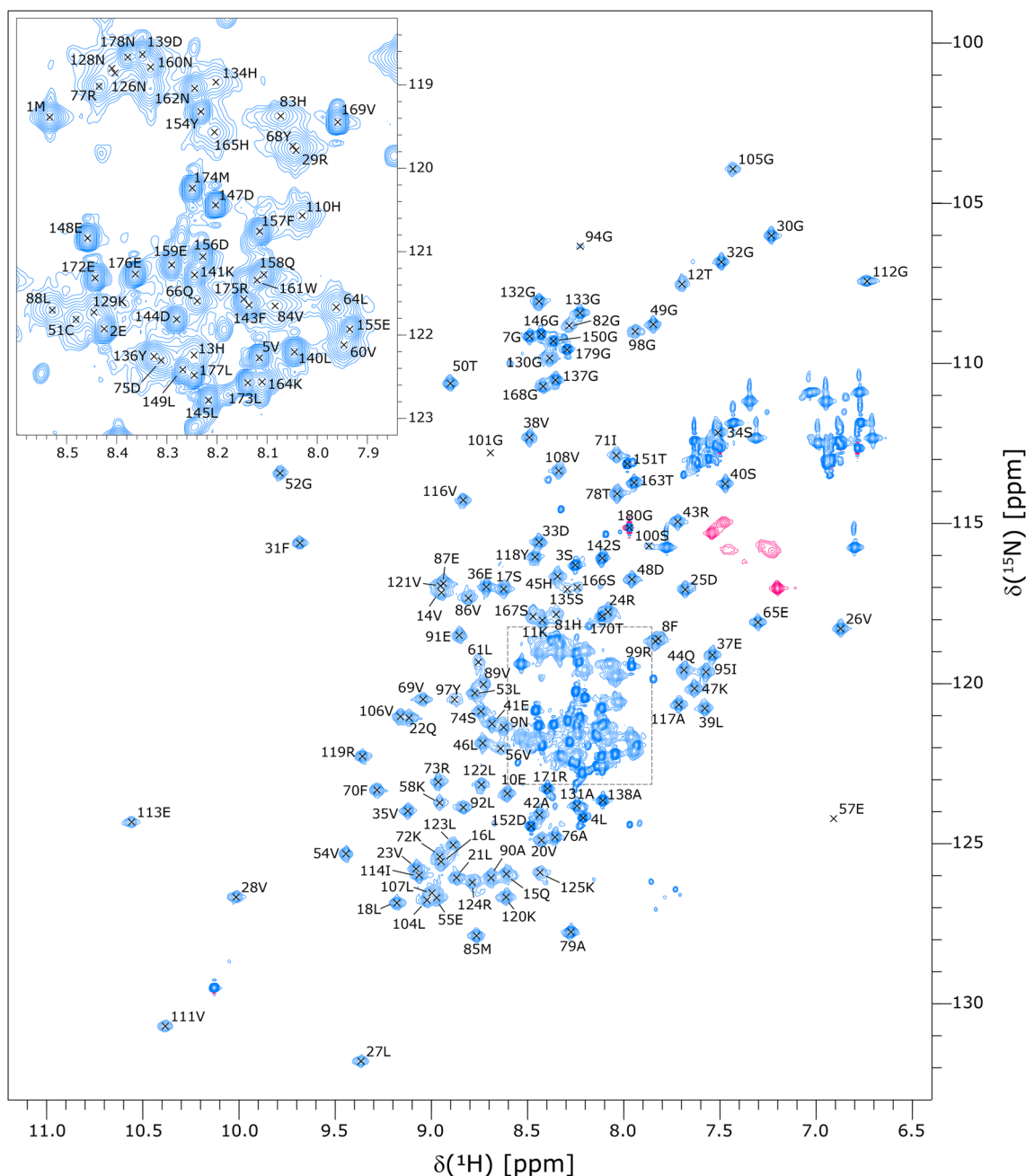


Fig. 1 ^{15}N -HSQC spectrum of SARS-CoV-2 Nsp1 with backbone amide assignments. The excerpt in the top-left corner is an expansion of the crowded central region (corresponding to the dashed box in the main spectrum). The unlabeled peak at $\delta(^1\text{H})=10.15$ ppm, $\delta(^{15}\text{N})=129.8$ ppm corresponds to the side-chain $\text{N}^{\epsilon 1}\text{-H}^{\epsilon 1}$ group of

residue 161W. The unlabeled pink peaks correspond to arginine side-chain $\text{N}^{\epsilon}\text{-H}^{\epsilon}$ groups, and are aliased from their true ^{15}N chemical-shift positions at $\sim 84\text{--}86$ ppm. The peaks of the asparagine and glutamine side-chain amide groups are also unlabeled. The spectrum was recorded at a ^1H field-strength of 850 MHz

from the C-terminal region are clustered more closely together with ^1H chemical-shifts in the range 7.9–8.5 ppm. In general, the peaks from the C-terminal region are much sharper and more intense than the more widely dispersed peaks from the globular domain, as would be expected for a disordered tail. Differential linewidths and intensities characteristic of exchange-mediated line-broadening

effects were observed within both the C-terminal region and the globular domain. Some residues in the C-terminal region and towards the C-terminal end of the globular domain appeared to give rise to multiple peaks, with one or two weak-intensity peaks in addition to a strong-intensity peak. The weak-intensity peaks are probably due to

low-populated alternate conformers associated with cis-geometry amide bonds at proline residues.

We assigned the backbone resonances to a high-degree of completeness, obtaining ^1H and ^{15}N assignments for 165 of the 172 assignable backbone amide groups (96.0%), and ^{13}C assignments for 96.1%, 97.2% and 97.5% of the C' , C^α and C^β nuclei, respectively. Amide assignments are missing for residues 59G, 63Q, 93E, 96Q, 102E, 103 T and 127G. These residues are located in four stretches for which the assigned amide peaks are significantly broader and weaker. The amide signals for the missing residues are probably so broadened that the corresponding peaks simply do not appear above the noise in the ^{15}N -HSQC spectrum, and certainly would not yield detectable peaks in the 3D triple-resonance spectra.

The assigned backbone chemical shifts were used to predict the secondary-structure with DANGLE (Cheung et al. 2010) and TALOS-N (Shen & Bax 2013), and also to calculate random-coil-index-derived order-parameters (RCI-S^2) (Berjanskii & Wishart 2005). Panel A of Fig. 2 shows the RCI-S^2 values and TALOS-N-derived helix/strand probabilities. Panel B shows the TALOS-N- and DANGLE-predicted secondary structures of SARS-CoV-2 Nsp1 together with the secondary-structure assignments of the two crystal structures of the folded globular domain (PDB codes 7k7p and 7k3n), the two cryo-EM structures of the C-terminus bound to the small ribosomal subunit (PDB codes 6zn5 (Thoms et al. 2020) and 7k5i), and the solution-NMR structure of Nsp1 from SARS-CoV-1 (PDB code 2hsx).

The well-folded globular domain of Nsp1 extends approximately from residue 10 to residue 125, and consists of six β -strands and two helices (labelled E1–E6 and H1–H2 in Fig. 2, respectively). The two long loops between strands E3 and E4, and between E4 and E5 show relatively low order-parameters, indicative of significant internal flexibility. Overall, the chemical-shift-derived secondary-structure predictions for the globular domain are very similar to the secondary structures of the two crystal structures and the solution-structure of Nsp1 from SARS-CoV-1. The short proline-containing 3_{10} helix (residues 61–63) and the short

β -sheet formed by residues 95–97 observed in the crystal structures are not predicted from the backbone chemical shifts, but some assignments were missing for residues in these two regions, compromising the ability to make accurate chemical-shift-based secondary-structure predictions.

In full-length Nsp1, the C-terminal region from residue 126 onwards is partially but not completely disordered, with RCI-S^2 values falling mostly in the range 0.3–0.6, and very little canonical secondary-structure propensity for residues 125–155. DANGLE predicts two additional helices at the very C-terminus (residues 157–162 and 171–176). While the TALOS-N helical assignments in the same sequence-stretch are limited to just two residues (171 & 173; labelled H3 in Fig. 2), the corresponding probabilities indicate a clear helical propensity for residues 171–175. Interestingly, the region of helical propensity at the very end of the C-terminus corresponds approximately to the two helices observed in the cryo-EM structure of Nsp1 interacting with the small ribosomal subunit, suggesting that these helices, which are partially formed in free Nsp1, become stabilized upon insertion into the mRNA tunnel of the small ribosomal subunit.

Based on the strong sequence-similarity with SARS-CoV-1 Nsp1, we also investigated the isolated Nsp1 NTD (residues 13–127). Unexpectedly, the protein showed little solubility and we were not able to achieve high concentrations in NMR-compatible buffers, although the domain boundaries are in agreement with the available crystal and NMR structures of SARS-CoV-1 Nsp1 (Fig. 2b). Nevertheless, we were able to record 2D ^1H , ^{15}N correlation spectra. An overlay of the ^{15}N -HSQC-TROSY spectrum of Nsp1 NTD with the ^{15}N -HSQC spectrum of full-length Nsp1 (Fig. 3) shows reasonable agreement for the residues of the NTD, although there are some chemical shift differences and also instances of relative line-broadening for the Nsp1 NTD peaks. It is conceivable that in the full-length protein, the disordered C-terminal domain interacts with the NTD, thereby stabilizing its fold and improving solubility. We are currently conducting further experiments to investigate this hypothesis.

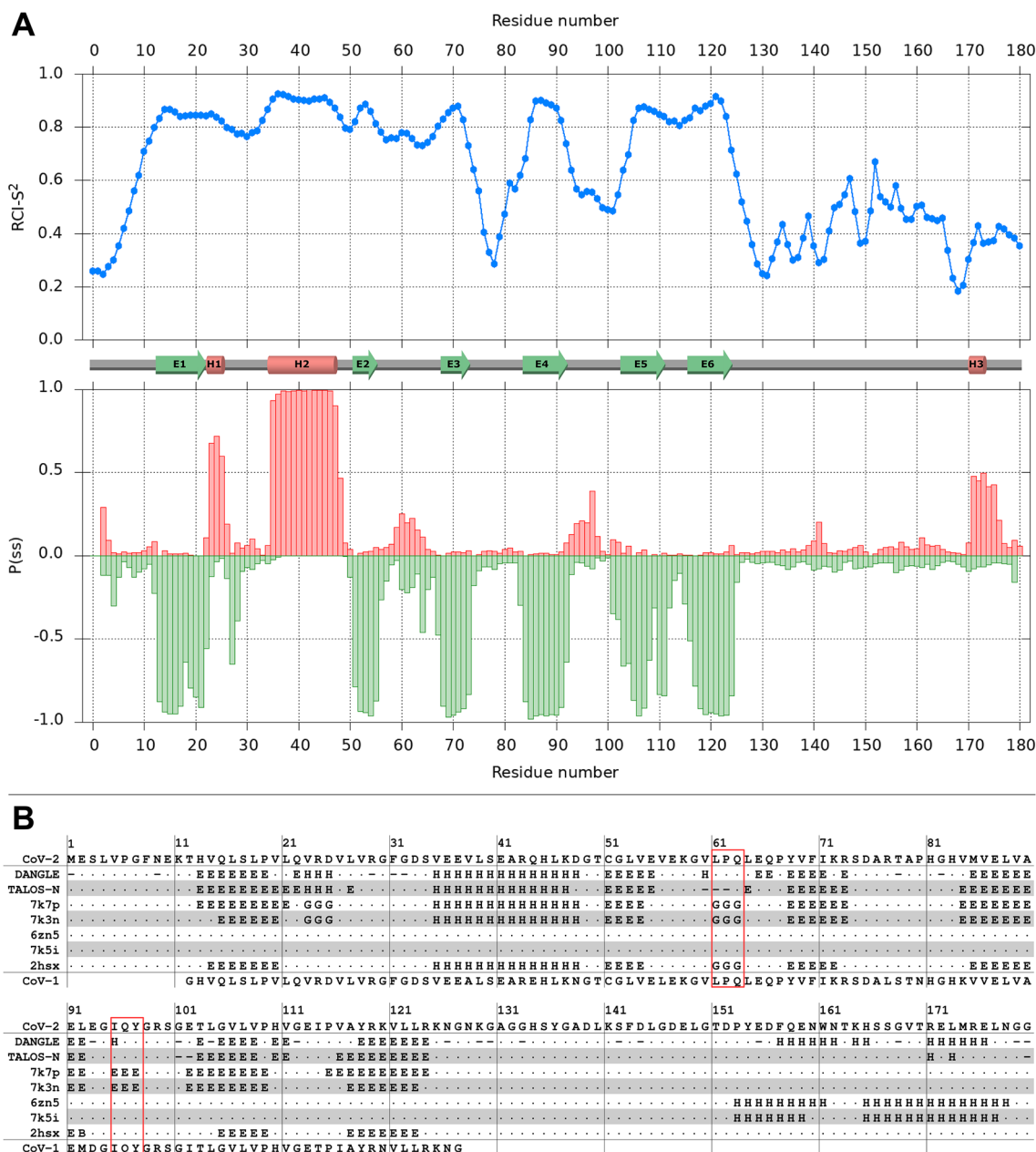


Fig. 2 Secondary-structure of SARS-CoV-2 Nsp1. **a.** Results from TALOS-N analysis of the assigned backbone chemical shifts. Top: random-coil-index-derived squared-order-parameters (RCI-S²) plotted by residue number. Bottom: probabilities of helical (positive, red) and strand (negative, green) secondary-structure plotted by residue number. The schematic secondary-structure shown between the two plots is based on the TALOS-N prediction. **b.** Comparison of the DANGLE- and TALOS-N-predicted secondary-structures for SARS-CoV-2 Nsp1 with those of the crystal structures for SARS-CoV-2 Nsp1 with those of the crystal structures of the globular domain (PDB codes 7k7p & 7k3n), the cryo-EM structures of the C-terminus in complex with the small ribosomal subunit (PDB codes 6zn5 & 7k5i) and the NMR structure of the globular domain

of SARS-CoV-1 Nsp1 (PDB code 2hsx). The two regions where secondary-structure elements present in the crystal structures are missing from the chemical-shift-based predictions are highlighted in red boxes. The secondary structures of the PDBs were calculated with STRIDE (Frishman & Argos 1995). The Nsp1 sequences from SARS-CoV-2 and SARS-CoV-1 (shown above and below the secondary-structure annotations, respectively) were aligned with ClustalW (Sievers et al. 2011). The codes for the secondary-structure annotations are as follows: 'E' denotes extended β -strand; 'H' denotes either generic helix (DANGLE & TALOS-N) or specific α -helix (STRIDE); 'G' denotes 3_{10} helix (STRIDE only)

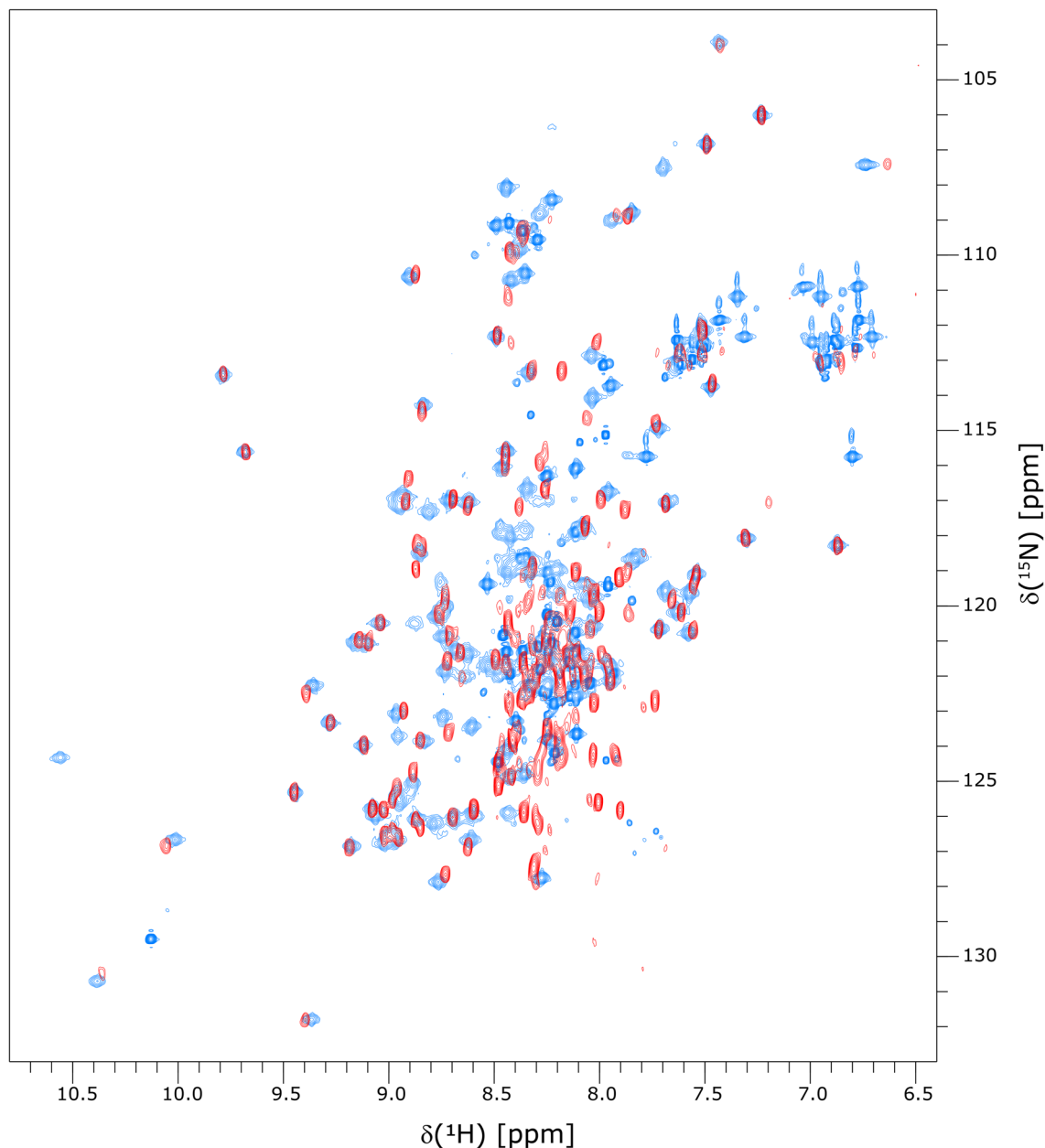


Fig. 3 Overlay of ^1H , ^{15}N correlation spectra of full-length Nsp1 and the Nsp1 NTD (residues 13–127). Overlay of the ^{15}N -HSQC-TROSY spectrum of Nsp1 NTD (red; recorded at 950 MHz) with the ^{15}N -HSQC spectrum of full-length Nsp1 (blue; recorded at

850 MHz). The signal-to-noise ratios of the two spectra are not quantitatively comparable due to differences in concentrations and acquisition parameters

Acknowledgements We acknowledge Dr Andrea Graziadei (TU Berlin) for the kind gift of the Nsp1 expression plasmid.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was funded by Goethe Corona Funds, the Deutsche Forschungsgemeinschaft through grant SFB902/B18 (to *covid19-nmr*), DFG large scale equipment grant 452632086, EU commission access programme iNEXT-discovery (grant agreement no. 871037), and grant CA294/16-1 to T.C.

Data availability The ^1H , ^{13}C and ^{15}N backbone chemical-shift assignments of SARS-CoV-2 Nsp1 have been deposited at the BioMagRes-Bank (<https://www.bmrb.wisc.edu>) under accession number 50620.

Declarations

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing,

adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Almeida MS, Johnson MA, Herrmann T, Geralt M, Wüthrich K (2007) Novel β -Barrel fold in the nuclear magnetic resonance structure of the replicase nonstructural protein 1 from the severe acute respiratory syndrome coronavirus. *J Virol* 81(7):3151–3161. <https://doi.org/10.1128/JVI.01939-06>
- Bax A, Ikura M (1991) An efficient 3D NMR technique for correlating the proton and ^{15}N backbone amide resonances with the α -carbon of the preceding residue in uniformly $^{15}\text{N}/^{13}\text{C}$ enriched proteins. *J Biomol NMR* 1(1):99–104. <https://doi.org/10.1007/BF01874573>
- Benedetti F, Snyder GA, Giovanetti M, Angeletti S, Gallo RC, Ciccozzi M, Zella D (2020) Emerging of a SARS-CoV-2 viral strain with a deletion in nsp1. *J Transl Med* 18(1):4–9. <https://doi.org/10.1186/s12967-020-02507-5>
- Berjanskii MV, Wishart DS (2005) A simple method to predict protein flexibility using secondary chemical shifts. *J Am Chem Soc* 127(43):14970–14971. <https://doi.org/10.1021/ja054842f>
- Bodenhausen G, Ruben DJ (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem Phys Lett* 69(1):185–189. [https://doi.org/10.1016/0009-2614\(80\)80041-8](https://doi.org/10.1016/0009-2614(80)80041-8)
- Cheung MS, Maguire ML, Stevens TJ, Broadhurst RW (2010) DAN-GLUE: a Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. *J Magn Reson* 202(2):223–233. <https://doi.org/10.1016/j.jmr.2009.11.008>
- Clubb RT, Thanabal V, Wagner G (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intrasite ^1H , ^{15}N , and ^{13}C chemical shifts in ^{15}N - ^{13}C -labelled proteins. *J Magn Reson*. [https://doi.org/10.1016/0022-2364\(92\)90252-3](https://doi.org/10.1016/0022-2364(92)90252-3)
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*. <https://doi.org/10.1007/BF00197809>
- Favier A, Brutscher B (2011) Recovering lost magnetization: polarization enhancement in biomolecular NMR. *J Biomol NMR* 49(1):9–15. <https://doi.org/10.1007/s10858-010-9461-5>
- Frishman D, Argos P (1995) Knowledge-based protein secondary structure assignment. *Proteins Struct Funct Genet* 23(4):566–579. <https://doi.org/10.1002/prot.340230412>
- Gorbalenya A, Baker S, Baric R, de Groot R, Drosten C, Gulyaeva A, Haagmans B, Lauber C, Leontovich A, Neuman B, Penzar D, Perlman S, Poon L, Samborskiy D, Sidorov I, Sola I, Ziebuhr J (2020) Severe acute respiratory syndrome-related coronavirus: the species and its viruses – a statement of the Coronavirus Study Group. *Nat Microbiol*. <https://doi.org/10.1101/2020.02.07.937862>
- Grzesiek S, Bax A (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J Magn Reson* 96(2):432–440. [https://doi.org/10.1016/0022-2364\(92\)90099-5](https://doi.org/10.1016/0022-2364(92)90099-5)
- Ikura M, Kay LE, Bax A (1990) A novel approach for sequential assignment of ^1H , ^{13}C , and ^{15}N spectra of larger proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy Application to Calmodulin. *Biochemistry* 29(19):4659–4667. <https://doi.org/10.1021/bi00471a022>
- Kay LE, Ikura M, Tschudin R, Bax A (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J Magn Reson* 89(3):496–514. [https://doi.org/10.1016/0022-2364\(90\)90333-5](https://doi.org/10.1016/0022-2364(90)90333-5)
- Kay LE, Xu GY, Yamazaki T (1994) Enhanced-sensitivity triple-resonance spectroscopy with minimal H_2O saturation. *J Magn Reson Ser A*. <https://doi.org/10.1006/jmra.1994.1145>
- Khailany RA, Safdar M, Ozaslan M (2020) Genomic characterization of a novel SARS-CoV-2. *Gene Rep* 19(January):100682. <https://doi.org/10.1016/j.genrep.2020.100682>
- Lescop E, Schanda P, Brutscher B (2007) A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J Magn Reson* 187(1):163–169. <https://doi.org/10.1016/j.jmr.2007.04.002>
- Lim Y, Ng Y, Tam J, Liu D (2016) Human coronaviruses: a review of virus-host interactions. *Diseases* 4(4):26. <https://doi.org/10.3390/diseases4030026>
- Masters PS (2006) The molecular biology of coronaviruses. *Adv Virus Res* 65(January):193–292. [https://doi.org/10.1016/S0065-3527\(06\)66005-3](https://doi.org/10.1016/S0065-3527(06)66005-3)
- Piotto M, Saudek V, Sklenář V (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR*. <https://doi.org/10.1007/BF02192855>
- Schubert K, Karousis ED, Jomaa A, Scaiola A, Echeverria B, Gurzeler LA, Leibundgut M, Thiel V, Mühlemann O, Ban N (2020) SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. *Nat Struct Mol Biol* 27(10):959–966. <https://doi.org/10.1038/s41594-020-0511-8>
- Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* 56(3):227–241. <https://doi.org/10.1007/s10858-013-9741-y>
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. <https://doi.org/10.1038/msb.2011.75>
- Solyom Z, Schwarten M, Geist L, Konrat R, Willbold D, Brutscher B (2013) BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *J Biomol NMR* 55(4):311–321. <https://doi.org/10.1007/s10858-013-9715-0>
- Thoms M, Buschauer R, Ameismeier M, Koepke L, Denk T, Hirschenberger M, Kratzat H, Hayn M, MacKens-Kiani T, Cheng J, Straub JH, Stürzel CM, Fröhlich T, Berninghausen O, Becker T, Kirchhoff F, Sparrer KMJ, Beckmann R (2020) Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science* 369(6508):1249–1256. <https://doi.org/10.1126/SCIENCE.ABC8665>
- Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED (2005) The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins Struct Funct Genet*. <https://doi.org/10.1002/prot.20449>
- Wittekind M, Mueller L (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the α - and β -carbon resonances in proteins. *J Magn Reson Ser B* 101(2):201–205. <https://doi.org/10.1006/jmrb.1993.1033>
- Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, Meng J, Zhu Z, Zhang Z, Wang J, Sheng J, Quan L, Xia Z, Tan W, Cheng G, Jiang T (2020a) Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* 27(3):325–328. <https://doi.org/10.1016/j.chom.2020.02.001>

Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, Hu Y, Tao Z-W, Tian J-H, Pei Y-Y, Yuan M-L, Zhang Y-L, Dai F-H, Liu Y, Wang Q-M, Zheng J-J, Xu L, Holmes EC, Zhang Y-Z (2020b) A new coronavirus associated with human respiratory disease in China. *Nature* 579(7798):265–269. <https://doi.org/10.1038/s41586-020-2008-3>

Yamazaki T, Muhandiranv DR, Kay LE, Lee W, Arrowsmith CH (1994) A suite of triple resonance NMR experiments for the backbone assignment of ^{15}N , ^{13}C , ^2H labeled proteins with high sensitivity. *J Am Chem Soc* 116(26):11655–11666. <https://doi.org/10.1021/ja00105a005>

Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, Si H-R, Zhu Y, Li B, Huang C-L, Chen H-D, Chen J, Luo Y, Guo H, Jiang

R-D, Liu M-Q, Chen Y, Shen X-R, Wang X, Xiao GF, Shi Z-L (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579(7798):270–273. <https://doi.org/10.1038/s41586-020-2012-7>

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