



# Backbone chemical shift spectral assignments of SARS coronavirus-2 non-structural protein nsp9

Erika F. Dudás<sup>1</sup> · Rita Puglisi<sup>1</sup> · Sophie Marianne Korn<sup>2,3</sup> · Caterina Alfano<sup>4</sup> · Maria Laura Bellone<sup>5</sup> · Fabrizio Dal Piaz<sup>5</sup> · Geoff Kelly<sup>6</sup> · Elisa Monaca<sup>4</sup> · Andreas Schlundt<sup>2,3</sup> · Harald Schwalbe<sup>2,3</sup> · Annalisa Pastore<sup>1</sup>

Received: 9 December 2020 / Accepted: 24 February 2021 / Published online: 23 March 2021  
© The Author(s) 2021

## Abstract

As part of an International consortium aiming at the characterization by NMR of the proteins of the SARS-CoV-2 virus, we have obtained the virtually complete assignment of the backbone atoms of the non-structural protein nsp9. This small (12 kDa) protein is encoded by ORF1a, binds to RNA and seems to be essential for viral RNA synthesis. The crystal structures of the SARS-CoV-2 protein and other homologues suggest that the protein is dimeric as also confirmed by analytical ultracentrifugation and dynamic light scattering. Our data constitute the prerequisite for further NMR-based characterization, and provide the starting point for the identification of small molecule lead compounds that could interfere with RNA binding and prevent viral replication.

**Keywords** Coronavirus · Covid-19 NMR · Protein · SARS-CoV-2 · Solution NMR · Structure

## Biological context

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection is already the third coronavirus infection that has occurred since the start of the third millennium. This third epidemic has plagued the world population over almost all of 2020 and is still actively on-going.

The current situation has thus underlined the importance of gaining a deep and lasting understanding of the rules that regulate viral internalization and reproduction and the necessity to translate this knowledge into a vaccine and/or other treatments. SARS-CoV-2 is an RNA virus comprising a large single-stranded positive polarity genome that acts as messenger RNA after entering the host (Wu et al. 2020; Masters and Perlman 2013). The 5' two-third of the genome encodes a long polyprotein that is translated in two open reading frames, ORF1a and ORF1b, through host ribosomal frameshifting. The viral RNA also encodes a few structural and accessory proteins within smaller ORFs. As in all viruses, the SARS-CoV-2 genome contains all the proteins necessary for host cell infection such as the RNA polymerase along with enzymes that facilitate RNA synthesis. These proteins are released by the action of two internally encoded proteases. The mature proteins are referred to as non-structural proteins (nsps) as they do not per se constitute the virion shield.

As a part of an International consortium aimed at the study of all SARS-CoV-2 proteins by nuclear magnetic resonance (NMR), this paper describes the nearly complete backbone NMR resonance assignment of the non-structural protein 9 (nsp9). This 12 kDa protein, which is encoded in ORF1a, is a replicase that has been shown to be essential for replication (Frieman et al. 2012; Littler et al. 2020).

Erika F. Dudás and Rita Puglisi have equally contributed.

✉ Annalisa Pastore  
annalisa.1.pastore@kcl.ac.uk

- <sup>1</sup> Department of Basic and Clinical Neuroscience, Maurice Wohl Institute, UK-DRI at King's College London, 5 Cutcombe Rd, London SE59RT, UK
- <sup>2</sup> Institute for Molecular Biosciences, Johann Wolfgang Goethe-University Frankfurt, Max-von-Laue-Str. 9, 60438 Frankfurt/M, Germany
- <sup>3</sup> Institute for Organic Chemistry and Chemical Biology, Center for Biomolecular Magnetic Resonance (BMRZ), Johann Wolfgang Goethe-University Frankfurt, Max-von-Laue-Str. 7, 60438 Frankfurt/M, Germany
- <sup>4</sup> Fondazione Ri.Med, 90133 Palermo, Italy
- <sup>5</sup> Department of Medicine and Surgery, University of Salerno, Via Giovanni Paolo II, 84081 Baronissi, SA, Italy
- <sup>6</sup> Francis Crick Institute, MRC Biomedical NMR Centre, 1 Midland Rd, London NW1 1AT, UK

This feature makes nsp9 a potential target for drug discovery aimed at inhibiting viral replication. Nsp9 forms discrete foci in the perinuclear region of infected cells and colocalizes with other components of the viral replication complex (Frieman et al. 2012). The structures of several homologues of nsp9 are available from SARS-CoV (Yang et al. 2003), the transmissible gastroenteritis (Anand et al. 2002), human corona 229E (Anand et al. 2003), avian infectious bronchitis, porcine epidemic diarrhea, and porcine delta viruses. The crystal structure of nsp9 from SARS-CoV-2 was also recently published (Littler et al. 2020). Availability of the crystal structure does not, nevertheless, reduce the interest of studying the protein in solution as this is the prerequisite to fragment based drug screening and other experimentally-based drug design strategies. Nsp9 proteins have a fold that vaguely resembles that of the oligonucleotide/oligosaccharide binding (OB) domain (Sutton et al. 2004). Nsp9 is an RNA-binding protein that interacts with and activates other proteins of the viral cascade (Sutton et al. 2004). The mechanism of RNA binding within the nsp9 protein family is not understood as these proteins have an unusual structural fold not previously seen in RNA-binding proteins (Egloff et al. 2004; Sutton et al. 2004). In all crystal structures, the protein consistently forms a dimer with an interface mediated by a conserved ‘‘GxxxG’’ motif in the C-terminal-helix. A dimeric form is thought to be critical for viral replication (Miknis et al. 2009). Disruption of key residues in the ‘GxxxG’ motif reduces both RNA binding (Sutton et al. 2004) and SARS-CoV viral replication (Frieman et al. 2012).

Our data provide the prerequisite for further studies of nsp9 by NMR and will help in the identification of small molecules able to interfere with RNA and other partner binding and, thus, to stop viral replication.

## Methods and experiments

### Construct design

The amino acid sequence of SARS-CoV-2 nsp9 was obtained from NCBI reference entry YP\_009725305.1 (Wu et al. 2020). Domain boundaries were defined according to the available crystal structures of SARS-CoV-2 nsp9 (PDB codes 6w4b and 6w9q). The sequence encoding for amino acids 1 to 113 (corresponding to full length nsp9) was codon-optimized for *E. coli* expression. The gene was obtained from GenScript Biotech (Netherlands), inserted into the pET3b-based vector pKM263, as well as pET21b(+)-based vector pET-TEV-Nco. Vector pKM263 encodes for an N-terminal His<sub>6</sub>-tag and a GST-tag, followed by a TEV cleavage site, while pET-TEV-Nco only included an N-terminal His<sub>7</sub>-tag and the TEV cleavage site. After

proteolytic cleavage (same for both vectors), the 12.4 kDa protein contained four artificial residues (Gly, Ala, Met and Gly), before the start of the native protein sequence.

### Protein production

SARS-CoV-2 nsp9 was cloned and recombinantly expressed in *E. coli*. Two independent preparations were carried out by different teams. In the first, uniformly <sup>13</sup>C,<sup>15</sup>N-labelled nsp9 was obtained from a BL21(DE3) culture (induction with 1 mM IPTG, for 20 h at 22 °C) adding to the medium 0.5 g/l <sup>15</sup>N ammonium chloride and 2 g/l <sup>13</sup>C D-glucose for labelling. It was purified by Immobilized Metal Affinity Chromatography (IMAC) on a Ni<sup>2+</sup>-NTA gravity flow column (Sigma-Aldrich) in 50 mM Tris-HCl, at pH 8, 300 mM NaCl, 10 mM Imidazole, 4 mM DTT and eluted between 150 and 500 mM imidazole. The GST-tag was cleaved overnight by TEV protease (0.5 mg of TEV protease per 1 L of culture) dialyzing in the same buffer used for IMAC, followed by a reverse Ni<sup>2+</sup>-NTA in the same buffer, as reported for SARS-CoV nsp9 (Sutton et al. 2004). Further purification was carried out by size exclusion (SEC) performed on a HiLoad 16/600 Sephadex 75 pg column (GE Healthcare) in SEC buffer (25 mM sodium phosphate, 150 mM NaCl, 2 mM TCEP, 0.02% NaN<sub>3</sub>, pH 7). Nsp9 eluted mainly as a dimer, as estimated by its elution volume. In addition, higher oligomeric species of nsp9 were observed, eluting earlier from the column, that showed significantly increased A260/280 ratios (> 1) and could not be highly concentrated (< 2 mg/mL). Pure nsp9 containing fractions (of the supposedly dimeric species) were determined by SDS-PAGE and pooled. A portion of this sample, called hereafter sample A, was kept at a low concentration of 180 μM to avoid potential induction of oligomerization. Another portion was concentrated to 410 μM, yielding sample B.

A similar protocol was followed in the second preparation scheme with minor adaptations. Both unlabelled and uniformly <sup>13</sup>C,<sup>15</sup>N-labelled samples were prepared. The protein was purified by IMAC in 25 mM sodium phosphate at pH 7.4, 150 mM NaCl, 1 mM DTT, and 20 mM imidazole and then eluted from the Nickel column with a 0% to 100% gradient of 25 mM sodium phosphate at pH 7.4, 150 mM NaCl, 1 mM DTT, and 400 mM Imidazole. TEV cleavage was obtained by overnight incubation of TEV at a protein:enzyme ratio of 50:1. The cleaved protein was recovered by reverse IMAC and further purified by SEC in 25 mM sodium phosphate, 150 mM NaCl, 2 mM TCEP, pH 7 using a HiLoad 16/600 Sephadex 75 pg column (GE Healthcare). For the unlabelled sample, SEC runs were carried out at different protein concentrations to check if the SEC profile changed as a function of concentration. No difference was observed. SEC fractions of the uniformly <sup>13</sup>C,<sup>15</sup>N-labelled

nsp9 sample were pulled together and concentrated from 100  $\mu$ M up to 650  $\mu$ M (sample C).

### Mass spectrometry-based analysis of amino acid sequence

The identity of nsp9 was verified by a mass spectrometry-based peptide mapping approach. Sample A and an unlabelled version of sample C were loaded on a 12.5% SDS polyacrylamide gel. The resulting band underwent trypsin-catalyzed in-gel digestion. NanoUPLC-hrMS/MS analysis of the sample was carried out on a Q-Exactive orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA), coupled with a nanoUltimate300 UHPLC system (Thermo Fisher Scientific). Peptides separation was performed on a capillary EASY-Spray PepMap column (0.075 mm  $\times$  50 mm, 2  $\mu$ m, Thermo Fisher Scientific) using aqueous 0.1% formic acid and CH<sub>3</sub>CN containing 0.1% formic acid as mobile phases and a linear gradient from 3 to 40% of B in 45 min and a 300 nL·min<sup>-1</sup> flow rate. Mass spectra were acquired over an m/z range from 375 to 1500. To achieve protein identification, MS and MS/MS data underwent Mascot software (v2.5, Matrix Science, Boston, MA, USA) analysis using the non-redundant UniprotKB/Swiss-Prot database (Release 2020\_03).

### NMR experiments

The spectra of the samples A, B and C from the two independent preparations were recorded and analysed. The samples were studied in 25 mM sodium phosphate buffer

at pH 7.0, 150 mM NaCl, 2 mM TCEP, 0.02% NaN<sub>3</sub>. For comparison, the assignment of a triply labelled <sup>13</sup>C, <sup>15</sup>N and <sup>2</sup>H protein of a close homologue (97% sequence identity) from SARS-CoV (BMRB entry 6501) had been obtained in 50 mM sodium phosphate at pH 6.8, 50 mM NaCl. By default, 10% D<sub>2</sub>O was added to the final NMR samples. Condition screenings were done in the range of temperatures 278–308 K and of pH 5–7. In the end, the best S/N ratio was observed at 298 K and pH 7.0. At lower pH we observed only a few additional HSQC cross-peaks. NMR spectra were recorded on Bruker spectrometers working at 700, 800 and 950 MHz and equipped with cryo-probes. Water suppression was achieved with the WATERGATE pulse sequence (Piotto et al. 1992). For the sequential backbone resonance assignment, a set of 3D NMR experiments was used: HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB and <sup>15</sup>N-edited NOESY-HSQC. Non-uniform sampling was employed, with the number of points in indirect dimensions set to  $\sim (0.66 * 2^N)$  and extended to 2<sup>N</sup> in SMILE (Ying et al. 2017). The schedules were exponentially weighted with a time constant equal to the acquisition time in that dimension. Selected parameters are listed in Table 1.

### Assignments and data deposition

The identity and integrity of the samples were tested by trypsin hydrolyses supported by high-resolution LC/MS–MS analysis. The backbone chemical shift assignment of nsp9 was carried out manually using standard double- and triple-resonance NMR experiments on uniformly labelled samples and recorded at 298 K. Despite the overall good quality of

**Table 1** Summary of the NMR experiments recorded

	TD			SW(ppm)			MHz	NS	NUS (%)
Sample A									
HSQC	2048	400		14	31		700	16	–
HNCO	2048	84	84	14	31	14	700	8	25
HN(CA)CO	2048	84	84	14	31	14	700	48	29
HNCA	2048	84	84	14	31	30	700	64	25
HN(CO)CA	2048	84	84	14	31	30	700	64	25
HNCACB	1024	60	140	14	31	74	800	16	–
N15-NOESY	2048	84	340	14	31	12	800	8	50
Sample B									
HSQC	3072	400		14	31		950	16	–
HNCO	1536	84	84	14	31	14	700	40	25
HN(CA)CO	1536	84	84	14	31	14	700	104	25
HNCA	1536	84	84	14	31	30	950	64	45
HN(CO)CA	1536	84	84	14	31	30	950	40	45
HNCACB	2048	84	168	14	31	80	950	64	28
N15-NOESY	1536	84	340	14	31	12	950	32	27

All NMR spectra were processed using the NMRPipe/SMILE (Delaglio et al. 1995; Ying et al. 2017) and analyzed using CCPNMR (Skinner et al. 2016)

the spectra obtained, assignment of nsp9 presented some challenge. The number of signals observed in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments for the less concentrated sample A was 112, including sidechain amides. This number was significantly lower than the expected resonances considering 111 non-proline residues, one indole of a tryptophan and twelve side chains of glutamines and asparagines. Conversely, the number of resonances was approximately 50% higher than expected for the more concentrated sample B (Fig. 1). We originally interpreted this difference with the presence of

an equilibrium between a monomeric and dimeric species in a slow exchange rate regime. However, this interpretation would disagree with previous literature on nsp9 both from SARS-CoV-2 and other coronaviruses (Littler et al. 2020; Sutton et al. 2004). These studies have all reported that the protein is present in solution as an obligate dimer above 100  $\mu\text{M}$  whose dimer interface involves a parallel packing of the C-terminal helices (Sutton et al. 2004).

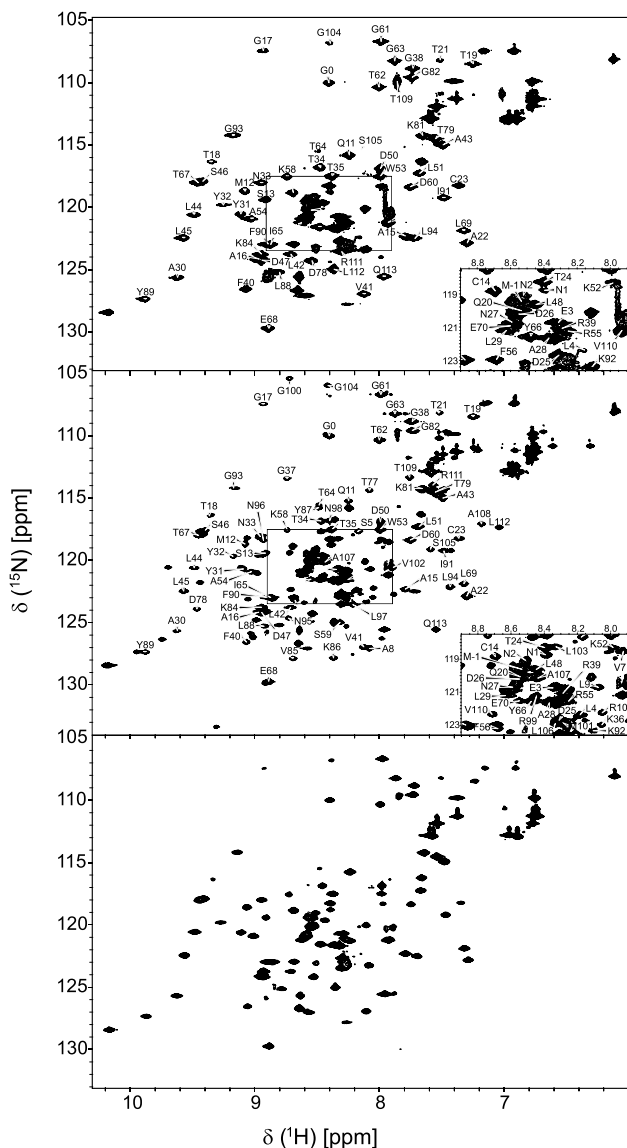
Later preparation of sample C clarified the issue: aliquots from the same preparation were increasingly concentrated from 100 to 650  $\mu\text{M}$  and HSQC spectra recorded. The resulting spectrum of sample C was independent of concentration and perfectly superimposable with that of sample A, indicating that this must be considered the spectrum representative of the dimeric form of this protein. Since the spectrum of this sample did not contain any new information it was kept only for comparison and not further analysed. Consequently, sample B must correspond to an alternative second conformation. Both the spectra of samples A and B, recorded at neutral pH, remained stable for at least four months.

$^1\text{H}$ ,  $^{15}\text{N}$  and  $^{13}\text{C}\alpha$  chemical shifts of a close homologue (three residues different) from SARS-CoV deposited in the BMRB (entry 6501) were transferred to  $^1\text{H}$ - $^{15}\text{N}$  HSQC, HNCA, and HN(CO)CA spectra recorded for sample A. Isolated resonances were identified first in the HSQC spectrum using the CCPNMR software, then assignment involving residues in the overlapping region was performed based on  $\text{C}\alpha$  shifts. Limited magnetization transfer constrained the use of the HNCACB strips only to resolve specific ambiguities of the HNCA.

Approximately 30 non-proline residues of sample B, most of them corresponding to residues in the C-terminal helix, were not observed in samples A and C despite mass spectrometry analysis had confirmed the presence and the correct molecular weight of all of the nine peptides expected from digestion, representing 99% of the protein sequence. In sample B, only residues Q49, C73, R74, F75, and V76 were not assigned. The C-terminal helix (N95-R111) could be fully assigned in this sample based on the HNCA spectra with the support of characteristic cross-peak patterns observed in strips in the  $^{15}\text{N}$ -edited NOESY-HSQC.

Consequently, we could assign in the end 73% of the nitrogen atoms, 73% of the amide protons, 77% of the  $\text{C}\alpha$ , 41% of the  $\text{C}\beta$  and 76% of the  $\text{C}'$  resonances for sample A. For sample B, 95% of the nitrogen atoms, 95% of the amide protons, 96% of the  $\text{C}\alpha$ , 63% of the  $\text{C}\beta$  and 96% of the  $\text{C}'$  resonances were assigned.

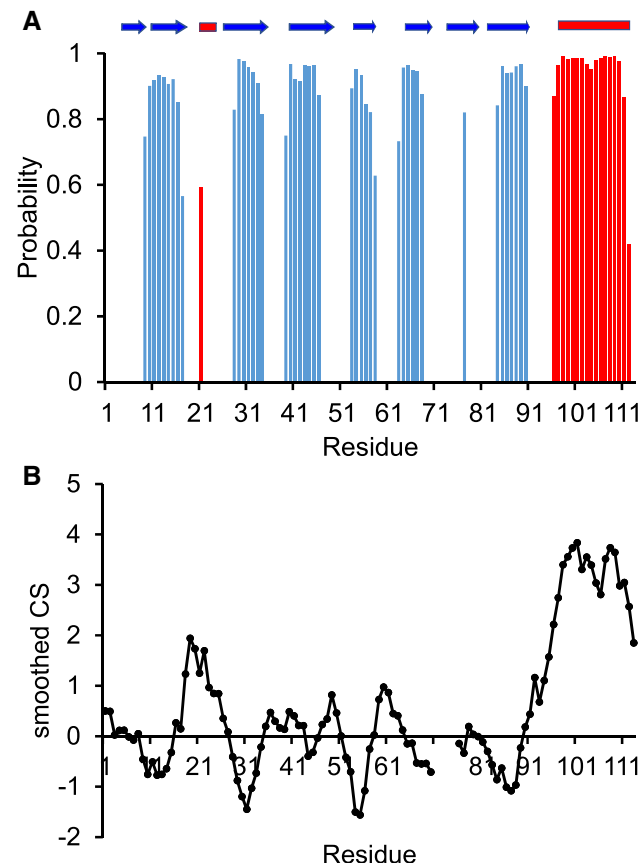
The assignments of both sample A and B were deposited in BMRB with accession numbers 50622 and 50621 respectively. They were overall in excellent agreement with each other except for the assignment of the C-terminus where main differences with sample B were evident at the few residues observable in sample A. Our assignment was also



**Fig. 1** Comparison of spectra recorded at three different protein concentrations. Spectra of samples **a** 180  $\mu\text{M}$ , **b** 410  $\mu\text{M}$ , middle and **c** 600  $\mu\text{M}$ , bottom recorded at 700 MHz, 950 MHz and 800 MHz respectively. Spectra of sample A and B were annotated with the assignment. Notice that, for clarity, the counter levels were chosen to exclude noise. Thus, not all of the duplicated peaks observed in the spectrum of sample B are visible

in good agreement with that of the SARS-CoV nsp9 homologue, especially for the C $\alpha$  carbons which are nuclei less

affected by the environmental conditions (pH, buffer content and perdeuteration). The larger differences were observed again in the assignment of the C-terminal helix.



**Fig. 2** Secondary-structure prediction of nsp9 based on the backbone chemical shifts of sample B. **a** TALOS-N prediction. On the top is, for reference, the schematic indication of the secondary structure elements observed in the X-ray structures (PDB codes 6w4b and 6w9q). **b** Prediction obtained by smoothing the secondary chemical shifts (observed chemical shifts – random coil chemical shifts) smoothed according to Pastore and Saudek (1990). The smoothing window was  $\pm 2$

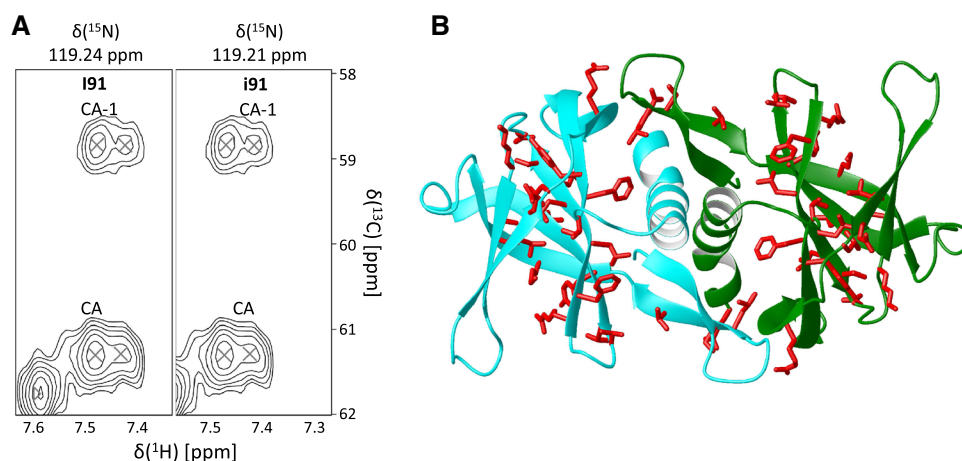
### Structural comparison of nsp9 from SARS-CoV2 and its crystal structure

The backbone resonance assignment was used to predict secondary structure elements of nsp9 by TALOS-N (Shen et al. 2013). We obtained a plot which clearly indicated the presence of seven beta strands and two helices (Fig. 2a). We also adopted for comparison the approach by Pastore and Saudek (1990) in which the secondary C $\alpha$  carbon chemical shifts are smoothed and plotted versus the protein sequence (Fig. 2b). The experimentally determined structural elements were compared with motifs obtained from the X-ray structures of SARS-CoV nsp9 (PDB codes 6w4b and 6w9q). Overall, the structural elements of the proteins are fully consistent reproducing seven beta strands out of the eight expected, a C-terminal helix and a short one-turn  $3_{10}$  helix at residues 22–24.

### Surface of dimerization

During the assignment process, it became clear that some residues could not be assigned in sample A. Most of the tracing could be done using the spectrum of sample B but some strips were duplicated (Fig. 3a). We could easily identify 30 residues for which splitting was evident. We reasoned that resonance duplication could either result from the co-presence of monomer and dimer, two different dimeric forms or a dimer and a tetramer (possibly in equilibrium). The monomer to dimer possibility was ruled out because it would not be consistent with all the previous literature. The presence of a tetramer is unlikely because the molecular weight would be at the limit of the NMR detection under these conditions (ca. 45 kDa). Investigation of the X-ray structure (6w9q)

**Fig. 3** Peak duplications in the spectrum of nsp9. **a** A representative duplicated strip (Ile91) from the HNCA spectrum. We indicated the main species in capital letters, the alternative one in lowercase. **b** Mapping of the duplicated resonances on the X-ray structure (PDB code 6w9q). None of the affected residues is directly involved in the dimer interface





showed the duplicated residues were distributed along the whole sequence and none of them was observed for residues directly in the interface (Fig. 3b). It is thus possible that resonance duplication reflected the presence of two dimers with slightly different structures or different interfaces, trapped during purification. Structural differences could for instance result from an asymmetrical assembly of the dimer in one of the two forms (Nooren et al. 1999), different interfaces (Sutton et al. 2004) or domain swapping (Liu and Eisenberg 2002). Domain swapping could also explain the differences observed between samples A and B and would be consistent with a rather different distribution of chemical shifts in the C-terminus of the two samples. Identification of the second species will require, in the future, an NMR high resolution structure determination backed up by independent evidence as it could come, for instance, from H/D exchange detected by mass spectrometry and/or filtered NOEs from labelled/unlabelled mixed samples. More work will be required in the future to clarify this point.

## Note in publication

After our manuscript was submitted, another paper was published describing the NMR spectrum assignment of SARS-CoV2 nsp9 (Buchko GW, Zhou M, Craig JK, Van Voorhis WC, Myler PJ. (2021) Backbone chemical shift assignments for the SARS-CoV-2 non-structural protein Nsp9: intermediate (ms–μs) dynamics in the C-terminal helix at the dimer interface. *Biomol NMR Assign.* Jan 4:1–10). On the whole, the data are in excellent agreement. The spectrum reported corresponds to that of our sample A.

**Acknowledgements** We would like to thank Woonghee Lee for his help in assignment validation at the early stages of this project.

**Funding** Work at BMRZ was supported by the state of Hesse. Work on Covid19-nmr was supported by the Goethe Corona Funds, by the IWB-EFRE-programme 20007375 of state of Hesse, and the DFG through CRC902: “Molecular Principles of RNA-based regulation” and through infrastructure funds (project numbers: 277478796, 277479031, 392682309, 452632086, 70653611). CA was supported by Patto per il Sud Regione Siciliana–Grant “CheMIST” (CUP G77B17000110001), and PO FESR Sicilia 2014/2020 Azione 1.5.1.—Grant “Potenziamento Infrastruttura di Ricerca “GMP Facility, Laboratori di Ricerca e Servizi Diagnostici e Terapeutici IRCCS-ISMETT” (CUP G76G17000130007), *Partnership IRCCS-ISMETT/Fondazione Ri.MED.* CA also acknowledges the ATeN Center of University of Palermo for infrastructures support. AP was supported by the Francis Crick Institute through provision of access to the MRC Biomedical NMR Centre. The Francis Crick Institute receives its core funding from Cancer Research UK (FC001029), the UK Medical Research Council (FC001029), and the Wellcome Trust (FC001029).

**Data availability** The assignments of both sample A and B were deposited in BMRB with Accession Numbers 50622 and 50621 respectively.

## 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

- Anand K, Palm GJ, Mesters JR, Siddell SG, Ziebuhr J, Hilgenfeld R (2002) Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *EMBO J* 21(13):3213–3224. <https://doi.org/10.1093/emboj/cdf327>
- Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R (2003) Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* 300(5626):1763–1767. <https://doi.org/10.1126/science.1085658>
- 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* 6(3):277–293. <https://doi.org/10.1007/BF00197809>
- Dosset P, Hus JC, Blackledge M, Marion D (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J Biomol NMR* 16(1):23–28. <https://doi.org/10.1023/a:1008305808620>
- Egloff MP, Ferron F, Campanacci V, Longhi S, Rancurel C, Dutartre H, Snijder EJ, Gorbalenya AE, Cambillau C, Canard B (2004) The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. *Proc Natl Acad Sci U S A* 101(11):3792–3796. <https://doi.org/10.1073/pnas.0307877101>
- Frieman M, Yount B, Agnihothram S, Page C, Donaldson E, Roberts A, Vogel L, Woodruff B, Scorpio D, Subbarao K, Baric RS (2012) Molecular determinants of severe acute respiratory syndrome coronavirus pathogenesis and virulence in young and aged mouse models of human disease. *J Virol* 86(2):884–897. <https://doi.org/10.1128/JVI.05957-11>
- Littler DR, Gully BS, Colson RN, Rossjohn J (2020) Crystal structure of the SARS-CoV-2 non-structural protein 9, nsp9. *iScience* 23(7):101258. <https://doi.org/10.1016/j.isci.2020.101258>
- Liu Y, Eisenberg D (2002) 3D domain swapping: as domains continue to swap. *Protein Sci* 11(6):1285–1299. <https://doi.org/10.1110/ps.0201402>
- Masters PS, Perlman S (2013) Coronaviridae. In: Knipe DM, Howley PM (eds) *Fields virology*. Lippincott Williams & Wilkins, Philadelphia, pp 825–858
- Miknis ZJ, Donaldson EF, Umland TC, Rimmer RA, Baric RS, Schultz LW (2009) Severe acute respiratory syndrome coronavirus nsp9 dimerization is essential for efficient viral growth. *J Virol* 83(7):3007–3018. <https://doi.org/10.1128/JVI.01505-08>
- Nooren IM, George AV, Kaptein R, Sauer RT, Boelens R (1999) NMR structure determination of the tetramerization domain of

- the Mnt repressor: an asymmetric alpha-helical assembly in slow exchange. *J Biomol NMR* 15(1):39–53. <https://doi.org/10.1023/A:1008312309535>
- Pastore A, Saudek V (1990) The relationship between chemical shift and secondary structure in proteins. *J Mag Res* 90(1):165–176. [https://doi.org/10.1016/0022-2364\(90\)90375-J](https://doi.org/10.1016/0022-2364(90)90375-J)
- Piotto M, Saudek V, Sklenár V (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* 2(6):661–665. <https://doi.org/10.1007/BF02192855>
- 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>
- Skinner SP, Fogh RH, Boucher W, Ragan TJ, Mureddu LG, Vuister GW (2016) CcpNmr analysis assign: a flexible platform for integrated NMR analysis. *J Biomol NMR* 66(2):111–124. <https://doi.org/10.1007/s10858-016-0060-y>
- Sutton G, Fry E, Carter L, Sainsbury S, Walter T, Nettleship J, Berrow N, Owens R, Gilbert R, Davidson A, Siddell S, Poon LL, Diprose J, Alderton D, Walsh M, Grimes JM, Stuart DI (2004) The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. *Structure* 12(2):341–353. <https://doi.org/10.1016/j.str.2004.01.016>
- Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G et al (2020) A new coronavirus associated with human respiratory disease in China. *Nature* 579(7798):265–269. <https://doi.org/10.1038/s41586-020-2008-3>
- Yang H, Yang M, Ding Y, Liu Y, Lou Z, Zhou Z, Sun L, Mo L, Ye S, Pang H, Gao GF, Anand K, Bartlam M, Hilgenfeld R, Rao Z (2003) The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc Natl Acad Sci U S A* 100(23):13190–13195. <https://doi.org/10.1073/pnas.1835675100>
- Ying J, Delaglio F, Torchia DA, Bax A (2017) Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. *J Biomol NMR* 68(2):101–118. <https://doi.org/10.1007/s10858-016-0072-7>

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