NMR reveals specific tracts within the intrinsically disordered regions of the SARS-CoV-2 Nucleocapsid protein involved in RNA encountering

Letizia Pontoriero,[a] Marco Schiavina,[a] Sophie M. Korn,[b] Andreas Schlundt,[b] Roberta Pierattelli,[a] and Isabella C. Felli,[a]

[a] *Department of Chemistry "Ugo Schiff" and CERM, University of Florence, Via della Lastruccia, 3, 50019 Sesto Fiorentino, Italy*

[b] *Institute for Molecular Biosciences, Center for Biomolecular Magnetic Resonance (BMRZ), Johann Wolfgang Goethe-University, Max-von-Laue-Str. 7+9, 60438 Frankfurt/M., Germany*

E-mail: letizia.pontoriero@unifi.it

The SARS-CoV-2 Nucleocapsid protein (N) is one of the four structural proteins of the virus and the most expressed one upon viral infection [1]. It is organized in an RNA binding N-terminal domain (NTD), a dimerization C-terminal domain (CTD) and three intrinsically disordered regions (namely IDR1, IDR2, and IDR3) that comprise almost 40% of the protein primary sequence. Thanks to its structural heterogeneous nature, N is involved in many crucial mechanisms for the infection cycle [2]. In this context, the interaction between N and several RNA constructs has been studied with various techniques. However, these studies focused on the N's structured domains due to the challenges posed by the absence of structure and the repetitive nature of the primary sequence of the disordered segments.

To obtain atomic resolution information on N disordered regions during RNA encountering, we exploited ^{13}C direct detected NMR experiments in combination with Multiple Receivers hardware [3].

The experiments were conducted on a 248 residue construct we designed comprising the folded N-terminal domain NTD and the flanking intrinsically disordered regions (IDR1 and IDR2) [4] and on a construct encompassing the NTD only, to identify the priming events of N binding to a structured regulatory SARS-CoV-2 5'-UTR RNA element. This also allowed us to unveil the behavior of the NTD, the domain responsible to bind genomic RNA, when connected to the disordered regions and to define the differences between the longer construct (referred to as NTR) and the NTD in isolation.

This is the first step to unravel the detailed molecular determinants of the N protein for its specificity for RNA interaction and to obtain topological information along the primary sequence of the IDRs, useful for the identification of any possible target site for the development of binding competitors for antiviral drug design.

[1] Giri, R.; Bhardwaj, T.; Shegane, M., Gehi, B.R.; Kumar, P., Gadhave, K., Oldfield, C.J., Uversky, V.N., *Cell. Mol. Life Sci.* **2021**, *78*, 1655–1688

[2] Matsuo, T. *Biology (Basel).* **2021**, *10*, 454

[3] Schiavina, M.; Murrali, M.G.; Pontoriero, L.; Sainati, V.; Kümmerle, R.; Bermel, W.; Pierattelli, R.; Felli, I.C. *Biophys. J.* **2019**, *117*, 46–55

[4] Schiavina, M.; Pontoriero, L.; Uversky, V.N.; Felli, I.C.; Pierattelli, R. *Biomol. NMR Assign.* **2021**, *15*, 219–227