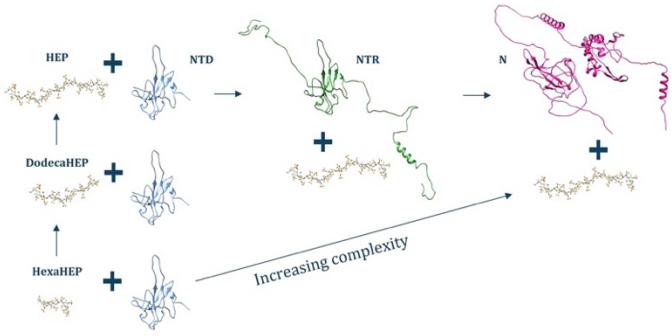
# NMR INVESTIGATION OF STRUCTURAL HETEROGENEITY AND LIGAND RECOGNITION IN THE SARS-COV-2 NUCLEOCAPSID PROTEIN

Tessa Bolognesi[a], Marco Schiavina[a], Cristina Ciabini[a], Stefano Elli[b], Michela Parafioriti[b], Marco Guerrini[b], Isabella C. Felli[a], Roberta Pierattelli[a]

[a] Department of Chemistry and Magnetic Resonance Center, University of Florence, Via L. Sacconi 6, Sesto F.no IT
[b] Istituto di Ricerche Chimiche e Biochimiche G. Ronzoni, Via G. Colombo 81, Milan IT
E-mail: tessa.bolognesi@unifi.it

Among the structural proteins of SARS-CoV-2, the nucleocapsid (N) protein is distinguished by its pronounced structural heterogeneity and multifunctionality throughout the viral life cycle, including RNA packaging, genome replication, and immune modulation. The 419-residue N protein comprises two folded domains, the N-terminal (NTD) and C-terminal (CTD) domains, interspersed with three intrinsically disordered regions (IDR1, IDR2, and IDR3). This mixed organization of ordered and disordered elements raises a central question: how can such a heterogeneous protein be characterized at atomic resolution, and how does structural disorder contribute to its molecular interactions?

To address this, advanced solution NMR spectroscopy was applied in a stepwise strategy [1, 2], progressing from the isolated NTD to the N-terminal region (NTR), which includes flanking disordered segments, and ultimately to the full-length protein. The NTD provided a structural and dynamic reference framework, while analysis of the NTR revealed that intrinsically disordered regions actively modulate domain flexibility and intersegment communication [2]. Despite experimental challenges, investigation of the full-length protein captured its modular organization and the coexistence of structured and highly flexible elements in solution.




Building on this characterization, protein–ligand interactions were examined using low molecular weight heparins of increasing oligosaccharide length. Heparins were selected as model ligands due to their high negative charge density, which mimics key electrostatic features of RNA, and their structural similarity to heparan sulfate in the extracellular matrix, a biologically relevant environment where the N protein has been detected and is expected to interact. NMR analyses revealed a clear correlation between ligand size and binding affinity, with longer saccharide chains inducing more extensive perturbations, particularly within positively charged regions. Inclusion of intrinsically disordered segments significantly enhanced ligand binding compared to the isolated NTD, while the full-length protein displayed distinct spectral behavior consistent with additional binding contributions and altered dynamics [3].

Overall, this work demonstrates the capacity of NMR spectroscopy to dissect the dynamic and multivalent nature of protein–polyanion interactions in structurally heterogeneous systems [1].

1. Schiavina, M., Bracaglia, L., Bolognesi, T., Rodella, M.A., Tagliaferro, G., Tino, A.S., Pierattelli, R., Felli, I.C., 2024. Intrinsically disordered proteins studied by NMR spectroscopy. J Magn Reson Open 18, 100143. https://doi.org/10.1016/j.jmro.2023.100143
2. Bolognesi T., Schiavina M., Felli I.C., Pierattelli R., 2025. NMR insights on multidomain proteins: the case of the SARS-CoV-2 nucleoprotein. Prog. Nucl. Magn. Reson., 148–149, 101577. https://doi.org/10.1016/j.pnmrs.2025.101577
3. Bolognesi T., Schiavina M., Ciabini C., Parafioriti M., Gardini C., Elli S., Guerrini M., Pierattelli R., Felli I.C., 2025. Exploring the role of structural and dynamic complexity in SARS-CoV-2 nucleocapsid protein–heparin interactions by NMR. J. Mol. Biol., in press, 169437. https://doi.org/10.1016/j.jmb.2025.169437

The support of the CERM/CIRMMP center of Instruct-ERIC is gratefully acknowledged. The Covid19-NMR consortium is greatly acknowledged for stimulating discussions. This work was supported in part by the project “Potentiating the Italian Capacity for Structural Biology Services in Instruct-ERIC" (ITACA.SB, Project no. IR0000009) and by the project ID-COVID (FISR2020IP_02112).