Check for updates

OPEN ACCESS

EDITED BY Peter Nagy, University of Debrecen, Hungary

REVIEWED BY

Narva Deshwar Kushwaha, Wayne State University, United States Arvind Ramanathan, Argonne National Laboratory (DOE), United States

*CORRESPONDENCE

Carla Isernia, i carla.isernia@unicampania.it Gaetano Malgieri, ig gaetano.malgieri@unicampania.it

RECEIVED 21 March 2025 ACCEPTED 30 May 2025 PUBLISHED 11 June 2025

comply with these terms.

CITATION

Shitaye G, Ventserova N, D'Abrosca G, Dragone M, Maina EW, Fattorusso R, Iacovino R, Russo L, Isernia C and Malgieri G (2025) The role of intrinsically disordered regions of SARS-CoV-2 nucleocapsid and non-structural protein 1 proteins. *Front. Chem.* 13:1597656. doi: 10.3389/fchem.2025.1597656

COPYRIGHT

© 2025 Shitaye, Ventserova, D'Abrosca, Dragone, Maina, Fattorusso, lacovino, Russo, Isernia and Malgieri. This is an open-access article distributed under the terms of the **Creative Commons Attribution License (CC BY)**. The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not

The role of intrinsically disordered regions of SARS-CoV-2 nucleocapsid and non-structural protein 1 proteins

Getasew Shitaye^{1,2}, Nataliia Ventserova¹, Gianluca D'Abrosca³, Martina Dragone¹, Eunice Wairimu Maina¹, Roberto Fattorusso¹, Rosa Iacovino¹, Luigi Russo¹, Carla Isernia^{1,4}* and Gaetano Malgieri^{1,4}*

¹Department of Environmental, Biological and Pharmaceutical Sciences and Technologies, University of Campania "Luigi Vanvitelli", Caserta, Italy, ²Department of Biomedical Sciences, College of Medicine and Health Sciences, Bahir Dar University, Bahir Dar, Ethiopia, ³Department of Human Sciences, Link Campus University, Roma, Italy, ⁴Interuniversity Research Centre on Bioactive Peptides, University of Naples "Federico II", Naples, Italy

Virus survival inside the host cell depends on the intricate mechanisms that recruit proteins involved in the arms race. Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) proteome exhibits important levels of structural order. However, some of the SARS-CoV-2 proteins, such as the Nucleocapsid (N) and Non-structural protein 1 (Nsp1), contain a considerably significant amount of intrinsically disordered regions (IDRs) that play indispensable roles in the intraviral and virus-host interaction. Here, focusing on proteins that contain a relevant percentage of IDRs, we discuss experimental and computational studies sought to support IDRs as a key player in the interplay with ordered domains, the biological role as potential origin for variants of SARS-CoV-2, and their association with virus transmissibility. Furthermore, we also highlight the potential involvement of IDRs in the viral-host protein interaction and host cellular machinery. Thus, shading lights on the dark proteome of the virus and looking for therapeutic approaches beyond the classic structure-function paradigm may contribute to the efforts sparking the quest for therapeutics.

KEYWORDS

intrinsically disordered proteins, coronaviruses, SARS-COV-2 variants, protein-protein interactions, host-viral interactions

1 Introduction

One of the central dogmas of molecular biology is the sequence-structure-function relationship: the primary sequence of amino acids determines the protein's threedimensional structure that, in turn, defines its function. In the past couple of decades, such a relationship in the field of structural biology has been questioned by the fact that naturally occurring unstructured proteins that deviate from this dogma are increasingly identified. The so-called Intrinsically Disordered Proteins (IDPs), which do not have well-defined secondary, tertiary and, in specific cases, quaternary structures but still exhibit many important biological processes (Dyson and Wright, 2005; Dunker et al., 2001; Kulkarni et al., 2025), have become hot topics in the "proteins world." Just like other functional molecules, disordered proteins play key roles in molecular and cellular events such as cellular signal transduction, cell cycle, transcriptional regulation, and translation (Tompa, 2002; Iakoucheva et al., 2002; Mishra et al., 2020). Moreover, a great number of proteins that are not commonly classified as IDPs encompass extensive segments characterized by an intrinsic disorder, often defined as intrinsically disordered regions (IDRs). These IDRs are widely distributed across the proteome of different organisms with significant variations within the entire protein in terms of composition and percentage content (Dunker et al., 2001; Peng et al., 2015).

Despite the fact that disordered regions act as mediators of cell signaling and other indispensable biological processes, generally, their essential role has a biological cost: IDRs, not protected by a structure, are more prone to involvement in mis-functional processes, i.e., aggregation above all.

Physico-chemical changes in the cellular environment can, in principle, affect their normal functions consequently leading to pathogenesis and diseases.

One of the hallmarks of many viruses is the effective utilization of their own IDPs/IDRs to perform essential cellular processes to escape from the host defense mechanism (Davey et al., 2011; Dyson and Wright, 2018). This outlines the importance of a thorough characterization of the role of IDRs to understand the molecular mechanism of diseases beyond the classic structure–function paradigm.

Even though the characterization of IDPs and IDRs is a major challenge in biology, since the outbreak of Coronavirus disease 2019 (COVID-19) tremendous efforts have been made to unravel the structural biology of IDRs in the SARS-CoV-2 proteome. A plethora of computational and experimental spectroscopic techniques such as Nuclear Magnetic Resonance (NMR), Circular Dichroism (CD) and Small-angle scattering of X-rays (SAXS), have been used to study these dancing proteins. Beyond reviewing the common knowledge on IDPs and IDRs, the aim of this review is summarizing and analyzing the intrinsically disordered regions of the SARS-CoV-2 proteome, the information available on their biological role, their interaction with globular domains, their importance as source of mutations and viral transmissibility, and the potential involvement with the host cellular machineries. Firstly, we highlighted the percentage of the intrinsic disordered regions within different proteins across the SARS-CoV-2 proteome and opted to summarize the role of IDRs in Nucleocapsid and Non-structural protein 1. We then discuss the role or association of IDRs with SARS-CoV-2 infectivity and transmissibility along with experimentally and computationally explored molecular driving forces acting as local determinants of IDRs conformational changes. Finally, focusing on the SARS-CoV-2 proteins that contain significant disordered regions, we highlighted the potential involvement of their IDRs in the viral-host protein interactions and the effects on the host cellular machinery.

2 Overview of genome organization and IDR content of coronaviruses

Coronaviruses (CoVs) are a widely varied family of enveloped positive-sense single-stranded RNA viruses. By virtue of their ubiquitous presence in nature, they are broadly spread among mammals and avian species. CoVs belong to the *Coronaviridaes* family of the *Nidovirales* order and can be classified into four genera: alphacoronavirus, betacoronavirus, gammacoronavirus and deltacoronavirus.

Betacoronavirus (β -CoV), such as Severe Acute Respiratory Syndrome Coronaviruses (SARS-CoV-1 and SARS-CoV-2) and Middle Eastern Respiratory Syndrome virus (MERS-CoV), are highly pathogenic viruses for humans that share similar genome and replication strategies (Enjuanes et al., 2006; V'kovski et al., 2021).

The SARS-CoV-2 genome, with an approximate length of 30 kb, encodes structural proteins, non-structural proteins (NSPs), and a number of accessory proteins. The four major structural proteins, namely, Spike (S), Nucleocapsid (N), Membrane (M), and Envelope (E) protein, are essential components required for the production of a structurally complete viral particle. Specific interactions between the viral S protein and host cell surface factors, serine protease TMPRSS2, and the receptor angiotensin-converting enzyme 2 (ACE2) promote viral uptake and cellular fusion. N protein is implicated in RNA encapsidation, while proteins E and M guarantee RNA genome (+ssRNA) incorporation into viral particles during the assembly process.

Upon infection, the incoming viral RNA is uncoated inside the host cytosol and two large viral open reading frames, ORF1a and ORF1b, are translated into two polyproteins, pp1a and pp1ab, that are further cleaved into 11 and 15 Nsps, respectively. The 16 Nsps thus produced form the viral replication and transcription complex (Figure 1) (V'kovski et al., 2021; Masters, 2006; Wang et al., 2017; McBride et al., 2014; Malone et al., 2022).

Bioinformatics studies analyzed the percentage of intrinsic disorder (PID), quantified as the ratio between residues predicted to be disordered and the total number of residues, in the MERS-CoV proteome and in individual proteins derived from the MERS-CoV genome. Predicted percentage of intrinsic disorder (PPID) is a metric calculated as the ratio of disordered residues (those predicted to be disordered) to the total number of residues, expressed as a percentage, which is used to estimate the degree of intrinsic disorder within a protein (Rajagopalan et al., 2011; Redwan et al., 2019). Thus, the levels of PPID were applied to categorize proteins as highly ordered (PPID <10%), moderately disordered (10% ≤ PPID <30%), and highly disordered (PPID ≥30%) (Rajagopalan et al., 2011; Giri et al., 2021). According to these examinations, N proteins are characterized by a higher value of PID (Goh et al., 2020; Goh et al., 2013; Alshehri et al., 2021).

In the scenario of SARS-CoV-2, a significant (refers to the percentage) level of disorder content was found for the S1 subunit of S protein and up to 51% for N protein. In these intrinsically disordered regions an abundance of mutations occurring in the 14 major variants were localized [*Variants of Concern (VOC)*]: Alpha, Beta, Gamma, Delta, Omicron BA.1 and BA.2; Variants of Interest (VOI) Lambda, Mu, Epsilon, Zeta, Eta, Theta, Iota and Kappa.) that could favor the virus with benefits for genetic and antigenic drift (Anjum et al., 2022; Quaglia et al., 2022). Another study buttresses that N protein possesses a predicted PID value of more than 60% and demonstrates moderate content of disorder for some of the nonstructural proteins and accessory proteins, namely, Nsp8, Nsp1, ORF6, ORF9b (Giri et al., 2021; Anjum et al., 2022; Rezaei et al., 2022). Interestingly, the



nonstructural proteins, including Nsp2-Nsp6, Nsp9, Nsp10, and Nsp12-Nsp16 of SARS-CoV-2 proteins, show a quite less content of disordered regions (Giri et al., 2021; Anjum et al., 2022).

2.1 The nucleocapsid protein

As previously discussed, within the SARS-CoV-2 proteome, the N protein stands out with a substantial amount of predicted disordered content. The SARS-CoV-2 N is a 419 amino acid protein that bears two ordered domains: the RNA-binding N terminal domain (NTD) (residues 48–175), which contains the RNA recognition motif (RRM), and the C terminal domain (CTD) (248–365) (Kang et al., 2020; Jayaram et al., 2006; Grossoehme et al., 2009; Huang et al., 2004), primarily associated with the formation of higher-ordered structures. The N protein also

contains three intrinsically disordered regions: IDR1 (residues 1-48), the central SR-IDR2 (residues 175-248), which is enriched with serines and arginines and separates the two ordered domains, and the C-terminal IDR. Interestingly, the N in complex with its partner proteins Nsp3a exhibits dynamic nature (Figure 2). For example, as clearly demonstrated from its solution NMR structure (PDB ID: 7PKU), the SR-IDR2 shows a tendency to form a helix (²¹⁹LALLLDRLNQL²³⁰) and disordered polar strand (²⁴³GQTVTKKSAAEAS²⁵⁵) (Bessa et al., 2022). N shows structural plasticity and its predicted structure could sample modest helical conformation based on the sequence 390 QTVTLLPAADLDDFSKQLQQSMSSADSTQA419 (IDR3 region) (Zhao et al., 2022).

A great number of studies have shown that several N proteins bind to multiple RNA molecules without exhibiting binding specificity to RNA sequences (Carlson et al., 2020; Iserm et al.,



FIGURE 2

Scheme of the overall structure of the SARS-CoV-2 N protein. N protein consists of the N-terminal arm (IDR1), the RNA-binding N-terminal domain (NTD), the central linker region (LKR) within a Ser/ Arg (SR)-rich motif, the C-terminal domain (CTD), associated with dimerization, and the C-terminal tail (C-tail). The up-left panel illustrates the surface of the amino-terminal (N-terminal) domain (PDB ID: 6M3M) Kang et al. (2020). The up-right panel illustrates the surface representation of the dimeric carboxy-terminal (C-terminal) domain (PDB ID: 6YUN)) (each monomer presented by separate color code). Zinzula et al. (2021). The lower panel represents AlphaFold 3 prediction of full-length N protein encompassing folded domain (surface representation) and IDRs Abramson et al. (2024).

2020; Perdikari et al., 2020; Savastano et al., 2020). It is assumed that a single molecule of genomic RNA is enveloped by multiple N proteins and packaged into virions (Yao et al., 2020; Cubuk et al., 2021; Klein et al., 2020).

In vivo studies showed that the recognition of RNA that has to be packed requires both the NTD and CTD domains of N protein (Woo et al., 2019); specifically, the N-terminal linker region and C-terminal segments are required for RNA recognition and binding. (Jayaram et al., 2006; Grossoehme et al., 2009; Cui et al., 2015; Yu et al., 2005; Nelson and Stohlman, 1993) (Cubuk et al., 2021; Chang et al., 2009; Wu et al., 2021). By means of charge distribution analysis of dimeric CTDs, Jia et al. identified in their surface a stretch of positively charged residues (between Lys257 and Arg262) involved in RNA-binding (Jia et al., 2022) and conserved among SARS-CoV-1, SARS-CoV-2, and MERS-CoV (Figure 3).

Mutations of these conserved positive residues can significantly decrease the binding affinity with RNA, confirming their role in RNA binding (Chen et al., 2007; Hsin et al., 2018; Yang et al., 2021). Upon binding to the RNA, the SARS-CoV-2 N protein undergoes a more compact conformation compared to its free form (Ribeiro-Filho et al., 2022). Interestingly, the stability of the small N protein structure seems to be dependent on the N-terminal RNA-binding domain (Sekine et al., 2023). Inhibiting the RNA binding specificity of N proteins in betacoronaviruses with antiviral drugs could combat a range of these viruses, aiding in future outbreaks.

2.2 Non-structural protein 1 (Nsp1)

Several research delves into the characterization of the fulllength SARS-CoV-2 Nsp1 (1–180) (Figure 4), which plays a dual role through initiating viral protein synthesis and inhibiting host translation; it engages with the host 40S ribosomal subunit, thereby prevents initiation of host protein translation. Crystal structures of Nsp1 proteins of SARS-CoV (aa 10–126) (Almeida et al., 2007) and SARS-CoV-2 (aa 12–127) (Clark et al., 2021; Semper et al., 2021) show nearly the same folded N-terminal domain architecture. Interestingly, most of the residues predicted to be disordered are localized in the C-terminal tail (125–179) that directly binds within the mRNA entry tunnel of the 40S (Agback et al., 2021; Ilyas et al., 2025).

2D ¹⁵N-{¹H} NOE NMR experiment on full-length SARS-CoV Nsp1 protein revealed that its residues 1 to 12 and 129 to 179 have small positive or negative ¹⁵N,¹H heteronuclear NOE values, characterizing these regions as short N-terminal and long C-terminal disordered tails (Almeida et al., 2007) with a remarkable conformational flexibility at sub-nanoseconds timescale. Using cryogenic electron microscopy, different groups of researchers were able to solve the structure of the C-terminal portion of SARS-CoV-2 Nsp1 in complex with the 40S ribosome





(Thoms et al., 2020; Schubert et al., 2020). These structures confirm the presence of a short α -helix (P153 to N160) and a second Cterminal α -helix spanning residues S166 to N178. Likewise, the study by Semper et al. combines SARS-CoV-2 Nsp1 (13–127) crystal structure (PDB ID:7K3N) with amino acid sequence analysis, data from prediction tools, and homology-based models, showing the presence of 12 N-terminal residues comprising disordered regions as well as a largely disordered linker between the end of the globular domain and the C-terminal helix-turn-helix (Semper et al., 2021).

Furthermore, a C-terminal helix-turn-helix motif is proposed in their homology-based model of full-length Nsp1 (Semper et al., 2021). While the study of Kumar and coworkers (Kumar et al., 2021), in line with previous studies, validated the C-terminal IDR region (131–180) in isolation, Thoms et al. have shown that after interaction with the 40S ribosome, such region adopts a helical structure (Thoms et al., 2020) suggesting its important role in regulating host mRNA translation.

Secondary structure analysis of the full-length SARS-CoV-2 Nsp1 protein by means of NMR (Agback et al., 2021; Wang Y. et al., 2021) revealed one folded domain that comprises residues T14–L125 and two disordered chains at the N-terminus (residues 1–13) and at the C-terminus (residues 126–180), respectively. As clearly demonstrated by Agback and coworkers (Agback et al., 2021), SARS-CoV-2 Nsp1 in solution shows different dynamics of N- and C-termini. They were also able to demonstrate that residues located between the folded domain and the disordered C-terminal exhibit different conformations. For example, the region at the beginning of the C-terminal domain (¹²²LLRK¹²⁵) has been shown to adopt two different conformations distinguishable on the NMR time scale even at high temperature (308K) (Agback et al., 2021).

3 Globular domain–IDRs interplay

Along with interplay between the structured part and the IDRs within each specific protein, the vast majority of disordered regions

in the SARS-CoV-2 are also involved in protein-protein interactions with other viral and host proteins. Protein-protein interaction (PPI) is a key biological process, established often in a specific and stable manner by a physical contact of two or more proteins in cells which form a complicated network termed "interactome" (Koh et al., 2012; Venkatesan et al., 2009).

Due to the dynamic nature of proteins, the environment in which they exist largely influences their motion, and it must be taken into proper consideration while discussing protein-protein interactions (PPI) (Figure 5).

IDPs are prone to posttranslational modifications (Darling and Uversky, 2018; Jin and Gräter, 2021; Niklas et al., 2015) and, therefore, are likely to exist as chemically heterogeneous intracellular populations. Compared to the structured proteins, IDPs are more susceptible to environmental changes that can determine the IDR/IDP structure and function. The term "environment" intends a broad context: physicochemical environment (pH, viscosity, pressure, salt concentration, etc.) (Lindsay et al., 2021; Verma et al., 2020), biomolecular environment, (interactors and location), physiological environment (cell-type, cell-cycle, stress, etc.) (Verma et al., 2020), and intramolecular interactions (Bugge et al., 2021; Bugge et al., 2020).

As well documented from physicochemical (Lin et al., 2018) and quantitative studies (Bowman et al., 2020), amino acid sequence encodes IDRs conformational behavior as well as their interaction and function (Das et al., 2015; Chong and Mir, 2021). Collectively, many studies evidence that polypeptides and proteins that carry high net charges are known to undergo extended conformations. Charged residues in intrinsically disordered proteins also drive the interaction with biomolecules (Bowman et al., 2020; Milorey et al., 2021).

Transient interactions are another attribute that characterize disordered proteins. Multiple interactions contribute to their dynamic behavior and to possible structural transitions. In many cases, hydrophobic interactions primarily play a role in the disorderstructure paradigm of the IDR systems. Abyzov et al. showed



interface Adapted from Spreitzer et al. (2020).

liquid-liquid phase separation (LLPS) derived as a result of intermolecular interactions. The interaction was observed to be responsible for modulating internal motions, slowing down the time scales of IDP motions, and restricting the amplitude of the local backbone conformational sampling (Abyzov et al., 2022). In their analysis, the authors noted that the locally accessible surface determines the interaction between IDP tails and the folded domains that in turn act as local determinants of IDR conformational behavior. In their study, Taneja and Holehouse exploited this behavior to investigate the role of the relative position and accessibility of charged residues across the surface of a folded domain. Simulations on homologs protein with variable charge distribution demonstrated that even slight changes in surface charge impact IDR behavior and interaction with folded domains (Taneja and Holehouse, 2021). Of note, due to the presence of multiple charge patches on naturally occurring proteins and their physical accessibility to local IDRs, these changes in the chemical environment are highly relevant to modulate the interaction (Wohl et al., 2025).

Mukherjee et al. used conformation-sensitive single-molecule Förster resonance energy transfer and CD to understand the effects of cellular components on the conformation of the SARS-CoV-2 RG-1 RNA G4Q. The authors confirmed that salts, co-solutes, crowders, and other IDPs affect the conformation of the SARS-CoV-2 RG-1 RNA G4Q. For instance, in the absence of monovalent K⁺ salt, the RG-1 SARS-CoV-2 RNA-G4Q exists as an equilibrium between folded and metastable conformation, whereas at a high concentration of K⁺, only the fully folded conformations are preserved. Recent reports of Parkinson's disease cases involving young patients after SARS-CoV-2 infections showed that amyloidogenic disordered a-Syn-derived peptides stabilized the folded state of RG-1 G4Q, although human islet amyloid polypeptide (hIAPP) stabilized the partially folded state, leading to the conclusion that these interactions may afflict expression profiles of disease-modifying genes (Mukherjee et al., 2022). Yesudhas et al. used molecular dynamics simulations to investigate the binding contributions of intrinsically disordered residues of S protein to human ACE2 proteins. They observed a disorder-to-order transition (DOT) profile upon binding that is strengthened mainly by the hydrophilic and aromatic residues localized in the region that undergoes such disorder-to-order transition. The authors hypothesized that this more energetically favorable binding of receptor binding domains with ACE2 is due to the positively correlated motion of the DOTs with its neighboring residues (Yesudhas et al., 2021; Akbayrak et al., 2022). Moreover, the conformational disorder of the N- and C-terminal residues and water accessibility of the SARS-CoV-2 envelop protein increased in acidic pH and high calcium ions concentration (Medeiros-Silva et al., 2022).

3.1 Nucleocapsid

While considerable emphasis has been placed on the S protein, numerous other proteins of SARS-CoV-2, such as N protein, also hold pivotal significance in viral physiology. However, our understanding of their structural and biophysical characteristics remains comparatively limited. N is an abundant RNA-binding protein involved in packaging and transcription facilitation. It is also one of the mutational hotspots in the SARS-CoV-2 genome.

As reported in Section 2.1, the SARS-CoV-2 N comprises three predicted intrinsically disordered regions, an N-arm (IDR1), a C-tail (IDR3), and a centrally located linker region (SR-IDR2) that connects the NTD and CTD. In their elegant work, Cubuk and co-authors using single-molecule fluorescence spectroscopy with all-atom simulations proved the structural heterogeneity of these three IDR domains (Cubuk et al., 2021).

In betacoronaviruses, the IDR site (the N protein primary phosphorylation site) that acts as a linker between the NTD and CTDdomains has arginine/serine rich region and can undergo phosphorylation, that in turn affects the protein function (Carlson et al., 2020; Peng et al., 2008; Lu et al., 2021). By means

of nanosecond fluorescence correlation spectroscopy (FCS), a slower reconfiguration time ($\tau_r = 170 \pm 30$ ns.) in comparison with other proteins with similar length and charge content was revealed, suggesting that the NTD transiently interacts with the RNAbinding domain. This linker limits the interaction between the NTD and CTD domains but enables them to have an independent function (Cubuk et al., 2021), which probably prevents the formation of NTD dimers. The unstructured central linker can critically affect the N proteins binding of IDR linker to both the CTD and NTD through intra and inter molecular interactions, (Różycki and Boura, 2022), the direct interaction of this linker to RNA *via* electrostatic interaction, and even the phosphorylation process (Carlson et al., 2020; Peng et al., 2008), can cause conformational changes in N protein.

Biomolecular condensates, essential for various cellular processes, can arise from the LLPS containing IDRs and RNAbinding domains. Then, a few studies have shown that, at the body temperature, the S/R-rich central disordered region significantly associates with the RNA-dependent phase separation of N protein where its phosphorylation regulates the condensate viscosity of cells (Iserm et al., 2020; Lu et al., 2021). Additionally, the recent study on the binding and compaction of SARS-CoV-2 full-length N protein and its truncated variants containing the globular domains with and without the linker region (Morse et al., 2023) confirms the pivotal role of the intrinsically disordered regions in maintaining the biological function of N protein. In detail, Morse et al. demonstrate that the N protein region IDR1-NTD-IDR2 missing the CTD and the disordered C-tail shows similar performance to compact the ssDNA as the full-length N protein, whilst the same portion without the IDR2 linker shows a dramatic lower activity. These binding differences to the ssDNA are due to the richness in serines and arginines of the IDR2 region that enables the preservation of the fast on/off kinetics and non-cooperative binding. It is also postulated that IDR2 may provide structural plasticity to the N protein (Zhao et al., 2022).

Experimental studies have quantified the affinity of N protein for RNA. For instance, Wu and his colleagues investigated contribution from each N protein domain to ssRNA binding (Wu et al., 2021). They found that N wild type binds the 20-nt ssRNA (sequence: UUU CACCUCCCUUUCAGUUU) with high affinity ($K_D = 0.007 \pm 0.001 \mu$ M). In contrast, the isolated N_{NTD} and N_{CTD} have low-affinity binding ($K_D = 20 \pm 10$ and $13 \pm 5 \mu$ M, respectively). However, inclusion of the LKR region increased RNA binding affinity significantly (0.50 \pm 0.08 and 0.35 \pm 0.04 μ M for N_{NTD}-LKR and N_{LKR-CTD}, respectively). Addition of CTD onto NTD-LKR in cis increases the binding affinity to the single-digit nM range (0.006 \pm 0.002 μ M), but not in trans. Overall this data quantitatively indicates that the wild-type N protein has high-affinity ssRNA binding, and unstructured regions contribute into ssRNA

The capacity of N protein to undergo LLPS is well-documented, defining starting N protein concentrations equals to 10 μ M. However, the droplet sizes did not grow with increased concentration, even at a protein concentration of 40 μ M. Interestingly, size of condensates was modulated by increasing RNAs concentration (at 1:0.25 protein: RNA ratio, the size and the sphericity of the droplets further increased. Since RNA is involved in the assembly of the viral to form a helical ribonucleoprotein (nucleocapsid), these condensates are proposed to enhance the efficiency of viral replication. Thus, these condensates could become a potential therapeutic target in order to disrupt cycles of SARS-CoV-2 replication (Wang J. et al., 2021).

Moreover, Morse et al. (Morse et al., 2023) suggest a description of a multistep process through which N proteins can bind and package viral RNA (vRNA). In particular, the N protein, which is abundantly expressed in the host cell, forms CTD-mediated dimers and quickly binds vRNA through NTDs and linkers, due to electrostatic interaction. This binding is non-cooperative and fully reversible. Over time, the binding increases resulting in substrate compaction as neighboring proteins multimerize. This assembly mechanism is modulated by the C-terminal domain, which can fully compact the substrate. Note that the ssDNAprotein complex compaction observed in this study is less extensive with respect to vRNA-N protein compaction inside the viral particle.

As reported above, the N-terminal and C-terminal IDRs contribute to the RNA-binding activity of the SARS-CoV-2 N protein. Using a biosafety level-2 cell culture system (Ju et al., 2021), Luo et al. found that deletion of IDR1 completely abolished SARS-CoV-2 production (Luo et al., 2021). Similarly, truncations of IDR1 weaken the RNA binding, as demonstrated by a chromatography study of the nucleoprotein both in RNA-bound states and RNA-free state (Wu et al., 2021), further suggesting the relevant biological role of IDRs cooperation with the folded domains.

Though both the NTD and CTD can bind RNA, practically no direct cooperation is observed between the two domains in RNA binding (Kubinski et al., 2025). Experimental data from combined coarse-grained molecular simulations with data from small-angle X-ray scattering (SAXS) on full-length SARS-CoV-2 N protein evidenced this scenario (Różycki and Boura, 2022). From the simulations, several intra- and inter-chain contacts between the formed hydrophobic α -helix within the IDR2 and the two ordered domains have been observed. This α -helix of IDR2 showed the most frequent degree of contact with the NTD of the same chain than the NTD of the other chain in the N dimer (Różycki and Boura, 2022).

IDR1 in SARS-CoV-2 N protein has been also found to be an indispensable mediator for the interaction between N protein and Ras-GTPase-activating protein SH3-domain binding (G3BP1), where the interaction leads to stress granule (SG) disassembly (Luo et al., 2021). As confirmed by the mutant, the N protein lacking IDR1 impaired its capability to colocalize with whereas deletion of the C-terminus containing G3BP1 homodimerization domain and IDR3 can facilitate the N protein diffusion within droplets. IDR1 initiates the LLPS of the N protein, whereas the C-terminal region stabilizes the condensate, likely by facilitating protein homodimerization. Meanwhile, SARS-CoV-2 N supposedly interacts with G3BP stress granule assembly factor 2, thereby consequently suppressing the integrated stress response and inhibiting type I interferon responses (Kruse et al., 2021; Huang et al., 2021; Biswal et al., 2022). The N-terminal intrinsically disordered tail is believed responsible for recruiting N protein into stress granules (Chang et al., 2014).

In the context of SARS-CoV-2 infection, Enoxaparin (EP) (a negatively charged, low molecular weight heparin) acted as an

anticoagulant medication. The 2D CON NMR experiments in the presence of EP showed considerable chemical shift perturbation within the arginine-rich region of IDR1 ³⁷SKQRRPQ⁴³ and influenced the resonances in the IDR2 segment mainly composed of leucine residues ²¹⁷AALALLLL²²⁴, by promoting this region to sample a helical conformation. Such changes do not describe this scenario as a direct interaction with a negative "ligand" but lead to the conclusion that interaction with EP disrupts intramolecular interactions, as also observed upon RNA binding (Schiavina et al., 2022).

Replication-transcription complex (RTCs) is one of the vital machinery that enables the virus to recruit proteins. In betacoronaviruses, N is the major cofactor of this complex that interacts with the amino-terminal ubiquitin-like domain of Nsp3 (Ubl1). The study of Bessa et al. revealed a bipartite interaction between Ubl1 and two motifs in the IDR2 linker of N a hydrophobic helix (²¹⁹LALLLLDRLNQL²³⁰) and a disordered polar strand (²⁴³GQTVTKKSAAEAS²⁵⁵). These interactions induce the folding of portions ofIDR2 around Ubl1, that in turn brings conformational changes in N protein and regulates its RNA binding (Bessa et al., 2022). Thus, the intrinsically disorder regions are hypothesized to be important players in complex mechanisms that underlie the regulation of ribonucleoprotein complex formation.

Serine and arginine residues are highly disorder-promoting residues (Campen et al., 2008), and are the most abundant in the flexible intrinsic disordered linker region of all types of coronaviruses (Barik, 2020). Coronavirus N protein phosphorylation is a pivotal process for the viral life cycle that predominantly occurs within distinct regions of intrinsically disordered proteins. Primarily, N protein phosphorylation is processed within the SR rich intrinsically disordered region, both in host cells and virions (Chen et al., 2005; White et al., 2007). However, the mechanisms and potential function of multiple SR rich motifs are still inadequately comprehended (Yaron et al., 2022; Bouhaddou et al., 2020; Davidson et al., 2020). N protein interacts with many viral proteins, nucleic acid, and the host protein factors through specific and nonspecific mechanisms of interaction. Such interaction may be facilitated, at least in part, by the phosphorylation process occurring mainly in the rich SR amino acid region of the IDRs.

It has been reported that phosphorylated N-protein strongly affects the RNA binding specificity (Wu et al., 2021) and nucleocytoplasmic shuttling of the N proteins (Bai et al., 2021). Hyper-phosphorylation of N protein also plays a role in promoting the interaction between replication-transcription complex and viral replicase subunit Nsp3 (Keane and Giedroc, 2013). Notably, phosphorylation offers a mechanism for the N protein to recruit host factors like RNA helicases that promote viral RNA template switching and subgenomic mRNA transcription (Wu et al., 2014). The N protein is presumably engaged to host mRNA and G3BP1, thus playing a role in suppressing the host innate immune response. Additionally, even though the underlying mechanisms are not clearly known, the phosphorylation of the central disordered region promotes the protein transcriptional function (Carlson et al., 2020). Indeed, the indispensable role of intrinsically disordered regions in the phosphorylation processes is worth mentioning and more should be explored on their regulatory roles and the phosphorylation itself in modulating and expanding the repertoire of biological activities of N protein in SARS-CoV-2.

Therefore, taken together all the above highlighted hallmarks make IDRs of the N protein a pivotal avenue for SARS-CoV-2 therapeutic advancement.

3.2 Non-structural protein 1

The biological function and mechanism of activity of regulation of Nsp1 seems dependent on the interplay between the N-terminal folded domain and the unstructured C-terminal domain, which comprises IDRs that bind to the host ribosome.

The role of the N-terminal domain in translation inhibition of Nsp1 and the binding of SARS-CoV-2 Nsp1 to the stem-loop structure of RNA 5'-untranslated region (SL1) without the ribosome are two important topics. These SL structures are components of RNA 5'-untranslated region (5'UTR) that plays a role in regulation of gene expression, translation initiation, RNA folding, and stability. The SL1 region of the SARS-CoV-2 5'UTR is a highly conserved sequence necessary to evade Nsp1-mediated translational suppression (Vora et al., 2022; Lan et al., 2022).

From solution NMR studies of SARS-CoV-2 Nsp1 (Agback et al., 2021), Agback et al. proposed intramolecular interactions between the disordered and folded Nsp1 domains. In fact, the Nsp1 H81P mutant shows chemical shifts perturbations within those parts of the protein that link structured/unstructured regions (amino acids 10-17 and 120-127). Similarly, other groups of researchers also suggested possible interactions between the disordered and folded Nsp1 domains (Figures 6A-E) (Wang Y. et al., 2021). These findings suggest that these domains are not fully independent units. Indeed, the folded NTD and the C-terminal intrinsically disordered region interact and their interaction is essential for the regulatory mechanism governing the activity of SARS-CoV-2 Nsp1. Molecular dynamics and simulation studies reported for SARS-CoV-2 the possible binding between SL1 and both the N-terminal globular and the C-terminal intrinsically disordered region of Nsp1 (Vankadari et al., 2020; Sakuraba et al., 2022): several interactions were observed between Nsp1 (residues including T12, Y118, R124, K125, N128, K129, L141, and D147) and SL1. In order to validate these findings, the residues K47, K58, R124/K125 were replaced with Alanine. Nsp1wt, K47A, and K58A mutants were co-immunoprecipitated with the luciferase RNAs containing the 5'UTR. Interestingly, coimmunoprecipitation of R124A/K125A with the luciferase RNAs containing the 5'UTR was not detected, suggesting the crucial role of Nsp1 residues R124, K125 for viral RNA recognition (Tanaka et al., 2012; Terada et al., 2017; Mendez et al., 2021). From previous observations, it was assumed that the C-terminal domain of SARS-CoV-2 Nsp1 is necessary and sufficient to shut off host gene expression (Schubert et al., 2020). Later on, other studies based on in vitro and in cell experiments, further underlined that the coordinated mechanism of action including the N-terminal and central domains of Nsp1 are essential for host translation inhibition (Mendez et al., 2021; Wang et al., 2023). This could further strength the hypothesis on the molecular interplay among the folded domains and disordered regions.



2.3 Å respectively. (PDB ID:8AOU). Adapted from Wang et al. (2023).

Wang and coworkers solved the solution NMR structure of SARS-CoV-2 full-length Nsp1 (PDB ID: 8AOU) (Wang et al., 2023). They demonstrated, via 3D NOESY experiments, the interactions between the tails and the globular domain. As a result, it was pinpointed that the interplay between the folded NTD and the disordered C-terminal tail is required for the Nsp1 mechanism at the base of its regulation activity. For example, long-range NOEs were observed between the side chains of L4 and L122 and between V5 and T78. The unambiguous NOEs which were reproduced in all structures of the NMR ensemble indicate that residue L177 of the CTD contacts V35 of the core domain while in the side chains of NTD (Figure 6C), V89 contacts A131 of the linker region (Figure 6E), W161 of the flexible C-terminal tail contacts residues R43 and K125 of the N-terminal domain. This latter residue belongs to the conserved amino acid motif (LRKxGKG) that was found to be important for the RNA-binding activity of Nsp1. Moreover, the residues in the CTD that form helices interacting with the 40S ribosomal subunit in the Cryo-EM structure of Nsp1 (Schubert et al., 2020) are mostly disordered in free Nsp1. Thus, in the free form, C-terminal residues tend to interact with NTD, as an interaction was observed at W161 and L177 on the surface of the Nsp1 NTD (Wang et al., 2023). Another concept that seems controversial is the binding of SARS-CoV-2 Nsp1 to SL1 without the ribosome. Some studies suggest a 40S ribosome-mediated association between Nsp1 and the 5'UTR of viral mRNA (Mendez et al., 2021; Tidu et al., 2021); others also observed a direct interaction between SARS-CoV-2 5'UTR RNA and the Nsp1-NTD (Vora et al., 2022). However, Wang et al. clearly demonstrated that free Nsp1 has no binding affinity for both the fulllength and stem-loop constructs of RNAs derived from the 5'UTR region of the SARS-CoV-2 genome. The authors further propose that the interaction of the C-terminal tail with the positively charged surface of the Nsp1 NTD may prevent non-specific RNA binding due to functional interplay between the two domains.

Wang et al. also uncovered the critical role of NTD in 40Smediated viral mRNA recognition and promotion of viral gene expression. Nsp1 NTD, the linker, and the C-terminal tail are all involved in recruiting mRNA, further implicating that the intramolecular interaction and communication between the domains is required for the biological activity of Nsp1. Moreover, to validate the role of the linker, Shi et al. made Nsp1 constructs with different lengths of the linker between the Nsp1-N terminal domain and the Nsp1-C terminal domain. Accordingly, Nsp1 constructs with a linker domain extended beyond 20 residues were incapable of supporting viral translation, underscoring the significance of this spacer in the regulation of Nsp1 activity (Vora et al., 2022). Furthermore, upon binding of viral 5' UTR to the Nsp1-NTD, the linker undergoes conformational changes that, in turn, significantly decrease the interaction of the CTD domain with the mRNA entry site of the 40S ribosome (Mendez et al., 2021; Tidu et al., 2021).

4 The role or association of IDP sites with variants of SARS-CoV-2 and infection

As discussed in the previous sections, the interaction of N-protein with the genomic RNA and other membrane proteins is a vital process for the assembly and release of new virions. To get insights into the evolution of drug resistance mutations, it is imperative to comprehend the mechanisms by which mutations are rectified by the virus. In this part, we summarize the contribution IDRs and IDPs across the SARS-CoV-2 genome in bearing mutations, modulating the protein functions and association with variants of SARS-Cov-2.

Genetic variations across the SARS-CoV-2 genome are associated with the frequency of variants, the transmissibility of the virus, and the anti-viral immune response of the host. Mutations in different proteins govern viral replication efficiency, transmissibility, and virulence properties. In the same scenario, mutations affect the properties of proteins and IDR patterns; they may increase as well the size of the IDRs, all of which have an essential effect on folding and SARS-CoV-2 pathogenesis. As IDRs mediate coronavirus transmission (Goh et al., 2020; Tomaszewski et al., 2020; Podder et al., 2021), offering also sequence and structural flexibility, they are believed critical for immune evasion and antibody escape (Quaglia et al., 2022; Li et al., 2024) through diversification and adaptation within its new host.

Several lines of experimental and computational studies were accomplished soon after the emergence of the pandemic, clearly demonstrating that expansion and frameshift of mutations, including the dominant variants of N and S proteins, mainly emerge across the intrinsic disorder regions in the SARS-CoV-2 proteome. This points out the vital role of IDPs in the emergence of new variants and higher transmissibility than previous SARS-coronaviruses (Quaglia et al., 2022; Feng et al., 2024; Fang et al., 2021).

The infection of SARS-CoV-2 in humans requires S proteins, which bind to the human ACE2 proteins. Particularly, several disordered regions in the receptor binding domains (RBDs) of Sprotein facilitate the interaction (Yesudhas et al., 2021). Multiple studies observed numerous pairwise interactions within unstructured regions while examining the atomic interaction between SARS-CoV-2 and ACE2 (Bhattacharjee et al., 2021; Celik et al., 2021; Michelitsch and Weissman, 2000; Campbell et al., 2021).

The mode of transmission and rigidity of the coronavirus outer shell appeared to be influenced by the unstructured region of Membrane and Envelope proteins (Goh et al., 2020), confirming the strong association between variants of SARS-CoV-2 and its intrinsically disordered regions (Quaglia et al., 2022; Gomari et al., 2023; Sinha and Roy, 2023; Ye et al., 2021).

From the analysis of 219.909 SARS-CoV-2 N sequences, most mutations localized in the intrinsically disordered regions, particularly in the SR-rich linker region (Zachrdla et al., 2022), that were as such believed to be the culprits in increasing the spread of the SARS-CoV-2 virus (Johnson et al., 2022; Rahman et al., 2021).

At the early stages of the SARS-CoV-2 outbreak, the increased spread of variants, including B.1.617.2 (Delta), which contains the R203M mutation has been noticed. Later on, this mutation (R203M) in N and D614G in S were documented along with the other six mutations (P323L in Nsp12, D614G in S, the T492I in Nsp4, R203M in N, T60A in Orf9b, and P1228L in Nsp3) that had greater than 50% frequency in the global SARS-CoV-2 mutations (Abbasian et al., 2023).

In a general context, the diversity of structural ensembles for proteins and protein-protein complexes can be explained by the major concepts of frustration and fuzziness. As illustrated in Figure 5, researchers used the terms frustration to describe conformational variations in proteins and fuzziness to describe how conformational heterogeneity occurs in protein complexes involving IDRs (Freiberger et al., 2021; Gianni et al., 2021). In this regard, Haque et al. analyzed 228 mutations from each domain of the N protein of the SARS-CoV-2. Accordingly, 11 mutations are predicted to be deleterious and destabilizing. Among these mutations, R32C, R191C, and R203 M mutations are localized in the disordered regions and show a significant change in frustration state (from neutral frustrated to major minimally frustrated state), which could raise the sampling of conformation space (Haque et al., 2023). It is worth noting that the R203M mutation improves RNA

packaging mechanism and production of higher titers of infectious virions (Haque et al., 2023; Syed et al., 2021).

Phosphorylation of the SARS-CoV-2 Nucleoprotein recruits human cytosolic 14-3-3 proteins, essential for virus replication. 14-3-3 isoforms have the highest expressed rate in a number of tissues, including the brain, lungs, and gastrointestinal system. A recent study predicts the effect of N mutations on the interaction of N protein with human cytosolic 14-3-3 proteins. The S202R mutation demonstrates the profound effect on the efficiency of N protein to package RNA. R203K/G204R mutations, that occurred in the IDRs, appear to enhance the interaction with the 14-3-3 binding motif due to a predicted Arg204:Asp225 salt bridge (Tugaeva et al., 2023). Moreover, it was observed that the arginine at site 203 decreased in prevalence by 8.33% (from 81.9% to 73.5%) as the incidence of a lysine variant increased by 8.2% (from 18.12% to 26.32%). The glycine residue of the directly adjacent 204 site was exchanged by an arginine, increasing of 8.2% the incidence of the mutation (from 18.1% to 26.3%) (Tomaszewski et al., 2020). These results underscore the importance of these specific mutations in contributing to the function of the disordered SR rich region of the protein. Notably, SR rich region in the IDR2 of the N protein that includes the R203K mutation was the hot spot harboring 7 of the top 10 most frequent mutations (Zhao et al., 2021). Mutant proteins bearing mutations in these regions resulted in noticeably weaker phase separation with respect to the wild-type N protein. As confirmed by deep sequence data analysis of SARS-CoV-2 from public databases and clinical samples, the R203K/G204R variant was also associated with novel transcription-regulating sequence-core sequence-dimerization domain RNA, that is supposed to have an increased effect on the expression of sub-genomic RNA (Leary et al., 2021).

Another important aspect of mutations is their role in modulating the viral-host protein-protein interaction. Interestingly, Del Veliz et al. suggested the presence of compensatory mutations in N disordered regions containing binding motifs for 14-3-3 proteins. From analysis of the variability of about 90, 000 sequences of the SARS-CoV-2 N protein, it has been observed that most residues were highly conserved within the 14-3-3 binding site. Moreover, the compensatory mutations, including half of the variants of concern, were found to maintain the interaction energy of N-14-3-3y which is suggested to be beneficial for the virus (Del et al., 2021).

Protein-protein interactions are initiated by molecular recognition features (MoRFs) (Mohan et al., 2006; Kotta-Loizou et al., 2013), the abundance of which is highly associated with the amount of intrinsic disorder in different forms of life (Yan et al., 2016). MoRFs are small disordered regions situated within intrinsic disorder proteins or regions that undergo a disorder-to-order structure arrangement during interaction with their partners (Gypas et al., 2013; Mishra et al., 2018).

Short linear motifs often located in intrinsically disordered regions offer a mechanism to hijack a large number of host processes (Van Roey et al., 2013; Van Roey et al., 2012). In SARS-CoV-2, nearly all proteins are predicted to have a significant percentage of MoRFs (Giri et al., 2021) suggesting the role of IDRs in protein-protein interaction networks. It has been observed that the replacement of hydrophobic to polar and charged

amino acids in S glycoproteins of SARS-CoV-2 creates an intrinsically disordered region that facilitates proteolysis and fusion with the host membrane (Podder et al., 2021). The antigenicity of the SARS-CoV-2 is due to the structural flexibility of its surface proteins. In S glycoproteins, mutation-driven accumulation of intrinsically disordered residues has been noticed and hypothesized to enhance viral transmissibility. This overlap between IDRs and major antigenic sites in the receptor-binding domain and N-terminal domain of S glycoprotein (S1 subunit) (Quaglia et al., 2022) is further implicated in the virus to provide genetic and antigenic drift.

While the mutations that originated in the intrinsically disordered regions improved the viral fitness and affected the interaction with the human proteins, it is possible that they influence protein folding and glycosylation leading to structural modifications. Studies by Fang et al. reported nonsynonymous single nucleotide variants (SNVs) on more than 60% of each S, ORF1ab, and ORF3a, with significantly higher SNV occurrence ratios in IDRs of S and ORF1ab. The authors observed mutations in the IDRs of S located near N-linked glycan sites, postulating potential impacts on glycosylation and protein folding, as well as on elevated levels of hydrophobicity (in S, ORF1ab, but not in ORF3a) (Fang et al., 2021).

Single mutations can affect the intrinsic disorder predisposition of viral proteins. In their per-residue disorder analysis, Hassan et al. used predisposition scores from 0 to one for the per-residual conditions, where 0 and one indicate residues entirely arranged and completely disordered, respectively. Accordingly, with different degrees of disorder predisposition, they observed that single mutations had a pronounced effect on the intrinsic disorder predisposition of ORF10 at the N- and C- terminal of its regions (Hassan et al., 2022). Mutations that emerge on the intrinsic disordered regions, could also lead to the development of drug and vaccine resistance. For example, as compared to wild type, receptor-binding domain mutants of Omicron, Delta AY.1, AY.2, and AY.3 were determined to be more disordered (Ranjan et al., 2022). This increase in disorder could perturb the interactions with antibodies, thus favoring the virus for antibody escape and, as a consequence leading to more pathogenicity.

It is also possible that mutations at the IDRs may lead to structure destabilization. For example, both SARS-CoV-2 and SARS-CoV-1 deletion and substitution mutations of Nsp1 in the C-terminal motif IDR significantly destabilize the structure of Nsp1 (Rezaei et al., 2022). In fact, point mutations at R99 and residues Arg124-Lys125 near the predicted unstructured regions of Nsp1 also lead to the loss of its ability to recognize viral RNA (Mendez et al., 2021; Lokugamage et al., 2012). Also interestingly, K129E and G130R mutations at the beginning of the intrinsically disordered C-terminal domain strongly reduced the ability of SARS-CoV-2 Nsp1 to inhibit cellular translation (Frolov et al., 2023).

5 SARS-CoV-2 IDRs and signaling pathways of the host cell

The unusually large genome of coronaviruses provides high structural complexity and intricate molecular interaction to efficiently replicate their genome. Once inside the infected cells, a spatiotemporal-oriented coordination between a cohort of enzymes, supporting proteins, and the host multiple factors, is required to produce viral mRNAs and new genomes. All stages of the life cycles of SARS-CoV-2 rely on host factors (Kim et al., 2020; Zhou et al., 2020), and hundreds of human proteins have been found to interact with viral proteins (Bouhaddou et al., 2020; Davidson et al., 2020; Gordon et al., 2020; Li et al., 2021).

In their large-scale proteomic analysis, Gordon et al. identified SARS-CoV-2 proteins interacting with human proteins involved in innate immune pathways, host translation machinery, cullin ubiquitin ligase, and bromodomain proteins (Gordon et al., 2020). Li and coworkers also identified 286 potential host targets using quantitative proteomic analysis approaches on SARS-CoV-2 intra-viral protein-protein interactions and virus-host interaction networks in human cells (Li et al., 2021). These findings further outline how the host cell infrastructure and metabolism ultimately determine the viral progeny.

In this section, we emphasize the importance of IDRs in SARS-CoV-2 by scrutinizing biochemical, structural, and virological experimental studies of selected viral proteins (Table 1), whose function includes IDRs/IDPs involvement in the host signaling machinery. However, the complete molecular mechanism of interaction of the virus protein with host proteins is not mentioned here because it is outside the scope of this review.

As was said earlier, intrinsically disordered protein regions evolve rapidly, promoting diversity and enabling interaction with multiple proteins among species (Ortiz et al., 2013; Chang et al., 2014) (Gitlin et al., 2014). In SARS-CoV-2, ORF6 is one of the second-highest IDR-containing accessory proteins. ORF6 hijacks nucleocytoplasmic mRNA transport factors Nucleoporin 98 (NUP98) and RNA export factor 1 (RAE1) (Miorin et al., 2020; Miyamoto et al., 2022; Kato et al., 2021). Using its C-terminal residues directly binds to transcription factors, signal transducers, and activators of transcription (STAT1) and inhibits its nuclear localization, therefore changing subcellular localization and resulting in downregulation of the interferon (IFN) pathway (Miyamoto et al., 2022; Lei et al., 2020). Importantly, this protein contains a long MoRF region (residues 26-61) near the C-terminal region (Giri et al., 2021), that matches with the residues 46-61 comprising trafficking motifs and with putative NUP98-RAE1 binding sequence (Gordon et al., 2020).

5.1 Nucleocapsid and host cell interaction

Several lines of experimental and computational studies revealed that the intrinsic disorder regions of N protein contain putative binding sites for different host proteins (Jia et al., 2022; Huang et al., 2021; Biswal et al., 2022; Del et al., 2021; El-Maradny et al., 2024). As a result of a recent investigation, 77 coronavirus N-interacting host proteins were examined, of which 33 are stress granule components, indicating a higher likelihood of IDRs to hijack SG proteins to the viral replication machinery (Moosa and Banerjee, 2020). It has been implicated that viral N-mediated host subversion is associated with π - π -interactions between π -rich disordered domains of coronavirus N and host proteins. This bioinformatics report is supported by experimental observations on the possible self-association of π -rich IDR-2 *via* π - π interactions (Cong et al., 2017). Likewise, IDR-1 also promotes self-interaction of SARS N (He et al., 2004). Furthermore, TABLE 1 Some selected examples of the potential involvement of intrinsically disordered regions of SARS-CoV-2 proteins in host signaling machinery.

Viral components/ IDR	Unstructured Region	IDR length/ location	MoRF motif sequence ^a	Experimental assay	Role in host machinery	Change/ response/effect on the host	References
N	Mostly in the N terminus	Residues 1–48, 175–248, 365–419	¹ MSDNGPQNQRNAPRITFGGPSDSTGSNQNGER ²⁰ , ⁸² DQIGYYRRATRRIRGGD ⁹⁸ , ¹⁰² KDLSPRWYFYYLGTGPE ¹¹⁸ , ⁴⁰³ FSKQLQQ ⁴⁰⁹	Analytical size- exclusion chromatography (SEC)	Interacts with seven human 14-3-3 isoforms ^c	Could hijack cellular pathways by 14-3- 3 sequestration	Del et al. (2021), Tugaeva et al. (2021)
				<i>In vitro</i> host-SARS- CoV-2 interactome, Y2H and TMT- AP-MS	Interacts LARP1 and RRP9	Downregulation of LARP1 phosphorylation Inhibition of protein synthesis	Gordon et al. (2020), Zhou et al. (2023)
ORF6	Mostly the C-terminal ^a	Residues 45–60	¹⁴ ILLIIMRT ²¹ , ²⁶ IWNLDYIINLIIKNLSKSLTENKYSQLDEEQPMEID ⁶¹	<i>In vitro</i> host-SARS- CoV-2 interactome	Interacts with NUP98/RAE complex	Prevents host mRNA export through the nuclear pore	Gordon et al. (2020)
ORF3a	N-terminus, and C-terminus ^a	Residues 1–19, 261–268	¹ MDLFMR ⁶ , ²⁶¹ EPIYDEPT ²⁶⁸	In vitro host-SARS- CoV-2 interactome, Y2H and TMT- AP-MS	Binds TRIM59 and other HOPS complex proteins	Endoplasmic reticulum stress	Gordon et al. (2020), Li et al. (2021); Zhou et al. (2023)
Nsp1	C-terminal	Residues 148–180	⁶⁴ LEQPYVFIKRSDA ⁷⁵ , ⁹³ EGIQ ⁹⁷ , ⁹⁹ RSGETLGVLVPHVGEIPVAYRKVLLRKNGNKGAGGHSYGAD LKSFDLGDELGTDPYEDFQENWNTKHSSGVTRELMRELNG ¹⁷⁹	Cryo-electron Microscopy, <i>In vitro</i> binding assay	Interacts with 40s ribosome	Mediates translation shutoff	Thoms et al. (2020), Frolov et al. (2023)
Nsp8	Mostly at the middle (residues 40–84 ^a)	Residues 40–84	¹⁸¹ AWPLI ¹⁸⁵	<i>In vitro</i> host-SARS- CoV-2 interactome	Interacts with SRPs components (SRP19, SRP54, SRP72)	Hijacks the Sec61- mediated protein translocation pathway for entry into the endoplasmic reticulum	Gordon et al. (2020)

^aThe disorder prediction results, and.

^b the predicted MoRF, motif sequences results are from Giri et al. (2021). It should be noted that Giri et al. used several MoRFs, predictors that may show some variation in the results, and the MoRF, regions presented here in Table 1 are drawn from the MoRFchibi_web predictor. The C-terminal sequence of SARS-CoV-2 ORF6, contains a trafficking motifs 46ENKYSQLDEEQPMEID61, and putative NUP98–RAE, binding sequence⁵⁶ QPMEID61 Gordon et al. (2020).

^cA study shows that greater than 90% of 14-3-3 protein partners contain disordered regions, suggesting that intrinsically disordered regions are favored by 14-3-3 proteins Bustos and Iglesias (2006).

Abbreviations: SRP, signal recognition particle; HOPS, homotypic fusion and protein sorting complex; LARP1, La-related protein 1; Y2H, yeast two-hybrid; TMT-AP-MS, tandem mass tag affinity purification followed by mass spectrometry.

there is also a tight interaction of the π -rich IDR-2 of coronavirus N with the π -rich disordered domain of heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) (Luo et al., 2005), all of which interactions contributes to suppress the host gene expression.

To fully understand the role of N-terminal IDR1, Huang et al., by means of GST pull-down and ITC assays, showed the key fragments for interaction between G3BP1/2 and N protein: the binding is mediated by N-IDR fragment (amino acids 11-24) and the NTF2 domain from the G3BP1 (G3BP1^{NTF2}) (Huang et al., 2021). Another important finding regards the sequence termed ITFG motif, found in the N-IDR (residues 1-48) region of the N protein. The motif shares similarity with that of a peptide known to interact with G3BP1NTF2 and fully conserved in SARS-CoV: interestingly, all four residues contribute to the binding, with Isoleucine and Phenylalanine paramount. Furthermore, from truncation of the N protein, it was shown that both the N-IDR1 and N-terminal RNA recognition motif RRM (residues 49-174) are required for inhibiting G3BP1-mediated LLPS (Huang et al., 2021). From another standpoint, the C-terminal domain (spanning residues from 248 to 365)-mediated inhibition of SG formation is not affected by disruption of the G3BP binding motif. On the other hand, alanine substitution in the C-terminal domain (248-365), as, for example, K257A, K261A, and their combination, disrupt dsRNA binding and abrogate suppression of SG formation (Aloise et al., 2023).

5.2 Nsp1 and host cell interaction

Nsp1 interactions with human proteins are indispensable for the virus life cycle. The immune evasion and the control of cellular response relies on host mRNA degradation and translation shutoff. The major virulence factor, Nsp1, governs this process *via* direct ribosomal association (Schubert et al., 2020; Vora et al., 2022; Lokugamage et al., 2012; Yuan et al., 2020).

Nsp1 N-terminus domain of both SARS-CoV-1 and SARS-CoV-2 binds the first stem-loop in the viral 5'leader sequence that guarantees continuous translation of the viral mRNA (Tanaka et al., 2012; Vora et al., 2022). Multiple structural and biochemical studies on coronaviruses demonstrated the C-terminal region as a key domain of Nsp1 that interacts with different ribosome subunits (Schubert et al., 2020; Kami et al., 2009; Huang et al., 2011; Nakagawa et al., 2016). Structural analysis by Cryo-electron microscopy evidenced that Nsp1 interacts with the 40S ribosomal subunit via residues conserved among coronaviruses, namely, Lys164 and His165. Thus, it leads to mRNA entry tunnel obstruction by C-terminus attachment. In consequence, the antiviral defense factor including RIG-I (retinoic acid-inducible gene-I) is suppressed by Nsp1. In fact, K164A and H165A mutants do not show binding, explaining the role of the two residues in the interaction with the 40S ribosome in both SARS-CoVs (Thoms et al., 2020; Schubert et al., 2020). Moreover, in the case of molecular interactions between Nsp1 and the 40S subunit, residues Y154, F157, and W161 in the first C-terminal helix (residues 153-160) interact through their multiple hydrophobic side chains with uS5 and uS3 proteins (Schubert et al., 2020).

It is also observed that the conserved KH motif (i.e., K164 and H165 in the short loop) experiences stacking interactions with helix

h18 of the 18S rRNA of the 40S subunit, thereby mediating the interaction in a conserved manner. The two C-terminal helices (i.e., the first helix (residues 153–160) and the second helix (residues 166–178) are connected by the KH motif. The other two conserved arginine R171 and R175 in the second helix make interactions with the phosphate backbone of h18 (Schubert et al., 2020). The uS3 subunit, which most likely corresponds to the Nsp1 unstructured linker, weakly interacts with h16. Furthermore, the C-terminal Nsp1 binds also 80S ribosomal complexes proving its role in the global translation inhibition (Schubert et al., 2020).

As highlighted in the previous sections, the efficient interplay between the NTD and C-terminal unstructured region is fundamental for Nsp1 interaction with host ribosomes (Wang et al., 2023).

6 IDR-targeted therapeutics and future perspectives

IDPs have gained great attention and opened new insights for understanding the molecular mechanism of diseases beyond the classic structure-function paradigm. Nowadays, the structural and functional characterization of IDPs and IDRs is becoming essential and thus unveiling the roles of protein disorder in protein-protein and protein-nucleic acid interactions is a promising field in understanding the biology of many diseases. Thus, understanding viral IDPs/IDRs and their molecular recognition features is essential for grasping their role in protein function and drug design, especially in light of the fact that SARS-CoV-2 variants arise within intrinsically disordered regions (Kulkarni et al., 2025; Gupta and Uversky, 2024).

One of the SARS-CoV-2 proteins bearing the highest content of intrinsically disordered regions is the multifunctional nucleoprotein N. Its major biological function depends on the interplay between the folded domains and the intrinsically disordered regions. Being the N protein a RNA packaging protein, the importance of its disordered regions in RNA binding and as major phosphorylation site, leads to consider the IDRs of the N protein potential areas for future studies to develop therapeutics. As such, though the SR-rich intrinsically disordered region is identified as a preferential binding site to human 14-3-3 protein (Del et al., 2021), little is known about the fundamental biology of this region.

Thus, the host-viral interacting IDPs/IDRs could be targeted for different therapeutic strategies.

Indeed, the survival of the virus inside the host cell is guaranteed by the dynamic molecular interactions of intrinsically disordered regions of Nsp1 and host ribosome subunits; this protein hijacks the host ribosome, through the intrinsically disordered C-terminal domain and thereby leads to host gene shutoff. The fully comprehension of the Nsp1 role in SARS-CoV-2 pathogenesis definitely needs the characterization of Nsp1 structural and dynamical properties along with its molecular interactions.

Gordon et al. reported a comprehensive SARS-CoV-2 PPI network, and the drug-human target network, suggesting that SARS-CoV-2 proteins, such as Nsp8 and ORF6, whose interaction sequence comprises the IDR, show potential druggable targets (Gordon et al., 2020).

Mutations can significantly affect the physicochemical properties of proteins. For instance, it has been shown that mutations in the IDRs of N protein can have nonlocal impact and modulate thermodynamic stability, secondary structure, particle formation, protein oligomeric state, and LLPS (Nguyen et al., 2024). From another perspective, experimental studies demonstrated the impact of the spike protein D614G mutation that induces IDR-like segments to form "order-disorder" transition (Zhang et al., 2021; Dokainish and Sugita, 2023). It has been further revealed that the order-disorder transition of IDRs in VOCs, which are associated with characteristic mutations such as D614G in the proximity of IDRs, controls the spike structure and dynamics (Sinha and Roy, 2023). This scenario opens an opportunity to target IDRs as remote operational switches for alternative therapeutic interventions. Moreover, as discussed above, IDRs show a helical propensity that possibly forms druggable conformational states. Interestingly enough, a recent study revealed that a small molecule to be evaluated in a clinical trial selectively interacts with an IDR oligomerization, further offering an insight that condensation and oligomerization could serve as key factors in the development of inhibitors for IDRs (Bielskutė et al., 2025).

Overall, although the presence of disordered regions could explain difficulties with crystallization of full-length N and Nsp1 proteins, deciphering the structural nuances of these proteins can provide important clues about potential therapeutic targets. Therefore, future studies should be oriented toward understanding the detailed mechanisms of intra- and intermolecular interactions of IDRs that may help to guide structure based therapeutics against SARS-CoV-2.

Author contributions

GS: Writing – original draft. NV: Writing – original draft. GD: Writing – original draft. MD: Writing – original draft. EWM:

References

Abbasian, M. H., Mahmanzar, M., Rahimian, K., Mahdavi, B., Tokhanbigli, S., Moradi, B., et al. (2023). Global landscape of SARS-CoV-2 mutations and conserved regions. *J. Transl. Med.* 21 (1), 152. doi:10.1186/s12967-023-03996-w

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 630 (8016), 493–500. doi:10.1038/s41586-024-07487-w

Abyzov, A., Blackledge, M., and Zweckstetter, M. (2022). Conformational dynamics of intrinsically disordered proteins regulate biomolecular condensate chemistry. *Chem. Rev.* 122 (6), 6719–6748. doi:10.1021/acs.chemrev.1c00774

Agback, T., Dominguez, F., Frolov, I., Frolova, E. I., and Agback, P. (2021). 1H, 13C and 15N resonance assignment of the SARS-CoV-2 full-length nsp1 protein and its mutants reveals its unique secondary structure features in solution. *PloS One* 16 (12), e0251834. doi:10.1371/journal.pone.0251834

Akbayrak, I. Y., Caglayan, S. I., Kurgan, L., Uversky, V. N., and Coskuner-Weber, O. (2022). Insights into the structural properties of SARS-CoV-2 main protease. *Curr. Res. Struct. Biol.* 4, 349–355. doi:10.1016/j.crstbi.2022.11.001

Almeida, M. S., Johnson, M. A., Herrmann, T., Geralt, M., and 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. doi:10.1128/jvi.01939-06

Aloise, C., Schipper, J. G., van Vliet, A., Oymans, J., Donselaar, T., Hurdiss, D. L., et al. (2023). SARS-CoV-2 nucleocapsid protein inhibits the PKR-mediated integrated stress response through RNA-binding domain N2b. *PLoS Pathogens*. 19(8):e1011582. doi:10. 1371/journal.ppat.1011582

Writing – original draft. RF: Writing – review and editing. RI: Writing – review and editing. LR: Writing – review and editing. CI: Writing – review and editing. GM: Writing – review and editing, Funding acquisition.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was supported by the PRIN2022 granted from Italian Research Minister, grant n. 2022MBK24T.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Alshehri, M. A., Manee, M. M., Alqahtani, F. H., Al-Shomrani, B. M., and Uversky, V. N. (2021). On the prevalence and potential functionality of an intrinsic disorder in the MERS-CoV proteome. *Viruses* 13 (2), 339. doi:10.3390/v13020339

Anjum, F., Mohammad, T., Asrani, P., Shafie, A., Singh, S., Yadav, D. K., et al. (2022). Identification of intrinsically disorder regions in non-structural proteins of SARS-CoV-2: new insights into drug and vaccine resistance. *Mol. Cell Biochem.* 477 (5), 1607–1619. doi:10.1007/s11010-022-04393-5

Bai, Z., Cao, Y., Liu, W., and Li, J. (2021). The SARS-CoV-2 nucleocapsid protein and its role in viral structure, biological functions, and a potential target for drug or vaccine mitigation. *Viruses* 13 (6), 1115. doi:10.3390/v13061115

Barik, S. (2020). Genus-specific pattern of intrinsically disordered central regions in the nucleocapsid protein of coronaviruses. *Comput. Struct. Biotechnol. J.* 18, 1884–1890. doi:10.1016/j.csbj.2020.07.005

Bessa, L. M., Guseva, S., Camacho-Zarco, A. R., Salvi, N., Maurin, D., Perez, L. M., et al. (2022). The intrinsically disordered SARS-CoV-2 nucleoprotein in dynamic complex with its viral partner nsp3a. *Sci. Adv.* 8 (3), eabm4034. doi:10.1126/sciadv.abm4034

Bhattacharjee, S., Bhattacharyya, R., and Sengupta, J. (2021). Dynamics and electrostatics define an allosteric druggable site within the receptor-binding domain of SARS-CoV-2 spike protein. *FEBS Lett.* 595 (4), 442–451. doi:10.1002/1873-3468. 14038

Bielskutė, S., Mateos, B., Awawdy, M., Garcia-Cabau, C., Niskanen, H., Sánchez-Zarzalejo, C., et al. (2025). Oligomerization enables the selective targeting of intrinsically disordered regions by small molecules. *bioRxiv Prepr. Serv. Biol.* 2025.03. 21.644603. doi:10.1101/2025.03.21.644603 Biswal, M., Lu, J., and Song, J. (2022). SARS-CoV-2 nucleocapsid protein targets a conserved surface groove of the NTF2-like domain of G3BP1. J. Mol. Biol. 434 (9), 167516. doi:10.1016/j.jmb.2022.167516

Bouhaddou, M., Memon, D., Meyer, B., White, K. M., Rezelj, V. V., Marrero, M. C., et al. (2020). The global phosphorylation landscape of SARS-CoV-2 infection. *Cell* 182 (3), 685–712.e19. doi:10.1016/j.cell.2020.06.034

Bowman, M. A., Riback, J. A., Rodriguez, A., Guo, H., Li, J., Sosnick, T. R., et al. (2020). Properties of protein unfolded states suggest broad selection for expanded conformational ensembles. *Proc. Natl. Acad. Sci. U. S. A.* 117 (38), 23356–23364. doi:10. 1073/pnas.2003773117

Bugge, K., Brakti, I., Fernandes, C. B., Dreier, J. E., Lundsgaard, J. E., Olsen, J. G., et al. (2020). Interactions by disorder – a matter of context. *Front. Mol. Biosci.* 7, 110. doi:10. 3389/fmolb.2020.00110

Bugge, K., Staby, L., Salladini, E., Falbe-Hansen, R. G., Kragelund, B. B., and Skriver, K. (2021). αα-Hub domains and intrinsically disordered proteins: a decisive combo. *J. Biol. Chem.* 296, 100226. doi:10.1074/jbc.rev120.012928

Bustos, D. M., and Iglesias, A. A. (2006). Intrinsic disorder is a key characteristic in partners that bind 14-3-3 proteins. *Proteins Struct. Funct. Bioinforma.* 63 (1), 35-42. doi:10.1002/prot.20888

Campbell, F., Archer, B., Laurenson-Schafer, H., Jinnai, Y., Konings, F., Batra, N., et al. (2021). Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. *Eurosurveillance* 26 (24), 2100509. doi:10.2807/1560-7917.es.2021.26. 24.2100509

Campen, A., Williams, R. M., Brown, C. J., Meng, J., Uversky, V. N., and Dunker, A. K. (2008). TOP-IDP-scale: a new amino acid scale measuring propensity for intrinsic disorder. *Protein Peptide Lett.* 15 (9), 956–963. doi:10.2174/092986608785849164

Carlson, C. R., Asfaha, J. B., Ghent, C. M., Howard, C. J., Hartooni, N., Safari, M., et al. (2020). Phosphoregulation of phase separation by the SARS-CoV-2 N protein suggests a biophysical basis for its dual functions. *Mol. Cell* 80 (6), 1092–103.e4. doi:10.1016/j. molcel.2020.11.025

Celik, I., Yadav, R., Duzgun, Z., Albogami, S., El-Shehawi, A. M., Idroes, R., et al. (2021). Interactions of the receptor binding domain of SARS-CoV-2 variants with hACE2: insights from molecular docking analysis and molecular dynamic simulation. *Biology* 10 (9), 880. doi:10.3390/biology10090880

Chang, C.-K., Hsu, Y.-L., Chang, Y.-H., Chao, F.-A., Wu, M.-C., Huang, Y.-S., et al. (2009). Multiple nucleic acid binding sites and intrinsic disorder of severe acute respiratory syndrome coronavirus nucleocapsid protein: implications for ribonucleocapsid protein packaging. *J. Virol.* 83 (5), 2255–2264. doi:10.1128/jvi. 02001-08

Chang, C.-k., Hou, M.-H., Chang, C.-F., Hsiao, C.-D., and Huang, T.-h. (2014). The SARS coronavirus nucleocapsid protein – forms and functions. *Antivir. Res.* 103, 39–50. doi:10.1016/j.antiviral.2013.12.009

Chen, C.-Y., Chang, C.-k., Chang, Y.-W., Sue, S.-C., Bai, H.-I., Riang, L., et al. (2007). Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. *J. Mol. Biol.* 368 (4), 1075–1086. doi:10.1016/j.jmb.2007.02.069

Chen, H., Gill, A., Dove, B. K., Emmett, S. R., Kemp, C. F., Ritchie, M. A., et al. (2005). Mass spectroscopic characterization of the coronavirus infectious bronchitis virus nucleoprotein and elucidation of the role of phosphorylation in RNA binding by using surface plasmon resonance. *J. Virol.* 79 (2), 1164–1179. doi:10.1128/jvi.79.2.1164-1179.2005

Chong, S., and Mir, M. (2021). Towards decoding the sequence-based grammar governing the functions of intrinsically disordered protein regions. *J. Mol. Biol.* 433 (12), 166724. doi:10.1016/j.jmb.2020.11.023

Clark, K., Karsch-Mizrachi, I., Lipman, D., Ostell, J., and Sayers, E. W. (2016). GenBank Nucleic Acids Res. 44, D67-D72. doi:10.1093/nar/gkv1276

Clark, L. K., Green, T. J., and Petit, C. M. (2021). Structure of nonstructural protein 1 from SARS-CoV-2. J. Virol. 95 (4), e02019-20. doi:10.1128/jvi.02019-20

Cong, Y., Kriegenburg, F., de Haan, C. A., and Reggiori, F. (2017). Coronavirus nucleocapsid proteins assemble constitutively in high molecular oligomers. *Sci. Rep.* 7 (1), 5740. doi:10.1038/s41598-017-06062-w

Cubuk, J., Alston, J. J., Incicco, J. J., Singh, S., Stuchell-Brereton, M. D., Ward, M. D., et al. (2021). The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. *Nat. Commun.* 12 (1), 1936. doi:10.1038/s41467-021-21953-3

Cui, L., Wang, H., Ji, Y., Yang, J., Xu, S., Huang, X., et al. (2015). The nucleocapsid protein of coronaviruses acts as a viral suppressor of RNA silencing in mammalian cells. *J. Virol.* 89 (17), 9029–9043. doi:10.1128/jvi.01331-15

Darling, A. L., and Uversky, V. N. (2018). Intrinsic disorder and posttranslational modifications: the darker side of the biological dark matter. *Front. Genet.* 9, 158. doi:10. 3389/fgene.2018.00158

Das, R. K., Ruff, K. M., and Pappu, R. V. (2015). Relating sequence encoded information to form and function of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* 32, 102–112. doi:10.1016/j.sbi.2015.03.008

Davey, N. E., Travé, G., and Gibson, T. J. (2011). How viruses hijack cell regulation. *Trends Biochem. Sci.* 36 (3), 159–169. doi:10.1016/j.tibs.2010.10.002 Davidson, A. D., Williamson, M. K., Lewis, S., Shoemark, D., Carroll, M. W., Heesom, K. J., et al. (2020). Characterisation of the transcriptome and proteome of SARS-CoV-2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. *Genome Med.* 12 (1), 68. doi:10.1186/s13073-020-00763-0

Del, V. S., Rivera, L., Bustos, D. M., and Uhart, M. (2021). Analysis of SARS-CoV-2 nucleocapsid phosphoprotein N variations in the binding site to human 14-3-3 proteins. *Biochem. Biophys. Res. Commun.* 569, 154–160. doi:10.1016/j.bbrc.2021. 06.100

Dokainish, H. M., and Sugita, Y. (2023). Structural effects of spike protein D614G mutation in SARS-CoV-2. *Biophysical J.* 122 (14), 2910–2920. doi:10.1016/j.bpj.2022. 11.025

Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., et al. (2001). Intrinsically disordered protein. *J. Mol. Graph. Model.* 19 (1), 26–59. doi:10. 1016/s1093-3263(00)00138-8

Dyson, H. J., and Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* 6 (3), 197–208. doi:10.1038/nrm1589

Dyson, H. J., and Wright, P. E. (2018). How do intrinsically disordered viral proteins hijack the cell? *Biochemistry*, 57 (28), 4045–4046. doi:10.1021/acs.biochem. 8b00622

El-Maradny, Y. A., Badawy, M. A., Mohamed, K. I., Ragab, R. F., Moharm, H. M., Abdallah, N. A., et al. (2024). Unraveling the role of the nucleocapsid protein in SARS-CoV-2 pathogenesis: from viral life cycle to vaccine development. *Int. J. Biol. Macromol.* 279, 135201. doi:10.1016/j.ijbiomac.2024.135201

Enjuanes, L., Almazán, F., Sola, I., and Zuñiga, S. (2006). Biochemical aspects of coronavirus replication and virus-host interaction. *Annu. Rev. Microbiol.* 60 (1), 211–230. doi:10.1146/annurev.micro.60.080805.142157

Fang, S., Liu, S., Shen, J., Lu, A. Z., Wang, A. K., Zhang, Y., et al. (2021). Updated SARS-CoV-2 single nucleotide variants and mortality association. *J. Med. Virol.* 93 (12), 6525–6534. doi:10.1002/jmv.27191

Feng, Y., Yi, J., Yang, L., Wang, Y., Wen, J., Zhao, W., et al. (2024). COV2Var, a function annotation database of SARS-CoV-2 genetic variation. *Nucleic acids Res.* 52 (D1), D701–D713. doi:10.1093/nar/gkad958

Freiberger, M. I., Wolynes, P. G., Ferreiro, D. U., and Fuxreiter, M. (2021). Frustration in fuzzy protein complexes leads to interaction versatility. *J. Phys. Chem. B* 125 (10), 2513–2520. doi:10.1021/acs.jpcb.0c11068

Frolov, I., Agback, T., Palchevska, O., Dominguez, F., Lomzov, A., Agback, P., et al. (2023). All domains of SARS-CoV-2 nsp1 determine translational shutoff and cytotoxicity of the protein. *J. Virol.* 97 (3), e01865-22. doi:10.1128/jvi.01865-22

Gadhave, K., Kumar, P., Kumar, A., Bhardwaj, T., Garg, N., and Giri, R. (2021). Conformational dynamics of 13 amino acids long NSP11 of SARS-CoV-2 under membrane mimetics and different solvent conditions. *Microb. Pathog.* 158, 105041. doi:10.1016/j.micpath.2021.105041

Gianni, S., Freiberger, M. I., Jemth, P., Ferreiro, D. U., Wolynes, P. G., and Fuxreiter, M. (2021). Fuzziness and frustration in the energy landscape of protein folding, function, and assembly. *Accounts Chem. Res.* 54 (5), 1251–1259. doi:10.1021/acs. accounts.0c00813

Giri, R., Bhardwaj, T., Shegane, M., Gehi, B. R., Kumar, P., Gadhave, K., et al. (2021). Understanding COVID-19 via comparative analysis of dark proteomes of SARS-CoV-2, human SARS and bat SARS-like coronaviruses. *Cell Mol. Life Sci.* 78 (4), 1655–1688. doi:10.1007/s00018-020-03603-x

Gitlin, L., Hagai, T., LaBarbera, A., Solovey, M., and Andino, R. (2014). Rapid evolution of virus sequences in intrinsically disordered protein regions. *PLoS Pathog.* 10 (12), e1004529. doi:10.1371/journal.ppat.1004529

Goh, G. K.-M., Dunker, A. K., Foster, J. A., and Uversky, V. N. (2020). Rigidity of the outer shell predicted by a protein intrinsic disorder model sheds light on the COVID-19 (Wuhan-2019-nCoV) infectivity. *Biomolecules* 10 (2), 331. doi:10.3390/biom10020331

Goh, G. K.-M., Dunker, A. K., and Uversky, V. (2013). Prediction of intrinsic disorder in MERS-CoV/HCoV-EMC supports a high oral-fecal transmission. *PLoS Curr.* 5. doi:10.1371/currents.outbreaks.22254b58675cdebc256dbe3c5aa6498b

Gomari, M. M., Tarighi, P., Choupani, E., Abkhiz, S., Mohamadzadeh, M., Rostami, N., et al. (2023). Structural evolution of Delta lineage of SARS-CoV-2. *Int. J. Biol. Macromol.* 226, 1116–1140. doi:10.1016/j.ijbiomac.2022.11.227

Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* 583 (7816), 459–468. doi:10.1038/s41586-020-2286-9

Grossoehme, N. E., Li, L., Keane, S. C., Liu, P., Dann, I. I. C. E., Leibowitz, J. L., et al. (2009). Coronavirus N protein N-terminal domain (NTD) specifically binds the transcriptional regulatory sequence (TRS) and melts TRS-cTRS RNA duplexes. *J. Mol. Biol.* 394 (3), 544–557. doi:10.1016/j.jmb.2009.09.040

Gupta, M. N., and Uversky, V. N. (2024). Protein structure–function continuum model: emerging nexuses between specificity, evolution, and structure. *Protein Sci.* 33 (4), e4968. doi:10.1002/pro.4968

Gypas, F., Tsaousis, G. N., and Hamodrakas, S. J. (2013). mpMoRFsDB: a database of molecular recognition features in membrane proteins. *Bioinforma. Oxf. Engl.* 29 (19), 2517–2518. doi:10.1093/bioinformatics/btt427

Haque, S., Khatoon, F., Ashgar, S. S., Faidah, H., Bantun, F., Jalal, N. A., et al. (2023). Energetic and frustration analysis of SARS-CoV-2 nucleocapsid protein mutations. *Biotechnol. Genet. Eng. Rev.* 39, 1234–1254. doi:10.1080/02648725.2023.2170031

Hassan, S. S., Lundstrom, K., Serrano-Aroca, Á., Adadi, P., Aljabali, A. A. A., Redwan, E. M., et al. (2022). Emergence of unique SARS-CoV-2 ORF10 variants and their impact on protein structure and function. *Int. J. Biol. Macromol.* 194, 128–143. doi:10.1016/j. ijbiomac.2021.11.151

He, R., Dobie, F., Ballantine, M., Leeson, A., Li, Y., Bastien, N., et al. (2004). Analysis of multimerization of the SARS coronavirus nucleocapsid protein. *Biochem. Biophys. Res. Commun.* 316 (2), 476–483. doi:10.1016/j.bbrc.2004.02.074

Hsin, W.-C., Chang, C.-H., Chang, C.-Y., Peng, W.-H., Chien, C.-L., Chang, M.-F., et al. (2018). Nucleocapsid protein-dependent assembly of the RNA packaging signal of Middle East respiratory syndrome coronavirus. *J. Biomed. Sci.* 25 (1), 47. doi:10.1186/ s12929-018-0449-x

Huang, C., Lokugamage, K. G., Rozovics, J. M., Narayanan, K., Semler, B. L., and Makino, S. (2011). SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to nsp1-induced RNA cleavage. *PLoS Pathog.* 7 (12), e1002433. doi:10.1371/journal.ppat.1002433

Huang, Q., Yu, L., Petros, A. M., Gunasekera, A., Liu, Z., Xu, N., et al. (2004). Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. *Biochemistry* 43 (20), 6059–6063. doi:10.1021/bi036155b

Huang, W., Ju, X., Tian, M., Li, X., Yu, Y., Sun, Q., et al. (2021). Molecular determinants for regulation of G3BP1/2 phase separation by the SARS-CoV-2 nucleocapsid protein. *Cell Discov.* 7 (1), 69. doi:10.1038/s41421-021-00306-w

Iakoucheva, L. M., Brown, C. J., Lawson, J. D., Obradović, Z., and Dunker, A. K. (2002). Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* 323 (3), 573–584. doi:10.1016/s0022-2836(02)00969-5

Ilyas, S., Manan, A., and Lee, D. (2025). Deep learning-based comparative prediction and functional analysis of intrinsically disordered regions in SARS-CoV-2. *Int. J. Mol. Sci.* 26 (7), 3411. doi:10.3390/ijms26073411

Iserman, C., Roden, C. A., Boerneke, M. A., Sealfon, R. S. G., McLaughlin, G. A., Jungreis, I., et al. (2020). Genomic RNA elements drive phase separation of the SARS-CoV-2 nucleocapsid. *Mol. Cell* 80 (6), 1078–91.e6. doi:10.1016/j.molcel.2020. 11.041

Jayaram, H., Fan, H., Bowman, B. R., Ooi, A., Jayaram, J., Collisson, E. W., et al. (2006). X-ray structures of the N-and C-terminal domains of a coronavirus nucleocapsid protein: implications for nucleocapsid formation. *J. Virol.* 80 (13), 6612–6620. doi:10.1128/jvi.00157-06

Jia, Z., Liu, C., Chen, Y., Jiang, H., Wang, Z., Yao, J., et al. (2022). Crystal structures of the SARS-CoV-2 nucleocapsid protein C-terminal domain and development of nucleocapsid-targeting nanobodies. *FEBS J.* 289 (13), 3813–3825. doi:10.1111/febs. 16239

Jin, F., and Gräter, F. (2021). How multisite phosphorylation impacts the conformations of intrinsically disordered proteins. *PLoS Comput. Biol.* 17 (5), e1008939. doi:10.1371/journal.pcbi.1008939

Johnson, B. A., Zhou, Y., Lokugamage, K. G., Vu, M. N., Bopp, N., Crocquet-Valdes, P. A., et al. (2022). Nucleocapsid mutations in SARS-CoV-2 augment replication and pathogenesis. *PLoS Pathog.* 18 (6), e1010627. doi:10.1371/journal.ppat.1010627

Ju, X., Zhu, Y., Wang, Y., Li, J., Zhang, J., Gong, M., et al. (2021). A novel cell culture system modeling the SARS-CoV-2 life cycle. *PLoS Pathog.* 17 (3), e1009439. doi:10. 1371/journal.ppat.1009439

Kamitani, W., Huang, C., Narayanan, K., Lokugamage, K. G., and Makino, S. (2009). A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nat. Struct. Mol. Biol.* 16 (11), 1134–1140. doi:10.1038/nsmb.1680

Kang, S., Yang, M., Hong, Z., Zhang, L., Huang, Z., Chen, X., et al. (2020). Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. *Acta Pharm. Sin. B* 10 (7), 1228–1238. doi:10.1016/j.apsb. 2020.04.009

Kato, K., Ikliptikawati, D. K., Kobayashi, A., Kondo, H., Lim, K., Hazawa, M., et al. (2021). Overexpression of SARS-CoV-2 protein ORF6 dislocates RAE1 and NUP98 from the nuclear pore complex. *Biochem. Biophys. Res. Commun.* 536, 59–66. doi:10.1016/j.bbrc.2020.11.115

Keane, S. C., and Giedroc, D. P. (2013). Solution structure of mouse hepatitis virus (MHV) nsp3a and determinants of the interaction with MHV nucleocapsid (N) protein. *J. Virol.* 87 (6), 3502–3515. doi:10.1128/jvi.03112-12

Kim, D., Lee, J.-Y., Yang, J.-S., Kim, J. W., Kim, V. N., and Chang, H. (2020). The architecture of SARS-CoV-2 transcriptome. *Cell* 181 (4), 914–21. e10. doi:10.1016/j.cell. 2020.04.011

Klein, S., Cortese, M., Winter, S. L., Wachsmuth-Melm, M., Neufeldt, C. J., Cerikan, B., et al. (2020). SARS-CoV-2 structure and replication characterized by *in situ* cryoelectron tomography. *Nat. Commun.* 11 (1), 5885. doi:10.1038/s41467-020-19619-7

Koh, G. C., Porras, P., Aranda, B., Hermjakob, H., and Orchard, S. E. (2012). Analyzing protein-protein interaction networks. *J. Proteome Res.* 11 (4), 2014–2031. doi:10.1021/pr201211w

Kotta-Loizou, I., Tsaousis, G. N., and Hamodrakas, S. J. (2013). Analysis of molecular recognition features (MoRFs) in membrane proteins. *Biochimica Biophysica Acta (BBA)-Proteins Proteomics.* 1834 (4), 798–807. doi:10.1016/j.bbapap.2013.01.006

Kruse, T., Benz, C., Garvanska, D. H., Lindqvist, R., Mihalic, F., Coscia, F., et al. (2021). Large scale discovery of coronavirus-host factor protein interaction motifs reveals SARS-CoV-2 specific mechanisms and vulnerabilities. *Nat. Commun.* 12 (1), 6761. doi:10.1038/s41467-021-26498-z

Kubinski, H. C., Despres, H. W., Johnson, B. A., Schmidt, M. M., Jaffrani, S. A., Turner, A. H., et al. (2025). Variant mutation G215C in SARS-CoV-2 nucleocapsid enhances viral infection via altered genomic encapsidation. *PLoS Biol.* 23 (4), e3003115. doi:10.1371/journal.pbio.3003115

Kulkarni, P., Porter, L., Chou, T.-F., Chong, S., Chiti, F., Schafer, J. W., et al. (2025). Evolving concepts of the protein universe. *iScience* 28 (3), 112012. doi:10.1016/j.isci. 2025.112012

Kumar, A., Kumar, A., Kumar, P., Garg, N., and Giri, R. (2021). SARS-CoV-2 NSP1 C-terminal (residues 131-180) is an intrinsically disordered region in isolation. *Curr. Res. Virological Sci.* 2, 100007. doi:10.1016/j.crviro.2021.100007

Lan, T. C., Allan, M. F., Malsick, L. E., Woo, J. Z., Zhu, C., Zhang, F., et al. (2022). Secondary structural ensembles of the SARS-CoV-2 RNA genome in infected cells. *Nat. Commun.* 13 (1), 1128. doi:10.1038/s41467-022-28603-2

Leary, S., Gaudieri, S., Parker, M. D., Chopra, A., James, I., Pakala, S., et al. (2021). Generation of a novel SARS-CoV-2 sub-genomic RNA due to the R203K/G204R variant in nucleocapsid: homologous recombination has potential to change SARS-CoV-2 at both protein and RNA level. *Pathogens Immun.* 6 (2), 27–49. doi:10.20411/pai.v6i2.460

Lei, X., Dong, X., Ma, R., Wang, W., Xiao, X., Tian, Z., et al. (2020). Activation and evasion of type I interferon responses by SARS-CoV-2. *Nat. Commun.* 11 (1), 3810. doi:10.1038/s41467-020-17665-9

Li, J., Guo, M., Tian, X., Wang, X., Yang, X., Wu, P., et al. (2021). Virus-host interactome and proteomic survey reveal potential virulence factors influencing SARS-CoV-2 pathogenesis. *Med* 2 (1), 99–112.e7. doi:10.1016/j.medj.2020.07.002

Li, Y., Yang, L., and Yang, L. Q. (2024). Effects of intrinsically disordered regions in gp120 underlying HIV neutralization phenotypes. *Biochem. biophysical Res. Commun.* 709, 149830. doi:10.1016/j.bbrc.2024.149830

Lin, Y.-H., Forman-Kay, J. D., and Chan, H. S. (2018). Theories for sequencedependent phase behaviors of biomolecular condensates. *Biochemistry* 57 (17), 2499–2508. doi:10.1021/acs.biochem.8b00058

Lindsay, R. J., Mansbach, R. A., Gnanakaran, S., and Shen, T. (2021). Effects of pH on an IDP conformational ensemble explored by molecular dynamics simulation. *Biophys. Chem.* 271, 106552. doi:10.1016/j.bpc.2021.106552

Lokugamage, K. G., Narayanan, K., Huang, C., and Makino, S. (2012). Severe acute respiratory syndrome coronavirus protein nsp1 is a novel eukaryotic translation inhibitor that represses multiple steps of translation initiation. *J. Virol.* 86 (24), 13598–13608. doi:10.1128/jvi.01958-12

Lu, S., Ye, Q., Singh, D., Cao, Y., Diedrich, J. K., Yates, J. R., et al. (2021). The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. *Nat. Commun.* 12 (1), 502. doi:10.1038/ s41467-020-20768-y

Luo, H., Chen, Q., Chen, J., Chen, K., Shen, X., and Jiang, H. (2005). The nucleocapsid protein of SARS coronavirus has a high binding affinity to the human cellular heterogeneous nuclear ribonucleoprotein A1. *FEBS Lett.* 579 (12), 2623–2628. doi:10.1016/j.febslet.2005.03.080

Luo, L., Li, Z., Zhao, T., Ju, X., Ma, P., Jin, B., et al. (2021). SARS-CoV-2 nucleocapsid protein phase separates with G3BPs to disassemble stress granules and facilitate viral production. *Sci. Bull.* 66 (12), 1194–1204. doi:10.1016/j.scib.2021.01.013

Malone, B., Urakova, N., Snijder, E. J., and Campbell, E. A. (2022). Structures and functions of coronavirus replication-transcription complexes and their relevance for SARS-CoV-2 drug design. *Nat. Rev. Mol. Cell Biol.* 23 (1), 21–39. doi:10.1038/s41580-021-00432-z

Masters, P. S. (2006). The molecular biology of coronaviruses. Adv. Virus Res. 66, 193–292. doi:10.1016/s0065-3527(06)66005-3

McBride, R., Van Zyl, M., and Fielding, B. C. (2014). The coronavirus nucleocapsid is a multifunctional protein. Viruses 6 (8), 2991–3018. doi:10.3390/v6082991

Medeiros-Silva, J., Somberg, N. H., Wang, H. K., McKay, M. J., Mandala, V. S., Dregni, A. J., et al. (2022). pH-and calcium-dependent aromatic network in the SARS-CoV-2 envelope protein. J. Am. Chem. Soc. 144 (15), 6839–6850. doi:10.1021/jacs.2c00973

Mendez, A. S., Ly, M., González-Sánchez, A. M., Hartenian, E., Ingolia, N. T., Cate, J. H., et al. (2021). The N-terminal domain of SARS-CoV-2 nsp1 plays key roles in suppression of cellular gene expression and preservation of viral gene expression. *Cell Rep.* 37 (3), 109841. doi:10.1016/j.celrep.2021.109841

Mészáros, B., Erdős, G., and Dosztányi, Z. (2018). IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic acids Res.* 46 (W1), W329–W337. doi:10.1093/nar/gky384

Michelitsch, M. D., and Weissman, J. S. (2000). A census of glutamine/asparaginerich regions: implications for their conserved function and the prediction of novel prions. Proc. Natl. Acad. Sci. U. S. A. 97 (22), 11910-11915. doi:10.1073/pnas.97.22. 11910

Milorey, B., Schweitzer-Stenner, R., Andrews, B., Schwalbe, H., and Urbanc, B. (2021). Short peptides as predictors for the structure of polyarginine sequences in disordered proteins. *Biophysical J.* 120 (4), 662–676. doi:10.1016/j.bpj.2020.12.026

Miorin, L., Kehrer, T., Sanchez-Aparicio, M. T., Zhang, K., Cohen, P., Patel, R. S., et al. (2020). SARS-CoV-2 Orf6 hijacks Nup98 to block STAT nuclear import and antagonize interferon signaling. *Proc. Natl. Acad. Sci. U. S. A.* 117 (45), 28344–28354. doi:10.1073/pnas.2016650117

Mishra, P. M., Uversky, V. N., and Giri, R. (2018). Molecular recognition features in Zika virus proteome. J. Mol. Biol. 430 (16), 2372–2388. doi:10.1016/j.jmb.2017.10.018

Mishra, P. M., Verma, N. C., Rao, C., Uversky, V. N., and Nandi, C. K. (2020). Intrinsically disordered proteins of viruses: involvement in the mechanism of cell regulation and pathogenesis. 174, 1–78. doi:10.1016/bs.pmbts.2020.03.001*Prog. Mol. Biol. Transl. Sci.*

Miyamoto, Y., Itoh, Y., Suzuki, T., Tanaka, T., Sakai, Y., Koido, M., et al. (2022). SARS-CoV-2 ORF6 disrupts nucleocytoplasmic trafficking to advance viral replication. *Commun. Biol.* 5 (1), 483. doi:10.1038/s42003-022-03427-4

Mohan, A., Oldfield, C. J., Radivojac, P., Vacic, V., Cortese, M. S., Dunker, A. K., et al. (2006). Analysis of molecular recognition features (MoRFs). J. Mol. Biol. 362 (5), 1043–1059. doi:10.1016/j.jmb.2006.07.087

Moosa, M. M., and Banerjee, P. R. (2020). Subversion of host stress granules by coronaviruses: potential roles of π -rich disordered domains of viral nucleocapsids. *J. Med. Virol.* 92 (12), 2891–2893. doi:10.1002/jmv.26195

Morse, M., Sefcikova, J., Rouzina, I., Beuning, P. J., and Williams Mark, C. (2023). Structural domains of SARS-CoV-2 nucleocapsid protein coordinate to compact long nucleic acid substrates. *Nucleic Acids Res.* 51 (1), 290–303. doi:10.1093/nar/gkac1179

Mukherjee, S. K., Knop, J. M., and Winter, R. (2022). Modulation of the conformational space of SARS-CoV-2 RNA quadruplex rg-1 by cellular components and the amyloidogenic peptides α-synuclein and hIAPP. *Chem. A Eur. J.* 28 (9), e202104182. doi:10.1002/chem.202104182

Nakagawa, K., Lokugamage, K., and Makino, S. (2016). Viral and cellular mRNA translation in coronavirus-infected cells. *Adv. Virus Res.* 96, 165–192. doi:10.1016/bs. aivir.2016.08.001

Nelson, G. W., and Stohlman, S. A. (1993). Localization of the RNA-binding domain of mouse hepatitis virus nucleocapsid protein. *J. General Virol.* 74 (9), 1975–1979. doi:10.1099/0022-1317-74-9-1975

Nguyen, A., Zhao, H., Myagmarsuren, D., Srinivasan, S., Wu, D., Chen, J., et al. (2024). Modulation of biophysical properties of nucleocapsid protein in the mutant spectrum of SARS-CoV-2. *eLife* 13. doi:10.7554/elife.94836

Niklas, K. J., Bondos, S. E., Dunker, A. K., and Newman, S. A. (2015). Rethinking gene regulatory networks in light of alternative splicing, intrinsically disordered protein domains, and post-translational modifications. *Front. Cell Dev. Biol.* 3, 8. doi:10.3389/ fcell.2015.00008

Ortiz, J. F., MacDonald, M. L., Masterson, P., Uversky, V. N., and Siltberg-Liberles, J. (2013). Rapid evolutionary dynamics of structural disorder as a potential driving force for biological divergence in flaviviruses. *Genome Biol. Evol.* 5 (3), 504–513. doi:10.1093/gbe/evt026

Peng, K., Radivojac, P., Vucetic, S., Dunker, A. K., and Obradovic, Z. (2006). Lengthdependent prediction of protein intrinsic disorder. *BMC Bioinforma*. 7, 208–217. doi:10. 1186/1471-2105-7-208

Peng, K., Vucetic, S., Radivojac, P., Brown, C. J., Dunker, A. K., and Obradovic, Z. (2005). Optimizing long intrinsic disorder predictors with protein evolutionary information. J. Bioinforma. Comput. Biol. 3 (01), 35–60. doi:10.1142/ s0219720005000886

Peng, T.-Y., Lee, K.-R., and Tarn, W.-Y. (2008). Phosphorylation of the arginine/ serine dipeptide-rich motif of the severe acute respiratory syndrome coronavirus nucleocapsid protein modulates its multimerization, translation inhibitory activity and cellular localization: phosphorylation of SARS CoV-N protein RS motif. *FEBS J.* 275 (16), 4152–4163. doi:10.1111/j.1742-4658.2008.06564.x

Peng, Z., Yan, J., Fan, X., Mizianty, M. J., Xue, B., Wang, K., et al. (2015). Exceptionally abundant exceptions: comprehensive characterization of intrinsic disorder in all domains of life. *Cell Mol. Life Sci.* 72 (1), 137–151. doi:10.1007/s00018-014-1661-9

Perdikari, T. M., Murthy, A. C., Ryan, V. H., Watters, S., Naik, M. T., and Fawzi, N. L. (2020). SARS-CoV-2 nucleocapsid protein phase-separates with RNA and with human hnRNPs. *EMBO J.* 39 (24), e106478. doi:10.15252/embj.2020106478

Podder, S., Ghosh, A., and Ghosh, T. (2021). Mutations in membrane-fusion subunit of spike glycoprotein play crucial role in the recent outbreak of COVID-19. *J. Med. Virol.* 93 (5), 2790–2798. doi:10.1002/jmv.26598

Quaglia, F., Salladini, E., Carraro, M., Minervini, G., Tosatto, S. C. E., and Le Mercier, P. (2022). SARS-CoV-2 variants preferentially emerge at intrinsically disordered protein sites helping immune evasion. *FEBS J.* 289 (14), 4240–4250. doi:10.1111/febs.16379

Rahman, M. S., Islam, M. R., Alam, A. R. U., Islam, I., Hoque, M. N., Akter, S., et al. (2021). Evolutionary dynamics of SARS-CoV-2 nucleocapsid protein and its consequences. *J. Med. Virol.* 93 (4), 2177–2195. doi:10.1002/jmv.26626

Rajagopalan, K., Mooney, S. M., Parekh, N., Getzenberg, R. H., and Kulkarni, P. (2011). A majority of the cancer/testis antigens are intrinsically disordered proteins. *J. Cell. Biochem.* 112 (11), 3256–3267. doi:10.1002/jcb.23252

Ranjan, P., Neha, D. C., Devar, K. A., and Das, P. (2022). The influence of new SARS-CoV-2 variant Omicron (B.1.1.529) on vaccine efficacy, its correlation to delta variants: a computational approach. *Microb. Pathog.* 169, 105619. doi:10.1016/j.micpath.2022. 105619

Redwan, E. M., AlJaddawi, A. A., and Uversky, V. N. (2019). Structural disorder in the proteome and interactome of Alkhurma virus (ALKV). *Cell Mol. Life Sci.* 76, 577–608. doi:10.1007/s00018-018-2968-8

Rezaei, S., Pereira, F., Uversky, V. N., and Sefidbakht, Y. (2022). Molecular dynamics and intrinsic disorder analysis of the SARS-CoV-2 Nsp1 structural changes caused by substitution and deletion mutations. *Mol. Simul.* 48 (13), 1192–1201. doi:10.1080/ 08927022.2022.2075546

Ribeiro-Filho, H. V., Jara, G. E., Batista, F. A. H., Schleder, G. R., Costa Tonoli, C. C., Soprano, A. S., et al. (2022). Structural dynamics of SARS-CoV-2 nucleocapsid protein induced by RNA binding. *PLoS Comput. Biol.* 18 (5), e1010121. doi:10.1371/journal. pcbi.1010121

Romero, P., Obradovic, Z., Li, X., Garner, E. C., Brown, C. J., and Dunker, A. K. (2001). Sequence complexity of disordered protein. *Proteins Struct. Funct. Bioinforma.* 42 (1), 38–48. doi:10.1002/1097-0134(20010101)42:1<38::aid-prot50>3.0.co;2-3

Różycki, B., and Boura, E. (2022). Conformational ensemble of the full-length SARS-CoV-2 nucleocapsid (N) protein based on molecular simulations and SAXS data. *Biophys. Chem.* 288, 106843. doi:10.1016/j.bpc.2022.106843

Sakuraba, S., Xie, Q., Kasahara, K., Iwakiri, J., and Kono, H. (2022). Extended ensemble simulations of a SARS-CoV-2 nsp1–5'-UTR complex. *PLoS Comput. Biol.* 18 (1), e1009804. doi:10.1371/journal.pcbi.1009804

Savastano, A., Ibáñez de Opakua, A., Rankovic, M., and Zweckstetter, M. (2020). Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerasecontaining condensates. *Nat. Commun.* 11 (1), 6041. doi:10.1038/s41467-020-19843-1

Schiavina, M., Pontoriero, L., Tagliaferro, G., Pierattelli, R., and Felli, I. C. (2022). The role of disordered regions in orchestrating the properties of multidomain proteins: the SARS-CoV-2 nucleocapsid protein and its interaction with Enoxaparin. *Biomolecules* 12 (9), 1302. doi:10.3390/biom12091302

Schubert, K., Karousis, E. D., Jomaa, A., Scaiola, A., Echeverria, B., Gurzeler, L.-A., et al. (2020). SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. *Nat. Struct. Mol. Biol.* 27 (10), 959–966. doi:10.1038/s41594-020-0511-8

Sekine, R., Tsuno, S., Irokawa, H., Sumitomo, K., Han, T., Sato, Y., et al. (2023). Inhibition of SARS-CoV-2 nucleocapsid protein–RNA interaction by guanosine oligomeric RNA. *J. Biochem.* 173 (6), 447–457. doi:10.1093/jb/mvad008

Semper, C., Watanabe, N., and Savchenko, A. (2021). Structural characterization of nonstructural protein 1 from SARS-CoV-2. *iScience* 24 (1), 101903. doi:10.1016/j.isci. 2020.101903

Sinha, A., and Roy, S. (2023). Intrinsically disordered regions function as a cervical collar to remotely regulate the nodding dynamics of SARS-CoV-2 prefusion spike heads. J. Phys. Chem. B 127 (39), 8393–8405. doi:10.1021/acs.jpcb.3c05338

Spreitzer, E., Usluer, S., and Madl, T. (2020). Probing surfaces in dynamic protein interactions. J. Mol. Biol. 432 (9), 2949-2972. doi:10.1016/j.jmb.2020.02.032

Syed, A. M., Taha, T. Y., Tabata, T., Chen, I. P., Ciling, A., Khalid, M. M., et al. (2021). Rapid assessment of SARS-CoV-2–evolved variants using virus-like particles. *Sci. (New York, NY)* 374 (6575), 1626–1632. doi:10.1126/science.abl6184

Tanaka, T., Kamitani, W., DeDiego, M. L., Enjuanes, L., and Matsuura, Y. (2012). Severe acute respiratory syndrome coronavirus nsp1 facilitates efficient propagation in cells through a specific translational shutoff of host mRNA. *J. Virol.* 86 (20), 11128–11137. doi:10.1128/jvi.01700-12

Taneja, I., and Holehouse, A. S. (2021). Folded domain charge properties influence the conformational behavior of disordered tails. *Curr. Res. Struct. Biol.* 3, 216–228. doi:10. 1016/j.crstbi.2021.08.002

Terada, Y., Kawachi, K., Matsuura, Y., and Kamitani, W. (2017). MERS coronavirus nsp1 participates in an efficient propagation through a specific interaction with viral RNA. *Virol.* 511, 95–105. doi:10.1016/j.virol.2017.08.026

Thoms, M., Buschauer, R., Ameismeier, M., Koepke, L., Denk, T., Hirschenberger, M., et al. (2020). Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Sci. (1979).* 369, 1249–1255. doi:10.1126/science.abc8665

Tidu, A., Janvier, A., Schaeffer, L., Sosnowski, P., Kuhn, L., Hammann, P., et al. (2021). The viral protein NSP1 acts as a ribosome gatekeeper for shutting down host translation and fostering SARS-CoV-2 translation. *Rna* 27 (3), 253–264. doi:10.1261/rna. 078121.120

Tomaszewski, T., DeVries, R. S., Dong, M., Bhatia, G., Norsworthy, M. D., Zheng, X., et al. (2020). New pathways of mutational change in SARS-CoV-2 proteomes involve regions of intrinsic disorder important for virus replication and release. *Evol. Bioinforma.* 16, 1176934320965149. doi:10.1177/1176934320965149

Tompa, P. (2002). Intrinsically unstructured proteins. *Trends biochem. Sci.* 27, 527–533. doi:10.1016/s0968-0004(02)02169-2

Tugaeva, K. V., Hawkins, D. E., Smith, J. L., Bayfield, O. W., Ker, D.-S., Sysoev, A. A., et al. (2021). The mechanism of SARS-CoV-2 nucleocapsid protein recognition by the human 14-3-3 proteins. *J. Mol. Biol.* 433 (8), 166875. doi:10.1016/j.jmb.2021.166875

Tugaeva, K. V., Sysoev, A. A., Kapitonova, A. A., Smith, J. L. R., Zhu, P., Cooley, R. B., et al. (2023). Human 14-3-3 proteins site-selectively bind the mutational hotspot region of SARS-CoV-2 nucleoprotein modulating its phosphoregulation. *J. Mol. Biol.* 435 (2), 167891. doi:10.1016/j.jmb.2022.167891

Vankadari, N., Jeyasankar, N. N., and Lopes, W. J. (2020). Structure of the SARS-CoV-2 Nsp1/5'-untranslated region complex and implications for potential therapeutic targets, a vaccine, and virulence. *J. Phys. Chem. Lett.* 11 (22), 9659–9668. doi:10.1021/acs.jpclett.0c02818

Van Roey, K., Dinkel, H., Weatheritt, R. J., Gibson, T. J., and Davey, N. E. (2013). The switches. ELM resource: a compendium of conditional regulatory interaction interfaces. *Sci. Signal.* 6 (269), rs7–rs. doi:10.1126/scisignal.2003345

Van Roey, K., Gibson, T. J., and Davey, N. E. (2012). Motif switches: decision-making in cell regulation. *Curr. Opin. Struct. Biol.* 22 (3), 378–385. doi:10.1016/j.sbi.2012.03.004

Venkatesan, K., Rual, J.-F., Vazquez, A., Stelzl, U., Lemmens, I., Hirozane-Kishikawa, T., et al. (2009). An empirical framework for binary interactome mapping. *Nat. Methods* 6 (1), 83–90. doi:10.1038/nmeth.1280

Verma, G., Singh, P., and Bhat, R. (2020). Disorder under stress: role of polyol osmolytes in modulating fibrillation and aggregation of intrinsically disordered proteins. *Biophys. Chem.* 264, 106422. doi:10.1016/j.bpc.2020.106422

V'kovski, P., Kratzel, A., Steiner, S., Stalder, H., and Thiel, V. (2021). Coronavirus biology and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol.* 19 (3), 155–170. doi:10.1038/s41579-020-00468-6

Vora, S. M., Fontana, P., Mao, T., Leger, V., Zhang, Y., Fu, T.-M., et al. (2022). Targeting stem-loop 1 of the SARS-CoV-2 5' UTR to suppress viral translation and Nsp1 evasion. *Proc. Natl. Acad. Sci. U. S. A.* 119 (9), e2117198119. doi:10.1073/pnas.2117198119

Vora, S. M., Fontana, P., Mao, T., Leger, V., Zhang, Y., Fu, T., et al. (2022). Targeting stem-loop 1 of the SARS-CoV-2 5' UTR to suppress viral translation and Nsp1 evasion. *Proc Natl Acad Sci U S A*. 119(9):e2117198119. doi:10.1073/pnas.2117198119

Wang, C., Zheng, X., Gai, W., Zhao, Y., Wang, H., Wang, H., et al. (2017). MERS-CoV virus-like particles produced in insect cells induce specific humoural and cellular imminity in rhesus macaques. *Oncotarget* 8 (8), 12686–12694. doi:10.18632/oncotarget.8475

Wang, J., Shi, C., Xu, Q., and Yin, H. (2021b). SARS-CoV-2 nucleocapsid protein undergoes liquid–liquid phase separation into stress granules through its N-terminal intrinsically disordered region. *Cell Discov.* 7 (1), 5. doi:10.1038/s41421-020-00240-3

Wang, Y., Kirkpatrick, J., Lage, Sz, and Carlomagno, T. (2023). Structural insights into the activity regulation of full-length non-structural protein 1 from SARS-CoV-2. *Structure* 31 (2), 128–37.e5. doi:10.1016/j.str.2022.12.006

Wang, Y., Kirkpatrick, J., zur Lage, S., Korn, S. M., Neißner, K., Schwalbe, H., et al. (2021a). 1H, 13C, and 15N backbone chemical-shift assignments of SARS-CoV-2 non-structural protein 1 (leader protein). *Biomol. NMR Assignments* 15 (2), 287–295. doi:10.1007/s12104-021-10019-6

White, T. C., Yi, Z., and Hogue, B. G. (2007). Identification of mouse hepatitis coronavirus A59 nucleocapsid protein phosphorylation sites. *Virus Res.* 126 (1-2), 139–148. doi:10.1016/j.virusres.2007.02.008

Wohl, S., Gilron, Y., and Zheng, W. (2025). Structural and functional relevance of charge-based transient interactions inside intrinsically disordered proteins. *ACS Phys. Chem. Au.* doi:10.1021/acsphyschemau.5c00005

Woo, J., Lee, E. Y., Lee, M., Kim, T., and Cho, Y.-E. (2019). An *in vivo* cell-based assay for investigating the specific interaction between the SARS-CoV N-protein and its viral RNA packaging sequence. *Biochem. Biophys. Res. Commun.* 520 (3), 499–506. doi:10. 1016/j.bbrc.2019.09.115

Wu, C., Qavi, A. J., Hachim, A., Kavian, N., Cole, A. R., Moyle, A. B., et al. (2021). Characterization of SARS-CoV-2 nucleocapsid protein reveals multiple functional consequences of the C-terminal domain. *iScience* 24 (6), 102681. doi:10.1016/j.isci.2021.102681

Wu, C.-H., Chen, P.-J., and Yeh, S.-H. (2014). Nucleocapsid phosphorylation and RNA helicase DDX1 recruitment enables coronavirus transition from discontinuous to

continuous transcription. *Cell Host Microbe* 16 (4), 462–472. doi:10.1016/j.chom.2014. 09.009

Xue, B., Williams R, W., Oldfield C, J., Kian-Meng Goh, G., Keith Dunker, A. N., and Uversky, V. (2010). Viral disorder or disordered viruses: do viral proteins possess unique features? *PPL* 17 (8), 932–951.

Yan, J., Dunker, A. K., Uversky, V. N., and Kurgan, L. (2016). Molecular recognition features (MoRFs) in three domains of life. *Mol. Biosyst.* 12 (3), 697–710. doi:10.1039/ c5mb00640f

Yang, M., He, S., Chen, X., Huang, Z., Zhou, Z., Zhou, Z., et al. (2021). Structural insight into the SARS-CoV-2 nucleocapsid protein C-terminal domain reveals a novel recognition mechanism for viral transcriptional regulatory sequences. *Front. Chem.* 8, 624765. doi:10.3389/fchem.2020.624765

Yao, H., Song, Y., Chen, Y., Wu, N., Xu, J., Sun, C., et al. (2020). Molecular architecture of the SARS-CoV-2 virus. *Cell* 183 (3), 730-8. e13. doi:10.1016/j.cell. 2020.09.018

Yaron, T. M., Heaton, B. E., Levy, T. M., Johnson, J. L., Jordan, T. X., Cohen, B. M., et al. (2022). Host protein kinases required for SARS-CoV-2 nucleocapsid phosphorylation and viral replication. *Sci. Signal.* 15 (757), eabm0808. doi:10.1126/ scisignal.abm0808

Ye, Q., Lu, S., and Corbett, K. D. (2021). Structural basis for SARS-CoV-2 nucleocapsid protein recognition by single-domain antibodies. *Front. Immunol.* 12, 719037. doi:10.3389/fimmu.2021.719037

Yesudhas, D., Srivastava, A., Sekijima, M., and Gromiha, M. M. (2021). Tackling Covid-19 using disordered-to-order transition of residues in the spike protein upon angiotensin-converting enzyme 2 binding. *Proteins Struct. Funct. Bioinforma*. 89 (9), 1158–1166. doi:10.1002/prot.26088

Yu, I.-M., Gustafson, C. L., Diao, J., Burgner, J. W., Li, Z., Zhang, J., et al. (2005). Recombinant severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein forms a dimer through its C-terminal domain. *J. Biol. Chem.* 280 (24), 23280–23286. doi:10.1074/jbc.m501015200

Yuan, S., Peng, L., Park, J. J., Hu, Y., Devarkar, S. C., Dong, M. B., et al. (2020). Nonstructural protein 1 of SARS-CoV-2 is a potent pathogenicity factor redirecting host protein synthesis machinery toward viral RNA. *Mol. Cell* 80 (6), 1055–66. e6. doi:10. 1016/j.molcel.2020.10.034

Zachrdla, M., Savastano, A., Ibáñez de Opakua, A., Cima-Omori, M. S., and Zweckstetter, M. (2022). Contributions of the N-terminal intrinsically disordered region of the severe acute respiratory syndrome coronavirus 2 nucleocapsid protein to RNA-induced phase separation. *Protein Sci.* 31 (9), e4409. doi:10.1002/pro.4409

Zhang, J., Cai, Y., Xiao, T., Lu, J., Peng, H., Sterling, S. M., et al. (2021). Structural impact on SARS-CoV-2 spike protein by D614G substitution. *Sci. (New York, NY)* 372 (6541), 525–530. doi:10.1126/science.abf2303

Zhao, D., Xu, W., Zhang, X., Wang, X., Ge, Y., Yuan, E., et al. (2021). Understanding the phase separation characteristics of nucleocapsid protein provides a new therapeutic opportunity against SARS-CoV-2. *Protein Cell* 12 (9), 734–740. doi:10.1007/s13238-021-00832-z

Zhao, H., Nguyen, A., Wu, D., Li, Y., Hassan, S. A., Chen, J., et al. (2022). Plasticity in structure and assembly of SARS-CoV-2 nucleocapsid protein. *PNAS Nexus* 1 (2), pgac049. doi:10.1093/pnasnexus/pgac049

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579 (7798), 270–273. doi:10.1038/s41586-020-2012-7

Zhou, Y., Liu, Y., Gupta, S., Paramo, M. I., Hou, Y., Mao, C., et al. (2023). A comprehensive SARS-CoV-2-human protein–protein interactome reveals COVID-19 pathobiology and potential host therapeutic targets. *Nat. Biotechnol.* 41 (1), 128–139. doi:10.1038/s41587-022-01474-0

Zinzula, L., Basquin, J., Bohn, S., Beck, F., Klumpe, S., Pfeifer, G., et al. (2021). Highresolution structure and biophysical characterization of the nucleocapsid phosphoprotein dimerization domain from the Covid-19 severe acute respiratory syndrome coronavirus 2. *Biochemical and biophysical research communications*. 538: 54–62. doi:10.1016/j.bbrc.2020.09.131