

## CORONAVIRUS

# A specific phosphorylation-dependent conformational switch in SARS-CoV-2 nucleocapsid protein inhibits RNA binding

**Maiia Botova†, Aldo R. Camacho-Zarco†, Jacqueline Tognetti, Luiza Mamigonian Bessa‡, Serafima Guseva§, Emmi Mikkola, Nicola Salvi¶, Damien Maurin, Torsten Herrmann, Martin Blackledge\***

The nucleocapsid protein of severe acute respiratory syndrome coronavirus 2 encapsidates the viral genome and is essential for viral function. The central disordered domain comprises a serine-arginine-rich (SR) region that is hyperphosphorylated in infected cells. This modification regulates function, although mechanistic details remain unknown. We use nuclear magnetic resonance to follow structural changes occurring during hyperphosphorylation by serine arginine protein kinase 1, glycogen synthase kinase 3, and casein kinase 1, that abolishes interaction with RNA. When eight approximately uniformly distributed sites have been phosphorylated, the SR domain binds the same interface as single-stranded RNA, resulting in complete inhibition of RNA binding. Phosphorylation by protein kinase A does not prevent RNA binding, indicating that the pattern resulting from physiologically relevant kinases is specific for inhibition. Long-range contacts between the RNA binding, linker, and dimerization domains are abrogated, phenomena possibly related to genome packaging and unpackaging. This study provides insight into the recruitment of specific host kinases to regulate viral function.

## INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the infectious agent responsible for the recent COVID-19 pandemic, is an enveloped positive-sense single-stranded RNA virus of the betacoronavirus genus that expresses its own replication machinery. Genome replication is achieved by the RNA-dependent RNA polymerase, whose structure has been investigated by cryo-electron microscopy (1–3), within viral replication organelles called double-membrane vesicles (DMVs) that are remodeled from the endoplasmic reticulum (4). DMVs have been shown to constitute principal sites of viral RNA replication (5, 6), containing narrow exit channels formed by nonstructural viral proteins (nsps) through which newly synthesized single-stranded RNA can exit before encapsidation.

The nucleocapsid protein is the most abundant protein in betacoronaviruses and an important cofactor of the replication machinery (7, 8) that is known to colocalize to the replication transcription complex (9–11) via an essential interaction with the N terminus of nonstructural protein 3 (nsp3) (12, 13). SARS-CoV-2 nucleocapsid protein (N) binds to, and encapsidates, the viral genome, shielding the RNA from the host innate immune system, as well as playing an essential but still poorly understood role in regulation of gene transcription (14). Electron tomography was used to identify N-RNA complexes within SARS-CoV-2 virions, displaying a morphology resembling beads on a string (15, 16).

N is highly disordered as revealed by nuclear magnetic resonance (NMR) spectroscopy (17, 18), comprising three intrinsically disordered domains (N1, N3, and N5) (Fig. 1), flanking the RNA binding

domain (N2) and the dimerization domain N4, whose structures have both been determined at atomic resolution (19–23). RNA binding to N2 has been described (24–32), including its potential role in liquid-liquid phase separation (LLPS) of N-RNA mixtures (26, 27, 33–39). N4 forms a highly stable dimeric structure involving two domain-swapped  $\beta$  hairpins (22, 40). Although N2 is known to represent the main RNA binding domain, secondary RNA binding sites have been proposed to exist in both N3 (26, 37, 41, 42) and N4 (40, 41, 43–46), an observation supported by the assembly of truncation mutants of N into large supramolecular complexes of the dimensions of ribonucleoproteins (RNPs) (47). N5 has also been associated with assembly into higher-order oligomeric states (48).

N3 is of particular interest for a number of reasons. It comprises three distinct linear segments, a serine-arginine-rich region (SR, S176 to S206), a hydrophobic, leucine-rich helix (H1), and a polar region, comprising a glutamine-rich strand and terminated by a lysine-rich region that forms a detached helix (H2) in the crystal structure of N4, and represents the junction between these two domains. N3 was recently shown to fold around the N-terminal, ubiquitin-like domain of nsp3 (Ubl1) (49). H1 binds tightly to a hydrophobic cleft on the surface of Ubl1, while the polyQ and H2 regions complete a bipartite folding-upon-binding interaction. The interaction results in a massive collapse in the conformational sampling of dimeric N, bringing N2 and N4 in close contact and forming a compact complex with Ubl1. Nsp3 forms the cytosolic extremity of the exit channel of DMVs, suggesting that this essential interaction between viral proteins plays a role in positioning N before encapsidation of the nascent RNA, a model supported by the observed accumulation of N in the vicinity of the surface of DMVs (4).

The functional importance of N3 has also been highlighted by the accumulation of important mutations, associated with dominant variants of concern in this region (50, 51) (for example, alpha, gamma, omicron <sup>203</sup>RG-KR, beta <sup>205</sup>T-I, delta <sup>203</sup>R-M, omicron <sup>215</sup>G-C, and BA.2.86 <sup>229</sup>Q-K). Most of these mutations are associated with the SR segment of N3. Indeed mini-replicon studies have

Copyright © 2024 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

Université Grenoble Alpes, CNRS, CEA, IBS, F-38000 Grenoble, France.

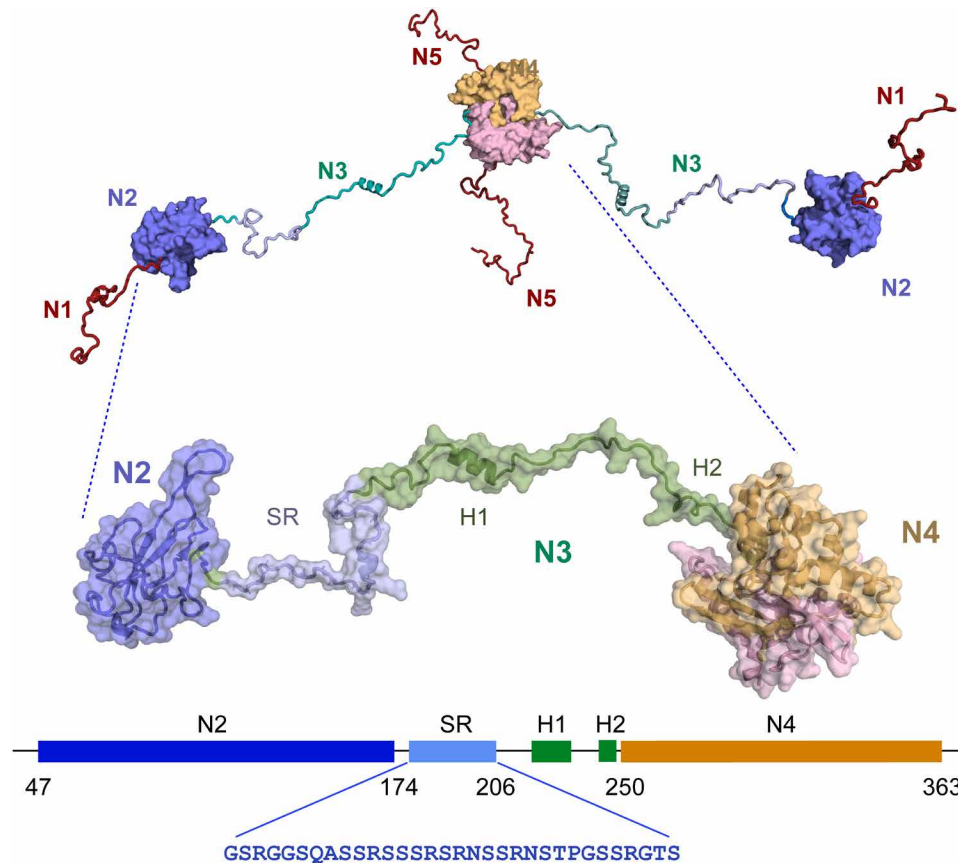
\*Corresponding author. Email: martin.blackledge@ibs.fr

†These authors contributed equally to this work.

‡Present address: Evotec (France) SAS, Toulouse, France.

§Present address: Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY, USA.

¶Present address: Sanofi R&D, Integrated Drug Discovery, 94403 Vitry-sur-Seine, France.



**Fig. 1. Figurative representation of SARS-CoV-2 nucleocapsid protein (N).** SARS-CoV-2 nucleocapsid protein comprises five domains (N1 to N5, top). The construct shown in the zoom comprises three of these domains, N2, the RNA binding domain; N3, the disordered central domain containing two helices H1 (219 to 231) and H2 (248 to 255) and the serine-arginine-rich (SR) region that is hyperphosphorylated in infected cells; and the dimerization domain N4. This construct is referred to as N234. The sequence of the SR region is shown in blue.

measured 10-fold increased mRNA delivery and protein expression for commonly found mutants, and a reverse genetics model revealed 50-fold more virus production for S202R and R203M mutations (52). Helix H1 has also been linked to assembly, with a recent study using analytical ultra-centrifugation and circular dichroism combined with molecular dynamics simulation, suggesting that the motif promotes higher-order assembly (53).

Perhaps most importantly, the SR region of N3 is found to be hyperphosphorylated in infected cells, although it is unphosphorylated during viral assembly and is found in the unphosphorylated form when bound to genomic RNA in infectious virions (54–56). It has been suggested that hyperphosphorylation or dephosphorylation may play a role as a functional switch in the replication cycle of the virus, either between transcription and replication, or to enable genome packaging or unpacking (14, 57–60). Phosphorylation has also been shown to affect the compactness of viral ribonucleoprotein (vRNP) complexes (61), and concurrent studies on this have shown that binding of the isolated N3 to long RNAs is modulated by phosphorylation (62), while fluorescence polarization measured nearly an order of magnitude weaker binding to a 10-nucleotide (nt) RNA following hyperphosphorylation (47).

N is also known to undergo LLPS upon mixing with RNA (26, 27, 33–38), possibly providing compartmentalization of some stage of

the viral replication cycle, as has recently been shown for negative-sense single-stranded RNA viruses, such as rabies (63) and measles (64). Hyperphosphorylation of the SR region (pSR) has also been shown to modulate the nature of membraneless organelles formed from N and RNA (27, 35, 37, 38). Although the specific function of SARS-CoV-2 viral compartments remains the subject of debate (65), it is clear that the storage of high concentrations of N in the vicinity of DMVs (4, 6, 16, 55) may provide an efficient mechanism for rapid encapsidation of newly synthesized RNA.

Despite the extensive work described above, little is known about the influence of phosphorylation on the structural and dynamic behavior of N and its impact on molecular function. The kinase cascade responsible for hyperphosphorylation of N was identified (60), implicating, sequentially, serine arginine protein kinase 1 (SRPK1), glycogen synthase kinase 3 (GSK-3), and casein kinase 1 (CK1). Here, we use time-dependent NMR spectroscopy, to characterize the conformational and functional implications of hyperphosphorylation of N at atomic resolution. This reveals that transient contacts between H1 and N2, and between N4 and N2 are abolished upon hyperphosphorylation and replaced by direct binding of pSR to the RNA binding surface of N2. pSR appears to interact with N2 via the same interface as single-stranded RNA, suggesting an auto-inhibitory mechanism involving direct competition with RNA. Notably, we

demonstrate that phosphorylation by SRPK1 and GSK-3 is sufficient for this inhibition, although not all phosphorylation sites are required and that *in vitro* phosphorylation by noncognate kinases such as protein kinase A (PKA), a commonly used proxy for hyperphosphorylation, does not result in inhibition of RNA binding. In particular, time-resolved NMR reveals the precise threshold of the specific phosphorylation pattern necessary to achieve the competitive interaction mechanism required to fully inhibit RNA binding, providing important and unique insight into the role of posttranslational modification in the SARS-CoV-2 replication cycle.

## RESULTS

### The disordered SR region of N234 is hyperphosphorylated by SRPK1, GSK-3, and CK1

A construct comprising domains N2, N3, and N4 (N234), including the RNA binding, dimerization, and hyperphosphorylation domains, was initially investigated (residues 45 to 375, lacking only the disordered N- and C-terminal domains). Incubation with SRPK1 resulted in the appearance of two phosphorylation sites that were assigned to sites pS188 and pS206. Published studies have demonstrated the sequential dependence of the three kinases, identifying that SRPK1 phosphorylates two sites that prime for subsequent phosphorylation of eight sites by GSK-3, which, in turn, primes for four more sites that are phosphorylated by CK1 (60). The addition of GSK-3 and then CK1 results in phosphorylation of eight and four additional sites, respectively, in agreement with recent reports (59) (Fig. 2). Under physiological conditions, the total theoretical charge of the 31-amino acid SR region thus evolves from +6, to +2, -14, and lastly -22 over the three phosphorylation steps [resulting in pN234(I), pN234(II), and pN234(III), corresponding to phosphorylation by SRPK1, GSK-3, and then CK1, respectively]. Each phosphorylation site was assigned using triple-resonance three-dimensional (3D) NMR, and real-time NMR was used to compare the rate of phosphorylation of different sites as a function of each kinase. In particular, this allows us to identify S176 and S186 that are phosphorylated notably slower by GSK-3 than S180, S184, S190, S194, S198, and S202 (Fig. 2E). Note that while the  $^{15}\text{N}$ - $^1\text{H}$  correlation peak of pS176 is resolved, and can therefore be accurately analyzed in terms of intensity buildup, pS186 is partially overlapped with pS184 and pS202 and notably slower than both, complicating further analysis (the nonphosphorylated peak is also strongly overlapped).

To investigate possible changes in backbone conformation,  $^{13}\text{C}^\alpha$  chemical shifts of pSR were compared to recently proposed random coil chemical shift values (66, 67) for unphosphorylated, pN234(II)-, pN234(III)-, and PKA-phosphorylated N234 (fig. S1), indicating that no obvious induction of secondary structure results from hyperphosphorylation, as has been previously observed for PKA-phosphorylated N (residues 1 to 209) (46).

### Hyperphosphorylation modifies the dynamic behavior of the disordered SR region

To investigate the impact of phosphorylation on the dynamic behavior of N234, we have measured  $^{15}\text{N}$  spin relaxation as a function of phosphorylation state (Fig. 2G).  $R_{1\rho}$  reports on local reorientational correlation times of internuclear  $^{15}\text{N}$ - $^1\text{H}$  bonds distributed along the primary sequence. This reveals that although N2 appears to maintain its hydrodynamic behavior, with comparable  $R_{1\rho}$  in the

presence and absence of phosphorylation, the overall correlation time of the SR region increases over the three steps. Similar observations have been made in phosphorylation studies (68, 69). The increase in pSR is more or less systematic for both pN234(II) and pN234(III), with local peaks observed around residues pS180, pS188, and 202 to 204. Comparison of  $R_{1\rho}$  measured from pN123(II) confirms the above observations and demonstrates that relaxation of the helical region H1 is not measurably affected by phosphorylation (resonances from this helix are too broadened in N234 to measure relaxation with sufficient precision).

### NMR maps the interaction of N234 with RNA from the viral 5' UTR

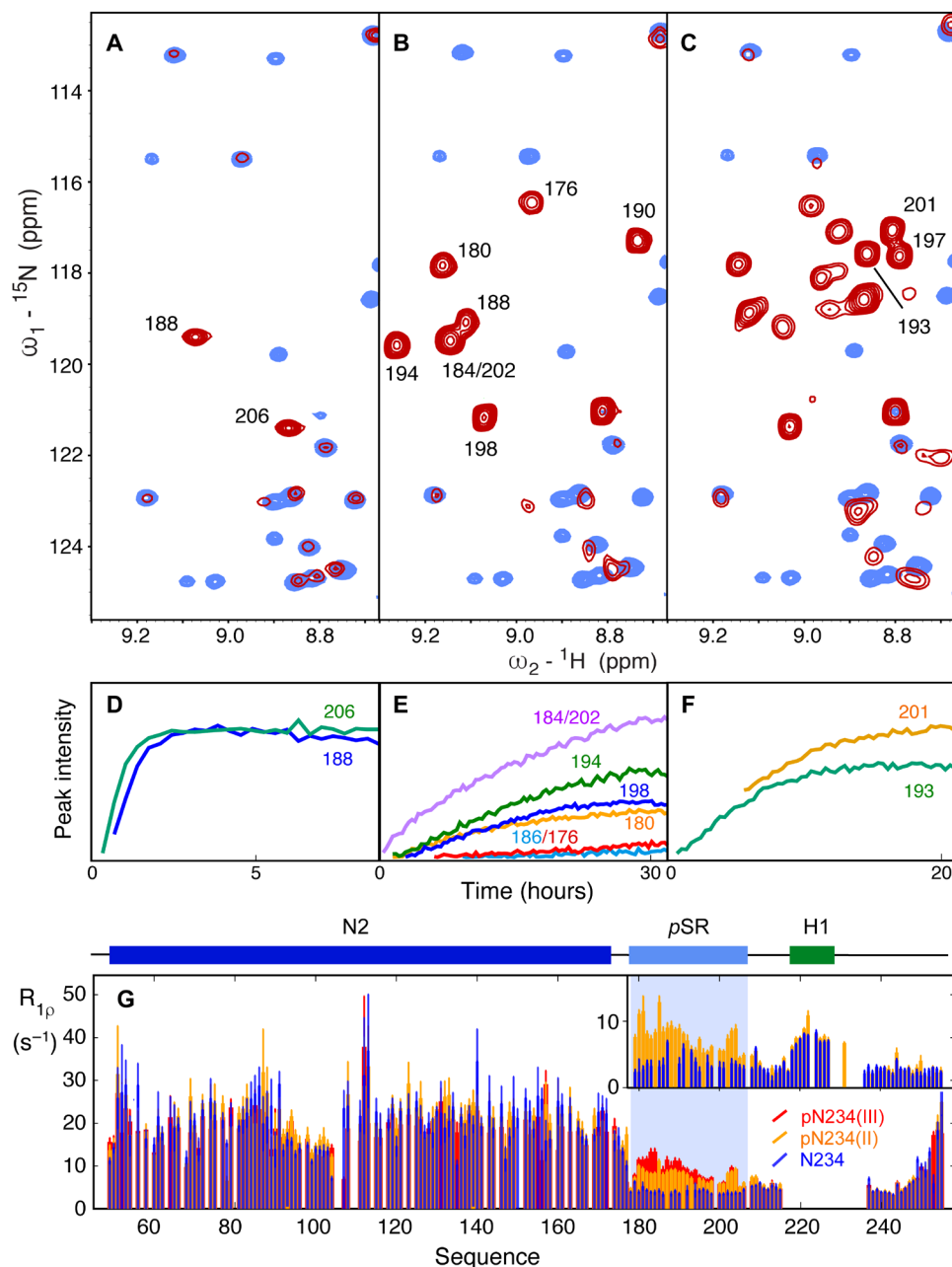
We have investigated the RNA binding profile of N234 using a 14-nt segment (UCUAAACGAACUUU) from the 5' UTR of SARS-CoV-2 and a 30-nt polyA.  $^{15}\text{N}$ - $^1\text{H}$  chemical shift perturbations (CSPs) in N2 measured in N234 (Fig. 3A) resemble those recently published by numerous authors (25, 28–31, 45), involving the RNA binding finger (residues 90 to 106), and the solvent-accessible surface of the baseplate formed by residues 48 to 66, 148 to 160, and 166 to 174. Additional shifts are seen throughout the SR region as well as in the boundary between N3 and N4 (residues 244 to 257), a lysine-rich region that folds as an  $\alpha$  helix (H2) upon binding nsp3, and forms a detached  $\alpha$  helix in the crystal structure of N4 (20, 21). This region was also shown to interact with RNA in the context of isolated N4 (44), as well as in SARS-CoV (41). Apart from H2, CSPs in N4 are small, except for the solvent-exposed C-terminal helix comprising residues 345 to 357.

The binding interface is confirmed by  $^{13}\text{C}$ - $^1\text{H}$  methyl CSPs (Fig. 3G). Chemical shift titration of six methyl group signals as a function of concentration of added RNA reveals an affinity of  $(230 \pm 10)$   $\mu\text{M}$  (Fig. 3D). Titration with the two different RNAs resulted in shifts of the same amino acids in N2 and appears upon the same linear CSP trajectories (fig. S2); the exact position is dependent on the titration admixture and exchange regime that differs between RNAs ranging from fast to slow exchange on the NMR time scale, likely related to different affinities (the longer RNA was found to be in slower exchange).

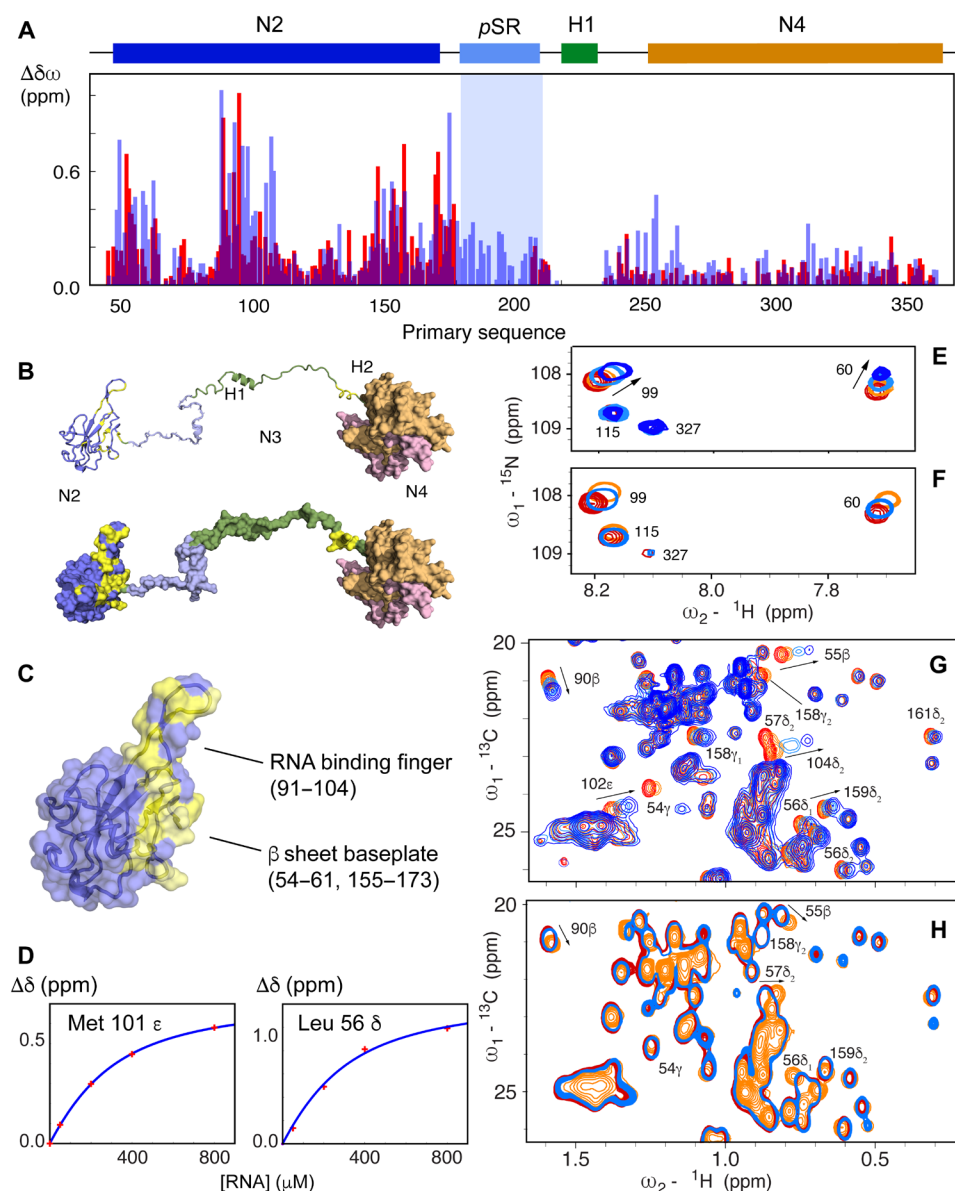
### Hyperphosphorylated SR inhibits N:RNA interaction via the RNA binding surface of N2

A comparison of the  $^{15}\text{N}$ - $^1\text{H}$  and  $^{13}\text{C}$ - $^1\text{H}$  CSPs due to binding of diverse RNAs, and those resulting from sequential phosphorylation by SRPK1, GSK-3, and CK1 reveals remarkable correspondences (Fig. 3A). While the SRPK1 priming step induces small shifts (see fig. S3), the addition of 12 phosphorylation sites in pN234(III) results in a very similar distribution of CSPs in N2 compared to those observed as a result of RNA binding [chemical shifts in pN234(II) and pN234(III) show a similar profile, fig. S4]. This indicates that hyperphosphorylation induces a change in conformation of N2, probably due to binding of N2 to the pSR region, that mirrors the impact of RNA binding. We note that although the CSPs are distributed similarly, they can be different in sign to those observed upon binding RNA, demonstrating that while the same sites on the surface of N2 are affected, the bound state chemical shifts (RNA-bound and pN3-bound) are different (Fig. 3, G and H).

We then tested the RNA binding properties of the different forms of pN234. While phosphorylation due to SRPK1 has no observable effect on RNA binding (Fig. 4A), the additional phosphorylation



**Fig. 2. Phosphorylation of N by SRPK1, GSK-3, and CK1.** (A) Comparison of  $^{15}\text{N}$ - $^1\text{H}$  HSQC of unphosphorylated (blue) and phosphorylated (red) N234 following incubation with SRPK1 and phosphorylation buffer [referred to as pN234(I)]. Two resonances appear that have been assigned to pS188 and pS206. (B) Comparison of  $^{15}\text{N}$ - $^1\text{H}$  HSQC of unphosphorylated (blue) and phosphorylated (red) N234 following incubation of pN234(I) with GSK-3 and phosphorylation buffer [referred to as pN234(II)]. Eight additional resonances appear (seven are shown here) that have been assigned to pS176, pS180, pS184, pS186, pS190, pS194, pT198, and pS202. (C) Comparison of  $^{15}\text{N}$ - $^1\text{H}$  HSQC of unphosphorylated (blue) and phosphorylated (red) N234 following incubation of pN234(II) with CK1 and phosphorylation buffer [referred to as pN234(III)]. Four additional resonances appear (three are shown) that have been assigned to pS193, pS197, and pS201. All spectra were buffer exchanged into standard NMR buffer before recording these spectra. (D to F) Evolution of peak intensity as a function of time for selected resonances that remain sufficiently resolved throughout the kinetic series to accurately measure peak intensities in each step (in phosphorylation buffer). (D) pN234(I), (E) pN234(II), and (F) pN234(III). (G) Phosphorylation modifies the dynamic behavior of N. Main panel: Rotating frame relaxation ( $R_{1\rho}$ ) of N234 at 850 MHz as a function of the phosphorylation state [blue, nonphosphorylated; orange, pN234(II); red, pN234(III)]. Data from N4 are not shown due to low signal-to-noise ratio. Inset: Rotating frame relaxation ( $R_{1\rho}$ ) of N123 at 950 MHz [blue, nonphosphorylated N123; orange, pN123(II)]. Helix H1 is visible in N123, probably due to the lower molecular weight of this construct, allowing comparison of the dynamic behavior of this region of N3. This comparison reveals that hyperphosphorylation strongly impacts the dynamic properties of the SR region.

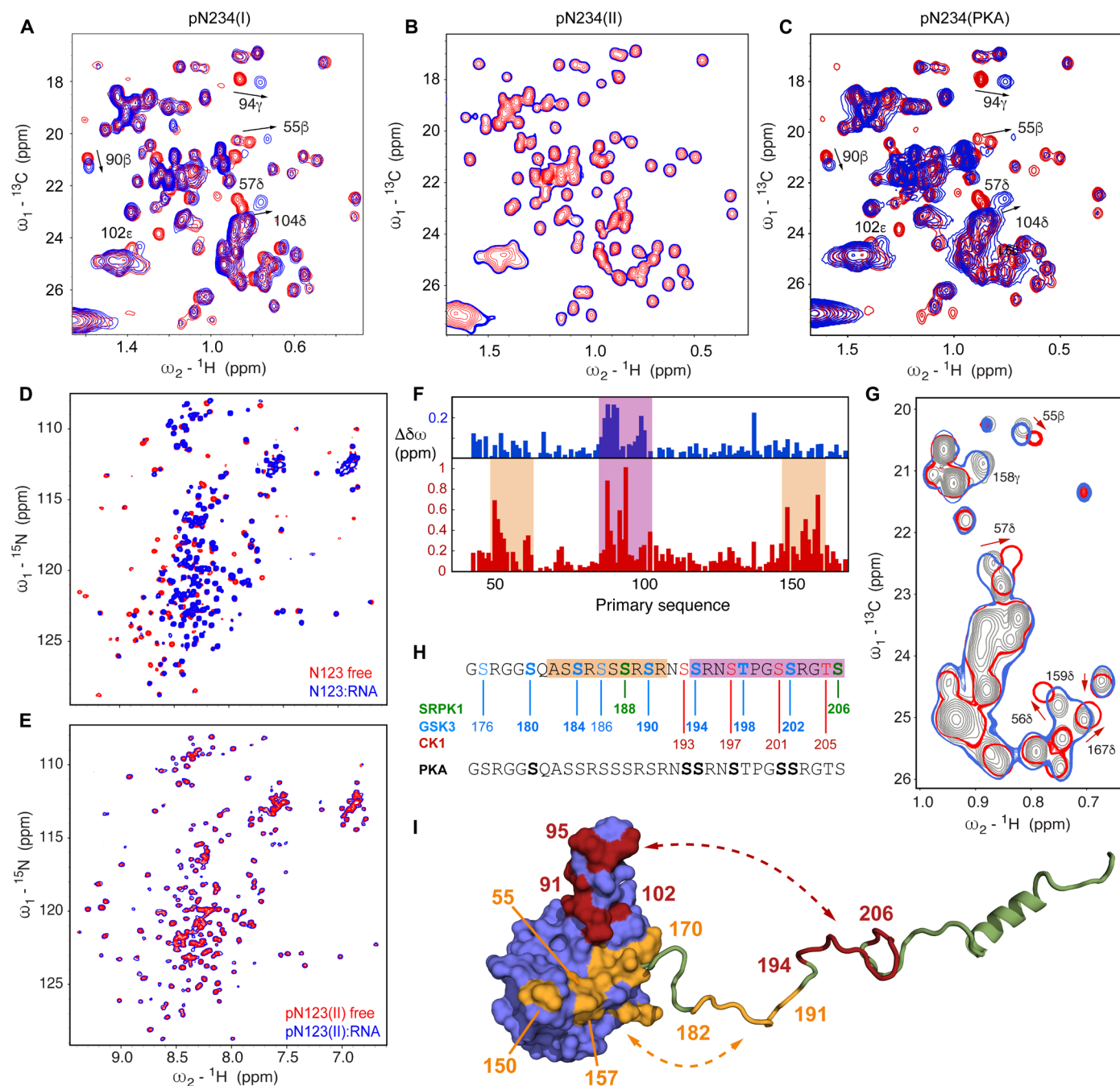


**Fig. 3. Hyperphosphorylation mimics RNA binding to N2.** (A) CSPs measured in the  $^{15}\text{N}$ - $^1\text{H}$  TROSY spectrum of N234 following addition of single-stranded 14-mer RNA (UCUAAACGAACUUU) (blue) in comparison with chemical shift differences between nonphosphorylated and SRPK1/GSK-3/CK1 phosphorylated N234 (red) (the extreme shifts from the pSR region due to direct phosphorylation of the residue of interest are not shown for reasons of dynamic range). CSPs due to RNA binding were scaled by a factor of 1.5 for optimal comparison with phosphorylation CSPs. (B) Mapping of the largest chemical shift differences onto N2 and N3 (yellow), shown in ribbon and surface representation. (C) Location of the RNA binding finger and baseplate  $\beta$  sheet on the RNA binding domain, N2. Yellow surface as in (B). (D) Examples of chemical shift titrations derived from  $^{13}\text{C}$ - $^1\text{H}$  HMQC of 150  $\mu\text{M}$  N234 in the presence of increasing concentrations of 14-mer RNA. Residues in the RNA binding finger (Met<sup>101</sup>) and baseplate  $\beta$  sheet (Leu<sup>56</sup>) on the RNA binding domain are shown. The association constant derived from these titrations is  $(230 \pm 10) \mu\text{M}$ . (E) CSPs in  $^{15}\text{N}$ - $^1\text{H}$  TROSY spectra of N234 (concentration, 150  $\mu\text{M}$ ) as a function of 14-mer single-stranded RNA concentration (red, 0%; orange, 25%; light blue, 100%; dark blue, 200%) (see fig. S8 for full spectra). (F) CSPs in  $^{15}\text{N}$ - $^1\text{H}$  TROSY spectra of N234 (concentration, 150  $\mu\text{M}$ ) for the same residues as (E) as a function of phosphorylation state [red, unphosphorylated N234; light blue, pN234(I); orange, pN234(II)]. (G) CSPs in  $^{13}\text{C}$ - $^1\text{H}$  HMQC spectra of N234 (concentration, 150  $\mu\text{M}$ ) for selected residues as a function of 14-mer single-stranded RNA concentration (red, 0%; orange, 25%; light blue, 100%; dark blue, 200%). (H) CSPs in  $^{13}\text{C}$ - $^1\text{H}$  HMQC spectra of N234 (concentration, 150  $\mu\text{M}$ ) for the same residues as (E) as a function of phosphorylation state [red, unphosphorylated N234; light blue, pN234(I); orange, pN234(II)].

sites resulting from incubation with GSK-3 and CK1 result in complete abrogation of RNA binding (Fig. 4B). This was found to be the case for both RNA molecules that were tested. Both pN234(II) and pN234(III) fail to bind RNA, demonstrating that phosphorylation by SRPK1/GSK-3 is necessary and sufficient to inhibit RNA binding. We note that it has been previously observed that

phosphorylation by SRPK1 is an essential step for phosphorylation by GSK-3 (60).

The broadly phosphorylating PKA has been exploited to phosphorylate N in a single step, including via coexpression with N (46), and notably to investigate the impact on binding of the host factor 14-3-3 (70). Incubation of N with PKA in vitro results in



**Fig. 4. Hyperphosphorylation by SRPK1, GSK-3, and CK1 inhibits RNA binding.** (A)  ${}^{13}\text{C}$ - ${}^1\text{H}$  HMQC spectra of N234 (concentration, 150  $\mu\text{M}$ ) for increasing concentrations of 14-mer single-stranded RNA (red, 0%; dark blue, 200%). (B) Chemical shifts in  ${}^{13}\text{C}$ - ${}^1\text{H}$  HMQC spectra of SRPK1/GSK-3 phosphorylated N234 [150  $\mu\text{M}$  pN234(II)], for selected residues as a function of 14-mer single-stranded RNA concentration (red, 0%; dark blue, 200%). No notable shifts were observed [compare to (A)]. The 200% spectrum is presented as a single contour to highlight the lack of binding. (C) Chemical shifts in  ${}^{13}\text{C}$ - ${}^1\text{H}$  HMQC spectra of N234 phosphorylated by PKA (150  $\mu\text{M}$ ), for selected residues as a function of 14-mer single-stranded RNA concentration (red, 0%; dark blue, 200%). Shifts are highly comparable with CSPs measured in unphosphorylated N234 [compare to (A)]. (D)  ${}^{15}\text{N}$ - ${}^1\text{H}$  TROSY spectra of nonphosphorylated N123 (concentration, 200  $\mu\text{M}$ ) in the presence and absence of equimolar 14-mer single-stranded RNA (red, 0%; blue, 100%). (E)  ${}^{15}\text{N}$ - ${}^1\text{H}$  TROSY spectra of SRPK1-GSK3 phosphorylated N123 (concentration, 200  $\mu\text{M}$ ) in the presence and absence of equimolar 14-mer single-stranded RNA (red, 0%; blue, 100%). The 100% spectrum is presented as a single contour to highlight the lack of binding. (F) CSPs (blue) in  ${}^{15}\text{N}$ - ${}^1\text{H}$  TROSY of N234 phosphorylated by PKA, and (red) pN234(II). (G) Section of the  ${}^{13}\text{C}$ - ${}^1\text{H}$  HMQC spectrum of unphosphorylated N234 (gray), with N234 phosphorylated by PKA (blue) and pN234(II) (red). (H) Sequence of SR region showing *in vitro* phosphorylation sites (bold) following incubation with PKA (gray/black) and sequentially SRPK1 (green), GSK-3 (blue), and CK1 (red). (I) Sketch illustrating putative differential binding sites of N-terminal (182 to 191) and C-terminal (194 to 206) regions of pSR on the surface of N2. Red shading on N2 indicates CSPs induced by both PKA and SRPK1/GSK-3 phosphorylation, while orange indicates regions that are only shifted by SRPK1/GSK-3 phosphorylation.

phosphorylation of six sites in the SR region of N234 (fig. S5) whose backbone resonances were assigned, namely, pS180, pS193, pS194, pS197, pS201, and pS202, although broader phosphorylation was observed by coexpression with PKA (46, 70). This sixfold phosphorylated form of N234 nevertheless maintains its binding to RNA (Fig. 4C), with essentially identical spectral changes compared to the nonphosphorylated form in interaction with the same RNA (Figs. 3G and 4C and fig. S6). The specific pattern resulting from SRPK1/GSK-3 phosphorylation is therefore critical for abrogating RNA binding to N234, while N234 phosphorylated by PKA *in vitro* is unable to reproduce this functional phenotype.

## N2 and N3 are the minimum necessary elements for RNA binding inhibition

A construct comprising domains N1, N2, and N3 (N123, residues 1 to 267), including the N-terminal intrinsically disordered region, RNA binding, and hyperphosphorylation domains, was also investigated. The same signature of RNA binding is observed in N2 in N123 as in N234 (fig. S7), demonstrating that any impact of the dimerization domain, N4, or the disordered N-terminal domain, N1, on binding of these RNA segments is too weak to be observed by  $^{13}\text{C}$ - $^1\text{H}$  and  $^{15}\text{N}$ - $^1\text{H}$  CSPs. The impact of hyperphosphorylation on RNA binding is the same as that observed for N234 (see Fig. 4 and fig. S8 for  $^{13}\text{C}$ - $^1\text{H}$  and  $^{15}\text{N}$ - $^1\text{H}$  spectra, respectively), with complete abrogation of binding in the SRPK1/GSK-3 phosphorylated form of the construct (Fig. 4, D and E), again identifying N2 and N3 as the minimum necessary elements for this inhibition. We cannot exclude the idea that N5, which we have not studied here and which contains a lysine-rich region, plays a role in RNA binding or assembly.

The possible origin of this differentiation between the impact of PKA phosphorylation and SRPK1/GSK-3 phosphorylation is revealed by comparison of  $^{15}\text{N}$ - $^1\text{H}$  and  $^{13}\text{C}$ - $^1\text{H}$  CSPs in the two cases (Fig. 4, F and G). Although much smaller shifts due to PKA phosphorylation are seen in general, the main interacting regions appear to involve the RNA binding finger (residues 90 to 106). CSPs involving the baseplate (residues 48 to 66 and 148 to 174) are considerably smaller. Notably, the region of pSR situated between positions 182 and 191 comprises four phosphorylation sites in pN234(II), but is unmodified by *in vitro* PKA phosphorylation (Fig. 4H). Together, these observations are consistent with a model whereby phosphorylation of the N-terminal strand of pSR (residues 182 to 191) is responsible for interaction with the baseplate, while phosphorylation of the C-terminal strand (residues 194 to 206), present in both pN234(II) and PKA-phosphorylated N234, would be responsible for binding the RNA binding finger (Fig. 4I). It is tempting to speculate that occupation of both interaction sites is required to abrogate interaction with RNA. Note that the specific differences in phosphorylation pattern of residues 194 to 206 may also contribute to stronger and more efficient binding of N2.

## Hyperphosphorylated SR sites specifically abrogate RNA binding

To further identify the specific phosphorylation steps that are essential for abrogating RNA binding to pN234(II), modulation of the RNA binding equilibrium was investigated during the phosphorylation process. Having established that the minimum phosphorylation level for abolishing RNA interaction occurs as a result of phosphorylation by SRPK1 and GSK-3, the 14-mer RNA was added to pN234(I), to which GSK-3 and ATP were added, to

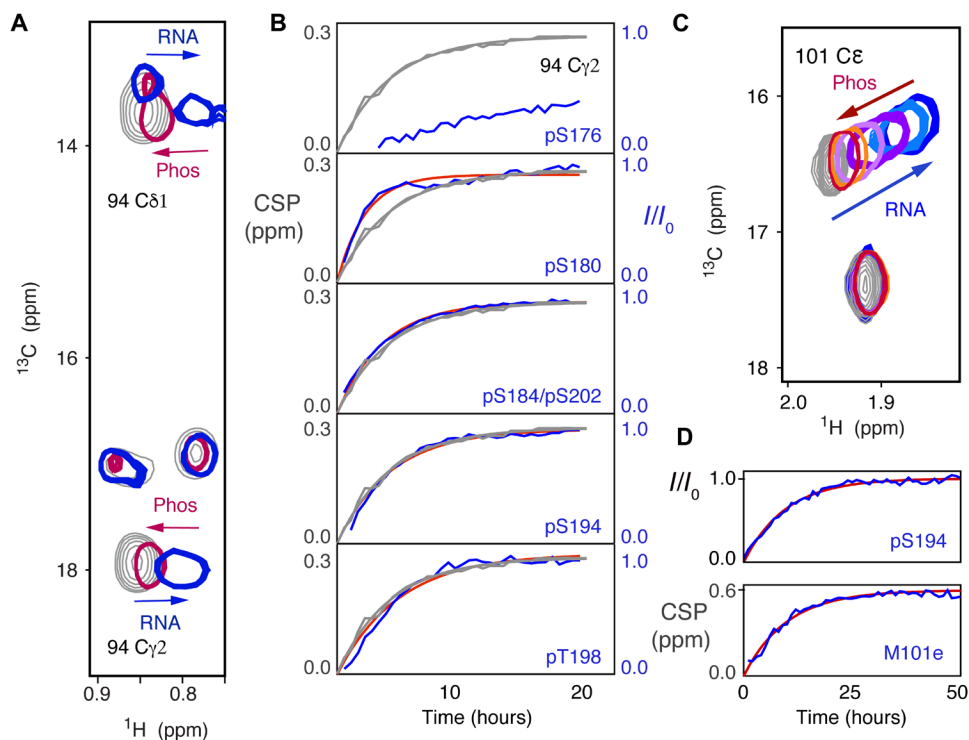
simultaneously observe the level of phosphorylation and RNA binding. Signature peaks reporting on single-stranded 14-mer RNA binding were followed using  $^{13}\text{C}$ - $^1\text{H}$  HMQC methyl spectra, and the level of phosphorylation was followed by measuring intensities of interleaved  $^{15}\text{N}$ - $^1\text{H}$  SOFAST HSQC, as described above. This allows us to follow the population shift between free and bound RNA by measuring sites that shift upon RNA binding (for example, the methyl group of I94, Fig. 5A), and compare this with the rate of appearance of the eight sites that are phosphorylated by GSK-3, where the covalent nature of the modification provides a direct measure of the extent of the phosphorylation.

The rates of phosphorylation are indicated in Fig. 5B, showing time constants of  $(0.50 \pm 0.07)$  hours for pS180,  $(0.30 \pm 0.04)$  hours for pS184/pS202,  $(0.26 \pm 0.07)$  hours for pS194, and  $(0.24 \pm 0.05)$  hours for pT198. pS176 and pS186 were too slow to be accurately fitted. Note that the level of phosphorylation of individual sites was investigated by also observing the intensity of the nonphosphorylated peak following the experiment (fig. S9), and where possible, both the increase in intensity of phosphorylated and the reduction in intensity of nonphosphorylated peaks were simultaneously fitted to the same time constant. Examples (fig. S9) indicate that with respect to experimental noise, phosphorylation is eventually complete.

Increases in intensity of resolved phosphorylation peaks are compared to the apparent chemical shift evolution of the methyl groups (Fig. 5, A and B). This comparison reveals that while S180 phosphorylates faster than the observed shift reporting on the RNA binding equilibrium, a number of sites show rates of phosphorylation that are very similar to the rate of RNA unbinding (S184, S194, T198, and S202), and two sites have slower phosphorylation rates (S176 and S186), suggesting that the latter two sites are not essential for the inhibition of RNA binding, at least under the conditions we have tested *in vitro*.

The experiment was repeated using 30-mer polyA RNA. In this case, we observe evolution of the chemical shift of Met 101ε, a peak that is sensitive to RNA binding, but relatively insensitive to phosphorylation, with a similar time constant compared to the intensity of the phosphorylation peak of pS194 (Fig. 5, C and D). Complete phosphorylation of pN234(I) with the eight additional sites necessary to reach pN234(II) requires intermediate creation of at least nine distinct species, comprising between zero and eight phosphorylation sites, that do not exchange with each other. The observed spectrum is therefore a superposition of differently phosphorylated states that each represents potentially different exchange regimes between free and RNA-bound protein, and as the population distribution changes from nonphosphorylated to fully phosphorylated, this spectrum evolves. We observe a distribution of peaks whose maximum shifts, from the known RNA-bound chemical shift of nonphosphorylated N234, to the chemical shift of free pN234(II), with a similar characteristic time constant as the buildup of phosphorylation. As the chemical shifts of pN234(I) and pN234(II) are very similar, the most likely explanation is that different transiently created phosphorylation states experience different exchange rates between free and RNA-bound forms, likely associated with a reduction in affinity, which gives rise to chemical shifts that are increasingly similar to the fully phosphorylated/unbound state.

Overall, simultaneous observation of the time dependence of RNA binding and SRPK1/GSK-3 phosphorylation reveals that increasing phosphorylation weakens RNA binding, until six of the eight phosphorylation sites are modified, at which point RNA can no longer bind.



**Fig. 5. Real-time observation of RNA binding inhibition by SR phosphorylation.** (A) CSPs in methyl  $^{13}\text{C}$ - $^1\text{H}$  HMQC of pN234(I), specifically 94C $\gamma$ 2, showing the impact of 14-mer RNA binding (gray, 0  $\mu\text{M}$ , blue 200  $\mu\text{M}$ ), and subsequent shift of the same peaks back toward the unbound form during incubation with GSK-3 (red). (B) Comparison of time course of intensity increase in the  $^{15}\text{N}$ - $^1\text{H}$  HSQC phosphorylation peaks from pSR (blue, experimental; red, single exponential fit) and the CSP of 94C $\gamma$ 2 (gray, experimental and fitted). Associated time constants of (0.50  $\pm$  0.07) hours (pS180), (0.30  $\pm$  0.04) hours (pS184/pS202), (0.26  $\pm$  0.07) hours (pS194), and (0.24  $\pm$  0.05) hours (pT198) compared with (0.24  $\pm$  0.03) hours and (0.26  $\pm$  0.03) hours for 94C $\gamma$ 2 and 94C $\delta$ 1, respectively. pS176 and pS186 (not shown) were too slow to be accurately fitted. (C) CSPs in methyl  $^{13}\text{C}$ - $^1\text{H}$  HMQC of pN234(I) (M101C $\epsilon$ ) showing the impact of 30-mer polyA RNA binding (gray, 0  $\mu\text{M}$ , blue 120  $\mu\text{M}$ ), and subsequent shift of the distribution of differently phosphorylated states toward the unbound, fully phosphorylated form during incubation with GSK-3 (light blue to red). (D) Comparison of (top) the time course of increase in intensity in the  $^{15}\text{N}$ - $^1\text{H}$  HSQC pS194 peak (blue, experimental; red, fitted to a single exponential) and (bottom) the shift of methyl  $^{13}\text{C}$ - $^1\text{H}$  CSP of 101C $\epsilon$  back from RNA-bound to unbound form during incubation with GSK-3 as shown in (C) (blue, experimental; red, single exponential fit). Associated time constants: (0.10  $\pm$  0.04) hours pS194 and (0.099  $\pm$  0.030) hours for 101C $\epsilon$ .

### Long-range order in N is modulated by hyperphosphorylation

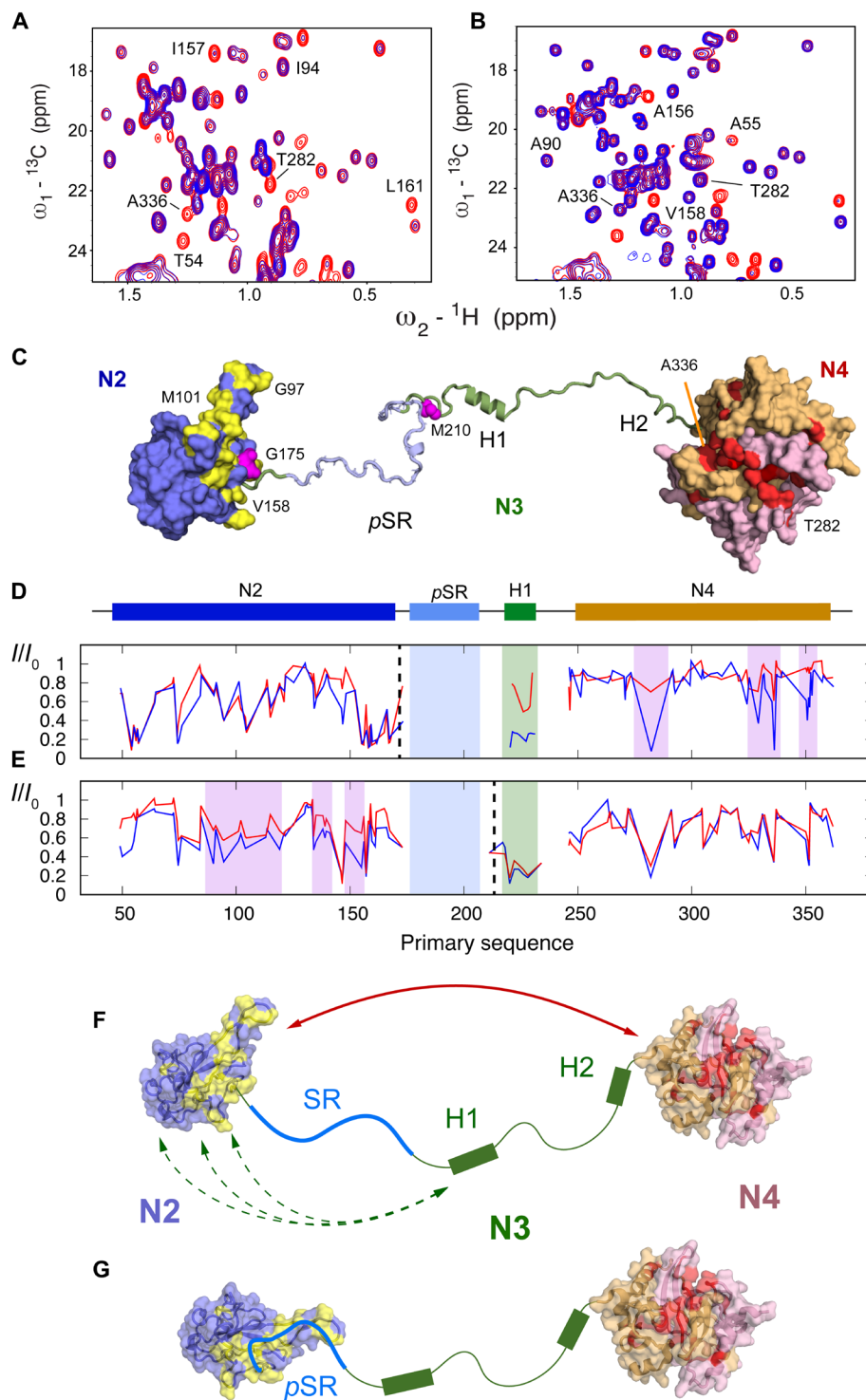
To develop further insight into the origin of CSPs observed upon phosphorylation, and to further investigate long-range interactions within N234, we have measured paramagnetic relaxation enhancements (PREs) of cysteine mutants of N234, labeled with 4-maleimido-TEMPO. Positions for cysteine mutation were selected on N2 (175) and N3 (210) (Fig. 6).

In nonphosphorylated N234, a continuous hydrophobic surface of N4 (Fig. 6), covering residues 333 to 340, 350, and 282, is broadened by the label positioned at residue 175, indicating that this region is involved in a long-range contact between the two folded domains. In addition, the linker region, in particular the peaks corresponding to the hydrophobic groups of H1, is strongly broadened. The suggestion that H1 interacts transiently with N2 is supported by data from the paramagnetic label at position 210, which induces broadening in N2, both in the RNA binding finger (residues 90 to 106) and in the C-terminal region (residues 145 to 170) and is supported by recent similar observations using PREs from neighboring sites in a shorter construct (30). Broadening is also observed on the hydrophobic patch on N4 due to the label at position 210.

$^{15}\text{N}$  chemical shifts of N2 were compared in the presence and absence of N1, N3, and N4. Small differences are observed between free N2 and the longer constructs (fig. S10), clustering mainly around the baseplate residues 60 to 70 and 160 to 170, located on a single continuous face of N2, suggesting that the weak interactions with N4 observed from PREs involve this surface.

To probe the impact of phosphorylation on long range order, two forms of the TEMPO-labeled protein (positions 175 and 210) were also hyperphosphorylated [pN234(III)] and PREs were again measured (Fig. 6). Note that the high quality of the PRE data is highlighted by the highly reproducible intra-domain PRE intensity ratios (between oxidized and reduced forms) within N2, for the TEMPO label positioned at residue 175. In terms of intradomain contacts, the broadening profile of the probe on N2 reveals systematic abrogation of the long-range contacts between N2 and the hydrophobic patch on N4. This surface participates in pSR interactions with N2, suggesting that the long-range contacts between N2 and N4 are removed due to the presence of bound pSR on the surface of N2. In addition, the enhanced broadening in H1 due to the label on N2 is weakened due to hyperphosphorylation, indicating that interactions of H1 with the surface of N2 are also abolished. This is supported by abrogation of PREs due to the 210 label that implicate the RNA





**Fig. 6. Modulation of paramagnetic relaxation enhancement due to hyperphosphorylation.** (A) Comparison of  $^{13}\text{C}$ - $^1\text{H}$  HMQC in oxidized (blue) and reduced (red) forms of TEMPO-labeled N234 (G175C). (B) Comparison of  $^{13}\text{C}$ - $^1\text{H}$  HMQC in oxidized (blue) and reduced (red) forms of TEMPO-labeled pN234(II) (G175C). Strong interdomain PREs highlighted in (A) (282 and 336) are considerably weakened due to phosphorylation. (C) Surface representation indicating selected effects on long-range order in N234. Red: residues on hydrophobic surface of N4 dimer (the two monomers are otherwise colored pink and beige) that broaden in the presence of TEMPO label (175), an interaction that is weakened upon phosphorylation. Yellow: residues on N2 that shift upon hyperphosphorylation. Sites of the TEMPO labels are shown in magenta. (D and E) Intensity ratios between paramagnetic and diamagnetic forms of the protein. Blue, free N234; red, pN234(II). Dashed lines indicate position of the TEMPO label (175 and 210). (F and G) In the free, nonphosphorylated form (A), long-range interactions between N2 and N4 and between H1 and N2 are revealed from PREs. Upon hyperphosphorylation, interactions are weakened (B) or completely suppressed. CSPs on N2 (yellow) [see (C)].

binding finger. The PRE of N4 due to 210 is not strongly affected by phosphorylation.

Small-angle x-ray scattering (SAXS) has also been measured on N234, pN234(III), and pN234(PKA). Although the SAXS curves show slight differences, in particular in the inverse distance ranges around 0.25 and 1.0 nm<sup>-1</sup>, indicating a possible redistribution of the adjacency of different domains of the protein, the overall dimensions are very similar (fig. S11). This is compatible with the observations presented above: Using PREs and CSPs, we identify contacts between the pSR region and N2 in the hyperphosphorylated protein, which would result in some compaction; however, at the same time, weak long-range contacts between N3/N4 and N2 are abrogated, resulting in larger distances between scattering centers. This combination appears to result in similar average radii of gyration, as estimated from analysis of the Guinier region of the SAXS curves. We also compare PKA-phosphorylated N234, whose scattering curve is again very similar to nonphosphorylated N234 (fig. S11).

The relevance of this abrogation of weak long-range contacts may be related to assembly or disassembly of nucleocapsids. In this respect, we note that we used negative staining electron microscopy to reveal the presence of cage-like particles upon addition of RNA (fig. S12). Such structural elements have been previously identified as substructures of the viral nucleocapsid (15, 16). The presence of these assemblies was not observed when RNA was added to hyperphosphorylated N234, confirming recent observations (47). In the absence of atomic resolution structures of assembled nucleocapsids, it is not yet clear whether this impact on the formation of larger assemblies is related to the effect of phosphorylation on long-range contacts in N.

## DISCUSSION

The SR region of SARS-CoV-2 nucleocapsid protein has been shown to be hyperphosphorylated in infected cells, and unphosphorylated during viral assembly and when bound to genomic RNA, and nucleocapsids present in the virion are not phosphorylated (54, 55). The level of phosphorylation of N is thought to act as a functional switch in the replication cycle, in genome unpacking (following infection) or packaging (before viral assembly) (14, 57–60). Phosphorylation was shown to affect compactness of RNP complexes (61), and binding of isolated N3 to long RNAs was also affected by phosphorylation (62). Fluorescence polarization also detected weaker binding to RNA as a result of hyperphosphorylation (47). A recent study proposed the sequential phosphorylation of N by host kinases SRPK1, GSK-3, and CK1 (60). Inhibition of SRPK1 was shown to reduce replication, while blocking the activity of GSK-3 reduced viral replication in cells and lowered infection in patients (71). Mass photometry and electron microscopy were recently combined to reveal that phosphorylated N forms less compact viral RNP particles (61).

Despite this intense interest, little is currently known about the molecular impact of hyperphosphorylation on the structural, dynamic, and functional behavior of N, and it is not yet clear where phosphorylation plays a role in the viral cycle of SARS-CoV-2. This study exploits the site-specific information provided by high-resolution NMR to describe the effects on local and long-range structure as a function of the proposed phosphorylation cycle, as well as the implications of sequential posttranslational modifications on molecular function, particularly RNA binding.

Our analysis substantiates the reported impact of phosphorylation by the previously identified putative triplet of host kinases of SRPK1, followed by GSK-3 and CK1 (59, 60). This confirms that the proposed sites are indeed phosphorylated (S206 and S188 by SRPK1, followed by S176, S180, S184, S186, S190, S194, T198, and S202 by GSK-3 and S193, S197, S201, and S205 by CK1). The order of phosphorylation by GSK-3 reveals that S180 is phosphorylated first, before near-simultaneous phosphorylation of S184, S190, S194, T198, and S202, followed by slower phosphorylation of S176 and S186. These observations are confirmed by monitoring the peaks reporting on the amide group of the directly phosphorylated amino acid, as well as of the immediately neighboring amino acids (fig. S13).

Phosphorylation affects the dynamic properties of N locally, in the SR hyperphosphorylation domain. The progressive increase of rotating frame relaxation rates (Fig. 2) as a function of increasing levels of phosphorylation may result from the rigidification of the local structure in pSR, owing to the increased presence of bulky phosphate groups, and/or from interaction with N2. The latter appears to be the case, with hyperphosphorylation inducing a very similar profile of CSPs compared to those observed due to RNA binding to N234 (Fig. 3). These changes are predominantly evident not only in N2 but also in the C terminus of N3, comprising a lysine-rich helical region (H2) that is formed in the presence of N4 and that folds into a helical conformation when N3 binds to nsp3 (49). It therefore appears that pSR binds via the same interface on N2 as that exploited by RNA. While phosphorylation by SRPK1 has no notable impact on the N:RNA interaction, phosphorylation by GSK-3 inhibits binding completely. Time-resolved NMR, at residue-specific and atomic resolution, allows us to identify which phosphorylation sites are essential for this inhibition, including most sites modified by GSK-3, with the exception of S176 and S186. The necessary phosphorylation pattern for inhibition of RNA binding is thereby revealed as a regularly spaced charge distribution with respect to the primary sequence: S180, S184, S188, S190, S194, T198, S202, and S206, underlining the probable role of electrostatics in this inhibition. These sites are all known to be highly conserved within the SR region.

Note that while a recent single-molecule characterization of the interaction of N123 with RNA suggested that N1 may play a regulatory role (32) in RNA binding, atomic resolution mapping reveals that the impact of N1 and N4 on the N2 RNA binding interface is negligible, at least for the RNAs used here, with essentially identical chemical shifts and apparent affinities observed in constructs comprising N123 and N234. This comparison also confirms N2 and N3 as the minimum necessary elements for phosphorylation-dependent inhibition of RNA binding.

Progressive phosphorylation of SR therefore appears to act as a conformational and functional switch for SARS-CoV-2 nucleocapsid protein, with time-resolved NMR providing crucial insight into the role of each phosphorylation site in the functional mechanism. It is tempting to suggest that recruitment of host kinases may thereby participate in releasing genomic RNA from encapsidation, upon unpacking from the infecting virion after cell entry. The observation that SRPK1 and GSK-3 are apparently sufficient to achieve this functional change also raises the question of the role of CK1 phosphorylation in viral function.

More insight into this putative interaction between pSR and the RNA binding interface of N2 is provided by comparison with the impact of *in vitro* phosphorylation by PKA. Under our

experimental conditions, this results in the phosphorylation of S180, S193, S194, S197, S201, and S202, generating a possible addition of 12 negative charges in the SR region, compared to the 16 that are required to abrogate RNA binding upon SRPK1/GSK-3 phosphorylation. Crucially, however, phosphorylation by PKA has no measurable impact on RNA binding, suggesting that the pattern resulting from the physiologically relevant kinases is indeed specific for inhibition, either by binding tighter to the surface of N2, by occupying more RNA binding sites on N2, or by ensuring a uniform electrostatic screening of the RNA binding surface. A recent study provided evidence that phosphorylation of N(1 to 209) with PKA resulted in chemical shifts on the RNA binding finger (90 to 106) of N2 (46). Our measurements confirm this observation. Differences in the chemical shifts induced by phosphorylation with PKA and SRPK1/GSK-3 suggest that the PKA phosphorylated SR does not interact with the baseplate regions (residues 48 to 66 and 148 to 174) as is the case for SRPK1/GSK-3 phosphorylated SR. Phosphorylated SR interacts with the RNA binding finger in both cases, although inducing much smaller CSPs in the case of PKA phosphorylation. This allows us to propose distinct interaction sites between the C- and N-terminal strands of pSR with the RNA binding finger and baseplate, respectively (Fig. 4). This occupation of additional RNA binding sites on N2 by SRPK1/GSK-3 phosphorylated N234 may explain why the pSR-N2 interaction necessary to abrogate RNA binding is effective in this case and not in the case of PKA phosphorylation.

Paramagnetic relaxation NMR was used to characterize the modulation of long-range order induced by phosphorylation. In the free, unmodified protein, helix H1 interacts transiently with the surface of N2, possibly mediated by hydrophobic interactions between the leucine-rich motif at the N terminus of H1 and hydrophobic sites on N2, and specifically involving the RNA binding finger (90 to 106) and the (145 to 170) region, also implicated in RNA binding. Helix H1 has previously been implicated in binding to N2 (30); indeed, we find some evidence of abrogation of transient contacts with N2 as a result of hyperphosphorylation, although the negligible effect of phosphorylation on transverse relaxation of H1 and the lack of CSP in N2 in the presence or absence of N3 suggest that such interactions must be weakly populated.

Stronger interactions are also observed between N2 and N4, specifically involving a hydrophobic patch on one side of the dimeric interface of N4, comprising residues 333 to 340, 350, and 282, and the baseplate of N2, suggesting a highly dynamic equilibrium between different states, as previously proposed (49). This equilibrium is modified upon hyperphosphorylation, probably because the N2 interaction surface is occupied by the pSR region. Clear differences can be identified, with notable weakening, even disappearance, of the interactions implicating the hydrophobic patch on N4. This abrogation of contacts within a single N dimer may correlate with the macroscopic observation of looser nucleocapsid structures derived from electron microscopy (61), or the reported modulation of the dynamics of membraneless organelles that can be formed upon mixing N and RNA under certain conditions (26, 35). We note that under our experimental conditions, cage-like structures are evident by negative staining electron microscopy upon mixing RNA with nonphosphorylated N234, but that such structures are absent for hyperphosphorylated N234/RNA mixtures. Nevertheless, the similarity of SAXS measured on nonphosphorylated and differently phosphorylated N234 suggests

that the binding of pSR to N2 compensates for the simultaneously abrogated weak long-range contacts between the three domains and that the overall dimensions, as measured by the average radius of gyration, do not change.

Phosphorylation of disordered peptides is known to facilitate rapid switching of biomolecular function of many physiological processes (72–74). Progressive phosphorylation of newly unpackaged N by host kinases likely maintains the protein in a form that is incapable of binding RNA, via an autoinhibitory strategy that protects from nonspecific binding to cytosolic host RNA, as well as premature binding to viral genomic RNA. The clear identification of a threshold of phosphorylation, achieved during GSK-3 phosphorylation, identifies specific molecular characteristics of this switching mechanism, a process that establishes phosphorylated N in an RNA-unbound form that may participate in additional functional activity such as regulation of transcription (14). Newly expressed N is apparently not phosphorylated before encapsidation, assembly, and packaging, suggesting that a strategy for inhibition of phosphorylation must be available. It has been suggested that replication occurs within membraneless organelles formed by LLPS of N (6, 26, 27, 33–36, 38, 65, 75, 76), and formation of organelles may provide some kind of protection from host kinases. Assuming encapsidation of newly synthesized RNA can occur within such organelles, as is the case for numerous negative strand RNA viruses (63, 64, 77–79), such a mechanism would provide an efficient means to escape from host kinase phosphorylation and facilitate efficient encapsidation of genomic RNA by unphosphorylated nucleocapsid protein.

In conclusion, we have used NMR spectroscopy to describe the impact of phosphorylation on the serine-arginine-rich region of the central disordered domain of SARS-CoV 2 nucleocapsid protein, a region that is hyperphosphorylated in infected cells. The proposed phosphorylation scheme, comprising SRPK1, GSK3, and CK1, is shown to abrogate binding of RNA. The phosphorylated SR domain interacts with the RNA binding domain of N via the same interface as single-stranded RNA, suggesting an auto-inhibitory mechanism involving direct competition with RNA. Long-range interactions between the folded domains of a single N protomer are also abrogated by phosphorylation, an observation potentially related to packaging or unpackaging of the viral genome. The observed impact of phosphorylation of SARS-CoV-2 N on RNA binding reports on a mechanism of inhibition of nonspecific binding of N to viral, or host RNA, while identifying a possible unpacking mechanism whereby host kinases can be recruited to extract RNA from the nucleocapsid after infection. The ability to determine the precise threshold of the highly specific phosphorylation pattern required to inhibit RNA binding provides crucial and unique insight into the role of post-translational modification in the viral cycle and identifies potential opportunities for inhibitory strategies for this class of viruses.

## MATERIALS AND METHODS

### Protein expression and purification

All proteins (N234 47 to 364, and N123 1 to 267) were expressed in *Escherichia coli* and purified as described previously (49). Single-point mutations of N234 (G175C and M210C) were made in-house by site-directed mutagenesis. 14-mer RNA (UCUAAACGAACUUU) and 30-mer polyA were purchased from Integrated DNA Technologies, San Diego.

## Phosphorylation of N123 and N234

N123, N234, and G175C and M210C mutations of N234 were incubated with SRPK1 (Sigma-Aldrich), GSK-3 (Promega), CK1 (Promega), or PKA (Promega) kinases in phosphorylation buffer: 50 mM Na-phosphate (pH 6.5), 250 mM NaCl, 2 mM DTT, supplemented with 10 mM ATP, and 10 mM MgCl<sub>2</sub> at 25°C. For phosphorylation cascade reaction, kinases (SRPK1, GSK-3, and CK1) were added to the sample sequentially or simultaneously.

## Nuclear magnetic resonance experiments

Unless stated otherwise, all NMR experiments were carried out in 50 mM Na-phosphate (pH 6.5), 250 mM NaCl, 2 mM DTT or 50 mM Na-phosphate (pH 6.5), 250 mM NaCl, and 1 mM TCEP at 25°C. Experiments were acquired on Bruker spectrometers with <sup>1</sup>H frequencies of 600, 700, 850, and 950 MHz. Spectra were processed with NMRPipe (80) and visualized using NMRFAM-Sparky (81).

<sup>15</sup>N R<sub>1ρ</sub> relaxation was measured at 298 K and a <sup>1</sup>H frequency of 850 MHz (N234 and pN234) and 950 MHz (N123 and pN123) using a spin lock of 1.5 kHz (82). Relaxation delays of 1, 10, 30, 70, 120, and 200 ms were used for <sup>15</sup>N R<sub>1ρ</sub> of N234, and 1, 10, 20, 50, 90, 140, 180, and 250 ms were used for <sup>15</sup>N R<sub>1ρ</sub> of N123 and included repetition of one delay. Relaxation rates were fitted using in-house software and errors estimated with noise-based Monte Carlo simulation.

## Resonance assignment

Backbone assignment of domains N1 to N4 in the nonphosphorylated forms of N123 and N234 has been recently published (17, 83). Assignment of the phosphorylated forms was performed using Band-selective Excitation Short-Transient (BEST)-type triple-resonance experiments (84). Side-chain assignment of N2 (25) and N4 from SARS CoV (41) has been published and served as references to complete the assignment of the methyl region of <sup>13</sup>C-<sup>1</sup>H HMQC of N234. <sup>13</sup>C<sup>α</sup> chemical shifts were compared to random coil values using the recently proposed random coil server for phosphorylated protein sequences (66). CSPs were calculated using the following expression (85).

$$\Delta\delta\omega = \sqrt{(\Delta\delta\omega_{N15})^2 + (6.5 * \Delta\delta\omega_{H1})^2} \quad (1)$$

## PRE experiments

PRE effects were measured from the peak intensity ratios by comparing <sup>15</sup>N-HSQC and HMQC 2D spectra recorded on samples labeled with TEMPO and samples reduced by the addition of 2 mM ascorbic acid. Briefly, purified cysteine mutants were reduced with 10 mM DTT at 4°C for 12 hours and dialyzed into 50 mM phosphate buffer, pH 7.0, containing 250 mM NaCl without DTT. Tenfold molar excess of 4-maleimido-TEMPO (Sigma-Aldrich) dissolved in dimethyl sulfoxide (DMSO) was added to a 100 μM stock of the reduced cysteine mutants. The reaction was incubated for 2 hours at room temperature and then left overnight at 4°C. To eliminate the excess of TEMPO, the samples were dialyzed into NMR buffer.

## Competition experiments

SRPK1-phosphorylated N234 in phosphorylation buffer was supplemented with 200% of 14-mer or 120% of polyA(30-mer) RNA. Kinases were added directly to the NMR tube, followed by interleaved, sequential <sup>15</sup>N and <sup>13</sup>C correlation spectra in the presence of GSK-3 kinase recorded using BEST-TROSY (transverse relaxation-optimized spectroscopy) (86) and SOFAST-HMQC (selective optimized flip-angle short-transient heteronuclear multiple quantum coherence) (87).

## Phosphorylation analysis

BEST-TROSY (86) and SOFAST-HMQC (87) phosphorylation peak intensity time-dependent variation  $I_i(t)$  was fitted to the following equation

$$I_i(t) = I_0(1 - e^{-k_i t}) + I_{i=0} \quad (2)$$

to extract the phosphorylation rate  $k_i$  for each site. Peak intensities were extracted using NMRFAM-Sparky (81).

## Electron microscopy

The particles used for negative-stain electron microscopy were obtained by adding 10 μM 14-mer RNA to 10 μM N234 nonphosphorylated or phosphorylated with the three kinases, in a final volume of 20 μl. Samples were incubated for 1 hour at room temperature, then applied to the clean side of carbon on mica (carbon/mica interface), and stained with 2% sodium silicotungstate (pH 7.0). Micrographs were taken with a T12 FEI microscope at 120 kV and a magnification of 30,000.

## Small angle x-ray scattering

Samples of nonphosphorylated, PKA-phosphorylated, and SRPK1-GSK3-CK1 [pN234(III)] were buffer exchanged into NMR buffer and diluted to 0.5, 1.0, and 2.0 mg/ml and measured in batch mode. SAXS measurements were recorded at the European Synchrotron Research Facility (ESRF) in Grenoble France on the BIOSAXS beamline BM29 (88). Data were analyzed using in-house software.

## Supplementary Materials

This PDF file includes:

Figs. S1 to S13

## REFERENCES AND NOTES

1. Y. Gao, L. Yan, Y. Huang, F. Liu, Y. Zhao, L. Cao, T. Wang, Q. Sun, Z. Ming, L. Zhang, J. Ge, L. Zheng, Y. Zhang, H. Wang, Y. Zhu, C. Zhu, T. Hu, T. Hua, B. Zhang, X. Yang, J. Li, H. Yang, Z. Liu, W. Xu, L. W. Guddat, Q. Wang, Z. Lou, Z. Rao, Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* **368**, 779–782 (2020).
2. H. S. Hillen, G. Kokic, L. Farnung, C. Dienemann, D. Tegunov, P. Cramer, Structure of replicating SARS-CoV-2 polymerase. *Nature* **584**, 154–156 (2020).
3. Q. Wang, J. Wu, H. Wang, Y. Gao, Q. Liu, A. Mu, W. Ji, L. Yan, Y. Zhu, C. Zhu, X. Fang, X. Yang, Y. Huang, H. Gao, F. Liu, J. Ge, Q. Sun, X. Yang, W. Xu, Z. Liu, H. Yang, Z. Lou, B. Jiang, L. W. Guddat, P. Gong, Z. Rao, Structural basis for RNA replication by the SARS-CoV-2 polymerase. *Cell* **182**, 417–428.e13 (2020).
4. G. Wolff, R. W. A. L. Limpens, J. C. Zevenhoven-Dobbe, U. Laugks, S. Zheng, A. W. M. de Jong, R. I. Koning, D. A. Agard, K. Grünwald, A. J. Koster, E. J. Snijder, M. Bárcena, A molecular pore spans the double membrane of the coronavirus replication organelle. *Science* **369**, 1395–1398 (2020).
5. B. W. Neuman, M. M. Angelini, M. J. Buchmeier, Does form meet function in the coronavirus replicative organelle? *Trends Microbiol.* **22**, 642–647 (2014).
6. E. J. Snijder, R. W. A. L. Limpens, A. H. de Wilde, A. W. M. de Jong, J. C. Zevenhoven-Dobbe, H. J. Maier, F. F. G. A. Faas, A. J. Koster, M. Bárcena, A unifying structural and functional model of the coronavirus replication organelle: Tracking down RNA synthesis. *PLoS Biol.* **18**, e3000715 (2020).
7. F. Almazán, C. Galán, L. Enjuanes, The nucleoprotein is required for efficient coronavirus genome replication. *J. Virol.* **78**, 12683–12688 (2004).
8. R. McBride, M. van Zyl, B. C. Fielding, The coronavirus nucleocapsid is a multifunctional protein. *Viruses* **6**, 2991–3018 (2014).
9. S. Stertz, M. Reichelt, M. Spiegel, T. Kuri, L. Martínez-Sobrudo, A. García-Sastre, F. Weber, G. Kochs, The intracellular sites of early replication and budding of SARS-coronavirus. *Virology* **361**, 304–315 (2007).
10. M. H. Verheije, M. C. Hagemeijer, M. Ulasli, F. Reggiori, P. J. M. Rottier, P. S. Masters, C. A. M. de Haan, The coronavirus nucleocapsid protein is dynamically associated with the replication-transcription complexes. *J. Virol.* **84**, 11575–11579 (2010).
11. P. V'kovski, M. Gerber, J. Kelly, S. Pfaender, N. Ebert, S. B. Lagache, C. Simillion, J. Portmann, H. Stalder, V. Gaschen, R. Bruggmann, M. H. Stoffel, M. Heller, R. Dijkman, V. Thiel,

- Determination of host proteins composing the microenvironment of coronavirus replicase complexes by proximity-labeling. *eLife* **8**, e42037 (2019).
12. K. R. Hurst, C. A. Koetzier, P. S. Masters, Characterization of a critical interaction between the coronavirus nucleocapsid protein and nonstructural protein 3 of the viral replicase-transcriptase complex. *J. Virol.* **87**, 9159–9172 (2013).
  13. Y. Cong, M. Ulasli, H. Schepers, M. Mauthe, P. V'kovski, F. Kriegenburg, V. Thiel, C. A. M. de Haan, F. Reggiori, Nucleocapsid protein recruitment to replication-transcription complexes plays a crucial role in coronaviral life cycle. *J. Virol.* **94**, e01925-19 (2020).
  14. C.-H. Wu, P.-J. Chen, S.-H. Yeh, Nucleocapsid phosphorylation and RNA helicase DDX1 recruitment enables coronavirus transition from discontinuous to continuous transcription. *Cell Host Microbe* **16**, 462–472 (2014).
  15. H. Yao, Y. Song, Y. Chen, N. Wu, J. Xu, C. Sun, J. Zhang, T. Weng, Z. Zhang, Z. Wu, L. Cheng, D. Shi, X. Lu, J. Lei, M. Crispin, Y. Shi, L. Li, S. Li, Molecular architecture of the SARS-CoV-2 virus. *Cell* **183**, 730–738.e13 (2020).
  16. S. Klein, M. Cortese, S. L. Winter, M. Wachsmuth-Melm, C. J. Neufeldt, B. Berikan, M. L. Stanifer, S. Boulant, R. Bartenschlager, P. Chlanda, SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. *Nat. Commun.* **11**, 5885 (2020).
  17. S. Guseva, L. M. Perez, A. Camacho-Zarco, L. M. Bessa, N. Salvi, A. Malki, D. Maurin, M. Blackledge, <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N backbone chemical shift assignments of the N-terminal and central intrinsically disordered domains of SARS-CoV-2 nucleoprotein. *Biomol. NMR Assign.* **15**, 255–260 (2021).
  18. M. Schiavina, L. Pontoriero, V. N. Uversky, I. C. Felli, R. Pierattelli, The highly flexible disordered regions of the SARS-CoV-2 nucleocapsid N protein within the 1–248 residue construct: Sequence-specific resonance assignments through NMR. *Biomol. NMR Assign.* **15**, 219–227 (2021).
  19. S. Kang, M. Yang, Z. Hong, L. Zhang, Z. Huang, X. Chen, S. He, Z. Zhou, Z. Zhou, Q. Chen, Y. Yan, C. Zhang, H. Shan, S. Chen, Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. *Acta Pharm. Sin. B* **10**, 1228–1238 (2020).
  20. Y. Peng, N. du, Y. Lei, S. Dorje, J. Qi, T. Luo, G. F. Gao, H. Song, Structures of the SARS-CoV-2 nucleocapsid and their perspectives for drug design. *EMBO J.* **39**, e105938 (2020).
  21. Q. Ye, A. M. V. West, S. Silletti, K. D. Corbett, Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. *Protein Sci.* **29**, 1890–1901 (2020).
  22. L. Zinzula, J. Basquin, S. Bohn, F. Beck, S. Klumpe, G. Pfeifer, I. Nagy, A. Bracher, F. U. Hartl, W. Baumeister, High-resolution structure and biophysical characterization of the nucleocapsid phosphoprotein dimerization domain from the Covid-19 severe acute respiratory syndrome coronavirus 2. *Biochem. Biophys. Res. Commun.* **538**, 54–62 (2021).
  23. S. Sarkar, B. Runge, R. W. Russell, K. T. Movellan, D. Calero, S. Zeinalilathori, C. M. Quinn, M. Lu, G. Calero, A. M. Gronenborn, T. Polenova, Atomic-resolution structure of SARS-CoV-2 nucleocapsid protein N-terminal domain. *J. Am. Chem. Soc.* **144**, 10543–10555 (2022).
  24. Q. Huang, L. Yu, A. M. Petros, A. Gunasekera, Z. Liu, N. Xu, P. Hajduk, J. Mack, S. W. Fesik, E. T. Olejniczak, Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. *Biochemistry* **43**, 6059–6063 (2004).
  25. D. C. Dinesh, D. Chalupska, J. Silhan, E. Koutna, R. Nencka, V. Veverka, E. Boura, Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein. *PLoS Pathog.* **16**, e1009100 (2020).
  26. A. Savastano, A. Ibáñez de Opakua, M. Rankovic, M. Zweckstetter, Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerase-containing condensates. *Nat. Commun.* **11**, 6041 (2020).
  27. C. Iserman, C. A. Roden, M. A. Boerneke, R. S. G. Sealfon, G. A. McLaughlin, I. Jungreis, E. J. Fritch, Y. J. Hou, J. Ekena, C. A. Weidmann, C. L. Theesfeld, M. Kellis, O. G. Troyanskaya, R. S. Baric, T. P. Sheahan, K. M. Weeks, A. S. Gladfelter, Genomic RNA elements drive phase separation of the SARS-CoV-2 nucleocapsid. *Mol. Cell* **80**, 1078–1091.e6 (2020).
  28. J. S. Redzic, E. Lee, A. Born, A. Issaian, M. A. Henen, P. J. Nichols, A. Blue, K. C. Hansen, A. D'Alessandro, B. Vögeli, E. Z. Eisenmesser, The inherent dynamics and interaction sites of the SARS-CoV-2 nucleocapsid N-terminal region. *J. Mol. Biol.* **433**, 167108 (2021).
  29. H. M. Forsythe, J. Rodriguez Galvan, Z. Yu, S. Pinckney, P. Reardon, R. B. Cooley, P. Zhu, A. D. Rolland, J. S. Prell, E. Barbar, Multivalent binding of the partially disordered SARS-CoV-2 nucleocapsid phosphoprotein dimer to RNA. *Biophys. J.* **120**, 2890–2901 (2021).
  30. L. Pontoriero, M. Schiavina, S. M. Korn, A. Schlundt, R. Pierattelli, I. C. Felli, NMR reveals specific tracts within the intrinsically disordered regions of the SARS-CoV-2 nucleocapsid protein involved in RNA encountering. *Biomolecules* **12**, 929 (2022).
  31. S. M. Korn, K. Dhamotharan, C. M. Jeffries, A. Schlundt, The preference signature of the SARS-CoV-2 nucleocapsid NTD for its 5'-genomic RNA elements. *Nat. Commun.* **14**, 3331 (2023).
  32. J. Cubuk, J. J. Alston, J. J. Incicco, A. S. Holehouse, K. B. Hall, M. D. Stuchell-Breton, A. Soranno, The disordered N-terminal tail of SARS-CoV-2 nucleocapsid protein forms a dynamic complex with RNA. *Nucleic Acids Res.* **52**, 2609–2624 (2024).
  33. H. Chen, Y. Cui, X. Han, W. Hu, M. Sun, Y. Zhang, P. H. Wang, G. Song, W. Chen, J. Lou, Liquid-liquid phase separation by SARS-CoV-2 nucleocapsid protein and RNA. *Cell Res.* **30**, 1143–1145 (2020).
  34. T. M. Perdikari, A. C. Murthy, V. H. Ryan, S. Watters, M. T. Naik, N. L. Fawzi, SARS-CoV-2 nucleocapsid protein phase-separates with RNA and with human hnRNPs. *EMBO J.* **39**, e106478 (2020).
  35. C. R. Carlson, J. B. Asfaha, C. M. Ghent, C. J. Howard, N. Hartooni, M. Safari, A. D. Frankel, D. O. Morgan, Phosphoregulation of phase separation by the SARS-CoV-2 N protein suggests a biophysical basis for its dual functions. *Mol. Cell* **80**, 1092–1103.e4 (2020).
  36. J. Cubuk, J. J. Alston, J. J. Incicco, S. Singh, M. D. Stuchell-Breton, M. D. Ward, M. I. Zimmerman, N. Vithani, D. Griffith, J. A. Wagoner, G. R. Bowman, K. B. Hall, A. Soranno, A. S. Holehouse, The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. *Nat. Commun.* **12**, 1936 (2021).
  37. S. Lu, Q. Ye, D. Singh, Y. Cao, J. K. Diedrich, J. R. Yates III, E. Villa, D. W. Cleveland, K. D. Corbett, The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. *Nat. Commun.* **12**, 502 (2021).
  38. A. Jack, L. S. Ferro, M. J. Trnka, E. Wehri, A. Nadgir, X. Nguyenla, D. Fox, K. Costa, S. Stanley, J. Schaletzky, A. Yildiz, SARS-CoV-2 nucleocapsid protein forms condensates with viral genomic RNA. *PLoS Biol.* **19**, e3001425 (2021).
  39. M. Zachrdla, A. Savastano, A. Ibáñez de Opakua, M. S. Cima-Omori, M. Zweckstetter, 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**, e4409 (2022).
  40. C.-Y. Chen, C. K. Chang, Y. W. Chang, S. C. Sue, H. I. Bai, L. Rieng, C. D. Hsiao, T. H. Huang, Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. *J. Mol. Biol.* **368**, 1075–1086 (2007).
  41. M. Takeda, C. K. Chang, T. Ikeya, P. Güntert, Y. H. Chang, Y. L. Hsu, T. H. Huang, M. Kainosho, Solution structure of the c-terminal dimerization domain of SARS coronavirus nucleocapsid protein solved by the SAIL-NMR method. *J. Mol. Biol.* **380**, 608–622 (2008).
  42. C.-K. Chang, Y. L. Hsu, Y. H. Chang, F. A. Chao, M. C. Wu, Y. S. Huang, C. K. Hu, T. H. Huang, Multiple nucleic acid binding sites and intrinsic disorder of severe acute respiratory syndrome coronavirus nucleocapsid protein: Implications for ribonucleocapsid protein packaging. *J. Virol.* **83**, 2255–2264 (2009).
  43. P.-K. Hsieh, S. C. Chang, C. C. Huang, T. T. Lee, C. W. Hsiao, Y. H. Kou, I. Y. Chen, C. K. Chang, T. H. Huang, M. F. Chang, Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent. *J. Virol.* **79**, 13848–13855 (2005).
  44. G. Padroni, M. Bikaki, M. Novakovic, A. C. Wolter, S. H. Rüdiger, A. D. Gossert, A. Leitner, F. H. T. Allain, A hybrid structure determination approach to investigate the druggability of the nucleocapsid protein of SARS-CoV-2. *Nucleic Acids Res.* **51**, 4555–4571 (2023).
  45. A. B. Estelle, H. M. Forsythe, Z. Yu, K. Hughes, B. Lasher, P. Allen, P. N. Reardon, D. A. Hendrix, E. J. Barbar, RNA structure and multiple weak interactions balance the interplay between RNA binding and phase separation of SARS-CoV-2 nucleocapsid. *PNAS Nexus* **2**, pgad333 (2023).
  46. E. Lee, J. S. Redzic, A. J. Saviola, X. Li, C. C. Ebmeier, T. G. Kutateladze, K. C. Hansen, R. Zhao, N. Ahn, N. N. Sluchanko, E. Eisenmesser, Molecular insight into the specific interactions of the SARS-coronavirus-2 nucleocapsid with RNA and host protein. *Protein Sci.* **32**, e4603 (2023).
  47. A. N. Adly, M. Bi, C. R. Carlson, A. M. Syed, A. Ciling, J. A. Doudna, Y. Cheng, D. O. Morgan, Assembly of SARS-CoV-2 ribonucleosomes by truncated N<sup>+</sup> variant of the nucleocapsid protein. *J. Biol. Chem.* **299**, 105362 (2023).
  48. H. Luo, J. Chen, K. Chen, X. Shen, H. Jiang, Carboxyl terminus of severe acute respiratory syndrome coronavirus nucleocapsid protein: Self-association analysis and nucleic acid binding characterization. *Biochemistry* **45**, 11827–11835 (2006).
  49. L. M. Bessa, S. Guseva, A. R. Camacho-Zarco, N. Salvi, D. Maurin, L. M. Perez, M. Botova, A. Malki, M. Nanao, M. R. Jensen, R. W. H. Ruigrok, M. Blackledge, The intrinsically disordered SARS-CoV-2 nucleoprotein in dynamic complex with its viral partner nsp3a. *Sci. Adv.* **8**, eabm4034 (2022).
  50. B. A. Johnson, Y. Zhou, K. G. Lokugamage, M. N. Vu, N. Bopp, P. A. Crocquet-Valdes, B. Kalveram, C. Schindewolf, Y. Liu, D. Scharton, J. A. Plante, X. Xie, P. Aguilar, S. C. Weaver, P. Y. Shi, D. H. Walker, A. L. Routh, K. S. Plante, V. D. Menachery, Nucleocapsid mutations in SARS-CoV-2 augment replication and pathogenesis. *PLoS Pathog.* **18**, e1010627 (2022).
  51. J. D. Bloom, A. C. Beichman, R. A. Neher, K. Harris, Evolution of the SARS-CoV-2 mutational spectrum. *Mol. Biol. Evol.* **40**, msad085 (2023).
  52. A. M. Syed, T. Y. Taha, T. Tabata, I. P. Chen, A. Ciling, M. M. Khalid, B. Sreekumar, P. Y. Chen, J. M. Hayashi, K. M. Soczek, M. Ott, J. A. Doudna, Rapid assessment of SARS-CoV-2-evolved variants using virus-like particles. *Science* **374**, 1626–1632 (2021).
  53. H. Zhao, D. Wu, S. A. Hassan, A. Nguyen, J. Chen, G. Piszczek, P. Schuck, A conserved oligomerization domain in the disordered linker of coronavirus nucleocapsid proteins. *Sci. Adv.* **9**, eadg6473 (2023).

54. P. V'kovski, A. Kratzel, S. Steiner, H. Stalder, V. Thiel, Coronavirus biology and replication: Implications for SARS-CoV-2. *Nat. Rev. Microbiol.* **19**, 155–170 (2021).
55. K. M. Scherer, L. Mascheroni, G. W. Carnell, L. C. S. Wunderlich, S. Makarchuk, M. Brockhoff, I. Mela, A. Fernandez-Villegas, M. Barysevich, H. Stewart, M. Suau Sans, C. L. George, J. R. Lamb, G. S. Kaminski-Schierle, J. L. Heeney, C. F. Kaminski, SARS-CoV-2 nucleocapsid protein adheres to replication organelles before viral assembly at the Golgi/ERGIC and lysosome-mediated egress. *Sci. Adv.* **8**, eabl4895 (2022).
56. Y. Ino, M. Nishi, Y. Yamaoka, K. Miyakawa, S. S. Jeremiah, M. Osada, Y. Kimura, A. Ryo, Phosphopeptide enrichment using Phos-tag technology reveals functional phosphorylation of the nucleocapsid protein of SARS-CoV-2. *J. Proteomics* **255**, 104501 (2022).
57. T.-Y. Peng, K.-R. Lee, W.-Y. Tarn, 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. *FEBS J.* **275**, 4152–4163 (2008).
58. C.-H. Wu, S. H. Yeh, Y. G. Tsay, Y. H. Shieh, C. L. Kao, Y. S. Chen, S. H. Wang, T. J. Kuo, D. S. Chen, P. J. Chen, Glycogen synthase kinase-3 regulates the phosphorylation of severe acute respiratory syndrome coronavirus nucleocapsid protein and viral replication. *J. Biol. Chem.* **284**, 5229–5239 (2009).
59. M. Bouhaddou, D. Memon, B. Meyer, K. M. White, V. V. Rezelj, M. Correa Marrero, B. J. Polacco, J. E. Melnyk, S. Ulferts, R. M. Kaake, J. Batra, A. L. Richards, E. Stevenson, D. E. Gordon, A. Rojic, K. Obernier, J. M. Fabius, M. Soucheray, L. Miorin, E. Moreno, C. Koh, Q. D. Tran, A. Hardy, R. Robinot, T. Vallet, B. E. Nilsson-Payant, C. Hernandez-Armenta, A. Dunham, S. Weigang, J. Knerr, M. Modak, D. Quintero, Y. Zhou, A. Dugourd, A. Valdeolivas, T. Patil, Q. Li, R. Hüttenhain, M. Cakir, M. Muralidharan, M. Kim, G. Jang, B. Tutuncuoglu, J. Hiatt, J. Z. Guo, J. Xu, S. Bouhaddou, C. J. P. Mathy, A. Gaulton, E. J. Manners, E. Félix, Y. Shi, M. Goff, J. K. Lim, T. McBride, M. C. O'Neal, Y. Cai, J. C. J. Chang, D. J. Broadhurst, S. Klippsten, E. de Wit, A. R. Leach, T. Kortemme, B. Shoichet, M. Ott, J. Saez-Rodriguez, B. R. tenOever, R. D. Mullins, E. R. Fischer, G. Kochs, R. Grosse, A. García-Sastre, M. Vignuzzi, J. R. Johnson, K. M. Shokat, D. L. Swaney, P. Beltrao, N. J. Krogan, The global phosphorylation landscape of SARS-CoV-2 infection. *Cell* **182**, 685–712.e19 (2020).
60. T. M. Yaron, B. E. Heaton, T. M. Levy, J. L. Johnson, T. X. Jordan, B. M. Cohen, A. Kerelsky, T. Y. Lin, K. M. Liberatore, D. K. Bulaon, S. H. van Nest, N. Koundouros, E. R. Kastenhuber, M. N. Mercadante, K. Shobana-Ganesh, L. J. E. Schwartz, S. Chen, H. Weinstein, O. Elemento, E. Piskounova, B. E. Nilsson-Payant, G. Lee, J. D. Trimarco, K. N. Burke, C. E. Hamele, R. R. Chaparian, A. T. Harding, A. Tata, X. Zhu, P. R. Tata, C. M. Smith, A. P. Possemato, S. L. Tkachev, P. V. Hornbeck, S. A. Beausoleil, S. K. Anand, F. Aguet, G. Getz, A. D. Davidson, K. Heesom, M. Kavanagh-Williamson, D. A. Matthews, B. R. tenOever, L. C. Cantley, J. Blenis, N. S. Heaton, Host protein kinases required for SARS-CoV-2 nucleocapsid phosphorylation and viral replication. *Sci. Signal.* **15**, eabm0808 (2022).
61. C. R. Carlson, A. N. Adly, M. Bi, C. J. Howard, A. Frost, Y. Cheng, D. O. Morgan, Reconstitution of the SARS-CoV-2 ribonucleosome provides insights into genomic RNA packaging and regulation by phosphorylation. *J. Biol. Chem.* **298**, 102560 (2022).
62. P. N. Reardon, H. Stuwe, S. Shah, Z. Yu, K. Hughes, E. J. Barbar, Modifications to the SR-rich region of the SARS-CoV-2 nucleocapsid regulate self-association and attenuate RNA interactions. *bioRxiv* 542392 [Preprint] (2023). <https://doi.org/10.1101/2023.05.26.542392>.
63. J. Nikolic, R. le Bars, Z. Lama, N. Scrima, C. Lagaudrière-Gesbert, Y. Gaudin, D. Blondel, Negri bodies are viral factories with properties of liquid organelles. *Nat. Commun.* **8**, 58 (2017).
64. S. Guseva, S. Milles, M. R. Jensen, N. Salvi, J.-P. Klemm, D. Maurin, R. W. H. Ruigrok, M. Blackledge, Measles virus nucleocapsid and phosphoproteins form liquid-like phase-separated compartments that promote nucleocapsid assembly. *Sci. Adv.* **6**, eaaz7095 (2020).
65. S. M. Cascarina, E. D. Ross, Phase separation by the SARS-CoV-2 nucleocapsid protein: Consensus and open questions. *J. Biol. Chem.* **298**, 101677 (2022).
66. R. Hendus-Altenburger, C. B. Fernandes, K. Bugge, M. B. A. Kunze, W. Boomsma, B. B. Kragelund, Random coil chemical shifts for serine, threonine and tyrosine phosphorylation over a broad pH range. *J. Biomol. NMR* **73**, 713–725 (2019).
67. A. C. Conibear, K. J. Rosengren, C. F. W. Becker, H. Kaehlig, Random coil shifts of posttranslationally modified amino acids. *J. Biomol. NMR* **73**, 587–599 (2019).
68. F.-X. Theillet, H. M. Rose, S. Liokatis, A. Binolfi, R. Thongwichian, M. Stuver, P. Selenko, Site-specific NMR mapping and time-resolved monitoring of serine and threonine phosphorylation in reconstituted kinase reactions and mammalian cell extracts. *Nat. Protoc.* **8**, 1416–1432 (2013).
69. S. Milles, N. Salvi, M. Blackledge, M. R. Jensen, Characterization of intrinsically disordered proteins and their dynamic complexes: From in vitro to cell-like environments. *Prog. Nucl. Magn. Reson. Spectrosc.* **109**, 79–100 (2018).
70. K. V. Tugayeva, D. E. D. P. Hawkins, J. L. R. Smith, O. W. Bayfield, D. S. Ker, A. A. Sysoev, O. I. Klychnikov, A. A. Antson, N. N. Sluchanko, The mechanism of SARS-CoV-2 nucleocapsid protein recognition by the human 14-3-3 proteins. *J. Mol. Biol.* **433**, 166875 (2021).
71. X. Liu, A. Verma, G. Garcia Jr., H. Ramage, A. Lucas, R. L. Myers, J. J. Michaelson, W. Coryell, A. Kumar, A. W. Charney, M. G. Kazanietz, D. J. Rader, M. D. Ritchie, W. H. Berrettini, D. C. Schultz, S. Cherry, R. Damoiseaux, V. Arumugaswami, P. S. Klein, Targeting the coronavirus nucleocapsid protein through GSK-3 inhibition. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2113401118 (2021).
72. M. A. Pufall, G. M. Lee, M. L. Nelson, H. S. Kang, A. Velyvis, L. E. Kay, L. P. McIntosh, B. J. Graves, Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region. *Science* **309**, 142–145 (2005).
73. F.-X. Theillet, A. Binolfi, T. Frembgen-Kesner, K. Hingorani, M. Sarkar, C. Kyne, C. Li, P. B. Crowley, L. Gierasch, G. J. Pielak, A. H. Elcock, A. Gershenson, P. Selenko, Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs). *Chem. Rev.* **114**, 6661–6714 (2014).
74. A. Bah, R. M. Vernon, Z. Siddiqui, M. Krzeminski, R. Muhandiram, C. Zhao, N. Sonenberg, L. E. Kay, J. D. Forman-Kay, Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **519**, 106–109 (2015).
75. C. A. Roden, Y. Dai, C. A. Giannetti, I. Seim, M. Lee, R. Sealfon, G. A. McLaughlin, M. A. Boerneke, C. Iserman, S. A. Wey, J. L. Ekena, O. G. Troyanskaya, K. M. Weeks, L. You, A. Chilkoti, A. S. Gladfelter, Double-stranded RNA drives SARS-CoV-2 nucleocapsid protein to undergo phase separation at specific temperatures. *Nucleic Acids Res.* **50**, 8168–8192 (2022).
76. B.-A. Chau, V. Chen, A. W. Cochrane, L. J. Parent, A. J. Mouland, Liquid-liquid phase separation of nucleocapsid proteins during SARS-CoV-2 and HIV-1 replication. *Cell Rep.* **42**, 111968 (2023).
77. B. S. Heinrich, Z. Maliga, D. A. Stein, A. A. Hyman, S. P. J. Whelan, Phase transitions drive the formation of vesicular stomatitis virus replication compartments. *mBio* **9**, e02290-17 (2018).
78. M. Ringel, A. Heiner, L. Behner, S. Halwe, L. Sauerhering, N. Becker, E. Dietzel, B. Sawatsky, L. Kolesnikova, A. Maisner, Nipah virus induces two inclusion body populations: Identification of novel inclusions at the plasma membrane. *PLoS Pathog.* **15**, e1007733 (2019).
79. Y. Zhou, J. M. Su, C. E. Samuel, D. Ma, Measles virus forms inclusion bodies with properties of liquid organelles. *J. Virol.* **93**, e00948-19 (2019).
80. F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, Nmrpipe—A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
81. W. Lee, M. Tonelli, J. L. Markley, NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).
82. N.-A. Lakomek, J. Ying, A. Bax, Measurement of <sup>15</sup>N relaxation rates in perdeuterated proteins by TROSY-based methods. *J. Biomol. NMR* **53**, 209–221 (2012).
83. S. M. Korn, R. Lambert, B. Fürtig, M. Hengesbach, F. Löh, C. Richter, H. Schwalbe, J. E. Weigand, J. Wöhnert, A. Schlundt, <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N backbone chemical shift assignments of the C-terminal dimerization domain of SARS-CoV-2 nucleocapsid protein. *Biomol. NMR Assign.* **15**, 129–135 (2021).
84. E. Lescop, P. Schanda, B. Brutscher, A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J. Magn. Reson.* **187**, 163–169 (2007).
85. A. Ayed, F. A. A. Mulder, G. S. Yi, Y. Lu, L. E. Kay, C. H. Arrowsmith, Latent and active p53 are identical in conformation. *Nat. Struct. Mol. Biol.* **8**, 756–760 (2001).
86. Z. Solyom, M. Schwarten, L. Geist, R. Konrat, D. Willbold, B. Brutscher, BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *J. Biomol. NMR* **55**, 311–321 (2013).
87. P. Schanda, B. Brutscher, Very fast two-dimensional nmr spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J. Am. Chem. Soc.* **127**, 8014–8015 (2005).
88. P. Pernot, A. Round, R. Barrett, A. de Maria Antolinos, A. Gobbo, E. Gordon, J. Huet, J. Kieffer, M. Lentini, M. Mattenet, C. Morawe, C. Mueller-Dieckmann, S. Ohlsson, W. Schmid, J. Surr, P. Theveneau, L. Zerrad, S. McSweeney, Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **20**, 660–664 (2013).

**Acknowledgments:** IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA). This work used the BM29 Bio-SAXS beamline at the ESRF with the aid of S. Hutin. **Funding:** This work was supported by the European Research Council Advanced Grant DynamicAssemblies under the European Union's Horizon 2020 research and innovation program (grant agreement number 835161) (M.B.I.) and ANR-19-CE15-0024. This work used the platforms of the Grenoble Instruct-ERIC center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-0005-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). A.R.C.-Z. received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 796490 and postdoctoral HFSP fellowship

LT001544/2017. **Author contributions:** M.Bo., A.R.C.-Z., N.S., E.M. T.H., L.M.B. and M.Bl. conceived the project. M.Bo., A.R.C.-Z., L.M.B., S.G., E.M., N.S., and J.T. ran NMR experiments. D.M., M.Bo., A.R.C.-Z., S.G., and L.M.B. prepared protein samples. M.Bo., A.R.C.-Z., L.M.B., S.G., J.T., T.H., and M.Bl. analyzed spectra, validated analysis, and curated data. T.H. and M.Bl. oversaw software utilization. E.M. ran and analyzed electron microscopy experiments. M.Bl. supervised the study and administered the project. All authors contributed to writing the article.

**Competing interests:** The authors declare that they have no competing interests.

**Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 31 January 2024

Accepted 27 June 2024

Published 2 August 2024

10.1126/sciadv.aax2323