Poster Reprint

**ASMS 2023**
**Poster number MP 674**

# Revealing non-canonical phosphorylation of the SARS-CoV-2 nucleocapsid using top-down electron capture dissociation and sequence tag generation

Rachel Franklin[1,2], Phillip Zhu[2], Adrian Guthals[1], Joseph Meeuwsen[1], Steve Madden[1], Michael C. Hare[1], Richard Cooley[2], Joseph S. Beckman[2]

[1]- e-MSion - A part of Agilent Technologies, Corvallis OR 97322

[2]- Oregon State University, Corvallis OR 97322

## Introduction

### Host kinases facilitate SARS-CoV-2 infection

Several host kinases phosphorylate the viral nucleocapsid protein at a central serine/arginine (SR)-rich domain. [1]

Phosphorylations in positions Ser188 and Ser206 are primers which activate downstream hyper-phosphorylation by GSK-3β, which is critical for viral replication.[1] (Figure 1)

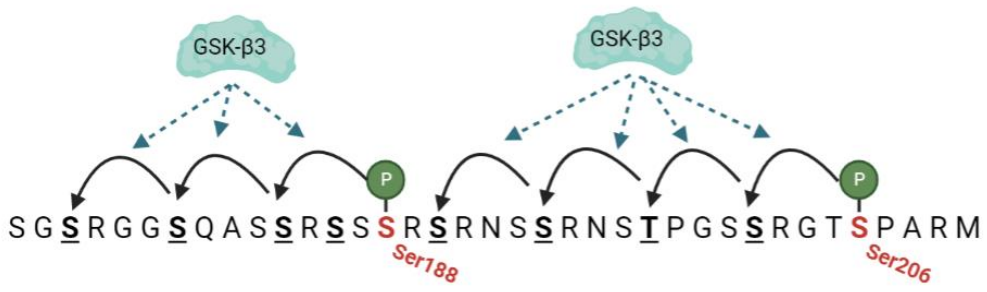


Figure 1- Primer-specific GSK-3β phosphorylation patterns in the nucleocapsid SR-rich domain. Priming sites are in red. Predicted GSK-3β phosphorylation sites are underlined and in bold.

Traditional bottom-up mass spectrometry techniques fail to accurately characterize hyper-phosphorylated proteins. SR-rich proteins are a worst-case scenario for bottom-up MS because of their repetitive sequences and high number of potential phosphorylation sites.

Our previous work utilized top-down electron capture dissociation mass spectrometry (ECD-MS) to characterize primer-specific GSK-3β phosphorylation pathways in the SR-rich domain.

### Primer independent phosphorylation by GSK-3β

Here, we identify a non-canonical GSK-3β phosphorylation pathway in the SR-rich domain which does not require an activating primer.

We use deconvolution of top-down MS/MS data to identify sequence tags, therefore, avoiding the use of data-base searches to score the combinatorial explosion of possible phospho-site combinations in the SR-rich domain.

## Experimental

### Sample Preparation

An 8 kDa construct containing the nucleocapsid SR-rich domain was reacted with GSK-3β. Phosphoserine was genetically encoded into position Ser206 to directly test primer-specific phosphorylation patterns. GSK-3β was also reacted with the unmodified SR-rich domain which lacks activating primer sites.

### Top-down Mass Spectrometry

Top-down experiments were performed using an Agilent 6545XT AdvanceBio LC/ Q-TOF equipped with an ExD cell for electron fragmentation. (Figure 2)

All samples were introduced via direct infusion. The most abundant charge state of each proteoform was isolated for electron capture dissociation (ECD).

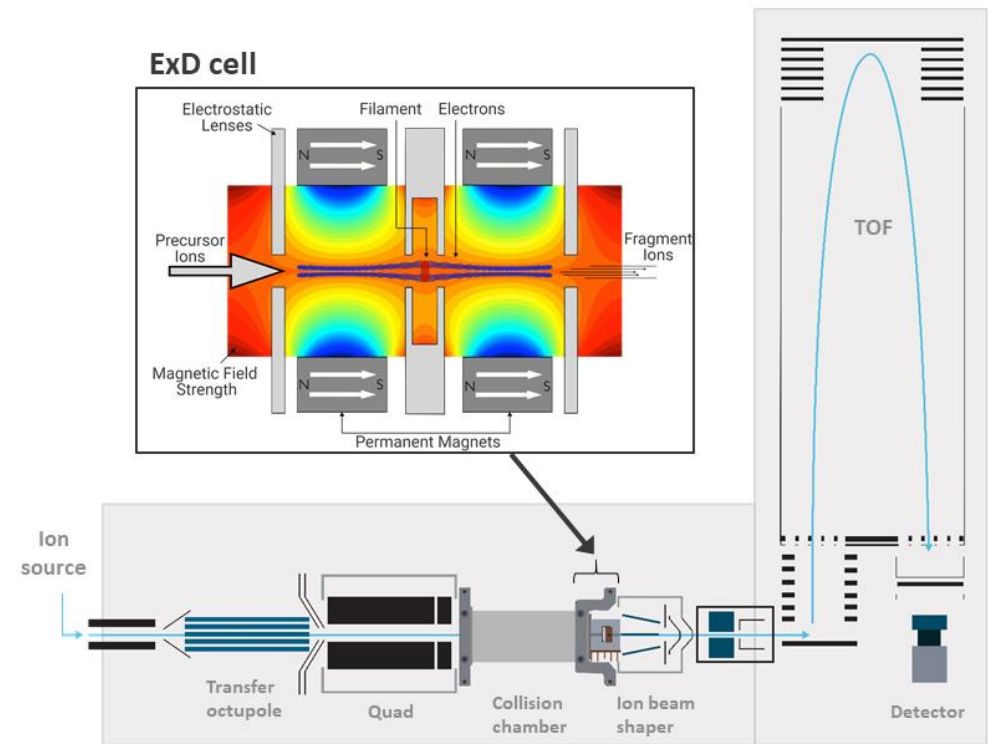


Figure 2- A diagram of an Agilent 6545XT AdvanceBio LC/Q-TOF equipped with an ExD cell

### ExDViewer – An approachable and interactive software for top-down data analysis

BioConfirm was used to determine intact masses. De-charged spectra were then re-plotted using MATLAB.

ExDViewer was used for identification of fragment ions and to deconvolute raw MS/MS profile data which avoids tedious file conversions and minimizes pre-processing steps such as peak-picking and the application of intensity thresholds that lose potentially valuable fragment information.

## Top-down analysis reveals proteoforms of the nucleocapsid GSK-3β phosphorylation cascade

Genetically encoded phosphoserine primers were used to interrogate primer specific GSK-3β phosphorylation patterns. [2]

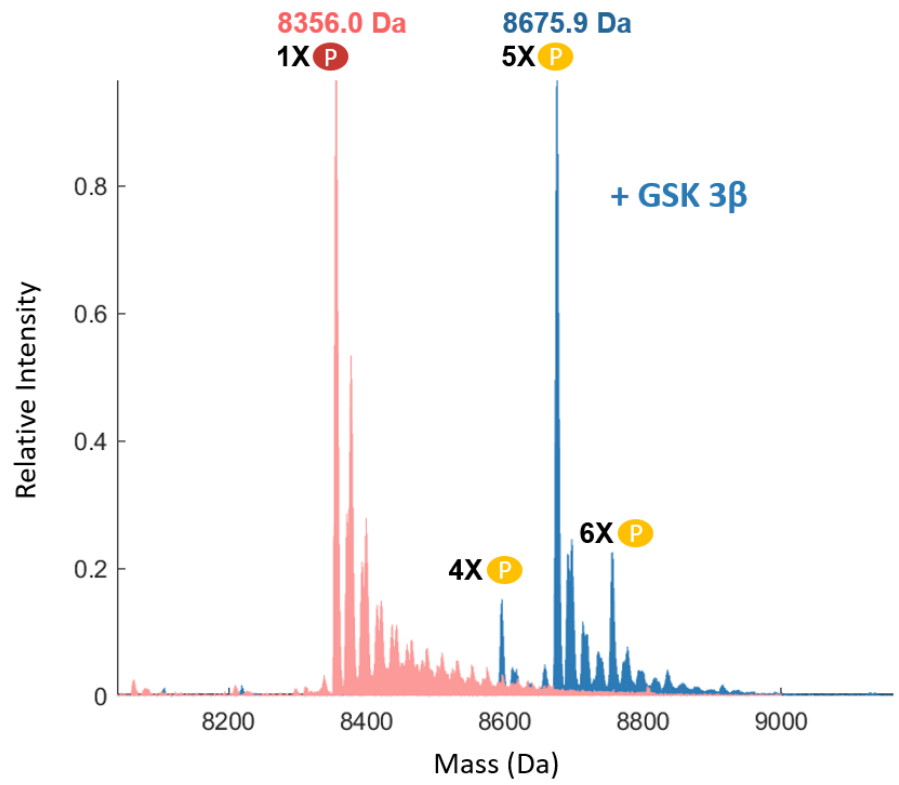


Figure 3- Intact mass spectra of the Ser206 primed SR-rich domain before (pink) and after (blue) reacting with GSK-3β.

It is predicted that GSK-3β adds 5 phospho-groups in an n+4 pattern in response to the Ser206 primer. However, top-down analysis reveals that GSK-3β mostly adds 4 phosphorylation groups in response to the Ser206 primer. (Figure 3)

Isolating the 5X phosphorylated protein confirmed the n+4 directionality and that Ser186 (a predicted site) remains unmodified in the most abundant proteoform.

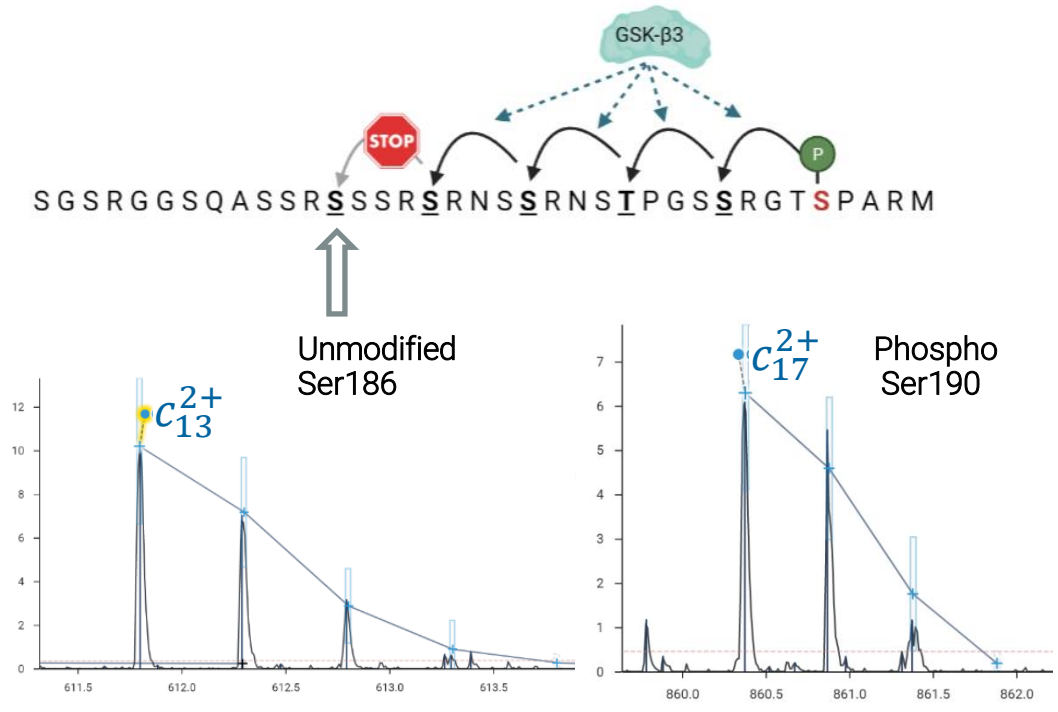


Figure 4- Site-specific c-ions which correspond to a phosphoserine 190 and unmodified serine 186.

## Discovering non-canonical GSK-3β phosphorylation in the SR-rich domain

In the absence of an activating phospho-priming site, GSK-3β is not predicted to modify the nucleocapsid SR-rich region.

However, treating the non-primed SR-rich domain with GSK-3β, results in a small peak corresponding to a triply phosphorylated protein (Figure 5).

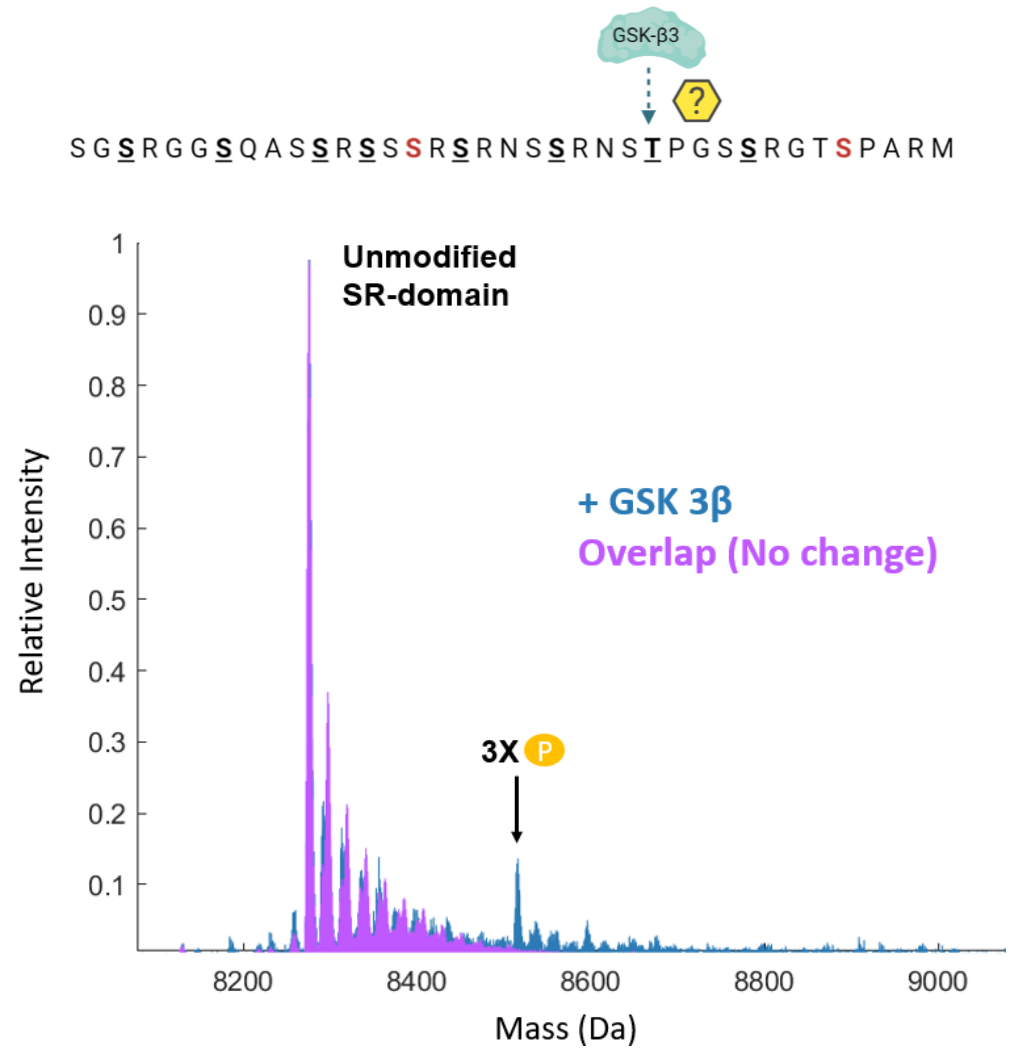


Figure 5- Intact mass spectra of the non-primed SR-rich domain before and after reacting with GSK-3β. Regions in purple are masses that did not change after the GSK-3β reaction.

The triply phosphorylated SR-rich region has 680 different possible combinations of phospho-sites. To avoid searching for every possible combination, deconvoluted MS/MS spectra were used to assign sequence ladders based on amino acid masses.

MS/MS deconvolution enabled the straight-forward identification Ser190 phosphorylation evidenced by the observation of a +166 Da mass shift in the c-ion series which corresponds to the mass of a phosphoserine. (Figure 6).

## Results and Discussion

### Sequence tags generated using ExDViewer enables the identification of phosphorylation sites in the non-canonical GSK-3β pathway

...S G **S** R G G **S** Q A S **S** R **S** S S R **S** R N S **S** R N S **T** P G...

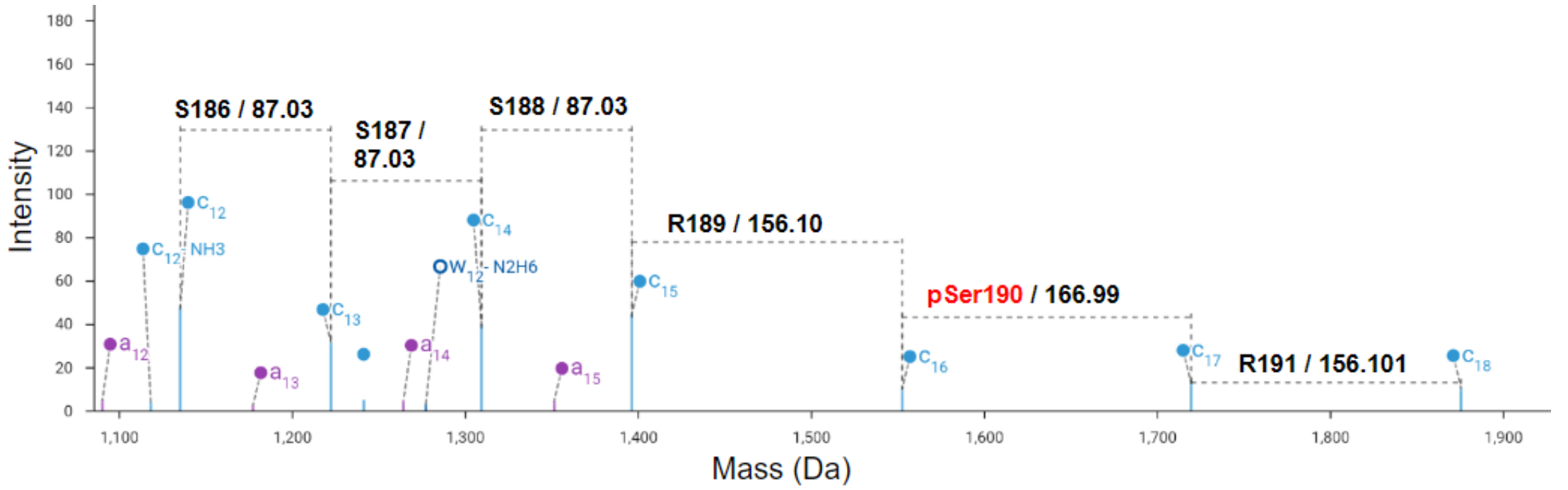


Figure 6- A deconvoluted MS2 spectrum showing c-ion sequence ladders that were used to identify an unknown phosphosite in the SR-rich region. The mass shift corresponding to phosphoserine 190 is highlighted in red.

### Primer independent phosphorylation may activate additional phosphorylation by CPK-1

Phosphorylation of Ser190 is known to prime another sequential kinase, CPK-1, which finishes the phosphorylation cascade after the canonical GSK-3β pathway. [1]

Primer-independent phosphorylation by GSK-3β could play a unique role in the viral life cycle by providing an alternate mechanism for CPK-1 activation in the SR-rich domain that is independent of upstream signaling events.

### Future work will test how mutations in the SR-rich domain influence GSK-3β phosphorylation

Mutations within the SR-rich region have monumental influence over viral fitness and are present in nearly all SARS-CoV-2 variants of concern.

Future planned experiments will test how the N protein P199L mutation effects the GSK-3β phosphorylation cascade.

## Conclusions

- ECD-MS reveals proteoform specific information from sequential kinase reactions
- Top-down analysis identified discrepancies in the experimental vs. predicted GSK-3β mechanisms, including a primer-independent pathway in the SR-rich domain
- The creation of sequence tags from ECD-MS data is a powerful tool for the analysis of heavily phosphorylated SR-rich proteins

## References

[1]Tomer M. Yaron et al. ,Host protein kinases required for SARS-CoV-2 nucleocapsid phosphorylation and viral replication.Sci.Signal.(2022).DOI:10.1126/scisignal.abm0808

[2]Phillip Zhu et al. ,Autonomous Synthesis of Functional, Permanently Phosphorylated Proteins for Defining the Interactome of Monomeric 14-3-3ζ. ACS Cent. Sci. 2023 9 (4), 816-835. DOI: 10.1021/acscentsci.3c00191

https://www.agilent.com/en/promotions/asms



RA45058.4471990741



Published in USA, May 31,2023

Agilent

Trusted Answers