analytical chemistry

pubs.acs.org/ac Letter

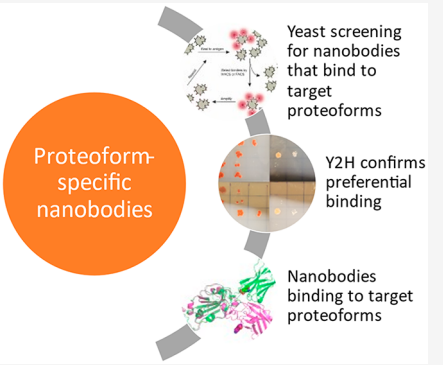
# Shaping Nanobodies and Intrabodies against Proteoforms

Bojana Leonard, Vincent Danna, Leo Gorham, Michelle Davison, William Chrisler, Doo Nam Kim, and Vincent R. Gerbasi*

Cite This: https://doi.org/10.1021/acs.analchem.3c00958 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** Proteoforms expand genomic diversity and direct developmental processes. While high-resolution mass spectrometry has accelerated characterization of proteoforms, molecular techniques working to bind and disrupt the function of specific proteoforms have lagged behind. In this study, we worked to develop intrabodies capable of binding specific proteoforms. We employed a synthetic camelid nanobody library expressed in yeast to identify nanobody binders of different SARS-CoV-2 receptor binding domain (RBD) proteoforms. Importantly, employment of the positive and negative selection mechanisms inherent to the synthetic system allowed for amplification of nanobody-expressing yeast that bind to the original (Wuhan strain RBD) but not the E484 K (Beta variant) mutation. Nanobodies raised against specific RBD proteoforms were validated by yeast-2-hybrid analysis and sequence comparisons. These results provide a framework for development of nanobodies and intrabodies that target proteoforms.






## INTRODUCTION

Proteoforms are individual modified forms of proteins to include genetic variants, truncations, and post-translational modifications. As discussed previously by Smith and Kelleher, proteoforms greatly expand the potential functional diversity of the genome.[1] Additionally, recent studies suggest that proteoforms play key roles in the structure of oncoproteins,[2,3] hematopoiesis,[4] and transplant rejection.[5]

Characterization and detection of proteoforms has expanded in recent years, with mass spectrometry breaking molecular weight barriers for proteoform analysis and diving deep into the human proteome[6−11].[12] While these studies are critical for characterizing the proteome, we lack a set of molecular tools to study the function of individual proteoforms. This limitation is due to the nature of recombinant DNA techniques, which largely seek to delete or modify genes at the protein level. While these techniques have accelerated our understanding of gene function, they lack the precision required to ascribe function to individual proteoforms.

Toward the ultimate goal of developing reagents that can bind to, and potentially disrupt specific proteoforms, we tested the ability of a synthetic yeast surface display nanobody library to distinguish different SARS-CoV-2 Spike protein receptor-binding-domain proteoforms. Our results show that after two steps of selection, including yeast surface display followed by yeast-2-hybrid screening, distinct nanobody (Nb) populations are identified that can distinguish proteoforms that vary at only a single amino acid position. With modifications to this experimental design, we propose that this screening process could result in the rational design of nanobodies against transient, post-translationally modified forms of proteins. Importantly, these nanobodies could be used to isolate specific proteoform populations and determine the contribution of specific proteoforms to cell and organismal phenotypes.

## RESULTS AND DISCUSSION

We sought to characterize proteoforms that varied at one amino acid position. The SARS-CoV-2 receptor binding-domain proteoforms are >90% pure, readily available, and have been characterized in multiple studies.[13−15] RBDs expressed and purified from mammalian cells display a repertoire of proteoforms, many of which contain glycosylated residues, variable processing of the N-terminus, and other unassigned modifications.[13−15] We analyzed RBD proteoforms with single amino acid substitutions at position 484 of the Spike protein. The original (Wuhan) strain of the RBD contained a glutamic acid at position 484 while the Beta version of the virus mutated to a lysine at position 484 (E484 K). An AlphaFold model of the Wuhan and Beta forms of the RBD, predicted structural changes within the protein (Figure 1A). Consistent with these major structural shifts, native mass spectrometry of the Wuhan and E484 K RBDs showed a clear modification of charge states between the two proteoforms (Figure S1) with the Wuhan RBD showing a dominant +10 charge state and the E484 K

Received: March 3, 2023
Accepted: May 23, 2023

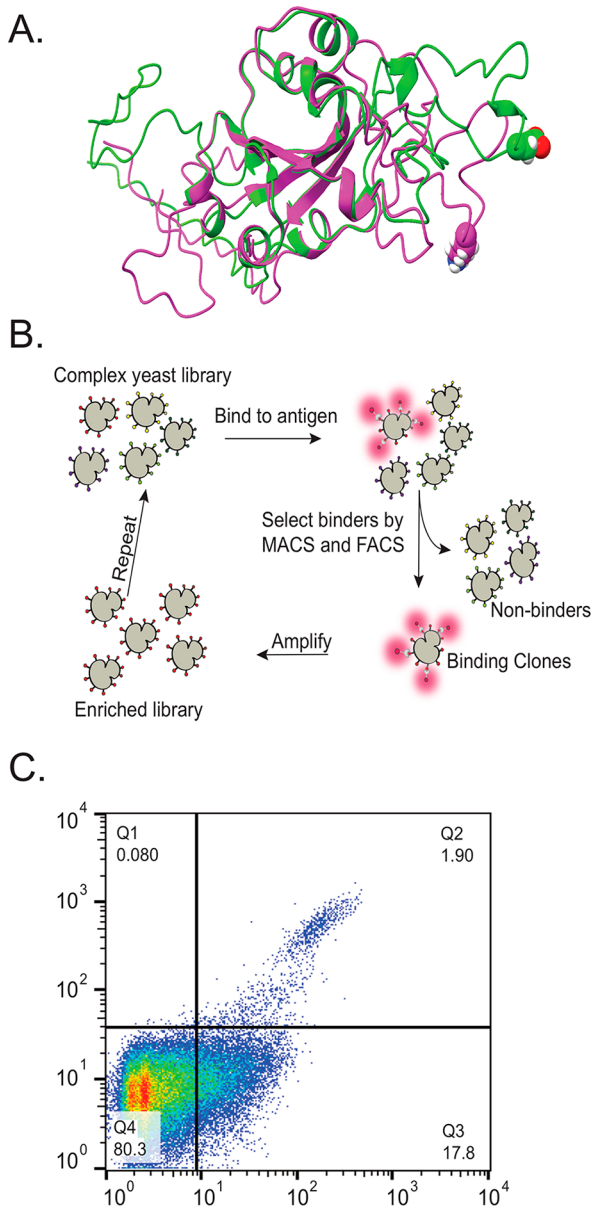



**Figure 1.** (A) AlphaFold predicted structure of the Wuhan SARS-CoV-2 Spike RBD (in green) with the E residue labeled as spheres on the green protein and the 484 K RBD (in pink) with the K residue labeled as a sphere. Notice the predicted changes in the Beta sheet and alpha helical content as well as the position of the disordered loop region between the two proteoforms. (B) Schematic representation of negative and positive selections of different RBD proteoforms through multiple rounds of MACS and FACS. (C) Final FACS analyses of yeast that were positively selected for E484 K RBD proteoforms (incubated here with E484 K RBD-A647 and Wuhan RBD-A488). The FACS shows a clear population of E484-only binding yeast (Q3) and yeast that bind to both the Wuhan and E484 K RBD proteoforms (Q2).

showing a dominant +11 charge state (Figure S2). Consistent with previous studies,[13−15] deconvolution of the RBD native mass spectra confirmed mass heterogeneity of the Wuhan and E484 K proteoforms (Figures S1 and S2). Importantly, these results suggest that the E-to-K mutation at position 484 induces proteoforms with static genetic differences and observably different biophysical properties. We reasoned that these two proteoform populations might serve as target antigens to develop proteoform-specific nanobodies.

We worked to develop nanobodies that would bind to either the Wuhan or E484 K proteoforms, but not both. Toward this goal, we employed a recently developed yeast (*S. cerevisiae*) cell surface display system expressing $10^8$ different nanobodies[16] (Figure 1B). Yeast were trained to bind to either the Wuhan or the E484 K soluble proteoforms through five rounds of metal-activated cell sorting and two rounds of fluorescence-activated cell sorting (FACS). The final FACS analysis showed a clear enrichment of E484 K-binding yeast expressing trained nanobodies against the E484 K proteoforms that showed few Wuhan proteoform-binding nanobodies (Figure 1C). Position 484 of the Spike protein is a mutational hotspot in the RBD. After the virus mutated from glutamic acid to lysine (in the beta version of the virus), the 484 position was later mutated to glutamine (Q). The E484Q mutation serves as an additional control to screen for proteoform-specific nanobodies. We similarly screened for nanobody-expressing yeast that would bind to RBDs with the E484Q mutation, but not the original Wuhan proteoform.

Our FACS analysis showed that we had successfully enriched for populations of nanobody-expressing yeast with a preference for a specific population of RBD proteoforms. We sought to characterize these populations of proteoform-specific nanobodies. We subjected each yeast population enriched against either the Wuhan RBD, E484 K RBD, or E484Q RBD proteoforms to next-generation DNA sequencing (NGS). All DNA sequences were trimmed and translated into nanobody amino acid sequences as described in the Supporting Information. We used Venn diagrams and heatmaps to identify unique nanobody populations trained to bind to each of the proteoforms. A focused comparison of the Wuhan vs E484 K RBD showed 56 shared nanobodies, 21,253 sequences unique to the Wuhan-enriched nanobodies, and 14,849 unique nanobody sequences enriched against the E484 K proteoform (Figure 2A). These results were encouraging, but showed that certain nanobodies would persist—those that bind to both, the Wuhan and E484 K proteoforms, despite multiple rounds of positive and negative selections to segregate different proteoform binders (Figure 2B). Many of the unique nanobody proteoforms had few total reads and were likely to be less

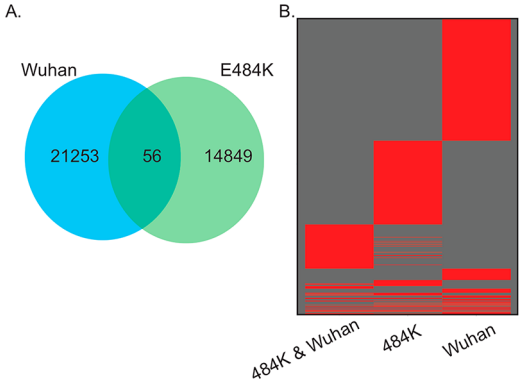



**Figure 2.** Analysis of nanobody populations from yeast surface display enrichment. (A) Venn diagram analysis of Nb populations trained to bind to either the Wuhan or the E484 K RBD proteoform. (B) Heatmap of yeast nanobodies trained to bind either the Wuhan, E484 K, or both RBD proteoforms. Red areas indicate the presence of a given Nb. Gray areas indicate the absence of a Nb.


https://doi.org/10.1021/acs.analchem.3c00958
Anal. Chem. XXXX, XXX, XXX−XXX

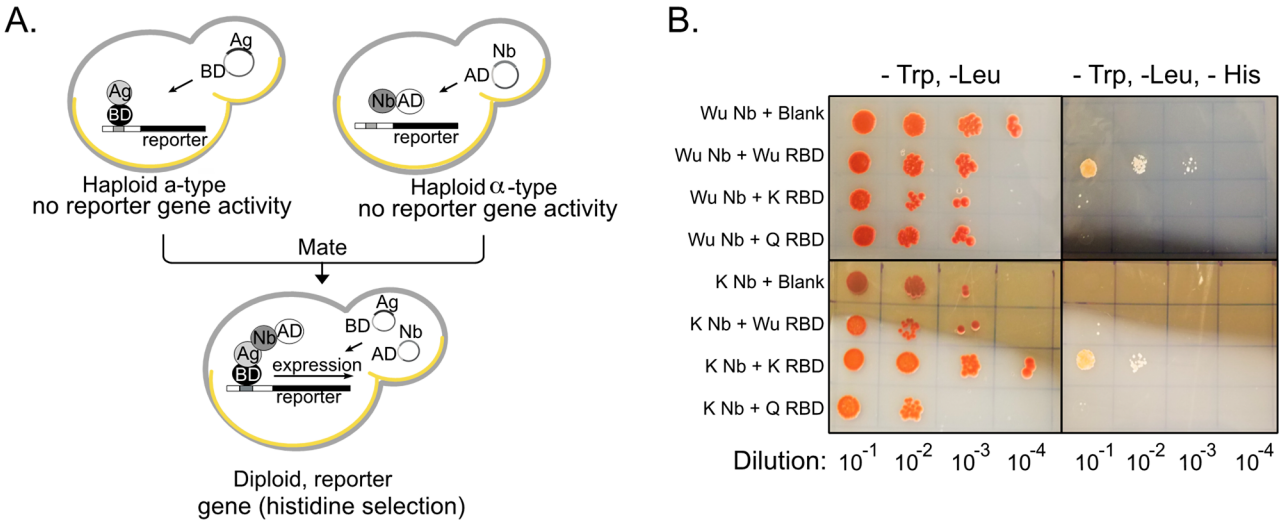



**Figure 3.** (A) Description of yeast-two-hybrid assay for validation of RBD proteoform-binding nanobodies. The RBD antigen (Ag) is fused to the GAL-4 transcription factor DNA binding domain (BD) to make an Ag-BD fusion protein. The pool of surface-display-enriched Nbs were cloned in-frame with the GAL-4 transcription factor activation domain (AD) to make a Nb-AD fusion. Haploid yeast carrying both fusion protein types were mated to produce a diploid yeast test for growth selection on SC-HIS media (with and without 3-AT). Yeast were also grown in the absence of SC-Leu and SC-Trp to select for the presence of the GAL-4 fusion plasmids. (B) (Left) Y2H growth analysis on SC-Trp and SC-Leu media (selects for the presence of the GAL-4 fusion plasmids). (Right) Y2H growth analysis on SC-Trp, SC-Leu, and SC-HIS media (selects for interactions between the GAL4-BD and GAL4-AD fusions). Labeling key: Wuhan (Wu) Nb + Blank = negative control with the Wuhan RBD-trained Nb library mated to the GAL-4 DNA binding domain (absent an RBD proteoform fusion). Wuhan Nb + Wuhan RBD = Nbs trained against the Wuhan RBD mated to the Wuhan RBD in the Y2H. Wuhan Nb + E484 K RBD = Nbs trained against the Wuhan RBD mated against the E484 K RBD. Wuhan Nb + E484Q RBD = Nbs trained against the Wuhan RBD mated against the E484Q RBD. E484 K Nb + Blank = negative control with the E484 K RBD-trained Nb library mated to the GAL-4 DNA binding domain (absent an RBD proteoform). E484 K Nb + Wuhan RBD = Nbs trained against the E484 K RBD mated to the Wuhan RBD. E484k Nb + E484 K RBD = Nbs trained against the E484 K RBD mated to the E484 K RBD. E484 K Nb + E484Q RBD = Nbs trained against the E484 K RBD mated to the E484Q RBD. The dilution series of the yeast is indicated below each growth plate.

abundant (Supplemental Table 1). These results suggested that the yeast surface display platform could clearly select for nanobodies that were proteoform specific.

The cameloid-style nanobodies (used in these studies) adjust their shapes to different proteoform antigen geometries by modifying their complementarity-determining regions (CDRs). We compared the CDRs of the top five most abundant proteoform-specific nanobodies (by read count) from our screen. The most abundant E484 K-specific nanobodies had CDRs that were different than those most abundant nanobodies specific for the Wuhan RBD (Figure S3). However, we observed a significant similarity between the CDRs of the most abundant Wuhan- and E484Q-trained nanobody CDRs, suggesting that there may be some limitations to the selectivity of the nanobody surface display system when two proteoform antigens have more conservative differences. Consistent with this possibility, nearly 6.5%−9.5% of all nanobody sequences directed against the Wuhan and E484Q proteoforms were shared (Figure S4).

Analysis of our proteoform-specific nanobody populations suggested that we might have achieved nanobodies with proteoform-level specificity. We sought a facile technique that would provide independent validation of nanobody specificity. We chose to test the specificity of each proteoform-trained nanobody library against the Wuhan, E484 K, or E484Q proteoforms using the yeast-2-hybrid (Y2H) assay. In this system, the RBD proteoform remains static as the "bait" and the pool of surface-display enriched nanobodies is the "prey".[17] Interaction between the Nb and proteoform antigen is required for yeast to conduct auxotrophic growth. Toward this goal, each of the RBD antigens was cloned in-frame with the GAL-4 DNA binding domain (DB plasmid) and transformed into the *Mat a* Y8800 haploid yeast strain. The entire pool of each proteoform-specific nanobody library was cloned in-frame with the GAL-4 activation domain (AD plasmid) and transformed into the *Mat α* Y8930 yeast strain. Interaction between the Nb-RBD proteoform brings the two GAL4 domains into proximity, making a functional transcription factor that drives the expression of the HIS gene and allows growth in the absence of histidine (Figure 3A). A negative control plasmid containing the GAL-4 BD without an RBD was included to test for the presence of spurious nanobody autoactivators of histidine synthesis (Figure 3B). The transformed yeast strains were mated, subjected to a 10-fold dilution series, and replica plated on a synthetic complete media selecting for either the AD plasmid or the BD plasmid. To select for RBD-Nb interactions, the same dilution series was plated on SC-TRP, SC-LEU, and SC-HIS media. Consistent with our sequencing results of the enriched libraries, a focused comparison of the performance of the Nb libraries trained against the Wuhan or the E484 K proteoforms in the Y2H confirmed a clear preference for the cognate proteoform antigen in the yeast dilution growth assay (Figure 3B). The test for off-target binders, screening for GAL-4 BD (alone) bound to Nbs enriched against RBDs, showed no growth on the histidine-lacking media. We noted that specificity was imperfect with some E484 K-trained Nb clones interacting with Wuhan and E484Q RBDs, albeit at a reduced growth rate in the Y2H when the trained Nbs were paired with off-target RBD proteoforms (Figure 3B). Collectively, these results suggested that the surface display enrichment step


https://doi.org/10.1021/acs.analchem.3c00958
Anal. Chem. XXXX, XXX, XXX−XXX

selected Nbs with significant proteoform preference or specificity.

We were encouraged by the apparent Nb preference of each of the proteoform-trained libraries from the Y2H results. In our Y2H experiments, we employed the HIS gene as our reporter for detecting interaction between trained Nb libraries and specific RBD proteoforms. The HIS gene is susceptible to competitive inhibition by the drug 3-amino triazole (3-AT), and this drug is routinely used to increase selection pressure in the Y2H to identify stronger protein−protein interactions.[18,19] To this end, we grew all Y2H Nb-RBD proteoform diploid yeast in liquid SC-TRP, SC-LEU, and SC-HIS media at 0.5 mM 3-AT to increase selection pressure on Nb-RBD proteoform interactions. To characterize the Nbs that had gone through this dual-selection process, we subjected the Y2H Nb-RBD proteoform clones to NGS. The resulting Nb sequences from each Y2H experiment were subjected to Venn diagram analysis (Figure S5). Importantly, this dual selection process identified a refined population of proteoform-specific nanobodies that were unique or shared within each proteoform-trained yeast population.

## CONCLUSION

The yeast synthetic nanobody library generated by Kruse and colleagues[16] can distinguish proteoforms that vary by a single amino acid position. Utilization of the surface display screening platform combined with Y2H provides a method for validation of Nb binders, determination of preference, and identification of nanobodies that bind their targets intracellularly (intrabodies). We suspect that this combined screening platform can be used in future studies to develop affinity reagents to track or purify specific proteoforms. Additionally, our results suggest that this platform can be used to develop intrabodies as tools to ascertain specific proteoform functions. Modifications of the experimental design presented here (including screens that utilize single, homogeneous proteoform antigens) could yield Nbs with specificity for post-translationally modified proteoforms.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.3c00958.

Nanobody sequences trained to bind either Wuhan RBD, E484K, or E484Q RBD proteoforms (XLSX)

Detailed experimental methods (PDF)

## AUTHOR INFORMATION


### Corresponding Author

**Vincent R. Gerbasi** − *Pacific Northwest National Laboratory, Richland, Washington 99354, United States;* orcid.org/0000-0001-5573-0578; Email: robertvince.gerbasi@pnnl.gov

### Authors

**Bojana Leonard** − *Pacific Northwest National Laboratory, Richland, Washington 99354, United States*

**Vincent Danna** − *Pacific Northwest National Laboratory, Richland, Washington 99354, United States*

**Leo Gorham** − *Pacific Northwest National Laboratory, Richland, Washington 99354, United States*

**Michelle Davison** − *Pacific Northwest National Laboratory, Richland, Washington 99354, United States*

**William Chrisler** − *Pacific Northwest National Laboratory, Richland, Washington 99354, United States*

**Doo Nam Kim** − *Pacific Northwest National Laboratory, Richland, Washington 99354, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.3c00958


### Author Contributions

V.R.G., B.L., and M.D. contributed to plan the experiments, generate the figures, and write the paper. V.R.G., W.C., D.N.K., B.L., V.D., and L.G. performed experiments and/or analyzed the data. V.R.G., V.D., B.L., and M.D. supported and designed the project. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This project was funded by the Congressionally Directed Medical Research Program award CDMRPL-20-0-PR201356 to V.G. and the PNNL Predictive Phenomics Initiative LDRD to V.G. PNNL is a multiprogram national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830. This research used computational resources provided by Research Computing at the Pacific Northwest National Laboratory. We thank Dr. Kurt Glaesemann for his assistance with AlphaFold installation. We thank Dr. Haiyuan Yu of Cornell University for providing yeast-2-hybrid reagents. We thank Dr. Jordan Smith of PNNL for reading the manuscript.

## REFERENCES


(1) Smith, L. M; Kelleher, N. L *Nat. Methods* **2013**, *10* (3), 186−187.

(2) Adams, L. M.; DeHart, C. J.; Drown, B. S.; Anderson, L. C.; Bocik, W.; Boja, E. S.; Hiltke, T. M.; Hendrickson, C. L.; Rodriguez, H.; Caldwell, M.; Vafabakhsh, R.; Kelleher, N. L. *J. Biol. Chem.* **2023**, *299* (1), 102768.

(3) Ntai, I.; Fornelli, L.; DeHart, C. J.; Hutton, J. E.; Doubleday, P. F.; LeDuc, R. D.; van Nispen, A. J.; Fellers, R. T.; Whiteley, G.; Boja, E. S.; Rodriguez, H.; Kelleher, N. L. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (16), 4140−4145.

(4) Melani, R. D.; Gerbasi, V. R.; Anderson, L. C.; Sikora, J. W.; Toby, T. K.; Hutton, J. E.; Butcher, D. S.; Negrao, F.; Seckler, H. S.; Srzentic, K.; Fornelli, L.; Camarillo, J. M.; LeDuc, R. D.; Cesnik, A. J.; Lundberg, E.; Greer, J. B.; Fellers, R. T.; Robey, M. T.; DeHart, C. J.; Forte, E.; Hendrickson, C. L.; Abbatiello, S. E.; Thomas, P. M.; Kokaji, A. I.; Levitsky, J.; Kelleher, N. L. *Science* **2022**, *375* (6579), 411−418.

(5) Toby, T. K.; Abecassis, M.; Kim, K.; Thomas, P. M.; Fellers, R. T.; LeDuc, R. D.; Kelleher, N. L.; Demetris, J.; Levitsky, J. *Am. J. Transplant* **2017**, *17* (9), 2458−2467.

(6) Gerbasi, V. R.; Melani, R. D.; Abbatiello, S. E.; Belford, M. W.; Huguet, R.; McGee, J. P.; Dayhoff, D.; Thomas, P. M.; Kelleher, N. L. *Anal. Chem.* **2021**, *93* (16), 6323−6328.

(7) Huguet, R.; Mullen, C.; Srzentic, K.; Greer, J. B.; Fellers, R. T.; Zabrouskov, V.; Syka, J. E. P.; Kelleher, N. L.; Fornelli, L. *Anal. Chem.* **2019**, *91* (24), 15732−15739.

(8) Kafader, J. O.; Melani, R. D.; Durbin, K. R.; Ikwuagwu, B.; Early, B. P.; Fellers, R. T.; Beu, S. C.; Zabrouskov, V.; Makarov, A. A.; Maze,

J. T.; Shinholt, D. L.; Yip, P. F.; Tullman-Ercek, D.; Senko, M. W.; Compton, P. D.; Kelleher, N. L. *Nat. Methods* **2020**, *17* (4), 391−394.
(9) Lai, S. H.; Tamara, S.; Heck, A. J. R. *iScience* **2021**, *24* (11), 103211.
(10) McGee, J. P.; Melani, R. D.; Yip, P. F.; Senko, M. W.; Compton, P. D.; Kafader, J. O.; Kelleher, N. L. *Anal. Chem.* **2021**, *93* (5), 2723−2727.
(11) Worner, T. P.; Snijder, J.; Bennett, A.; Agbandje-McKenna, M.; Makarov, A. A.; Heck, A. J. R. *Nat. Methods* **2020**, *17* (4), 395−398.
(12) Doubleday, P. F.; Fornelli, L.; Kelleher, N. L. *J. Proteome Res.* **2020**, *19* (2), 938−948.
(13) D'Ippolito, R. A.; Drew, M. R.; Mehalko, J.; Snead, K.; Wall, V.; Putman, Z.; Esposito, D.; DeHart, C. J. *J. Proteome Res.* **2021**, *20* (9), 4427−4434.
(14) Roberts, D. S.; Mann, M.; Melby, J. A.; Larson, E. J.; Zhu, Y.; Brasier, A. R.; Jin, S.; Ge, Y. *J. Am. Chem. Soc.* **2021**, *143* (31), 12014−12024.
(15) Wilson, J. W.; Bilbao, A.; Wang, J.; Liao, Y.-C.; Velickovic, D.; Wojcik, R.; Passamonti, M.; Zhao, R.; Gargano, A. F. G.; Gerbasi, V. R.; Pasa-Tolic, L.; Baker, S. E.; Zhou, M. *Anal. Chem.* **2022**, *94* (15), 5909−5917.
(16) McMahon, C.; Baier, A. S.; Pascolutti, R.; Wegrecki, M.; Zheng, S.; Ong, J. X.; Erlandson, S. C.; Hilger, D.; Rasmussen, S. G. F.; Ring, A. M.; Manglik, A.; Kruse, A. C. *Nat. Struct Mol. Biol.* **2018**, *25* (3), 289−296.
(17) Fields, S.; Song, O. *Nature* **1989**, *340* (6230), 245−6.
(18) Fragoza, R.; Das, J.; Wierbowski, S. D.; Liang, J.; Tran, T. N.; Liang, S.; Beltran, J. F.; Rivera-Erick, C. A.; Ye, K.; Wang, T. Y.; Yao, L.; Mort, M.; Stenson, P. D.; Cooper, D. N.; Wei, X.; Keinan, A.; Schimenti, J. C.; Clark, A. G.; Yu, H. *Nat. Commun.* **2019**, *10* (1), 4141.
(19) Zhou, Y.; Liu, Y.; Gupta, S.; Paramo, M. I.; Hou, Y.; Mao, C.; Luo, Y.; Judd, J.; Wierbowski, S.; Bertolotti, M.; Nerkar, M.; Jehi, L.; Drayman, N.; Nicolaescu, V.; Gula, H.; Tay, S.; Randall, G.; Wang, P.; Lis, J. T.; Feschotte, C.; Erzurum, S. C.; Cheng, F.; Yu, H. *Nat. Biotechnol.* **2023**, *41* (1), 128−139.


https://doi.org/10.1021/acs.analchem.3c00958
*Anal. Chem.* XXXX, XXX, XXX−XXX