

# Reviving Protein-Observed $^{19}\text{F}$ Lineshape Analysis for Deep Insight into Protein-Ligand Binding Events

Scott A. Showalter<sup>1,2,\*</sup>

<sup>1</sup>Department of Chemistry and <sup>2</sup>Center for Eukaryotic Gene Regulation, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania

When I was a graduate student at Washington University in St. Louis, Professor Carl Frieden took delivery of what was, to my knowledge, the first cryogenically cooled NMR probe for  $^{19}\text{F}$ -detection. This probe was put to brilliant use investigating the folding kinetics of the *Escherichia coli* periplasmic protein PapD (1). Indeed, the Frieden lab had a history among the pioneers studying protein folding kinetics by  $^{19}\text{F}$  NMR in which their expertise in kinetic simulation naturally led to the use of NMR lineshape analysis to extract intermediate state populations during the course of urea denaturation experiments (2). Both of these visionary studies required use of auxotrophic *E. coli* strains to allow  $^{19}\text{F}$ -enriched tyrosine and tryptophan incorporation, respectively. Despite the utility of detecting  $^{19}\text{F}$  in proteins for folding studies or investigation of ligand binding, as will be discussed here, these methods are rarely applied because of the perceived complexity and/or cost associated with the available methods for fluorine labeling (3). With an article appearing in this issue of the *Biophysical Journal*, Stadmiller and colleagues remind us of an incredibly cheap and simple method to

achieve 5-fluorotryptophan labeling in proteins and demonstrate the ability of  $^{19}\text{F}$  NMR lineshape analysis to simultaneously determine both equilibrium binding constants and the kinetic rates of association and dissociation for protein-ligand binding events (4). This study demonstrates that  $^{19}\text{F}$  lineshape analysis achieves high quantitative accuracy comparable to that of two-dimensional (2D)  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC lineshape analysis but with a more straightforward postacquisition analysis procedure. It is clear that the methods described will have a broad impact for applications to protein-ligand interactions, protein folding, and potentially in small-molecule library screening.

Kinetic analysis has long been a staple of molecular biophysics through applications to protein interactions with other macromolecules or small-molecule ligands, protein folding, and enzymatic catalysis. Although stopped-flow spectroscopy, surface plasmon resonance, and, more recently, single-molecule fluorescence measurements are the workhorses of this enterprise, NMR spectroscopy has inarguably been a key methodology as well. To provide one example, NMR spin relaxation and lineshape analysis allowed the construction of a model for folding of the kinase inducible domain (KID) from CREB, which is coupled to binding the KIX domain of the CREB binding protein (5).

This seminal study spurred great interest in folding-upon-binding interactions and established the KID/KIX system as a key case study in the disordered protein literature. In one follow-up study, the Clarke lab demonstrated through stopped-flow fluorescence measurements that phosphorylation of the KID region does not speed up binding, as initially hypothesized, but instead increases the lifetime of complexes with KIX by decreasing the kinetic dissociation rate constant (6). These studies demonstrate that NMR spectroscopy is both synergistic with optical methods for kinetic analysis and extremely useful in its own right because of the ability of NMR to report on dynamics covering a wide range of timescales associated with biomolecular processes. The problem is that most of the NMR dynamics methods available require specialized pulse programs that are difficult to implement outside of highly trained laboratories or require challenging and expensive protein isotope-labeling schemes.

In their current work, Stadmiller and colleagues demonstrate an exciting experimental approach to kinetic analysis through protein NMR that relies on simple and inexpensive incorporation of 5-fluorotryptophan, read out by one-dimensional NMR experiments. The key to this method is that *E. coli* tryptophan synthase readily incorporates 5-fluoroindole into tryptophan, allowing production of specifically

Submitted March 30, 2020, and accepted for publication April 1, 2020.

\*Correspondence: sas76@psu.edu

Editor: Jane Dyson.

<https://doi.org/10.1016/j.bpj.2020.04.003>

© 2020 Biophysical Society.



5-fluorotryptophan-labeled protein from the growth of widely used cell lines, such as BL21, in minimal media supplemented with 5-fluoroindole (7). Stadmiller and colleagues take great care to control for perturbations in protein stability due to 5-fluorotryptophan labeling and to demonstrate quantitative agreement between  $^{19}\text{F}$  lineshape analysis and more broadly applied methods. In particular, labeling with additional isotopes such as  $^{15}\text{N}$  to permit backbone  $^1\text{H},^{15}\text{N}$ -HSQC analysis is a straightforward extension of this method.  $^1\text{H},^{15}\text{N}$ -HSQC analysis is state of the art, as recently demonstrated for substrate binding by thymidylate synthase in the *Biophysical Journal* (8). The key to success with 2D  $^1\text{H},^{15}\text{N}$ -HSQC (or TROSY-HSQC) lineshape analysis in both studies was application of the TITAN (TITration Analysis) software package (9). In this method, one simulates not only the final NMR spectrum but also the evolution of spin operators throughout the pulse sequence, allowing for better accounting of lineshape effects from spin relaxation that occurs during the course of complex experiments. To end users unfamiliar with the spin physics that underpin the calculation, the key result is an improved fit quality because the state of the system entering data acquisition is more accurately modeled. Although care must always be taken to process spectra appropriately, most notably in terms of the window function selected, this method is approachable and available through the NMRbox software suite ([www.nmrbox.org](http://www.nmrbox.org)), which should facilitate broad adoption.

Stadmiller and colleagues apply both  $^{19}\text{F}$  and  $^1\text{H},^{15}\text{N}$ -HSQC lineshape analysis to the N-terminal SH3 domain from *Drosophila melanogaster* Drk, which has been shown to interact with three short proline-rich motifs from son of sevenless. Each proline-rich motif investigated varies slightly in sequence, with concomitant variation in their binding affinities, which is a common feature when multiple intrinsically disordered motifs bind to a single folded protein domain. For

example, the N-terminal SH3 domain of CrkII binds to three proline-rich motifs in cAbl; relaxation-dispersion NMR spectroscopy was used to demonstrate that differences in the equilibrium binding constants for these three motifs derive mostly from differences in the dissociation rate constant (10), similar to the results for KID/KIX discussed previously. In contrast, the results for Drk binding to son of sevenless indicate that significant changes in both association and dissociation rates contribute to the range of equilibrium constants observed. Most importantly, the best-fit rate constants are in quantitative agreement whether one-dimensional  $^{19}\text{F}$  lineshapes or 2D  $^1\text{H},^{15}\text{N}$ -HSQC lineshapes are analyzed. Although SH3 domains are an extensively studied class of modular protein-protein interaction domains, they are hardly unique in displaying weak but specific binding to multiple partners. Many-to-one interactions within cellular signaling cascades and in transcription are ubiquitous; more recently, they have also been recognized as a key feature of liquid-liquid phase separation. The ability to interrogate the kinetics of these interactions while maintaining thermal equilibrium is a key advantage to lineshape analysis, as this study beautifully demonstrates.

Incorporation of fluorine into protein for lineshape analysis is clearly useful, but not required; fluorine is commonly found in drug-like small molecules because of its pharmacological benefits, allowing for ligand-directed interrogation of relevant interactions. As an example,  $^{19}\text{F}$  NMR revealed details of coregulator binding to both orthosteric and allosteric sites on the nuclear receptor PPAR $\gamma$ , resolving a debate in the literature generated by prior observation of multiple agonist binding modes in the orthosteric site (11). The benefits of  $^{19}\text{F}$  lineshape analysis for small-molecule library screening or for mechanistic analysis as performed for PPAR $\gamma$  may be one of the greatest benefits to reviving the technique.

In summary,  $^{19}\text{F}$  spectroscopy has been applied in biophysics for decades, although difficulty incorporating the isotope into protein has throttled adoption of the technique. With the labeling strategy adopted by Stadmiller and colleagues, the perceived difficulty of  $^{19}\text{F}$ -labeling proteins should no longer concern those interested in adopting the method. The advantages to  $^{19}\text{F}$  NMR are clear: the gyromagnetic ratio is near to that of  $^1\text{H}$ , imparting good sensitivity, and there is almost zero background signal in biomolecular  $^{19}\text{F}$  spectra because of the element's absence from most natural biomolecules. Currently, there is great interest in performing NMR in cells, where background suppression is absolutely critical. The Pielak laboratory was among the earliest groups to recognize this benefit, reporting  $^{19}\text{F}$  spectra recorded in *E. coli* cells in 2010 (12) before the simplification of the labeling strategy to what is reported in their study. As briefly reviewed throughout this New and Notable,  $^{19}\text{F}$  lineshape analysis will readily extend to the investigation of protein folding, allosteric ligand binding, and enzyme mechanism in general. Thus, the approach demonstrated by Stadmiller and colleagues will generalize broadly, making quantitative NMR spectroscopy accessible to a wide range of biophysical laboratories.

## ACKNOWLEDGMENTS

Grace Usher is gratefully acknowledged for discussion and assistance preparing this manuscript.

This work was supported by U.S. National Science Foundation grant MCB-1932730.

## REFERENCES

1. Bann, J. G., J. Pinkner, ..., C. Frieden. 2002. Real-time and equilibrium ( $^{19}\text{F}$ )-NMR studies reveal the role of domain-domain interactions in the folding of the chaperone PapD. *Proc. Natl. Acad. Sci. USA.* 99:709–714.
2. Ropson, I. J., and C. Frieden. 1992. Dynamic NMR spectral analysis and protein folding: identification of a highly populated folding intermediate of rat intestinal fatty acid-

- binding protein by  $^{19}\text{F}$  NMR. *Proc. Natl. Acad. Sci. USA.* 89:7222–7226.
- Divakaran, A., S. E. Kirberger, and W. C. K. Pomerantz. 2019. SAR by (protein-observed)  $^{19}\text{F}$  NMR. *Acc. Chem. Res.* 52:3407–3418.
  - Stadmler, S. S., J. S. Aguilar, ..., G. J. Pielak. 2020. Rapid quantification of protein-ligand binding via  $^{19}\text{F}$  NMR lineshape analysis. *Biophys. J.* 118:2537–2548.
  - Sugase, K., H. J. Dyson, and P. E. Wright. 2007. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature.* 447:1021–1025.
  - Dahal, L., S. L. Shammass, and J. Clarke. 2017. Phosphorylation of the IDP KID modulates affinity for KIX by increasing the lifetime of the complex. *Biophys. J.* 113:2706–2712.
  - Crowley, P. B., C. Kyne, and W. B. Monteith. 2012. Simple and inexpensive incorporation of  $^{19}\text{F}$ -tryptophan for protein NMR spectroscopy. *Chem. Commun. (Camb.)*. 48:10681–10683.
  - Bonin, J. P., P. J. Sapienza, ..., A. L. Lee. 2019. Positive cooperativity in substrate binding by human thymidylate synthase. *Biophys. J.* 117:1074–1084.
  - Waudby, C. A., A. Ramos, ..., J. Christodoulou. 2016. Two-dimensional NMR lineshape analysis. *Sci. Rep.* 6:24826.
  - Zeng, D., V. S. Bhatt, ..., J. H. Cho. 2016. Kinetic insights into the binding between the nSH3 domain of CrkII and proline-rich motifs in cAbl. *Biophys. J.* 111:1843–1853.
  - Hughes, T. S., J. Shang, ..., D. J. Kojetin. 2016. Probing the complex binding modes of the PPAR $\gamma$  partial agonist 2-chloro-N-(3-chloro-4-(5-chlorobenzo[d]thiazol-2-yl)thio)phenyl)-4-(trifluoromethyl)benzenesulfonamide (T2384) to orthosteric and allosteric sites with NMR spectroscopy. *J. Med. Chem.* 59:10335–10341.
  - Li, C., G. F. Wang, ..., G. J. Pielak. 2010. Protein ( $^{19}\text{F}$ ) NMR in *Escherichia coli*. *J. Am. Chem. Soc.* 132:321–327.