A CURE Biochemistry Laboratory Module to Study Protein-Protein Interactions by NMR Spectroscopy

William M. Marsiglia[1], Rohini Qamra[1], Kimberly M. Jackson[2], and Nathaniel J. Traaseth[1*]

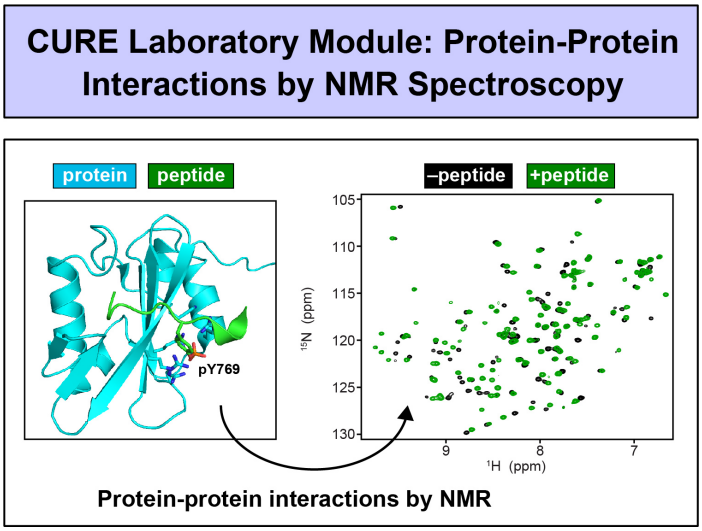
[1]Department of Chemistry, New York University, New York, NY 10003

[2]Department of Chemistry and Biochemistry, Spelman College, Atlanta, GA 30314

ABSTRACT

Design of undergraduate laboratory courses that provide meaningful research-based experiences enhance undergraduate curricula and prepare future graduate students for research careers. In this article, a Course-based Undergraduate Research Experience (CURE) laboratory module was designed for upper-division undergraduate biochemistry and chemistry students. The laboratory module enabled students to build upon recently published data in the literature to decipher atomistic insight for an essential protein-protein interaction in human biology through the use of biomolecular NMR spectroscopy. Students compared their results with published data with the goal of identifying specific regions of the protein-protein interaction responsible for triggering an allosteric conformational change. The laboratory module introduced students to basic and advance laboratory techniques, including protein purification, NMR spectroscopy, and analysis of protein structure using molecular visualization software.

GRAPHICAL ABSTRACT



KEYWORDS

Upper-Division Undergraduate, Biochemistry, Inquiry-Based/Discovery Learning, Biophysical Chemistry, NMR Spectroscopy, Proteins/Peptides.

INTRODUCTION

Motivating and preparing undergraduate students for Science, Technology, Engineering and Mathematics (STEM) careers is one of the central missions of academic life. Faculty are either directly involved in teaching undergraduate courses and/or indirectly exposed to the quality of undergraduate curricula through the mentorship of graduate students. The undergraduate teaching laboratory offers a unique opportunity for instructors to closely interact with students in a more hands-on manner than through lecture-based classes. The primary challenge in laboratory courses is ensuring experiments encourage student excitement, while balancing the need to teach laboratory skills and expose students to *real-life* research. Implementation of outdated experiments represents a lost opportunity for engaging the intellectual curiosity of students and faculty alike and can demotivate undergraduate students for research-based careers and faculty to be engaged in teaching [1]. The type of a laboratory course can be categorized as skill-based [2-5], project-based [6,7] or inquiry-based [8-10]. A combined strategy

integrating each type is referred to as Course-based Undergraduate Research Experiences (CUREs) [11,12]. Goals of CUREs include exposing students to the scientific method and ensuring students have relevant experiences that resemble those found in a research laboratory [13]. Both are important objectives since they help establish a foundation for future graduate students to better understand the significance of a research project and to identify and frame unanswered questions in a field of study. Although no universal strategy exists for motivating every student, successful CUREs place an emphasis on: (i) the importance of rationale and significance in the scientific method, (ii) data analysis and contextualizing findings relative to existing scientific literature, and (iii) incorporating basic and advanced laboratory techniques, including state-of-the-art methods.

A key consideration for CURE laboratory modules is to place an emphasis on the significance of an experiment as it relates to a scientific field. Many beginning Ph.D. students pursuing biomedically relevant research often think that significance is strictly related to how a particular research project will cure or lead to therapeutics to treat diseases. While contributing knowledge or therapeutics to human disease is certainly a motivating factor for biomedical research, it is important to emphasize that basic research is essential [14]. Therefore, it is critical that students are taught to think broadly about how an approach or experimental finding will advance a scientific field. Indeed, significance has been clearly defined by the National Institutes of Health: "*Does the project address an important problem or a critical barrier to progress in the field?*" [15]. Advanced undergraduate laboratories offer an excellent place to teach students about significance. This can occur during a pre-laboratory lecture or video when the rationale for the experiment is presented (**Figure 1**). Topics include discussing knowledge that was established in the field over years or decades, and then articulating what are the unanswered questions in the field. Choosing a topic with controversy in the literature can also be an effective way to motivate student experiments. This approach is underscored in the philosophy of CUREs and is compatible with advancements in graduate student training that places a focus on developing a range of skills [16].

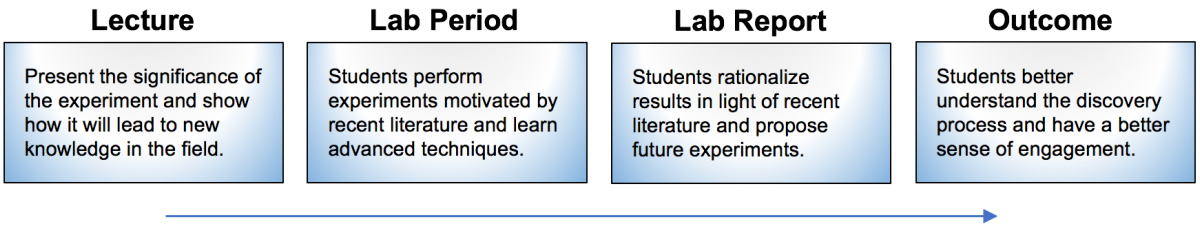


Figure 1. Workflow of a CURE laboratory module based on current literature, which exposes students to relevant basic and advanced techniques and methods. The figure depicts the experience of students and the goals and outcomes of the laboratory module.

With the goals of emphasizing significance in the undergraduate laboratory and teaching advanced techniques to students, a new CURE laboratory module was developed for upper-division undergraduate students majoring in biochemistry or chemistry. This laboratory module, which requires two, four-hour sessions to complete, was successfully implemented at New York University and Spelman College, a Ph.D.-granting research institution and a primarily undergraduate college, respectively, where students investigated a protein-protein interaction using two-dimensional biomolecular NMR spectroscopy. The specific protein-protein interaction studied occurs within the phospholipase Cγ (PLCγ) cell signaling cascade and involves a phosphorylated peptide from the fibroblast growth factor (FGF) receptor with a Src homology 2 (SH2) domain from PLCγ. The results enabled students to decipher novel details of binding reactions including allosteric conformational changes propagating from protein-protein interfaces. This laboratory module represents the first published CURE to use NMR spectroscopy to study protein-protein interactions and demonstrates the feasibility of applying NMR to study other protein-protein interactions using NMR spectrometers that are available at most institutions (e.g., 400 MHz). Finally, perspectives are shared for future laboratory design based on the instructors' experience in developing the laboratory module.

LABORATORY MODULE

Motivation for the Laboratory Module: SH2 Domain of PLCγ Binding to a Phosphopeptide from the FGF Receptor

The broader themes of the laboratory module focused on protein-protein interactions and intracellular signaling through receptor tyrosine kinases (RTKs). These topics represent important

material covered during biochemistry lecture courses and other classes such as molecular and cellular biology. A key feature within cellular signaling networks is the presence of protein-protein interactions that are required for propagating the signal, which ultimately leads to a functional response. To emphasize how protein-protein interactions can be studied using atomistic techniques, this laboratory module used NMR spectroscopy to probe a molecular interaction between the FGF receptor (an example RTK) and PLCγ (a substrate of the FGF receptor) that results in activation of PLCγ and formation of secondary messengers diacylglycerol and inositol 1,4,5-trisphosphate from the substrate phosphatidylinositol 4,5-bisphosphate [17,18]. While this specific laboratory module focused on interactions involving the FGF receptor with PLCγ, the approach described in this work to study protein-protein interactions is generalizable to other complexes based on the interests of the instructor and students.

To give students the necessary background on the interaction between PLCγ and the FGF receptor, two structure-based papers were introduced to the class through a pre-laboratory module lecture. These articles proposed alternative binding interactions occurring between a phosphotyrosine within the C-terminal tail of the FGF receptor and an SH2 domain within PLCγ. Namely, Bae *et al.* [19] solved a crystal structure (PDB ID 3GQI) of a tandem SH2 domain composed of N- and C-terminal SH2 domains (nSH2 and cSH2, respectively) where the nSH2 domain was bound to the receptor. Based on these data and others, the authors concluded the nSH2 interactions led to autophosphorylation by the same kinase (i.e., in *cis*). Subsequent to this article, Huang *et al.* [20] published a crystal structure showing that the cSH2 domain binds to the receptor and is phosphorylated by a second phosphorylating kinase (i.e., in *trans*). Thus, the latter work proposed a 2:1 RTK:PLCγ complex and proposed a reason why receptor dimerization is required beyond kinase activation (**Figure 2A**). The introductory class period took place during the lecture session that is associated with the laboratory module. The lecture at New York University was delivered by the graduate student working on this area of research as part of his Ph.D. thesis (W. M. Marsiglia), and a separate lecture at Spelman College was given by the faculty investigator (N. J. Traaseth). Lectures were delivered a few days prior to the laboratory period and allowed students to review the material presented and assigned literature

relevant to the laboratory module. Note that if a dedicated lecture period is unavailable for the laboratory course, it is recommended to record a lecture and make it available to students prior to the laboratory module, which is commonly employed in flipped classrooms.

Following introduction to the PLCγ and FGF receptor system, students were taught how a biomolecular complex can be studied at atomic detail using NMR spectroscopy. One of the most common ways this is accomplished is to perform chemical shift perturbation (CSP) analysis, which a method to quantify the extent of chemical shift changes upon binding to a molecule (e.g., substrate, ligand, protein, nucleic acid, etc.). A schematic of the CSP experiment is shown in **Figure 3A**. A series of spectra can be acquired in a sequential fashion to infer residues involved in binding or to determine a binding constant. The student laboratory module was designed to build upon previously published findings in the Huang *et al.* [20] study by having students perform CSP analysis between the cSH2 domain of PLCγ and a C-terminal region from the FGF receptor containing a phosphotyrosine (**Figure 2B**). The published work reported CSP values from an experiment involving ^{15}N-labeled cSH2 domain of PLCγ (i.e., NMR active) and natural abundance FGF receptor comprised of the kinase domain with a phosphotyrosine in the C-terminal region (i.e., NMR-inactive). It was reported that CSPs extended beyond the canonical SH2 phosphotyrosine binding pocket to the C-terminal region of the cSH2 domain of PLCγ. These long-range changes supported a role for allostery that may lead to increased accessibility of tyrosine residues within PLCγ that are subsequently phosphorylated by the FGF receptor kinase. The student laboratory module involving the shorter construct of the FGF receptor allowed students to determine whether the allosteric structural changes could be induced at the C-terminal region of the cSH2 domain with a smaller peptide or whether the kinase domain was essential to inducing these changes.

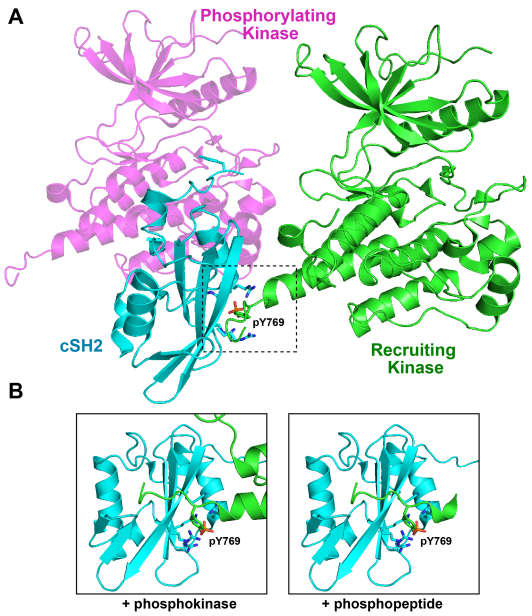


Figure 2. Protein-protein interactions studied in the laboratory module. (A) View of the crystal structure of the cSH2 domain from PLCγ (cyan) binding to the C-terminal tail containing the phosphotyrosine of the recruiting kinase (green) (PDB ID 5EG3). In addition, Tyr771 of the C-terminal tail of the cSH2 domain is bound to the phosphorylating kinase (purple) that occurs through crystal contacts within the lattice (PDB ID 5EG3). (B) Zoomed-in view of the interaction between the recruiting kinase and the cSH2 domain (left) and a schematic view of the C-terminal tail peptide of the FGF receptor used in the laboratory module (right).

Specifics of the Laboratory Module

Protein expression and purification are widely utilized in biochemical experiments and these skillsets can be broadly applied to other proteins of interest. In order to purify cSH2 domain from PLCγ, students were provided with bacterial lysate corresponding to the expression of this protein in ^{15}N minimal media in *Escherichia coli*. Each student group of two or three was given the equivalent of ~50 mL bacterial culture, which amounted to 50 mg of $^{15}NH_4Cl$ at a cost of $0.82 per student group. Lysate was provided to minimize the time demands of bacterial growths and the need to accommodate other experiments during the semester. Starting from the lysate, students purified the cSH2 domain using cation-exchange chromatography in a manual fashion with a syringe attached to the column. A detailed procedure of the protein purification for the laboratory module is provided in the Supporting Information. The cost of each column was $28 and can be reused several times. Students analyzed SDS-PAGE results to determine the fractions containing cSH2 to concentrate for NMR spectroscopy. The efficient expression of cSH2 samples resulted in NMR samples of 0.5 mL at a concentration of approximately 0.5 mM. $^{1}H/^{15}N$ HSQC experiments were acquired in the absence and presence of a 10-residue phosphopeptide corresponding to the C-terminal tail of the FGF receptor (TNEEpYLDLSQ). The cost of the synthetic peptide was ~$3.72 per student group.

Across all semesters, students successfully acquired high quality 2D HSQC spectra with one failed experiment resulting from improper usage of the cation-exchange column (i.e., no purified protein). Note that the laboratory module was implemented for three different semesters with class sizes varying from 7 to 10 students at Spelman College and for four different semesters with class sizes varying from 4 to 49 students at New York University. At Spelman College, spectra were acquired with a 400 MHz NMR spectrometer equipped with a room temperature probe that enabled 2D data sets to be collected in 15 min (**Figure 3B**, top panel). At New York University, spectra were acquired with a 600 MHz NMR spectrometer equipped with a cryogenic probe that enabled 2D data sets to be collected in 5 min (**Figure 3B**, bottom panel). Peaks within the 2D spectra had average signal to noise ratios of ~28/1 and ~47/1 at 400 MHz and 600 MHz, respectively. These data sets were generally of publishable quality and enabled students to analyze the majority of non-proline peaks in the spectrum.

It is important to emphasize that the biomolecular NMR approach described in this laboratory module can be applied to other protein-protein interactions. The key considerations when selecting a suitable complex are the total size of the system to be studied and the protein concentrations that are achievable. The former is limited to ~25 kDa using a similar isotopic labeling approach described in this work. The latter is an important consideration for acquiring spectra using NMR spectrometers available at most undergraduate institutions (e.g., 400 MHz and without cryogenic probes) and stems from the need to achieve sufficient signal to noise to acquire 2D HSQC spectra. It is recommended that a minimum concentration of ~0.3 mM of the isotopically enriched protein is obtained in order to study a protein-protein complex using a 400 MHz spectrometer in the absence of a cryogenic probe.

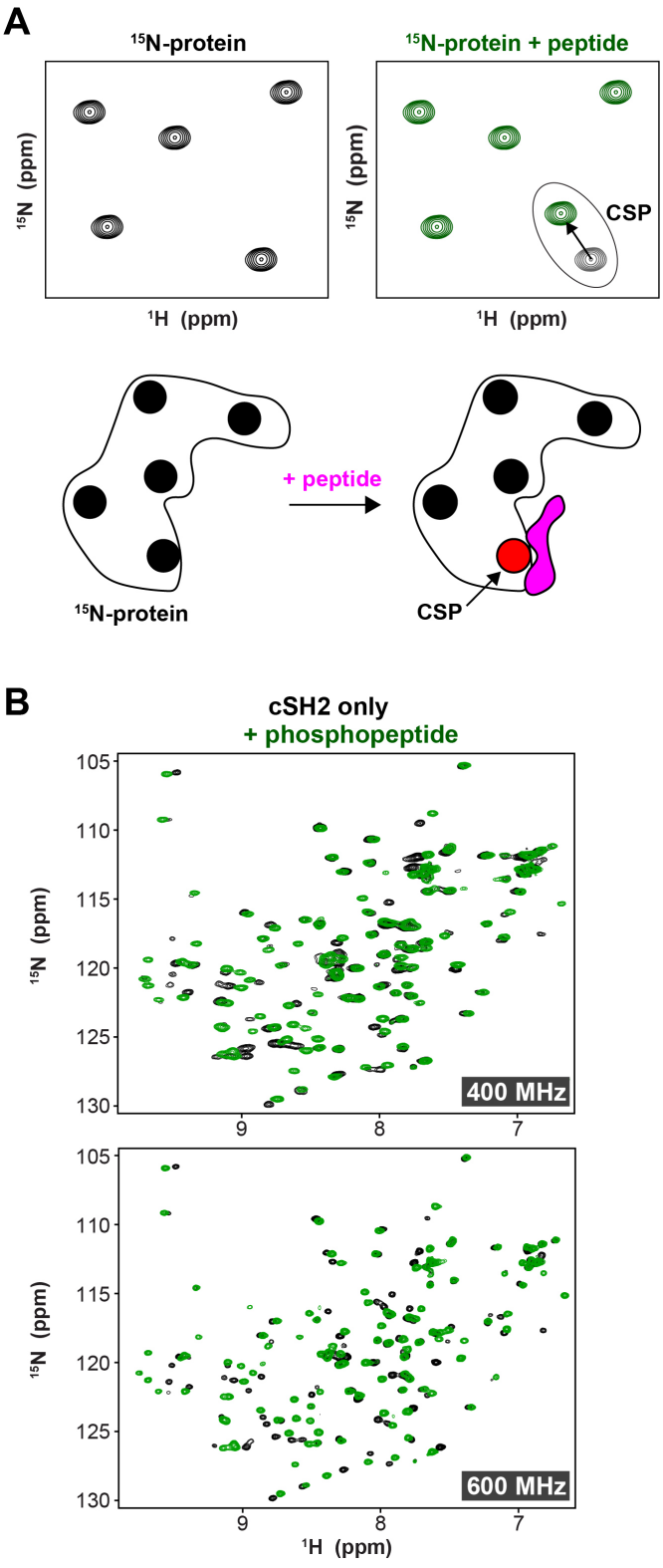


Figure 3. (A) Schematic of the protein-protein CSP experiment involving addition of natural abundance peptide into a ^{15}N labeled protein. The peak with the indicated CSP is near the location of peptide binding. (B) Overlay of experimental ^{15}N-HSQC spectra at 400 MHz (top) and 600 MHz (bottom) comparing chemical shifts for the free cSH2 domain (in black) and the cSH2 domain bound to the phosphopeptide (in green).

Data Interpretation from NMR Experiments

The primary data sets that students analyzed were HSQC spectra of cSH2 in the absence and presence of the phosphopeptide (**Figure 3B**). Spectral data obtained on student cSH2 samples were consistent among various student groups within the class and across different semesters. Using these HSQC spectra, students quantified the CSPs and plotted these on the available crystal structure (PDB ID 5EG3) [20]. Students compared their calculated CSP values with published results that used an FGF receptor construct that contained both the kinase domain and the C-terminal tail. It was underscored to students that the peptide used in their experiments did not contain N-terminal or C-terminal modifications such as acetylation, which might give rise to additional CSPs relative to the published results. To facilitate the CSP analysis, chemical shift assignments were provided to students for the cSH2 domain in the apo state [20] (see Supporting Information). A Sparky tutorial and additional insights into NMR data interpretation were delivered by the principal investigator (N. J. Traaseth), who has extensive experience in NMR spectroscopy. This offered students an opportunity to interact with and learn from someone directly in the scientific field. Evaluation of student lab reports revealed that the majority of students correctly identified CSPs induced upon addition of the phosphopeptide corresponding to the C-terminal tail of FGF receptor. Note that tutorials to interpret NMR spectra using Sparky [21] and subsequent analysis with PyMOL [22] is provided in the Supporting Information. These tutorials are applicable to other protein-protein interactions studied using biomolecular NMR spectroscopy.

The laboratory module incorporated a variety of skills, including protein purification using ion-exchange chromatography, evaluation of protein purity using gel electrophoresis, concentration of protein samples for structural biology, biomolecular NMR spectroscopy, and analysis of protein structure with molecular visualization software.

Perspectives about the Laboratory Module and Future CUREs

The design of the laboratory module was approached with the following central paradigm: *what motivates basic research can also inspire students in the teaching laboratory classroom*. This approach

served two primary goals: (i) to motivate undergraduate students to participate in the process of scientific discovery[1,11,13,23] and (ii) to motivate instructors' natural passion for scientific discovery which leads to a mutual excitement from students [24]. The instructors found that the laboratory module encouraged a greater sense of participation from the students relative to more traditional laboratory experiences where the results were well-established. Students commented to the instructors that they enjoyed working directly with researchers in the field and showed a high level of enthusiasm for learning how NMR data sets were collected and used to study protein-protein interactions. Positive responses from student course evaluations included the following: "*Conducting larger experiments throughout the course of the semester is one of the strong points of this class, as it mimics the multi-step nature of "real research" and allowed us to explore questions that have not been answered yet (as opposed to most lab courses, where you know exactly what outcome to expect)";* and, *"If I were not graduating, I would definitely be interested in another class or lab similar to this one."*

Based on the instructors' experience in developing the laboratory module, the following design principles were viewed as important for incorporating into future modules. Note that these points have been supported by academic literature and suggest improved student knowledge retention and participation [23,25], which was also observed in the classrooms at New York University and Spelman College.

- **Emphasis on the scientific rationale for carrying out the experiment**. Evaluation of CUREs shows that students benefit from the process of motivating a scientific question by understanding literature and why carrying out an experiment will advance a field of study [11]. The pre-laboratory lecture delivered to students is provided in the Supporting Information.
- **Incorporate techniques as part of the laboratory, but not as the goal for the laboratory.** Introducing students to techniques forms an integral component of undergraduate experiments. The laboratory module was designed to teach several skills while at the same time placing the emphasis on a scientific question.
- **Engage students with ongoing research within the institution.** It is our opinion that undergraduate students in the natural sciences should be as familiar with ongoing science within

the institution as their sports teams. This knowledge and exposure can build a sense of pride within the research enterprise.

- **Introduce students to advanced techniques**. Undergraduate lecture-based courses in biochemistry and molecular biology routinely teach students about classical experiments and state-of-the-art techniques that form the basis for understanding protein and nucleic acid structure-function relationships [26]. It is often more challenging to incorporate advanced techniques into the lab [3-5]. However, exposure to these techniques can better prepare aspiring Ph.D. students for graduate school research in interdisciplinary fields and can reduce the *fear factor* surrounding advanced techniques that can be generated in lower-division courses.
- **Involve graduate students and post-doctoral associates in the design and implementation of laboratory modules (if possible).** Recent evidence suggests that involving graduate students and post-doctoral associates in evidence-based teaching strategies benefit research production and improves overall communication skills [27].
- **Incorporate new laboratory modules/directions.** Similar to basic research, classroom experiments can be designed such that student results seed new experiments [28]. This further underscores to students the discovery process and provides an opportunity to contextualize findings relative to scientific literature. Of course, one aspect of new experiments is the possibility of unanticipated hurdles and failed experiments. Nevertheless, similar to academic research, negative results are a part of science and students often learn as much from failures as successes [29].

With regard to future laboratory modules, signal transduction pathways offer the opportunity to conduct several structure-function studies due to the abundance of enzymes and protein-protein interactions. Future years will also explore the possibility of diversifying the classroom by giving individual student groups separate projects. If selected carefully, student groups can interpret data collectively to derive conclusions beyond what would be achievable in a single laboratory module.

CONCLUSION

A CURE undergraduate laboratory module was developed to obtain atomistic insights into a protein-protein interaction using biomolecular NMR spectroscopy. In the process, students gained knowledge of protein purification, NMR spectroscopy, and analysis of structural data by using molecular visualization software. The overall design principle was to pursue objectives in the same manner as that of an academic laboratory. This focus on significance harnessed the natural tendency from research-active instructors to be energized by acquisition of new experiments with the goal of obtaining new knowledge. Class evaluations and student feedback suggested that students enjoyed learning advanced techniques and being involved in the discovery of new knowledge.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available on the ACS Publications website at DOI: 10.1021/acs.jchemed.XXXXXXX.

Appendix describing the experiment in a protocol-style manner including a tutorial for NMR data analysis (PDF).

Pre-laboratory lecture that illustrates the background and motivation for the laboratory module (PDF).

AUTHOR INFORMATION

Corresponding Author

*E-mail: traaseth@nyu.edu

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (MCB1506420) (to N.J.T). The NMR data collected at NYU were collected on a cryogenic probe made available by funding from the National Institutes of Health (S10OD016343). W.M.M. acknowledges predoctoral funding from the National Institutes of Health (F99CA212474). We are grateful to Moosa Mohammadi for generously

providing the plasmid for expressing the SH2 domain and to Shanina Sanders Johnson for providing NMR support at Spelman College.

REFERENCES

(1) Brownell, S. E.; Kloser, M. J. Toward a conceptual framework for measuring the effectiveness of course-based undergraduate research experiences in undergraduate biology. *Stud. High. Educ.* **2015**, *40* (3), 525-544.

(2) Boyer, R. Concepts and skills in the biochemistry/molecular biology lab. *Biochem. Mol. Biol. Educ.* **2003**, *31* (2), 102-105.

(3) Garrett, E.; Wehr, A.; Hedge, R.; Roberts, D. L.; Roberts, J. R. A novel and innovative biochemistry laboratory: Crystal growth of hen egg white lysozyme. *J. Chem. Educ.* **2002**, *79* (3), 366-368.

(4) Wright, N. T. Heteronuclear Multidimensional Protein NMR in a Teaching Laboratory. *J. Chem. Educ.* **2016**, *93* (2), 287-291.

(5) Rovnyak, D.; Thompson, L. E. An accessible two-dimensional solution nuclear magnetic resonance experiment on human ubiquitin. *Biochem. Mol. Biol. Educ.* **2005**, *33* (2), 117-122.

(6) Beier, M. E.; Kim, M. H.; Saterbak, A.; Leautaud, V.; Bishnoi, S.; Gilberto, J. M. The effect of authentic project-based learning on attitudes and career aspirations in STEM. *J. Res. Sci. Teach.* **2019**, *56* (1), 3-23.

(7) Pontrello, J. K. Bringing Research into a First Semester Organic Chemistry Laboratory with the Multistep Synthesis of Carbohydrate-Based HIV Inhibitor Mimics. *Biochem. Mol. Biol. Educ.* **2015**, *43* (6), 417-427.

(8) Beck, C.; Butler, A.; da Silva, K. B. Promoting Inquiry-Based Teaching in Laboratory Courses: Are We Meeting the Grade? *CBE-Life Sci. Educ.* **2014**, *13* (3), 444-452.

(9) Weaver, G. C.; Russell, C. B.; Wink, D. J. Inquiry-based and research-based laboratory pedagogies in undergraduate science. *Nat. Chem. Biol.* **2008**, *4* (10), 577-580.

(10) Tomasik, J. H.; Cottone, K. E.; Heethuis, M. T.; Mueller, A. Development and Preliminary Impacts of the Implementation of an Authentic Research-Based Experiment in General Chemistry. *J. Chem. Educ.* **2013**, *90* (9), 1155-1161.

(11) Auchincloss, L. C.; Laursen, S. L.; Branchaw, J. L.; Eagan, K.; Graham, M.; Hanauer, D. I.; Lawrie, G.; McLinn, C. M.; Pelaez, N.; Rowland, S.; Towns, M.; Trautmann, N. M.; Varma-Nelson, P.; Weston, T. J.; Dolan, E. L. Assessment of Course-Based Undergraduate Research Experiences: A Meeting Report. *CBE-Life Sci. Educ.* **2014**, *13* (1), 29-40.

(12) Dolan, E. L. *Course-based Undergraduate Research Experiences: Current knowledge and future directions;* Commissioned Paper by the Committee on Strengthening Research Experiences for Undergraduate STEM Students for the National Research Council: Washington, DC, 2016.

(13) Corwin, L. A.; Graham, M. J.; Dolan, E. L. Modeling Course-Based Undergraduate Research Experiences: An Agenda for Future Research and Evaluation. *CBE-Life Sci. Educ.* [Online] **2015**, *14* (1), Article es1. https://doi.org/10.1187/cbe.14-10-0167 (accessed Nov 25, 2019).

(14) Fang, F. C.; Casadevall, A. Lost in translation--basic science in the era of translational research. *Infect. Immun.* **2010**, *78* (2), 563-566.

(15) National Institutes of Health. Definitions of Criteria and Considerations for Research Project Grant (RPG/R01/R03/R15/R21/R34) Critiques. https://urldefense.proofpoint.com/v2/url?u=https-3A__grants.nih.gov_grants_peer_critiques_rpg.htm&d=DwIFaQ&c=slrrB7dE8n7gBJbeO0g-IQ&r=GwZFjvFDPtYAfWCmd94Sbg&m=AKF7KdqYefv2RBcQ6-gRLhSYVYlJl5L5yDvfxtAF0nY&s=FCbiK1Cbmyw5f9MD3Hr0q2nobd7rtg7oox9nvzuMK3U&e= (accessed Nov 25, 2019).

(16) Parker, R. Skill development in graduate education. *Mol. Cell* **2012**, *46* (4), 377-381.

(17) Majerus, P. W. Inositol phosphate biochemistry. *Annu. Rev. Biochem.* **1992**, *61*, 225-250.

(18) Schlessinger, J. Phospholipase Cγ activation and phosphoinositide hydrolysis are essential for embryonal development. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (7), 2798-2799.

(19) Bae, J. H.; Lew, E. D.; Yuzawa, S.; Tome, F.; Lax, I.; Schlessinger, J. The selectivity of receptor tyrosine kinase signaling is controlled by a secondary SH2 domain binding site. *Cell* **2009**, *138* (3), 514-524.

(20) Huang, Z.; Marsiglia, W. M.; Basu Roy, U.; Rahimi, N.; Ilghari, D.; Wang, H.; Chen, H.; Gai, W.; Blais, S.; Neubert, T. A.; Mansukhani, A.; Traaseth, N. J.; Li, X.; Mohammadi, M. Two FGF Receptor Kinase Molecules Act in Concert to Recruit and Transphosphorylate Phospholipase Cγ. *Mol. Cell* **2016**, *61* (1), 98-110.

(21) Goddard, T. D.; Kneller, D. G. Sparky—NMR Assignment and Integration Software, UCSF Sparky/SPARKY 3. https://urldefense.proofpoint.com/v2/url?u=https-3A__www.cgl.ucsf.edu_home_sparky_&d=DwIFaQ&c=slrrB7dE8n7gBJbeO0g-IQ&r=GwZFjvFDPtYAfWCmd94Sbg&m=AKF7KdqYefv2RBcQ6-gRLhSYVYlJl5L5yDvfxtAF0nY&s=Yq4s4tW1ZvtpkHQ0e85Gc3p4T__zYR3YPagfSbUEUW4&e= (accessed Nov 25, 2019).

(22) DeLano, W. L. The PyMOL Molecular Graphics System. https://urldefense.proofpoint.com/v2/url?u=http-3A__pymol.sourceforge.net_overview_index.htm&d=DwIFaQ&c=slrrB7dE8n7gBJbeO0g-IQ&r=GwZFjvFDPtYAfWCmd94Sbg&m=AKF7KdqYefv2RBcQ6-gRLhSYVYlJl5L5yDvfxtAF0nY&s=ZIsbw2QsjUFVefb2eL77rM0gcHalVePeJpWBLR37MeM&e= (accessed Nov 3, 2019).

(23) Cooper, K. M.; Soneral, P. A. G.; Brownell, S. E. Define Your Goals Before You Design a CURE: A Call to Use Backward Design in Planning Course-Based Undergraduate Research Experiences. *J. Microbiol. Biol. Educ.* [Online] **2017**, *18* (2). doi:10.1128/jmbe.v18i2.1287 (accessed Nov 25, 2019).

(24) Shortlidge, E. E.; Bangera, G.; Brownell, S. E. Faculty Perspectives on Developing and Teaching Course-Based Undergraduate Research Experiences. *Bioscience* **2016**, *66* (1), 54-62.

(25) Thompson, S. K.; Neill, C. J.; Wiederhoeft, E.; Cotner, S. A Model for a Course-Based Undergraduate Research Experience (CURE) in a Field Setting. *J. Microbiol. Biol. Educ.* **2016**, *17* (3), 469-471.

(26) White, H. B.; Benore, M. A.; Sumter, T. F.; Caldwell, B. D.; Bell, E. What Skills Should Students of Undergraduate Biochemistry and Molecular Biology Programs Have Upon Graduation? *Biochem. Mol. Biol. Educ.* **2013**, *41* (5), 297-301.

(27) Shortlidge, E. E.; Eddy, S. L. The trade-off between graduate student research and teaching: A myth? *PLoS One* [Online] **2018**, *13* (6), Article e0199576. https://doi.org/10.1371/journal.pone.0199576 (accessed Nov 25, 2019).

(28) Jeffery, E.; Nomme, K.; Deane, T.; Pollock, C.; Birol, G. Investigating the Role of an Inquiry-Based Biology Lab Course on Student Attitudes and Views toward Science. *CBE-Life Sci. Educ.* [Online] **2016**, *15* (4), Article ar61. https://doi.org/10.1187/cbe.14-11-0203 (accessed Nov 25, 2019).

(29) Craig, P. A. Lessons from my undergraduate research students. *J. Biol. Chem.* **2018**, *293* (27), 10447-10452.

SUPPORTING INFORMATION

A CURE Biochemistry Laboratory Module to Study Protein-Protein Interactions by NMR Spectroscopy

William M. Marsiglia[1], Rohini Qamra[1], Kimberly M. Jackson[2], and Nathaniel J. Traaseth[1*]

[1]Department of Chemistry, New York University, New York, NY 10003

[2]Department of Chemistry and Biochemistry, Spelman College, Atlanta, GA 30314

*Corresponding author: traaseth@nyu.edu

Purification and NMR analysis of an SH2 domain

A. Cation-Exchange Chromatography

1. Obtain specified amounts of the following buffers:
   a. 25 mM HEPES pH 6.5 — 30 mL
   b. 25 mM HEPES pH 7.5 — 10 mL
   c. 25 mM HEPES + 150 mM NaCl, pH 7.5 — 15 mL
   d. 25 mM HEPES + 1 M NaCl pH 6.5 — 10 mL
2. Place 7 borosilicate test tubes in a rack and label them: FT, 0 mM pH 6.5 #1, 0 mM pH 6.5 #2, 0 mM pH 7.5, 150 mM pH 7.5 #1, 150 mM pH 7.5 #2, 150 mM pH 7.5 #3.
3. Prepare the cation exchange column (GE Healthcare fast-flow sepharose, part number 17-5054-01) by unscrewing the caps from the top and bottom of the column.
4. Screw on a small adaptor to a 20 mL Luer lock syringe and pass 5 mL of 25 mM HEPES through the column. The flow rate should not exceed 5 drops per second.
5. Pass 5 mL of 25 mM HEPES + 1 M NaCl pH 6.5 through the column to charge it.
6. Pass 10 mL of 25 mM HEPES pH 6.5 through the column to equilibrate.
7. Obtain 5 mL of cell lysate from your instructor and take a 10 µL sample into a 1.5 mL microcentrifuge tube and mix with 10 µL of 2x sample buffer. Set this sample to the side.
8. Pass the remaining lysate through the column using a syringe and collect the flow-through into the tube labeled “FT.”
9. Pass the following buffers through the column with increasing ionic strength. **Collect each in separate labeled tubes.**

5 mL	25 mM HEPES pH 6.5,	**0 mM** NaCl
5 mL	25 mM HEPES pH 6.5,	**0 mM** NaCl
5 mL	25 mM HEPES pH 7.5,	**0 mM** NaCl
5 mL	25 mM HEPES pH 7.5,	**150 mM** NaCl
5 mL	25 mM HEPES pH 7.5,	**150 mM** NaCl
5 mL	25 mM HEPES pH 7.5,	**150 mM** NaCl

B. SDS-PAGE Gel

1. Take 10 µL samples of the 150 mM pH 7.5 #1, 150 mM pH 7.5 #2, 150 mM pH 7.5 #3 elution into separate labeled microcentrifuge tubes.
2. Add 10 µL of 2x sample buffer to each microcentrifuge tube making sure to change pipette tips in between.
3. Load 10 µL of the lysate, and 20 µL each of the elution fraction samples onto a gel.
4. Run the gel under constant voltage until the gel front approaches the bottom of the gel.
5. Place the gel in a suitable container with Coomassie brilliant blue G and microwave the gel for 45 sec. Remove the stain and add ddH_2O. Place in the microwave for an additional four minutes. Additional cycles can be used until the gel is sufficiently destained such that protein bands can be resolved.
6. Analyze the gel by identifying the fractions corresponding to the cSH2 domain that will appear at 13.4 kDa relative to a protein ladder.

C. NMR Sample Preparation

1. Combine fractions of 25 mM HEPES + 150 mM NaCl (pH 7.5) in a 15 mL 3kDa centrifuge concentrator. Add 5 µL of 100 mM PMSF dissolved in isopropanol. (* see Hazards).

2. Spin in 15 min increments at 2,488 x g at 4°C until the volume is reduced to 500 µL. Gently invert the concentrator after every spin.
3. Resuspend the protein after the final spin by pipetting up-and-down while being careful not to touch the membrane in the concentrator. Transfer the sample from the concentrator into a microcentrifuge tube on ice.
4. Determine the protein concentration using absorbance at 280nm (ε=15,930 M^{-1} cm^{-1}). Make sure to use the 25 mM HEPES + 150 mM NaCl (pH 7.5) supplemented with PMSF as a blank.
5. Add 10 uL of D_2O into the microcentrifuge tube.
6. (**DAY OF NMR**) Pipette the sample into a 5 mm rounded-bottom NMR tube. Place a rubber cap on the top of the tube.

D. Cation exchange column cleaning

1. Pass over the column using a syringe, 5 mL 1 M NaOH, 5 mL ddH_2O, and 5 mL 20 % ethanol. (* see Hazards)
2. Replace the screw caps on both ends of the column and return to your instructor.

E. NMR Spectroscopy

1. Acquire an HSQC on the sample. (* see Hazards)
2. Add 1.5 molar-equivalent (~10 uL) of phosphopeptide (from 30 mM stock in 25 mM HEPES + 150 mM NaCl, pH 7.5) to the inside of the NMR tube (near the top).
3. Replace the cap and invert a few times to mix.
4. Acquire a second HSQC.

F. Analysis of Chemical Shift Perturbations

1. Analyze chemical shifts using Topspin and Sparky software for Bruker and Agilent/Varian spectrometers, respectively. Note that students should be introduced to these software packages by a tutorial (see Instructor Notes).
2. Chemical shift assignments in the phosphopeptide bound form will be transferred based on the assignment of the unbound cSH2 domain [1]. Calculate chemical shift perturbations (CSPs) using Equation 1, where $\Delta\delta_N$ and $\Delta\delta_H$ are the change in ^{15}N and ^{1}H chemical shifts, respectively.

$$CSP = \sqrt{(0.154\,\Delta\delta_N)^2 + \Delta\delta_H^2} \qquad (1)$$

*** Hazards**

PMSF, NaOH, and acrylamide are hazardous chemicals used in the lab. The acrylamide is applicable if students polymerize their own SDS-PAGE gels. If PMSF powder is weighed out, students should wear a mask. All other personal protective equipment such as gloves, lab coats, and goggles should be worn as mandated in undergraduate laboratories. For NMR experiments, high magnetic fields can be harmful to those with implanted devices such as pacemakers. Students with these devices should maintain a safe distance from the magnet as instructed by the manager of the NMR equipment.

Instructor Notes

- Preparation of ^{15}N labeled E. coli lysate. Residues N661 to L774 of rat PLCγ (Accession code: NP037319-1) was cloned in the pET41b+ plasmid (gift from Moosa Mohammadi). E. coli BL21(DE3) cells were transformed with the plasmid. Single colonies were used to inoculate a 1 L LB culture and subsequently grown at 37 °C. When the optical density at 600 nm (OD600) reached 1.5, the cells were spun down for 10 min at 3,000 x g and resuspended in 3 L ^{15}N labeled M9 media with ^{15}N NH_4Cl. The cells were grown at 37 °C until an OD600 at 0.5. At this point, the cells were induced with a final IPTG concentration of 0.1 mM and grown for 16 hr at 20 °C. Following completion of the growth, the cells were harvested by centrifuging for 15 min at 6,000 x g. The pelleted cells were resuspended in 240 mL lysate buffer (20 mM sodium phosphate dibasic pH 6.5, 1.25 mM EDTA, 400 µM DTT, 1.45 µM pepstatin A, 2.3 µM leupeptin, 8.8 µM lysozyme, 5% glycerol, 600 µM PMSF) and passed through a French pressure cell four times (1,000 psi). The cell lysate was spun down for 20 min at 70,600 x g. The supernatant of the lysate was stored at -80 °C until the day of the laboratory. Each group of students received 5 mL of this lysate supernatant.
- The day of the purification, stock buffers of HEPES and HEPES/NaCl can be prepared for students to further save time and avoid mistakes in buffer calculations.
- Use of the cation exchange column should follow the manufacturer guidelines. Students should be instructed not to pull volume into the syringe when connected to the column. This will pull the cation exchange resin from the column and will effectively break the column and reduce or eliminate the yield of the cSH2 protein.
- CAS numbers for chemicals used:

HEPES	7365-45-9
Sodium chloride	7647-14-5
Coomassie	6104-58-1
PMSF	329-98-6
Deuterium oxide	7789-20-0
Sodium hydroxide	1310-73-2
Ethanol	64-17-5

- The tutorial for NMR analysis in Sparky and plotted the chemical shift perturbations onto the structure in Pymol is given in the subsequent pages.

NMR Tutorial

1. Download Sparky software:

   https://www.cgl.ucsf.edu/home/sparky

2. Install Sparky and download NMR data to your computer. *MAC users need to install Xquartz first (https://www.xquartz.org).

3. Launch Sparky and open the two .ucsf files (i.e. +/- peptide).

Sparky Commands

1. In the Pointer window, click "select”.

2. Click in the SH2 spectrum with no peptide and type "ct". Change the settings in “Lowest” to display the signals. Change the “Positive” “Levels” to 10. Repeat for the peptide HSQC spectrum, except use the color yellow (positive).

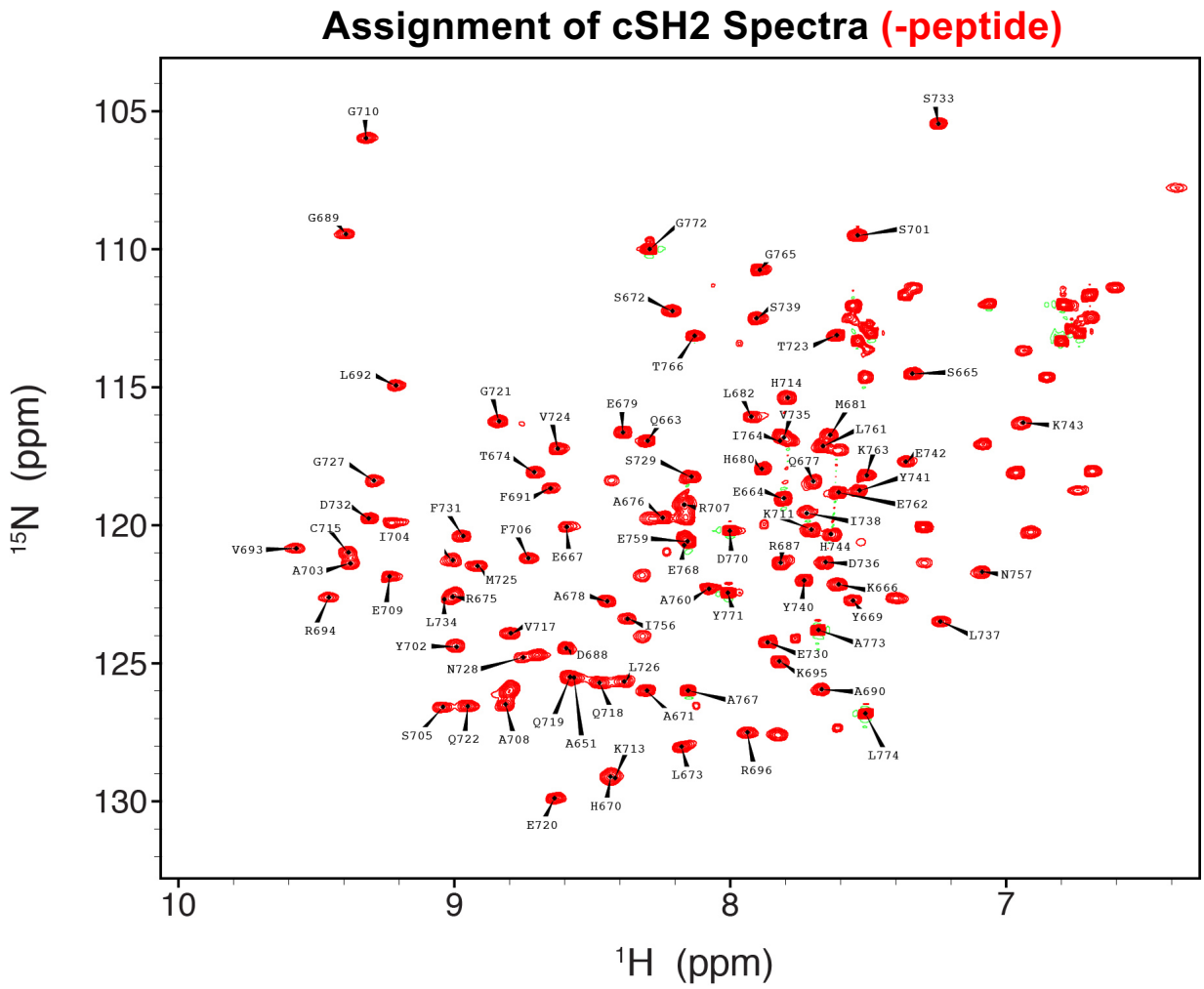
3. In the Pointer window, select "find/add peak". Drag your left cursor over the peaks. Several boxes should appear over the peaks.

4. Click "select" in the Pointer window.

5. Type "pa", then "oz". Adjust the “label size” to 0.500 and “peak size” to 0.250, then click "apply". This controls the font size of the peak labels.

6. Save the Sparky project by clicking "File" then under "Project", click "Save as". Type in your project: SH2-experiment.

7. Label all peaks in the apo spectrum using the assignments below. Click one peak label, then type "at". Name the peak by number; do not include the letter. For example: for G710, label it "710". After labeling, click "apply" and the peak will be labeled in your spectrum. Periodically save your work by clicking "Save" under "File" and "Project". Note that Sparky does not automatically save.



8. Once you finished labeling peaks in the apo spectrum, you can copy and paste the labels onto the +peptide spectrum. To do this, click the apo spectrum and type “pa”, then “oc”. Click the +peptide spectrum and type “op”. You will need to manually move the peaks to find your “best guess” movement for each peak. By clicking on each peak label, you can type “pc” which will automatically center the peak label on the peak – this works best when the peak label is closest to the peak of interest. To make the labels larger, use the instructions in step #5.

9. Create a spectral overlay for the two datasets. In the "select" mode in the pointer, click in one of spectra and type "ol". In the "From View", select your cSH2. In the "Onto View", select your peptide spectrum. Click "add". If you want to remove the spectrum, click "remove".

10. To export a spectrum, click in that spectrum and type “pt". Save the file as a postscript.

11. After you have finished labeling all peaks, export the peak list. In each spectrum, type "lt" then "Save". This will be a text file that can be importing in Excel.

12. Open Excel and import the two peak lists (one for –peptide and one for +peptide).

13. Carry out the following math to calculate the chemical shift perturbations for each residue:

$$\mathrm{CSP} = \Delta\delta = \sqrt{(0.154 \times \Delta\delta_N)^2 + \Delta\delta_H^2}$$

$\Delta\delta_N = \delta_{N,\ \text{no peptide}} - \delta_{N,\ +\ \text{peptide}}$
$\Delta\delta_H = \delta_{H,\ \text{no peptide}} - \delta_{H,\ +\ \text{peptide}}$

14. Note the residues with CSP values greater than 0.07 ppm. These will be highlighted onto the crystal structure using PyMOL.

PyMOL Software

1. Register and download PyMOL software

   https://pymol.org/edu/?q=educational

2. Install PyMOL

3. Open PyMOL

4. Download the PDB file 5EG3.pdb from the link below:

   https://www.rcsb.org

PyMOL Commands

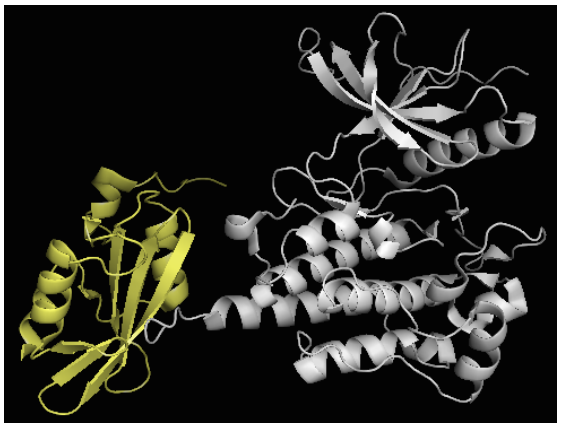
1. Open the 5EG3.pdb file in PyMOL. To display the protein structure in a "ribbon" format (sometimes called cartoon), type the following commands into the command line.

a. Type "hide"
b. Type "show cartoon"

2. To distinguish the kinase from the SH2 domain by color, type the following into the command line.

a. Type "select kinase, chain A"
b. Type "select SH2, chain B"
c. Type "color grey90, kinase"
d. Type "color paleyellow, SH2"

3. Orient the structure as shown below. Note that kinases are typically shown in the "Pac-Man" orientation. To rotate the structure, left click and drag. To zoom-in for a closer representation of the structure, right click and drag.



SH2 (yellow)

kinase (white)

4. Under the “Display” pull down menu, select “Sequence on”
Move the scroll bar to the right to see the SH2 numbering in yellow.
These numbers correspond to the NMR peak assignments.

5. To select residues with CSP values greater than 0.07 (calculations you previously performed), carry out the following steps:

a. Click residues on the sequence with CSP values > 0.07.
b. Type “color red, sele”
c. Then go to File → Save Image As ... → PNG

REFERENCES

(1) Huang, Z.; Marsiglia, W. M.; Basu Roy, U.; Rahimi, N.; Ilghari, D.; Wang, H.; Chen, H.; Gai, W.; Blais, S.; Neubert, T. A.; Mansukhani, A.; Traaseth, N. J.; Li, X.; Mohammadi, M. *Mol. Cell* **2016**, *61*, 98-110.