JOURNAL OF CHEMICAL EDUCATION

Laboratory Experiment

pubs.acs.org/jchemeduc

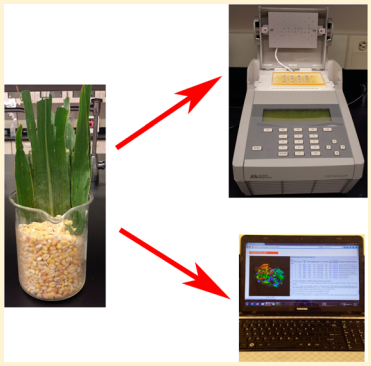
Detection of the *cp4 epsps* Gene in Maize Line NK603 and Comparison of Related Protein Structures: An Advanced Undergraduate Experiment

Nicole K. Swope,* Patrick J. Fryfogle, and Tami L. Sivy

Department of Chemistry, Saginaw Valley State University, University Center, Michigan 48710, United States

Supporting Information

ABSTRACT: A flexible, rigorous laboratory experiment for upper-level biochemistry undergraduates is described that focuses on the Roundup Ready maize line. The work is appropriate for undergraduate laboratory courses that integrate biochemistry, molecular biology, or bioinformatics. In this experiment, DNA is extracted and purified from maize kernel and leaf samples collected from a Roundup Ready maize grower's field. A small segment of DNA (108 base pairs) specific to the Roundup Ready transgene that codes for CP4 5-enolpyruvylshikimate-3-phosphate synthase (CP4 EPSPS) is amplified with polymerase chain reaction (PCR) to detect the presence of the gene in the maize samples. Students additionally choose a protein closely related to CP4 EPSPS as determined by amino acid sequence alignments. The selected amino acid sequences are submitted to an online protein modeling program where students compare their protein with the herbicide-resistant enzyme found in Roundup Ready crops. This experimental paradigm gives students a physical appreciation for the central dogma of biology, as they are exposed to products derived from the replication, transcription, and translation events belonging to a genetically modified crop. The PCR portion of the laboratory allows students to perform an in vitro replication of a portion of the NK603 transgene and identify the segment via ultraviolet radiation. In direct connection to the PCR portion, the protein structure elucidation gives rise to the central ideas of evolution—that slight changes in the genetic code of DNA, translated into proteins, produce novel protein structures with significantly different function.



KEYWORDS: *Upper-Division Undergraduate, Biochemistry, Interdisciplinary/Multidisciplinary, Computer-Based Learning, Hands-On Learning/Manipulatives, Molecular Biology, Molecular Modeling*

With the controversy surrounding consumption and growth of genetically modified crops around the world, biochemistry and molecular biology students have expressed increased interest in exploring genetically modified vegetables at the molecular level. Previous undergraduate laboratory experiments have introduced students to methods, such as standard polymerase chain reaction (PCR), for detecting the presence of a common transgene promoter (cauliflower mosaic virus) in corn-meal items[1] and soybean powder.[2] Some computer-based bioinformatics work is integrated into the corn-meal experiment, which explores weed resistance to herbicides in addition to identifying transgenic proteins that are potential allergens. While more in-depth bioinformatics exercises exploring PCR with various genes[3,4] and computational modeling of proteins[5−8] can be found in this *Journal*, these experiments often lack comprehensive gene-to-protein studies from the same genetic source. The present experiment offers students the unique opportunity to detect not only the transgene of a genetically modified crop, but also to assess the structural characteristics of the protein translated from the gene of interest using a combination of bioinformatics and computational-modeling techniques.

In this experiment, students study the popular Roundup Ready (RR) maize line developed by Monsanto Co., which is resistant to Roundup (N-phosphonomethyl glycine, glyphosate) through the stable insertion of a gene from *Agrobacterium* sp. strain CP4.[9] The inserted gene sequence expresses glyphosate-tolerant CP4 5-enolpyruvylshikimate-3-phosphate synthase (CP4 EPSPS, CP4 EPSP synthase), which is also referred to as CP4 3-phosphoshikimate 1-carboxyvinyltransferase. Plants and microorganisms naturally contain EPSP synthase, an enzyme that aids in the biosynthesis of aromatic amino acids. Glyphosate irreversibly binds to EPSP synthases found in many plants and microorganisms, which ultimately prevents proper protein synthesis and leads to the organism's death.[10] Class II EPSP synthases, including CP4 EPSPS, are considered resistant to glyphosate most notably due to the amino acid mutation Gly100Ala (numbering based on CP4 EPSPS) in which the methyl group of alanine prevents glyphosate from binding effectively to the enzyme.[11] Farmers that grow RR vegetables containing CP4 EPSPS are, therefore, able to treat their fields with Roundup, eliminating nonresistant weeds while having no adverse effects on their crops.

This laboratory experiment aims to utilize a combination of PCR, amino acid sequence databases, and protein modeling

DOI: 10.1021/ed500655j
J. Chem. Educ. XXXX, XXX, XXX−XXX

software to provide students with the opportunity to study the RR maize line at the molecular level. The PCR portion of this experiment acquaints students with methods to detect the *cp4 epsps* gene, while computational modeling of its translated protein allows students to visualize the structure of CP4 EPSPS as well as the characteristics that distinguish CP4 EPSPS from EPSP synthases found in wild-type species.

Upon completion of this experiment, students will be able to describe the importance of PCR for identifying RR maize such as its detection in farmers' fields and food products. Aside from performing traditional wet lab molecular biology, students will be able to navigate the widely used National Center for Biotechnology Information (NCBI) databases and analyze amino acid sequences. This experiment last focuses on teaching students how to use programs for predicting protein structure and function, such as the free online server I-TASSER,[12] to form a visual tie-in to sequence comparisons. Overall, the student learning during this experiment is assessed with questions provided in the student handout section of the Supporting Information in addition to in-class discussions and participation.

EXPERIMENTAL DETAILS

This laboratory experiment is conducted over four, 3 h class periods. Advanced biochemistry undergraduate students performed this experiment four times in a laboratory course during the 2013−2014 academic year. It is recommended that students work in groups of two or three for this experiment.

As a prelaboratory assignment, students choose a protein or enzyme that contains a related amino acid sequence to CP4 EPSPS (GenBank accession no. AAO17037.1) by using the NCBI Protein BLAST feature. Once students choose a protein whose sequence is 25−85% identical to CP4 EPSPS, the selected protein's amino acid sequence in FASTA format is submitted to the I-TASSER server, an online program that predicts protein structure and function. They are notified when the predictions are finished; however, in-depth model analyses are set aside until the PCR portion of this experiment is complete.

On the first day, students randomly select either one kernel or a leaf sample to extract and purify its DNA using the *Qiagen* DNeasy Blood and Tissue kit. Only RR maize samples were provided by the instructor for this experiment. During the second laboratory period, students use standard PCR with primers validated by Monsanto Co. to amplify a 108 bp DNA sequence spanning the 3′ end of the RR gene insert to the maize genome.[13] After beginning the PCR program, students take advantage of the remaining class time to explore I-TASSER. The third laboratory period includes analyzing the PCR samples with agarose gel electrophoresis. Following electrophoresis, students are free to continue working with I-TASSER. If predictions are complete, students begin to analyze their models and address any confusing elements at this time.

During the final day, students perform the majority of their work with I-TASSER to compare the 3D structures of their selected proteins with CP4 EPSPS, including analyses of the core, binding sites, and functional residues of each protein. The major similarities that all students should observe while comparing their proteins are overall structural characteristics such as domain number and exposed or buried secondary structures. Likewise, students locate conserved and mutated residues and assess their importance in the proteins' structure and function. The ultimate goal is for students to conclude whether or not their selected protein may be resistant to glyphosate based on their analyses. In-depth experimental details can be found in both the student handout and instructor notes provided in the Supporting Information.

HAZARDS

When handling the components used for DNA extraction and PCR, such as buffers, proteinase K, ethanol, PCR mix, and agarose, goggles and gloves should be worn at all times. Ethanol is flammable and should be handled with care. If used for agarose gel staining, ethidium bromide is a hazardous mutagen that should be handled with extra precaution;[14] however, instructors are encouraged to use safer alternatives if they are available. Whenever working with ethidium bromide in the laboratory, goggles, gloves, and lab coats should be worn. If using ultraviolet (UV) technology, protective eyewear should be worn in the event that the instrument does not have a shield or viewing screen.

RESULTS AND DISCUSSION

The PCR results obtained from the *cp4 epsps* gene, along with studying the translated protein, provide a snapshot view of the central dogma of biology. The key connection between the PCR experiment and the modeling exercise is that the gene associated with the PCR fragment for the CP4 enzyme is then transferred to an RNA transcript, available in the NCBI database,[15] from which the enzyme is translated. From the known CP4 EPSPS protein sequence, students modeled the resulting enzyme, thereby establishing a framework of the central dogma in transgenic species from first-principles: replication, transcription, and translation. In conjunction with establishing this framework, comparison of CP4 EPSPS with wild-type EPSP synthase allowed students to observe the changes in CP4 EPSPS that result in its resistance to Roundup.

Students completed the prelaboratory assignment prior to the first day of bench work and made standard Protein BLAST searches to find a comparable enzyme to CP4 EPSPS. The predictions took approximately 48−72 h to be completed by I-TASSER, after which students were able to access their protein models.

Example student results of the amplified 108 bp DNA segment specific to RR maize following PCR and agarose gel electrophoresis are shown in Figure 1. All three leaf samples and one kernel sample contain PCR products belonging to the target sequence, which are typical results for this experiment. The identity of the PCR product corresponding to the 500 bp marker in lanes 2, 3, and 4 is unknown. The relative intensities of the 108 bp bands in lanes 2−5 correspond to the success of PCR for each sample, with lane 5 giving the strongest amplification product. Overall, 17 of the 23 PCR reactions tested by students were successful in amplifying the desired gene.

The two samples that did not contain the desired PCR products were most likely due to lack of the target sequence in the DNA template or error during DNA extraction and PCR, as the other samples clearly show amplification of the target sequence. Successful PCR amplification from kernels was notably less compared to leaves, which may be from mechanical digestion that is not sufficiently thorough for intact kernels. To increase the concentration of extracted kernel DNA available for PCR, soaking of the maize kernels in phosphate buffer overnight followed by glass bead homogenization would likely

DOI: 10.1021/ed500655j
J. Chem. Educ. XXXX, XXX, XXX−XXX

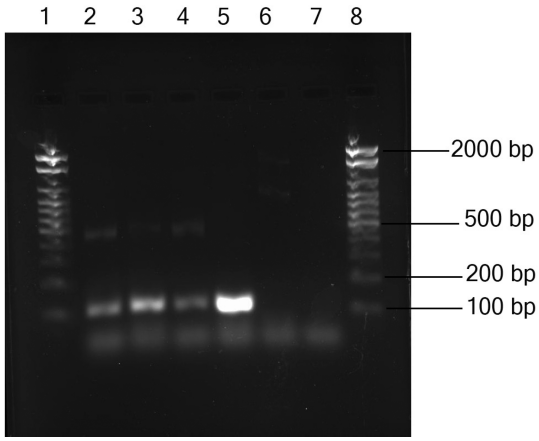


Figure 1. Representative student UV visualization of amplified NK603 PCR products on a 1.2% agarose gel stained with ethidium bromide using isolated DNA from maize leaf and kernel samples. Lanes 1 and 8, DNA ladder; lanes 2−4, maize leaf DNA samples; lanes 5−7, maize kernel DNA samples.

improve results. Student discussion of the results included an assessment of which samples were RR and analysis of differences in band brightness and base-pair size evident on stained agarose gels. It was suggested that unexpected bands could be due to primer dimers or nonspecific primers. In addition to answering these questions, students brainstormed examples of how they could use the DNA extraction and PCR protocols in this experiment in an agriculture or food industry setting.

After finishing the PCR experiment, students focused on completing their I-TASSER model analyses. One student chose to compare CP4 EPSPS (Figure 2a) with *Zea mays* EPSP synthase (Figure 2b). Other EPSP synthases investigated by students primarily belonged to bacteria species. Students observed that CP4 EPSPS, *Zea mays* EPSP synthase, and other similar proteins have two domains and fold with β-sheets covering α-helices in a distinct pattern based on the global views of the proteins. Each I-TASSER prediction, furthermore, includes amino acid residues that potentially play a role in substrate binding or catalytic activity. By isolating the binding and catalytic amino acid residues, students directly observed similarities and differences between the residues associated with the ligand binding sites. Less variation among these residues was seen in proteins that shared high sequence identities with CP4 EPSPS, while those that contained lower identities exhibited more variation. When comparing CP4 EPSPS to a class I EPSP synthase, like *Zea mays* EPSP synthase, students noted important mutations, such as Gly100Ala (Figure 2) and others, that possibly account for glyphosate resistance in class II EPSP synthases.

After aligning their amino acid sequences and completing protein model analyses, students described the relationship between the percent identities of the proteins. In the case of *Zea mays* EPSP synthase and CP4 EPSPS, it was found that the aligned amino acid sequences exhibited a 26% identity, with 44% of the total amino acids being at least similar. Using evidence from sequence alignments and I-TASSER models, most students determined that EPSP synthases naturally found in the organisms that they investigated possess functional and structural similarity to CP4 EPSPS despite their intolerance to glyphosate. Out of all models tested by students, 20 of the 23 attempts at protein modeling were successful. The three unsuccessful protein models were generated from 3-phosphoshikimate 1-carboxyvinyltransferases of the species *Rhizobium leguminosarum*, *Ochtrobactrum intermedium*, and *Agrobacterium rhizogenes*. Because these protein models were almost identical to CP4 EPSPS in terms of structure and predicted binding site and functional residues, students found it difficult to conclude that they were susceptible to glyphosate, even with the Gly100Ala mutation. To avoid protein models that are easily mistaken by students for being glyphosate-resistant, a suggestion for instructors implementing this experiment is to modify the percent identity range of student-selected proteins to 25−75%.

CONCLUSION

Overall, students that performed this experiment were pleased with its relevance to genetically modified crops and the biotechnology field. Students enjoyed completing the bio-

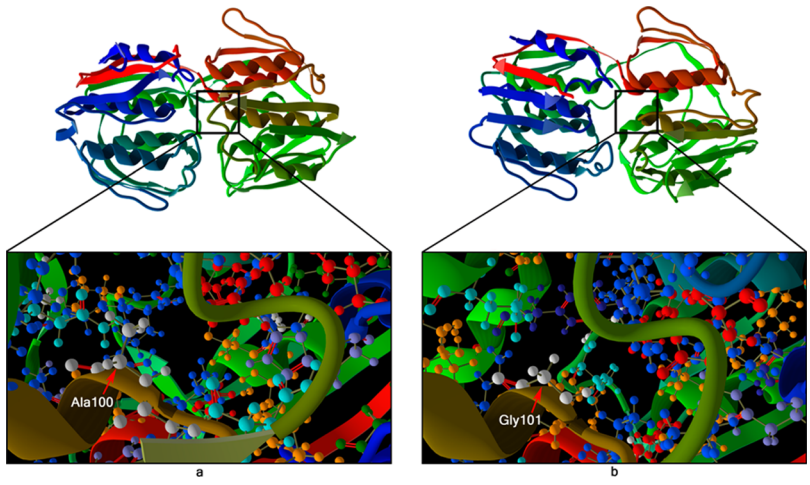


Figure 2. Overall protein structures with zoomed area of the binding sites for (a) CP4 EPSPS and (b) *Zea mays* EPSP synthase as predicted by the I-TASSER server. Visualizing the key mutation from Gly101 in *Zea mays* EPSP synthase to Ala100 in CP4 EPSPS, depicted on the models, in addition to analyzing the structural differences and associated residues in the binding site allowed students to determine why RR maize is resistant to glyphosate and wild-type maize is not.

DOI: 10.1021/ed500655j
J. Chem. Educ. XXXX, XXX, XXX−XXX

informatics and protein modeling exercises; they stated that their assignments with Protein BLAST and I-TASSER complemented the hands-on laboratory work. It was generally agreed that the visual aids and independently generated sequence comparisons offered a complete "picture" of RR maize that greatly enhanced their knowledge of the subject. One student explained that "Everybody always talks about issues with GMOs and DNA modification, but I rarely hear about differences or similarities in the proteins that are translated from the genes. This experiment was helpful because it explored both the gene and protein of a GMO, which are equally important."

The learning goals of this experiment were to (1) provide students with an introduction to molecular biology techniques used to detect genetically modified crops, specifically RR maize, (2) familiarize students with easily accessible bioinformatics and molecular modeling resources, such as the NCBI databases and I-TASSER, and (3) allow students to determine the differences between a protein found in a GMO and wild-type organisms. All three learning objectives were met with excellent results as demonstrated by the students that performed this experiment, and they would recommend this laboratory exercise to future biochemistry students.

■ ASSOCIATED CONTENT

Supporting Information

Materials, hazards, student handout, and instructor notes. This material is available via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: nkswope@svsu.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was funded by Saginaw Valley State University's Chemistry Department. The authors would like to thank David Karpovich for donating maize samples and David Stanton for providing expertise and laboratory equipment.

■ REFERENCES

(1) Taylor, A.; Sajan, S. Testing for Genetically Modified Foods Using PCR. *J. Chem. Educ.* **2005**, *82* (4), 597−598.

(2) Laurent, T.; Vossen, C.; Couderc, B.; Erard, M.; Clemençon, B. Detection of genetically modified organisms in food by DNA extraction and PCR amplification. *Biochem. Mol. Biol. Educ.* **2002**, *30* (1), 51−55.

(3) Millard, J. T.; Chuang, E.; Lucas, J. S.; Nagy, E. E.; Davis, G. T. Case-Study Investigation of Equine Maternity via PCR-RFLP: A Biochemistry Laboratory Experiment. *J. Chem. Educ.* **2013**, *90* (11), 1518−1521.

(4) Elkins, K. M. Designing Polymerase Chain Reaction (PCR) Primer Multiplexes in the Forensic Laboratory. *J. Chem. Educ.* **2011**, *88* (10), 1422−1427.

(5) Lowery, M. S.; Plesniak, L. A. Some Like It Cold: A Computer-Based Laboratory Introduction to Sequence and Tertiary Structure Comparison of Cold-Adapted Lactate Dehydrogenases Using Bioinformatics Tools. *J. Chem. Educ.* **2003**, *80* (11), 1300−1302.

(6) Peterson, R. R.; Cox, J. R. Integrating Computational Chemistry into a Project-Oriented Biochemistry Laboratory Experience: A New Twist on the Lysozyme Experiment. *J. Chem. Educ.* **2001**, *78* (11), 1551−1555.

(7) Yuriev, E.; Chalmers, D.; Capuano, B. Conformational Analysis of Drug Molecules: A Practical Exercise in the Medicinal Chemistry Course. *J. Chem. Educ.* **2009**, *86* (4), 477−478.

(8) Hayes, J. M. An Integrated Visualization and Basic Molecular Modeling Laboratory for First-Year Undergraduate Medicinal Chemistry. *J. Chem. Educ.* **2014**, *91* (6), 919−923.

(9) Lau, L.; Collins, R. A.; Yiu, S.; Xing, J.; Yu, A. C. Detection and Characterization of Recombinant DNA in the Roundup Ready Soybean Insert. *Food Control* **2004**, *15* (6), 471−478.

(10) Li, L.; Lu, W.; Han, Y.; Ping, S.; Zhang, W.; Chen, M.; Zhao, Z.; Yan, Y.; Jiang, Y.; Lin, M. A Novel RPMXR Motif among Class II 5-Enolpyruvylshikimate-3-phosphate Synthases Is Required for Enzymatic Activity and Glyphosate Resistance. *J. Biotechnol.* **2009**, *144* (4), 330−336.

(11) Eschenburg, S.; Healy, M. L.; Priestman, M. A.; Lushington, G. H.; Schönbrunn, E. How the Mutation Glycine96 to Alanine Confers Glyphosate Insensitivity to 5-Enolpyruvylshikimate-3 phosphate Synthase from *Escherichia coli*. *Planta* **2002**, *216* (1), 129−135.

(12) Zhang, Y. I-TASSER Server for Protein 3D Structure Prediction. *BMC Bioinf.* **2008**, *9* (1), No. 40.

(13) Recommended Procedure for Real-Time Quantitative TaqMan PCR for Roundup Ready Corn NK603. *Monsanto Biotechnology Regulatory Sciences*; Monsanto Company: St. Louis, MO, 2004. http://www.monsanto.com/products/documents/dna-detection/corn_dna_dmnk603.pdf (accessed January 2014).

(14) Ethidium bromide solution; SDS No. E1510. *Product Specification*; Sigma-Aldrich: Saint Louis, MO, 2014.

(15) *Synthetic Construct CP4EPSPS Protein (CP4EPSPS) Gene, Complete CDS*; The National Center for Biotechnology Information: Bethesda, MD, 2014. http://www.ncbi.nlm.nih.gov (accessed May 2014).

DOI: 10.1021/ed500655j
J. Chem. Educ. XXXX, XXX, XXX−XXX