

The Genetic Code Kit: An open-source cell-free platform for biochemical and biotechnology education

1 Layne C. Williams^{# 1,2}, Nicole E. Gregorio^{# 1,2}, Byungcheol So^{1,2}, Wesley Y. Kao^{1,2}, Alan L.
2 Kiste¹, Pratish A. Patel³, Katharine R. Watts^{1,2*}, Javin P. Oza^{1,2*}

3
4 ¹Chemistry and Biochemistry Department, California Polytechnic State University, San Luis Obispo,
5 CA, USA

6 ²Center for Applications in Biotechnology, California Polytechnic State University, San Luis Obispo,
7 CA, USA

8 ³Department of Finance, Orfalea College of Business, California Polytechnic State University, San
9 Luis Obispo, CA, USA

10
11 # These authors contributed equally to this work

12
13 * Correspondence:

14
15 Katharine R. Watts
16 krwatts@calpoly.edu

17 Javin P. Oza
18 joza@calpoly.edu

19
20 **Keywords:** Biochemical education¹, Learn by Doing² Cell-free protein synthesis (CFPS)³, In
21 vitro transcription and translation⁴, Synthetic biology (synbio)⁵, Central Dogma of Molecular
22 Biology (CDMB)⁶, Chemical education and teaching⁷, Augmented Reality (AR)⁸.

23
24 Words: 6187, Figures: 4, Tables: 2

25 **Abstract**

26 Teaching the processes of transcription and translation is challenging due to the intangibility
27 of these concepts and a lack of instructional, laboratory-based, active learning modules. Harnessing
28 the genetic code *in vitro* with cell-free protein synthesis (CFPS) provides an open platform that
29 allows for the direct manipulation of reaction conditions and biological machinery to enable inquiry-
30 based learning. Here, we report our efforts to transform the research-based CFPS biotechnology into
31 a hands-on module called the “Genetic Code Kit” for implementation into teaching laboratories. The
32 Genetic Code Kit includes all reagents necessary for CFPS, as well as a laboratory manual, student
33 worksheet, and augmented reality activity. This module allows students to actively explore
34 transcription and translation while gaining exposure to an emerging research technology. In our
35 testing of this module, undergraduate students who used the Genetic Code Kit in a teaching
36 laboratory showed significant score increases on transcription and translation questions in a post-lab
37 questionnaire compared with students who did not participate in the activity. Students also
38 demonstrated an increase in self-reported confidence in laboratory methods and comfort with CFPS,
39 indicating that this module helps prepare students for careers in laboratory research. Importantly, the
40 Genetic Code Kit can accommodate a variety of learning objectives beyond transcription and
41 translation and enables hypothesis-driven science. This opens the possibility of developing Course-
42 Based Undergraduate Research Experiences (CUREs) based on the Genetic Code Kit, as well as
43 supporting next-generation science standards in 8-12th grade science courses.

44 **1 Introduction**

45 Transcription and translation are fundamental cellular processes typically taught in high
46 school and undergraduate science courses and utilized extensively in research settings. As such,
47 students are expected to have an intimate grasp of these concepts to support both their academic and
48 career goals. However, there is evidence that misconceptions about transcription and translation often
49 persist for students even after they have completed these courses (Wright et al., 2014; Newman et al.,
50 2016; Queloz et al., 2017). This issue likely stems from the intangibility of the microscopic processes
51 of the “central dogma” when taught through lecture alone. In the absence of active learning modules,
52 students are unable to visualize and represent these processes for further learning (Kozma et al.,
53 2000; Duncan and Reiser, 2007). To address these limitations and allow students to interact with the
54 individual steps of transcription and translation in the classroom, a variety of model-, analogy-, and
55 virtual- based simulations have been developed (Pigage, 1991; Rotbain et al., 2008; Altiparmak and
56 Nakiboglu Tezer, 2009; Debruyn, 2012; Takemura and Kurabayashi, 2014; Marshall, 2017; Dorrell
57 and Lineback, 2019; Ibarra-Herrera et al., 2019; Chang et al., 2020) (Figure 1). Efforts to develop
58 such activities represent educators’ broad interest in providing students with active-learning modules
59 to improve student learning outcomes. However, chemistry and biology curricula generally rely on
60 laboratory practicals for active learning, as they help students connect scientific concepts and
61 practices. Unfortunately, current wet-lab procedures for teaching transcription and translation are
62 based on bacterial expression of fluorescent proteins, which precludes students from directly
63 accessing and manipulating the genetic code machinery (Ward et al., 2000; Bassiri, 2011; Newman
64 and Wright, 2013; Deutch, 2019) (Figure 1). While all these existing activities are generally low-cost
65 and useful learning tools to help students understand the broad scope and details of transcription and
66 translation, no single activity enables in-depth, hands-on, inquiry-based laboratory learning. The
67 limitations of existing approaches underscore the need for an active learning laboratory-based
68 module that allows students to interrogate transcription and translation in a learn-by-doing fashion.

69 Active learning has been demonstrated to increase student test scores and decrease the odds of
70 failing classes in science, technology, engineering, and mathematics (STEM) (Nogaj, 2013; Freeman
71 et al., 2014). In addition to these learning benefits, active learning is more engaging for students,
72 ultimately promoting positive attitudes towards their education (Armbruster et al., 2009). Prior work
73 also suggests that active learning may engage underrepresented students more than lecture-based
74 courses, helping to narrow the achievement gap in STEM courses (Haak et al., 2011; Theobald et al.,
75 2020). Curriculum design at our own university has led to the development of studio classrooms for
76 general chemistry courses, which integrate the laboratory and lecture portions of the course into one
77 space and time period. The studio classroom helps students to explicitly connect concepts taught in
78 lecture through experimentation, resulting in improved exam scores, more expert-like learning
79 attitudes, and positive assessments of the active learning environment from both students and
80 instructors (Kiste et al., 2017). In order to apply these findings and address the lack of active learning
81 opportunities for transcription and translation in our biochemistry curriculum, we sought to
82 incorporate cell-free protein synthesis (CFPS) into our classroom laboratories. Toward this end, we
83 developed the “Genetic Code Kit,” a classroom-ready, modular CFPS kit that is amenable to broad
84 dissemination. Importantly, we sought to determine whether implementing the Genetic Code Kit
85 improves student performance on content-based assessments, as well as students’ self-assessed
86 comfort and confidence with experimental procedures.

87 Advancements in the CFPS platform over the last few decades have enabled a multitude of
88 novel applications in biotechnology, including rapid prototyping for engineering biological systems
89 and easy-to-use point of care diagnostics and biosensors (Pardee et al., 2016; Salehi et al., 2017;
90 Benítez-Mateos et al., 2018; Bundy et al., 2018; Dopp and Reuel, 2018; Takahashi et al., 2018;
91 Gräwe et al., 2019; Gregorio et al., 2019; Silverman et al., 2019; Jung et al., 2020). CFPS generally
92 relies on a cell-extract containing the cellular machinery that supports transcription and translation *in*

93 *vitro* and is supplemented with additional reagents that provide the necessary energy and precursors.
94 The open nature of the CFPS system is one of the main advantages of the platform as it allows the
95 user to produce proteins on-demand without relying on living cells. Thus, CFPS permits the user to
96 directly manipulate the environment of protein synthesis to suit their needs without the limitation of
97 cellular viability constraints, as is the case for *in vivo* protein expression. The unique advantages of
98 CFPS are also what makes it well suited for active, inquiry-based learning in ways that can transform
99 biochemical and biotechnology education, while simultaneously exposing students to experimental
100 procedures associated with an emerging biotechnology. The pioneering work by BioBits and
101 myTXL have provided the proof-of-concept in adapting CFPS to classroom settings and engaging
102 students at various grade levels (Huang et al., 2018; Stark et al., 2018, 2019; Collias et al., 2019).
103 Additionally, CFPS remains robust in a variety of chemical environments (Yin and Swartz, 2004;
104 Seki et al., 2008; Dopp et al., 2019; Gregorio et al., 2019; Karim et al., 2020) providing extensive
105 flexibility in accommodating a broad range of learning objectives. These advantages make CFPS a
106 next-generation educational technology to help meet the next-generation science standards. Moving
107 beyond the proof-of-concept, we focus on using CFPS to teach the fundamental processes of
108 transcription and translation and assess the extent and context of learning gains at the undergraduate
109 level.

110 Transitioning the CFPS platform from a research-focused technology to one that is broadly
111 accessible to high school and university classrooms required extensive simplification, reduced costs,
112 and improved reagent stabilization. Our work to date has taken incremental steps toward these
113 milestones by reducing the number of pipetting steps in CFPS setup (Levine et al., 2019a), creating a
114 less-labor intensive cell extract preparation workflow (Levine et al., 2019b), and identifying a low-
115 cost formulation of additives that enables storage and transport of cell-free extract at room
116 temperature (Gregorio et al., 2020). These advances are part of a concerted effort by the research
117 community to make CFPS accessible to classrooms around the world (Huang et al., 2018; Stark et al.,
118 2018, 2019; Collias et al., 2019). As a result, instructors and institutions now have many options for
119 obtaining CFPS resources for implementation in their classrooms. Each option has its respective
120 advantages that allow instructors to support their learning objectives. Given these combined
121 advancements in accessibility, CFPS is becoming even easier to broadly implement in the teaching
122 laboratory, with the potential for supporting 100s to 1000s of students per quarter.

123 Here, we report the Genetic Code Kit, an implementation of CFPS used to teach transcription
124 and translation. This kit is intended to be low-cost and open source to support accessibility and broad
125 dissemination, especially to schools and programs with limited funding. To accommodate a variety
126 of curricular limitations, the Genetic Code Kit can be completed within a single 3-hour laboratory
127 period, and does not require instructors to dedicate time in a subsequent day to collect data. The kit
128 utilizes crude, *E. coli*-based extract and a DNA template encoding superfolder green fluorescent
129 protein (sfGFP), which together have been broadly demonstrated to support robust and reliable
130 protein expression (Park et al., 2017; Levine et al., 2019a, 2019b). Importantly, the sfGFP product
131 resulting from a successful CFPS reaction is easy to visualize in real-time with minimal equipment or
132 processing, and introduces students to a workhorse reporter broadly used in research and industry.
133 The Genetic Code Kit contains 4 components: 1) a tube containing cell extract in which the reaction
134 mixture is to be assembled, 2) the sfGFP DNA template, 3) “solution A” containing cofactors and
135 substrates, and 4) “solution B” containing the energy system. The liquid transfer of just three
136 reagents ranging from 4.2 μ L to 11.4 μ L allows students to gain micro-pipetting experience while
137 reducing the likelihood of failed reactions. In our implementation, this setup proved reliable and
138 forgiving, with all students able to obtain visible titers of sfGFP within 90 minutes. Requiring
139 students to manually add all reagents necessary for transcription and translation is an important
140 aspect of the Genetic Code Kit, as it provides the opportunity to identify and discuss the importance
141 of each class of reagent (e.g. DNA template, energy reagents, building blocks). This aspect of the kit

142 also provides the flexibility to modify the kit based on the desired learning objectives, allowing for
143 other inquiry-based learning opportunities, as well as course-based undergraduate research
144 experiences (CUREs).

145 We have also developed laboratory materials to accompany the Genetic Code Kit, which help
146 students connect the microscopic processes taking place inside their CFPS reactions to the
147 macroscopic outcome. This includes the laboratory manual and student worksheet (Supplementary
148 Data Sheet 1 and 2). Additionally, we created an augmented reality activity that allows students to
149 interrogate the structure function relationships of GFP to understand the basis for green fluorescence
150 as a function of protein synthesis in their tubes (Supplementary Data Sheet 3). In addition to these
151 specific pedagogical goals related to the central dogma, students also gain exposure to research
152 techniques such as pipetting, reagent handling, the importance of negative and positive controls in
153 experimental design, reaction setup, and data analysis. Importantly, we conducted a controlled study
154 to investigate improvements in student understanding of transcription and translation and their self-
155 assessed comfort with performing an emergent research technique as a function of their hands-on
156 experience with the Genetic Code Kit. Our work demonstrates that implementing CFPS as a hands-
157 on laboratory module leads to significant learning gains associated with transcription and translation
158 learning objectives, as well as positive self-assessment of comfort and confidence with research
159 techniques.

160 2 Materials and Methods

161 2.1 Extract Preparation

162 *E. coli* cell extract was generated using our previously reported CFAI workflow (Levine et
163 al., 2019b). A culture was prepared by inoculating a loopful of BL21* DE3 cells into a 2 L baffled
164 flask containing 1 L of Cell-free Autoinduction media (5.0 g of sodium chloride, 20.0 g of tryptone,
165 5.0 g of yeast, 14.0 g of potassium phosphate dibasic, 6.0 g of potassium phosphate monobasic, 6.0
166 mL of glycerol, 4.0 g of D-lactose, 0.5 g of D-glucose, and nanopure water to 1.0 L). The culture was
167 incubated at 30 °C and 200 rpm for approximately 15 hours. Subsequently, the culture was
168 centrifuged at 4 °C and 5,000 g for 10 min. Harvested cells were resuspended in 30 mL of S30 buffer
169 (10 mM Tris OAc, pH 8.2, 14 mM Mg(OAc)₂, 60 mM KOAc, 2 mM DTT) by vortexing, then spun
170 down at 4 °C and 5000 g for 10 min. Supernatant was removed and cell pellets were flash frozen and
171 stored at -80 °C or used immediately for extract preparation. Cell pellets were resuspended in 1 mL
172 of S30 buffer per 1 g of cells. 1.4 mL of resuspended cells were aliquoted into a 1.5 mL microfuge
173 tube. The resuspension was sonicated using a Qsonica Q125 Sonicator with a 3.175 mm probe, with
174 the cell resuspension surrounded by an ice water bath. Three pulses of 45 s on and 59 s off, at 50%
175 amplitude were carried out. Immediately after sonication, 4.5 µL of 1.0 M DTT was spiked into the
176 lysate and the tube was inverted several times to mix. Lysate was centrifuged at 4 °C and 18,000 g
177 for 10 min. The resulting supernatant is the cell extract. The mixture was flash frozen and stored at -
178 80 °C until Genetic Code Kit preparation.

179 2.2 DNA Purification

180 DNA template pJL1-sfGFP was purified from DH5 α cells using an Invitrogen PureLink
181 HiPure Plasmid Maxiprep Kit. DNA was eluted using warm molecular biology grade water instead
182 of the provided TE buffer for compatibility with the CFPS system. DNA plasmid was diluted with
183 molecular biology grade water to a concentration of 42.1 ng/µL, such that no additional water was
184 needed to prepare 30 µL CFPS reactions with a final DNA concentration of 16 ng/µL. DNA was
185 stored at -20 °C until Genetic Code Kit preparation.

186 **2.3 Solution A and B Preparation**

187 Solution A (containing cofactors and substrates) was prepared with the following specified
188 concentrations of reagents: 8.14 mM ATP, 5.77 mM GTP, 5.77 mM UTP, 5.77 mM CTP, 153.8
189 mg/mL folic acid, 771.9 mg/mL tRNA, 2.71 mM NAD, 1.81 mM CoA, 27.1 mM oxalic acid, 6.79
190 mM putrescine, 10.2 mM spermidine, 386.9 mM HEPES buffer. Solution B (containing the energy
191 system) was prepared with the following specified concentrations of reagents: 71.6 mM magnesium
192 glutamate, 71.6 mM ammonium glutamate, 930.8 mM potassium glutamate, 14.3 mM 20 amino
193 acids, and 238.1 mM phosphoenolpyruvate. All reagents were dissolved in molecular biology grade
194 water. Both solutions were stored at -80 °C until Genetic Code Kit preparation, however these
195 solutions are also stable at -20 °C for 3 months (Supplementary Figure 1).

196 **2.4 Genetic Code Kit Preparation and Reaction Setup**

197 Each kit contained the appropriate amount of pre-aliquoted reagents for the laboratory size
198 and was stored at -20 °C for up to 5 days until student use. Each pair of students was provided a strip
199 of four PCR tubes, each containing 10 µL of extract. Each group of 4 students shared a set of PCR
200 tubes containing molecular biology grade water, pJL1-sfGFP DNA plasmid, solution A, and solution
201 B. Students added 11.4 µL of water, 4.4 µL of solution A, and 4.2 µL of solution B to two tubes as
202 negative controls and 11.4 µL of DNA plasmid, 4.4 µL of solution A, and 4.2 µL of solution B to two
203 tubes as positive controls. All reagents were kept on ice throughout reaction setup. The completed
204 reactions were placed in a 37 °C incubator and checked intermittently for green fluorescence.
205 Necessary equipment includes a p20 pipette, pipette tips, an incubator, and a UV light. More details
206 can be found in the laboratory manual (Supplementary Data Sheet 1).

207 **2.5 Development of Lab Materials**

208 The lab manual and worksheet (Supplementary Data Sheet 1 & 2) for the Genetic Code Kit
209 were developed with the following student learning objectives as a framework: A) Illustrate and
210 describe the processes of transcription and translation; B) Identify the minimally necessary genetic
211 components, enzymes, and reagents necessary for transcription and translation *in vitro*; C) Predict
212 and visualize the outcomes of adding, or not adding, various components to CFPS reactions; D)
213 Define CFPS and its advantages over *in vivo* protein synthesis; E) Paraphrase how energy
214 metabolism sustains transcription and translation in a CFPS reaction. Background on CFPS, the
215 processes of transcription and translation, including the necessary components for each of these
216 processes, and the energy metabolism system operating in CFPS reactions was provided in the lab
217 manual (Supplementary Data Sheet 1).

218 The student worksheet contained open-ended questions corresponding to each of the learning
219 objectives; some questions also required students to draw a schematic to represent their
220 understanding of a topic (Supplementary Data Sheet 2). For example, for learning objective B,
221 students were asked to illustrate the templates for transcription and translation, including genetic
222 elements like a promoter and ribosomal binding site, and their relative locations to one another on a
223 DNA template. Students were asked to consider the outcome of the experiment if certain elements
224 were missing, such as dNTPs or a particular amino acid, in order to address learning objective C.
225 Questions related to learning objective E focused on steps that require energy input, and how the
226 levels of high-energy molecules like ATP change throughout the CFPS reaction.

227 The student questionnaire contained 16 content-based questions and 12 attitudinal questions
228 (Supplementary Data Sheet 4). All questions were multiple choice. The content-based section
229 contained three baseline questions that tested knowledge independent of the intervention's learning
230 objectives and were not expected to be impacted by this laboratory exercise. They acted as a control

231 for differences in baseline aptitudes between the pre- and post- questionnaires. Of the remaining 13
232 content-based questions, four questions tested transcription knowledge and nine tested translation
233 knowledge. Transcription questions focused on key enzymes and required genetic elements on the
234 DNA template for initiation and termination of transcription. Translation questions were focused on
235 the basic mechanism of the ribosome, including how tRNA and mRNA interact, and the required
236 genetic elements on the mRNA template for initiation and termination of translation. The 12
237 attitudinal questions asked students to rank their knowledge of transcription and translation
238 vocabulary and comfort with research techniques.

239 The augmented reality activity utilized Augment¹, a smart phone application, to project the
240 three-dimensional structure of sfGFP onto student benchtops for an exploration of protein structure,
241 structure-function relationships, and the structural basis for fluorescence (Supplementary Data Sheet
242 3). However, our pre- and post- questionnaire did not assess student understanding of sfGFP structure
243 or structure-function relationships, so the impacts of this activity on student learning cannot be
244 reported here.

245 2.6 Implementation of the Genetic Code Kit and Data Collection

246 The Genetic Code Kit and relevant assessments were implemented in the laboratory
247 component of our non-majors’ “Survey of Biochemistry and Biotechnology” course (CHEM 313)
248 taught by biochemistry faculty. The prerequisite for enrollment was the completion of an
249 introductory organic chemistry course. Our curriculum allows students to select either Organic
250 Chemistry I (CHEM 216), which is the first quarter of a year-long organic chemistry sequence or
251 Survey of Organic Chemistry (CHEM 312), which is a one-quarter survey of organic chemistry
252 (Table 1). The students involved in this study represent a breadth of educational backgrounds, with
253 diverse majors from four colleges at Cal Poly SLO (Table 1). All student data was used with written
254 consent of the participants in the study, based on Institutional Review Board (IRB) approval obtained
255 prior to execution.

256 Implementation occurred over a three-week period, with each lab section meeting once a
257 week for three hours. As a “pre-questionnaire” in week one, all students completed the questionnaire
258 described above (Supplementary Data Sheet 4). In week two, students in the intervention group used
259 the Genetic Code Kit in their regularly scheduled lab section (Supplementary Data Sheet 1, 2, & 3).
260 The control group did not meet and did not perform the experiment or augmented reality activity due
261 to a holiday. However, they were provided with the lab manual and completed the same post-lab
262 worksheet. In week three, all students repeated the same questionnaire administered in week one,
263 representing the “post-questionnaire.” A total of 69 students completed both pre- and post- laboratory
264 questionnaires, with 15 in the control group and 54 in the intervention group.

265 Intervention group students performed the Genetic Code Kit lab module in a single three-hour
266 lab period. They were provided the lab manual at least 3 days prior to performing the experiment.
267 After a brief introduction to the experiment in the lab period, students were asked to follow the
268 instructions for reaction setup described in the lab manual, commencing *in vitro* transcription and
269 translation. Reaction tubes were then placed in a 37 °C mini-incubator for 1 to 1.5 hours
270 (Supplementary Figure 2). During the incubation period, students completed the post-lab worksheet
271 and augmented reality activity (Supplementary Data Sheet 2 & 3) and listened to a short lecture from
272 instructors on the basics of transcription and translation. This brief lecture reviewed information on
273 transcription and translation that was also covered in the 4-hour per week lecture portion of the
274 course, and introduced the components of each of the solutions in the Genetic Code Kit that
275 correspond to these processes. This information was also available to students in the control group in

¹ www.augment.com

276 the form of the introduction in the lab manual, and in the course textbook. At the end of the
 277 incubation period, students visualized fluorescence with the naked eye, and enhanced visibility was
 278 achieved using a handheld black light before the lab period was over.

279 **2.7 Statistical Methods**

280 Student responses were collected and all anonymized assessment scores and responses can be
 281 found in Supplementary Table 1. Content-based questions were divided into baseline (1, 5, 6) and
 282 transcription and translation (2-4, 7-16) categories based on each question topic. Each category was
 283 analyzed by comparing the pre- and post- questionnaire scores for the control and intervention
 284 groups, visualized via box and whisker plots generated in SigmaPlot. Paired t-tests were run for both
 285 groups using JMP, and p-values were recorded with a significance level of 0.05. These categories
 286 were also analyzed by calculating the normalized learning gain and effect size for both student
 287 groups to understand the magnitude of the effect of the Genetic Code Kit. Normalized gain enables
 288 the comparison of groups that start at different levels of performance, as it calculates the score
 289 increases with respect to the window of potential learning based on pre-questionnaire scores (Hake,
 290 1998). Effect size provides an additional metric that accounts for the number of students tested and
 291 the variation in scores among the students (Cohen, 1988). Question-based normalized gain was
 292 calculated to determine student performance on each of the 16 questions individually. This metric
 293 uses the same equation as normalized gain, however the average pre- and post- scores are replaced by
 294 the percentage of students who answered the question correctly on the pre- and post- questionnaires.
 295 Additionally, the content-based data was matched to student major and previous course completion
 296 data in the form of an Excel dashboard that allows the user to analyze trends that occur within these
 297 subgroups (Supplementary Table 2). The dashboard also allows for a statistical comparison of the
 298 control group relative to the intervention group. Due to the different sample sizes, the comparison
 299 was performed using the Fisher's Z Test. Point biserial analysis was performed using the Akindi
 300 software².

301 Attitudinal questions were analyzed by comparing the trends in the percentage of students
 302 that selected each answer choice on the pre- and post- questionnaires. For statistical analysis, student
 303 answers were converted to numerical values, where A=1 and E=5. Paired t-tests comparing pre- and
 304 post- scores for each question were run using JMP and p-values were recorded with a significance
 305 level of 0.05.

306 **3 Results**

307 **3.1 Content-based Assessment of Student Learning**

308 The content-based section of the questionnaire contained 16 questions (3 baseline, 13
 309 transcription and translation). For baseline questions unrelated to the learning objectives, there was a
 310 minimal increase in the mean percentage of correct answers; the control group's mean score
 311 increased from 35.6% to 37.8% and the intervention group's mean score increased from 34.0% to
 312 41.3% (Figure 2A). However, a two-sided paired t-test showed that neither of these increases were
 313 significant (p-value > 0.05). Thus, we concluded that neither group became significantly better at
 314 answering the post-questionnaire as a result of previous exposure in the pre-questionnaire. On
 315 transcription and translation questions for the control group, we observed minimal increases in the
 316 mean score, from 41.5% to 48.7%. Comparatively, the intervention group had a larger increase in the
 317 average score on transcription and translation questions, from 49.6% to 63.8% (Figure 2B). One-
 318 sided paired t-tests within the control and intervention groups comparing pre- and post- student

² www.akindi.com

319 scores indicated no significant increase ($p\text{-value} > 0.5$) for the control group and a significant
320 increase ($p\text{-value} < 0.001$) for the intervention group. This indicates that completing the hands-on
321 Genetic Code Kit experiment significantly improves students' ability to correctly answer questions
322 regarding transcription and translation.

323 In addition to observing improvements in average assessment scores, we also wanted to better
324 understand the magnitude of the effect of the intervention on student learning gains. Toward this
325 goal, we evaluated both normalized learning gains and effect sizes, since both are commonly used
326 metrics in STEM education. The extent of normalized learning gains is categorized as low ($\text{gain} <$
327 0.3), medium ($0.7 > \text{gain} \geq 0.3$), and high ($\text{gain} \geq 0.7$) (Hake, 1998). On baseline questions, the
328 control and intervention groups demonstrated low gains of 0.03 and 0.11 , respectively as expected
329 (Figure 3A). For the transcription and translation questions, the control group demonstrated a
330 normalized gain of 0.12 while the intervention group demonstrated a gain of 0.28 . Effect sizes were
331 also calculated as an additional metric to understand the magnitude of learning gains, while
332 accounting for the student sample size and variation. Effect sizes are categorized as small ($\text{effect} =$
333 0.2), medium ($\text{effect} = 0.5$), and large ($\text{effect} = 0.8$) (Cohen, 1988). For the baseline questions, we
334 observed small effect sizes of 0.07 for the control group and 0.28 for the intervention group (Figure
335 3B). Effect sizes on the transcription and translation questions were 0.32 for the control (small-
336 medium) and 0.60 for the intervention (medium-large). As with the normalized gain analysis, the
337 intervention group's ability to correctly answer questions related to transcription and translation after
338 using the Genetic Code Kit module was much greater than the control group, who did not carry out
339 the activity.

340 Lastly, we analyzed the question-based normalized gains for each of the 16 questions
341 individually (Figure 3C). This analysis was intended to indicate student performance on individual
342 questions, allowing us to identify questions that were poorly designed or not well-addressed by the
343 Genetic Code Kit. The outcome of question-based normalized gain assessment was the identification
344 of questions 7 and 11 as particularly challenging for the intervention group. In fact, the control group
345 outperformed the intervention group on those two questions, and the normalized gain for the
346 intervention group was negative for question 11. Quantitatively, the point-biserial correlation
347 coefficient values for questions 7 and 11 were above 0.2 , suggesting that they are "fair" questions.
348 Qualitatively, it is possible that these questions were written ineffectively, were mismatched with our
349 learning objectives, or that CFPS was not able to resolve student misconceptions regarding the
350 macromolecular interactions involved in translation. In fact, non-covalent interactions involved in
351 translation were not explicitly covered in the pre-lab lecture, worksheet, or lab manual.

352 Given that we observed meaningful normalized learning gains and effect sizes upon
353 intervention despite questions 7 and 11, we remained curious about the learning gains observed in the
354 remaining questions. In a follow-up analysis (Supplementary Table 2 & 3), we removed questions 7
355 and 11 from the group of transcription and translation questions and used this narrower scope to
356 evaluate learning gains by student demographics. We observed that students who had previously
357 taken Ochem I, the first quarter in a year-long series of organic chemistry, had significantly higher
358 learning gains compared to the control group ($p\text{-value} < 0.05$), while those who had taken Survey of
359 Ochem did not significantly benefit ($p\text{-value} > 0.05$) from the Genetic Code Kit intervention
360 compared to the control group (Supplemental Table 3). The intervention group students that did not
361 significantly benefit were mostly from the College of Agriculture, Food, and Environmental Science,
362 who have historically underperformed in the Survey of Biochemistry and Biotechnology course.
363 While this observation is only suggestive when we removed questions 7 and 11 from the analysis, it
364 represents an intriguing starting point for using CFPS to consider preparation gaps and achievement
365 gaps within our student populations. These results suggest that if question design can be improved
366 and sample size can be increased, implementation of CFPS has the potential to explore the basis for
367 preparation and achievement gaps in biochemical education. Regardless, these additional findings are

368 contingent on solving the learning issues identified in questions 7 and 11, as these differences in
 369 prerequisite preparation only appear when they are removed from the analysis.

370 Overall, significant increases in the average scores on content-based questions (Figure 2), a
 371 normalized learning gain around 0.3, and an effect size of 0.6 for the intervention group (Figure 3)
 372 indicate that implementing the Genetic Code Kit improved students' ability to comprehend and
 373 answer questions relating to transcription and translation. As no significant increase (p -value > 0.05)
 374 in the performance on baseline questions was observed, we propose that the observed increase in
 375 assessment scores for transcription and translation questions was a result of the Genetic Code Kit
 376 rather than repeated exposure to the questionnaire.

377 3.2 Attitudinal-based Assessment of Student Learning

378 The pre- and post- questionnaires completed by both the control and intervention groups
 379 contained a total of 12 attitudinal questions. These questions prompted students to self-assess their
 380 recognition and knowledge of transcription and translation vocabulary, as well as their comfort with
 381 laboratory techniques used in CFPS. Prior work has documented students' deficiency in
 382 metacognitive skills and found that active learning pedagogies can strengthen these skills (National
 383 Research Council, 2000). Our attitudinal-based questions allow us to examine how students'
 384 perceptions of their learning correlate with their results on the content-based assessment
 385 (Supplementary Figure 3). We found that both the control and intervention groups showed positive
 386 correlations on pre- and post- questionnaires, with an increase in the slope from pre- to post-
 387 questionnaire. For the control group, the pre-questionnaire R^2 value was 0.02 and post-questionnaire
 388 was 0.36. For the intervention, the pre-questionnaire R^2 value was 0.10 and post-questionnaire was
 389 0.30. The relatively low pre-questionnaire R^2 is noteworthy: it shows that students' knowledge and
 390 attitudes are, effectively, uncorrelated. The increase in post-questionnaire R^2 indicates that
 391 knowledge and attitudes move in the same direction. Overall, this analysis indicates that students'
 392 self-reported confidence correlated with their performance on content-based questions. As a result,
 393 we pursued more detailed analysis of the attitudinal-based questions.

394 We first considered the possibility that improvements in students' self-assessment of their
 395 confidence were an outcome of their recognition of vocabulary terms through repeated exposure to
 396 the questionnaire rather than as a result of improved conceptual understanding of the terms. To
 397 address this concern, we chose to perform detailed per-question analysis for the attitudinal-based
 398 assessment on questions that involved comfort with CFPS as an indicator of how beneficial the
 399 activity was in introducing a novel biotechnology. For the intervention group, we observed
 400 significant increases (p -value < 0.05) between pre- and post- scores for questions 23 and 25-27 using
 401 a one-sided paired t-test (Figure 4). When prompted with "I know what cell-free protein synthesis is"
 402 (question 23) on the pre-questionnaire, over 50% of the intervention group students indicated that
 403 they had no idea what the term meant and ~11% indicated that they knew what the term meant
 404 (Figure 4A). After conducting the experiment, this changed to less than 5% and greater than 50%,
 405 respectively. The control group saw a similar, but less extensive shift in the trend with almost 40% of
 406 students reporting that they knew what the term meant in the post-questionnaire (Supplementary
 407 Figure 4). The comparable shift in the control and intervention groups is likely due to the background
 408 information that they received on CFPS through the lab manual alone. For question 25 (Figure 4B),
 409 "I am comfortable conducting experiments with enzymes," pre-questionnaire comfort was generally
 410 high for the intervention group, but only ~11% of students indicated that they "strongly agreed."
 411 However, the Genetic Code Kit increased intervention student confidence in working with enzymes,
 412 such that almost 30% of students said they "strongly agreed" on the post-questionnaire. This was a
 413 noteworthy observation, since the Genetic Code Kit was implemented at the end of the quarter, and
 414 students had worked with enzymes in numerous previous laboratory modules. The intervention

415 group's comfort with conducting experiments involving *in vitro* transcription and translation
416 (questions 26 and 27; Figure 4C & D) also showed notable improvement, with the number of
417 students answering "strongly agree" increasing to ~25% from less than 2%. Comparatively, the
418 control group had less than 8% of students say that they "strongly agreed" in response to questions
419 25-27 (Supplemental Figure 4). These data indicate that the intervention group's hands-on exposure
420 to the CFPS reaction improved their comfort with these laboratory skills over the control group.
421

422 4 Discussion

423 The cell-free protein synthesis (CFPS) platform has seen significant development and
424 widespread use as a biotechnology tool in recent years. CFPS harnesses the genetic code in a test-
425 tube, in a flexible and tunable biochemical milieu, making it poised to be a transformative
426 educational technology. Specifically, CFPS allows students to probe the processes of transcription
427 and translation in a way that improves their learning outcomes, while providing them the technical
428 skills for careers in biotechnology. Here, we report the implementation of our Genetic Code Kit, a
429 simplified, yet modular CFPS reaction, in college-level biochemistry curriculum. Importantly, the
430 Genetic Code Kit improved students' understanding of transcription and translation for
431 undergraduate students in a survey of biochemistry course. Our results suggest that the tactile process
432 of setting up a CFPS reaction by adding solutions containing the building blocks, energy system, and
433 DNA template to *E.coli* extract, and observing the real-time production of a fluorescent protein
434 increases students' comprehension of transcription and translation. Our observations are consistent
435 with the extensive literature on the benefits of a physical experience in student learning (Bopagedera,
436 2011; Zacharia et al., 2012; Kontra et al., 2015; Kiste et al., 2016). Moreover, the Genetic Code Kit
437 may help resolve common student misconceptions surrounding transcription and translation. For
438 example, physically supplementing the CFPS reaction vessel with amino acids may eliminate
439 potential confusion on the source of amino acids or the misconception that translation produces
440 amino acids (Fisher, 1985). Additionally, requiring students to add both DNA and nucleotides to the
441 CFPS reaction vessel could help resolve student misconceptions that DNA is converted into RNA via
442 a chemical reaction instead of being used as a template for a new nucleotide strand (Wright et al.,
443 2014).

444 The shifts in responses to attitudinal-based questions showcase the usefulness of the Genetic
445 Code Kit to prepare students for future careers in laboratory science. Notably, these benefits to
446 students extend beyond the learning gains in the content-based questions to support increased student
447 confidence with the laboratory techniques used for CFPS. This work suggests that improvements to
448 familiarity with biotechnologies and comfort in implementing biotechnology-based experiments
449 provide fundamental advances toward workforce development. Prior work has documented that
450 exposing students to research as part of science curriculum has improved student engagement in
451 research outside of the classroom (Lindsay and McIntosh, 2000). Furthermore, undergraduate
452 involvement in research experiences is known to increase student interest in obtaining a Ph.D. and
453 pursuing a STEM field, especially when students are invested and interested in their research
454 (Russell et al., 2007).

455 In order to enable all students to access these learning outcomes, the Genetic Code Kit is
456 designed to be a low-cost, easy to assemble and implement, highly tailorable platform for various
457 curricula and learning objectives, and requires minimal training and equipment. The Genetic Code
458 Kit costs \$4.08 per student, based on 4x CFPS reactions per student (Table 2). The cost of \$1.02 per
459 30 μ L reaction is inclusive of all materials, reagents, and labor at an estimated rate of \$25/hr for the
460 technician's efforts. The development of the previously reported CFAI workflow has allowed us to
461 significantly reduce the time required for cell extract preparation, reducing the cost associated with
462 labor (Levine et al., 2019b). For example, preparing kits for 375 students requires under 25 person-

463 hours. Notably, our kit preparation can be completed entirely by undergraduate students, as was done
464 in this work, which significantly reduces the cost of implementation. The Genetic Code Kit
465 preparation is also highly scalable. In fact, preparing larger quantities becomes more cost-effective.
466 After the cost of labor, the next largest expense is the energy reagents that drive the PANOxSP-based
467 CFPS reaction, but prior work has shown that this cost could be further reduced by leveraging
468 glucose metabolism (Calhoun and Swartz, 2005). Instructors and institutions now benefit from a
469 variety of CFPS options for their classrooms and Table 2 provides a list of options to choose from.
470 We include cost comparisons in Table 2, since this may be one possible driver for selecting a path to
471 implementing CFPS. However, we urge instructors to review the benefits of all listed options, as they
472 may outweigh costs, particularly for convenience of implementation or suitability to specific learning
473 objectives.

474 The Genetic Code Kit can be tailored to meet a variety of learning objectives beyond teaching
475 transcription and translation. The open nature of the system makes it poised to support inquiry-based
476 learning at a variety of grade levels and course-based undergraduate research experiences (CUREs)
477 through minor modifications to the reaction setup or DNA template described here. These
478 possibilities can help tailor the kit to the desired grade level and course learning objectives, and
479 include 1) the sequence-function relationships of various genetic elements such as promoters,
480 ribosome binding sites, and codon optimization, 2) riboswitches and aptamers, 3) genetic circuits, 4)
481 CRISPR, 5) probing the mechanisms of various antibiotics, such as protein synthesis inhibitors, and
482 many more. Some unique applications of CFPS for classroom instruction have already been
483 developed for the BioBits kits (Huang et al., 2018; Stark et al., 2018, 2019). Lastly, the Genetic Code
484 Kit can be implemented as a free-standing laboratory module to fit within a single 3-hour lab course,
485 but it can also be integrated into existing curricula. For example, this lab could be preceded by
486 molecular biology labs including PCR or CRISPR and followed by analysis of the protein product
487 via other traditional biochemical methods such as western blotting, ELISA, or SDS-PAGE.

488 Overall, this work represents the first controlled study of student learning gains resulting from
489 a hands-on, learn-by-doing intervention based on CFPS. While this study's findings are limited by a
490 small sample size and focus on undergraduate students from a single institution, we observed
491 significant gains for learning objectives relating to transcription and translation. Thus, the results of
492 this work provide the foundation to expand assessments of learning gains to various educational
493 levels, pursue multi-institutional efforts that include large student sample sizes, and iterate on the
494 design of the kit to further improve student learning gains for a broad range of learning objectives.
495 We propose that the expansion of this work will further validate the important role of CFPS in
496 biochemical education while supporting workforce development for the growing biotechnology
497 industry.

498 **Data Availability Statement**

499 The datasets [GENERATED/ANALYZED] for this study can be found in the [NAME OF
500 REPOSITORY] [LINK].

501 **Author Contributions**

502 LW and NG wrote the manuscript, performed statistical analysis, and generated the figures. LW, NG,
503 BS, and WK prepared all reagents for teaching lab implementation. BS and WK performed reagent
504 storage, and time-course testing. AK helped design the teaching lab implementation and guided the
505 analysis and visualization of assessments of student learning gains. PP helped with statistical analysis
506 of the data. JO conceived the project. KW and JO designed and executed the teaching lab
507 implementation. All authors helped revise the manuscript and agree to the accuracy of the work
508 reported.

509 **Funding**

510 Authors acknowledge funding support from the Bill and Linda Frost Fund for scholarships to LW,
 511 NK, BS, and WK, for supporting general research expenses, and for funding open access publication
 512 fees. Authors also acknowledge the Center for Applications in Biotechnology's Chevron
 513 Biotechnology Applied Research Endowment Grant, Cal Poly Research, Scholarly, and Creative
 514 Activities Grant Program (RSCA 2017), and the National Science Foundation (NSF-1708919).

515 **Conflict of Interest**

516 The authors declare that the research was conducted in the absence of any commercial or financial
 517 relationships that could be construed as a potential conflict of interest.

518 **Abbreviations**

519 Cell-free protein synthesis (CFPS); superfolder green fluorescent proteins (sfGFP); course-based
 520 undergraduate research experiences (CUREs)

521 **Acknowledgments**

522 Authors acknowledge Dr. Eric Jones, Dr. Anya Goodman, and Dr. Steven Wilkinson for supporting
 523 the implementation of the Genetic Code Kit in their teaching laboratories for this work; Andrea
 524 Laubscher for technical support; Dr. Derek Gragson for support in obtaining historical student data;
 525 Dr. Aaron Engelhart (University of Minnesota), Dr. Seth Bush, Max Levine, Philip Smith, and Logan
 526 Burrington for helpful discussions.

527 **Supplementary Material**

528 The Supplementary Material for this article can be found online at:

529 **References**

530 Altiparmak, M., and Nakiboglu Tezer, M. (2009). Hands on group work paper model for teaching
 531 DNA structure, central dogma and recombinant DNA. *US-China Educ. Rev.* 6, 19–23.

532 Armbruster, P., Patel, M., Johnson, E., and Weiss, M. (2009). Active learning and student-centered
 533 pedagogy improve student attitudes and performance in introductory biology. *CBE Life Sci.
 534 Educ.* 8, 203–213. doi:10.1187/cbe.09-03-0025.

535 Bassiri, E. A. (2011). pGLO mutagenesis: A laboratory procedure in molecular biology for biology
 536 students. *Biochem. Mol. Biol. Educ.* 39, 432–439. doi:10.1002/bmb.20538.

537 Benítez-Mateos, A. I., Llarena, I., Sánchez-Iglesias, A., and López-Gallego, F. (2018). Expanding
 538 One-Pot Cell-Free Protein Synthesis and Immobilization for On-Demand Manufacturing of
 539 Biomaterials. *ACS Synth. Biol.* 7, 875–884. doi:10.1021/acssynbio.7b00383.

540 Bopegedera, A. M. R. P. (2011). Putting the laboratory at the center of teaching chemistry. *J. Chem.
 541 Educ.* 88, 443–448. doi:10.1021/ed100045z.

542 Bundy, B. C., Porter Hunt, J., Jewett, M. C., Swartz, J. R., Wood, D. W., Frey, D. D., et al. (2018).
 543 Cell-free biomanufacturing. *Curr. Opin. Chem. Eng.* 22, 177–183.

544 Calhoun, K. A., and Swartz, J. R. (2005). Energizing cell-free protein synthesis with glucose
 545 metabolism. *Biotechnol. Bioeng.* 90, 606–613. doi:10.1002/bit.20449.

546 Chang, P.-S., Lee, S.-H., and Wen, M. L. (2020). Metacognitive inquiry activities for instructing the
 547 central dogma concept: 'button code' and 'beaded bracelet making.' *J. Biol. Educ.* 54, 47–62.
 548 doi:10.1080/00219266.2018.1546756.

549 Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Routledge.

550 Collias, D., Marshall, R., Collins, S. P., Beisel, C. L., and Noireaux, V. (2019). An educational
 551 module to explore CRISPR technologies with a cell-free transcription-translation system. *Synth.
 552 Biol.* 4. doi:10.1093/synbio/ysz005.

553 DeBruyn, J. M. (2012). Teaching the Central Dogma of Molecular Biology using Jewelry. *J.
554 Microbiol. Biol. Educ.* 13, 62–64. doi:10.1128/jmbe.v13i1.356.

555 Deutch, C. E. (2019). Transformation of *Escherichia coli* with the pGLO Plasmid: Going beyond the
556 Kit. *Am. Biol. Teach.* 81, 52–55. doi:10.1525/abt.2019.81.1.52.

557 Dopp, B. J. L., Tamiev, D. D., and Reuel, N. F. (2019). Cell-free supplement mixtures: Elucidating
558 the history and biochemical utility of additives used to support *in vitro* protein synthesis in *E.
559 coli* extract. *Biotechnol. Adv.* 37, 246–258. doi:10.1016/J.BIOTECHADV.2018.12.006.

560 Dopp, J. L., and Reuel, N. F. (2018). Process optimization for scalable *E. coli* extract preparation for
561 cell-free protein synthesis. *Biochem. Eng. J.* 138, 21–28. doi:10.1016/J.BEJ.2018.06.021.

562 Dorrell, M. I., and Lineback, J. E. (2019). Using Shapes & Codes to Teach the Central Dogma of
563 Molecular Biology: A Hands-On Inquiry-Based Activity. *Am. Biol. Teach.* 81, 202–209.
564 doi:10.1525/abt.2019.81.3.202.

565 Duncan, R. G., and Reiser, B. J. (2007). Reasoning across ontologically distinct levels: Students'
566 understandings of molecular genetics. *J. Res. Sci. Teach.* 44, 938–959. doi:10.1002/tea.20186.

567 Fisher, K. M. (1985). A misconception in biology: Amino acids and translation. *J. Res. Sci. Teach.*
568 22, 53–62. doi:10.1002/tea.3660220105.

569 Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014).
570 Active learning increases student performance in science, engineering, and mathematics. *Proc.
571 Natl. Acad. Sci. U. S. A.* 111, 8410–8415. doi:10.1073/pnas.1319030111.

572 Gräwe, A., Dreyer, A., Vornholt, T., Barteczko, U., Buchholz, L., Drews, G., et al. (2019). A paper-
573 based, cell-free biosensor system for the detection of heavy metals and date rape drugs. *PLoS
574 One* 14, 1–22. doi:10.1371/journal.pone.0210940.

575 Gregorio, N. E., Kao, W. Y., Williams, L. C., Hight, C. M., Patel, P., Watts, K. R., et al. (2020).
576 Unlocking Applications of Cell-Free Biotechnology through Enhanced Shelf Life and
577 Productivity of *E. coli* Extracts. *ACS Synth. Biol.* 9, 766–778. doi:10.1021/acssynbio.9b00433.

578 Gregorio, N. E., Levine, M. Z., and Oza, J. P. (2019). A User's Guide to Cell-Free Protein Synthesis.
579 *Methods Protoc.* 2, 24. doi:10.3390/mps2010024.

580 Haak, D. C., HilleRisLambers, J., Pitre, E., and Freeman, S. (2011). Increased structure and active
581 learning reduce the achievement gap in introductory biology. *Science (80-).* 332, 1213–1216.
582 doi:10.1126/science.1204820.

583 Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student
584 survey of mechanics test data for introductory physics courses. *Cit. Am. J. Phys.* 66, 64.
585 doi:10.1119/1.18809.

586 Huang, A., Nguyen, P. Q., Stark, J. C., Takahashi, M. K., Donghia, N., Ferrante, T., et al. (2018).
587 BioBits™ Explorer: A modular synthetic biology education kit. *Sci. Adv.* 4, eaat5105.
588 doi:10.1126/sciadv.aat5105.

589 Ibarra-Herrera, C. C., Carrizosa, A., Yunes-Rojas, J. A., and Mata-Gómez, M. A. (2019). Design of
590 an app based on gamification and storytelling as a tool for biology courses. *Int. J. Interact. Des.
591 Manuf.* 13, 1271–1282. doi:10.1007/s12008-019-00600-8.

592 Jung, J. K., Alam, K. K., Verosloff, M. S., Capdevila, D. A., Desmau, M., Clauer, P. R., et al. (2020).
593 Cell-free biosensors for rapid detection of water contaminants. *Nat. Biotechnol.*, 1–9.
594 doi:10.1038/s41587-020-0571-7.

595 Karim, A. S., Rasor, B. J., and Jewett, M. C. (2020). Enhancing control of cell-free metabolism
596 through pH modulation. *Synth. Biol.* 5. doi:10.1093/synbio/ysz027.

597 Kiste, A. L., Hooper, R. G., Scott, G. E., and Bush, S. D. (2016). Atomic Tiles: Manipulative
598 Resources for Exploring Bonding and Molecular Structure. *J. Chem. Educ.* 93, 1900–1903.
599 doi:10.1021/acs.jchemed.6b00361.

600 Kiste, A. L., Scott, G. E., Bukenberger, J., Markmann, M., and Moore, J. (2017). An examination of
601 student outcomes in studio chemistry. *Chem. Educ. Res. Pract.* 18, 233–249.

602 doi:10.1039/C6RP00202A.

603 Kontra, C., Lyons, D. J., Fischer, S. M., and Beilock, S. L. (2015). Physical experience enhances
604 science learning. *Psychol. Sci.* 26, 737–49. doi:10.1177/0956797615569355.

605 Kozma, R., Chin, E., Russell, J., and Marx, N. (2000). The Roles of Representations and Tools in the
606 Chemistry Laboratory and Their Implications for Chemistry Learning. *J. Learn. Sci.* 9, 105–143.
607 doi:10.1207/s15327809jls0902_1.

608 Levine, M. Z., Gregorio, N. E., Jewett, M. C., Watts, K. R., and Oza, J. P. (2019a). Escherichia coli-
609 Based Cell-Free Protein Synthesis: Protocols for a robust, flexible, and accessible platform
610 technology. *J. Vis. Exp.*, e58882. Available at: <https://www.jove.com/video/58882/escherichia-coli-based-cell-free-protein-synthesis-protocols-for> [Accessed August 15, 2019].

611 Levine, M. Z., So, B., Mullin, A. C., Watts, K. R., and Oza, J. P. (2019b). Redesigned upstream
612 processing enables a 24-hour workflow from *E. coli* cells to cell-free protein synthesis. *bioRxiv*
613 [Preprint], 729699. doi:10.1101/729699.

614 Lindsay, H. A., and McIntosh, M. C. (2000). Early Exposure of Undergraduates to the Chemistry
615 Research Environment: A New Model for Research Universities. *J. Chem. Educ.* 77, 1174–
616 1177. doi:10.1021/ed077p1174.

617 Marshall, P. A. (2017). A Hands-On Activity to Demonstrate the Central Dogma of Molecular
618 Biology Via a Simulated VDJ Recombination Activity. *J. Microbiol. Biol. Educ.* 18.
619 doi:10.1128/jmbe.v18i2.1277.

620 National Research Council (2000). *How People Learn*. Washington, D.C.: The National Academies
621 Press doi:10.17226/9853.

622 Newman, D. L., Snyder, C. W., Fisk, J. N., and Wright, L. K. (2016). Development of the Central
623 Dogma Concept Inventory (CDCI) Assessment Tool. *CBE—Life Sci. Educ.* 15, 1–14.
624 doi:10.1187/cbe.15-06-0124.

625 Newman, D. L., and Wright, L. K. (2013). Using PCR to Target Misconceptions about Gene
626 Expression. *J. Microbiol. Biol. Educ.* 14, 93–100. doi:10.1128/jmbe.v14i1.539.

627 Nogaj, L. A. (2013). Using Active Learning in a Molecular Biology. *J. Coll. Sci. Teach.* 42, 50–55.
628 Available at: <http://www.nsta.org/> [Accessed March 31, 2020].

629 Pardee, K., Green, A. A., Takahashi, M. K., Braff, D., Lambert, G., Lee, J. W., et al. (2016). Rapid,
630 Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell* 165,
631 1255–1266. doi:10.1016/J.CELL.2016.04.059.

632 Park, Y. J., Lee, K.-H., and Kim, D.-M. (2017). Assessing translational efficiency by a reporter
633 protein co-expressed in a cell-free synthesis system. *Anal. Biochem.* 518, 139–142.
634 doi:10.1016/J.AB.2016.11.019.

635 Pigage, H. K. (1991). The Central Dogma in Action. *Am. Biol. Teach.* 53, 436–437.
636 doi:10.2307/4449352.

637 Queloz, A., Klymkowsky, M. W., Stern, E., Hafen, E., and Köhler, K. (2017). Diagnostic of
638 students' misconceptions using the Biological Concepts Instrument (BCI): A method for
639 conducting an educational needs assessment. *PLoS One* 12, e0176906.
640 doi:10.1371/journal.pone.0176906.

641 Rotbain, Y., Marbach-Ad, G., and Stavy, R. (2008). Using a computer animation to teach high school
642 molecular biology. *J. Sci. Educ. Technol.* 17, 49–58. doi:10.1007/s10956-007-9080-4.

643 Russell, S. H., Hancock, M. P., and McCullough, J. (2007). Benefits of undergraduate research
644 experiences. *Science (80-)* 316, 548–549. doi:10.1126/science.1140384.

645 Salehi, A. S. M., Shakalli Tang, M. J., Smith, M. T., Hunt, J. M., Law, R. A., Wood, D. W., et al.
646 (2017). Cell-Free Protein Synthesis Approach to Biosensing hTR β -Specific Endocrine
647 Disruptors. *Anal. Chem.* 89, 3395–3401. doi:10.1021/acs.analchem.6b04034.

648 Seki, E., Matsuda, N., Yokoyama, S., and Kigawa, T. (2008). Cell-free protein synthesis system from
649 Escherichia coli cells cultured at decreased temperatures improves productivity by decreasing

651 DNA template degradation. *Anal. Biochem.* 377, 156–161. doi:10.1016/J.AB.2008.03.001.
652 Silverman, A. D., Akova, U., Alam, K. K., Jewett, M. C., and Lucks, J. B. (2019). Design and
653 optimization of a cell-free atrazine biosensor. *bioRxiv*, 779827. doi:10.1101/779827.
654 Stark, J. C., Huang, A., Hsu, K. J., Dubner, R. S., Forbrook, J., Marshalla, S., et al. (2019). BioBits
655 Health: Classroom Activities Exploring Engineering, Biology, and Human Health with
656 Fluorescent Readouts. *ACS Synth. Biol.* 8, 1001–1009. doi:10.1021/acssynbio.8b00381.
657 Stark, J. C., Huang, A., Nguyen, P. Q., Dubner, R. S., Hsu, K. J., Ferrante, T. C., et al. (2018).
658 BioBits™ Bright: A fluorescent synthetic biology education kit. *Sci. Adv.* 4, eaat5107.
659 doi:10.1126/sciadv.aat5107.
660 Takahashi, M. K., Tan, X., Dy, A. J., Braff, D., Akana, R. T., Furuta, Y., et al. (2018). A low-cost
661 paper-based synthetic biology platform for analyzing gut microbiota and host biomarkers. *Nat.*
662 *Commun.* 9, 1–12. doi:10.1038/s41467-018-05864-4.
663 Takemura, M., and Kurabayashi, M. (2014). Using analogy role-play activity in an undergraduate
664 biology classroom to show central dogma revision. *Biochem. Mol. Biol. Educ.* 42, 351–356.
665 doi:10.1002/bmb.20803.
666 Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Nicole Arroyo, E., Behling, S., et al. (2020).
667 Active learning narrows achievement gaps for underrepresented students in undergraduate
668 science, technology, engineering, and math. *Proc. Natl. Acad. Sci. U. S. A.* 117, 6476–6483.
669 doi:10.1073/pnas.1916903117.
670 Ward, W. W., Swiatek, G. C., and Gonzalez, D. G. (2000). Green fluorescent protein in
671 biotechnology education. *Methods Enzymol.* 305, 672–680. doi:10.1016/S0076-6879(00)05518-
672 X.
673 Wright, K., Fisk, J., and Newman, D. L. (2014). DNA→RNA: What do students think the arrow
674 means? *CBE Life Sci. Educ.* 13, 338–348. doi:10.1187/cbe.CBE-13-09-0188.
675 Yin, G., and Swartz, J. R. (2004). Enhancing multiple disulfide bonded protein folding in a cell-free
676 system. *Biotechnol. Bioeng.* 86, 188–195. doi:10.1002/bit.10827.
677 Zacharia, Z. C., Loizou, E., and Papaevripidou, M. (2012). Is physicality an important aspect of
678 learning through science experimentation among kindergarten students? *Early Child. Res. Q.* 27,
679 447–457. doi:10.1016/j.ecresq.2012.02.004.
680

681

Table 1. Student population distributions by major and completed courses.

College	Major	Control	Intervention	Ochem I	Survey of Ochem
College of Science and Mathematics	Biological Sciences	6	16	17	7
	Kinesiology	0	1	0	1
	Marine Science	0	1	1	0
	Microbiology	1	3	4	0
College of Agriculture, Food, and Environmental Science	Animal Science	2	7	2	7
	Food Science	0	3	0	3
	Nutrition	2	14	2	13
	Wine and Viticulture	2	5	0	6
College of Engineering	Biomedical Engineering	1	2	3	0
	Materials Engineering*	0	1	1	1
College of Liberal Arts	Psychology	1	1	1	1
	Total Students	15	54	31	39

682 *The materials engineering student took both Ochem I and Survey of Ochem

683

Table 2. CFPS reaction costs for in-house and commercially available kits.

Product	Vol/rxn (μ L)	Cost/rxn	Cost/ student	Cost/100 students	Reference
Genetic Code Kit	30	\$ 1.02	\$ 4.08	\$ 408	(Levine et al., 2019b)
miniPCR BioBits	7	\$ 2.97	\$ 11.88	\$ 1,235	(Stark et al., 2018)
Bioneer AccuRapid Midi	30	\$ 2.94	\$ 11.76	\$ 1,544	-
Promega S30 for Circular DNA	30	\$ 9.86	\$ 39.44	\$ 3,944	-
Arbor myTXTL	12	\$ 10.65	\$ 42.60	\$ 4,260	(Collias et al., 2019)
NEBExpress	30	\$ 10.20	\$ 40.80	\$ 5,100	-
Thermo Expressway Maxi	25	\$ 13.20	\$ 52.80	\$ 5,280	-
Sigma iPE-Quick Kit	30	\$ 12.42	\$ 49.68	\$ 5,400	-

684

685 **Figure 1.** Traditional central dogma teaching tools and the next-generation Genetic Code Kit. The
686 Genetic Code kit utilizes cell-free protein synthesis and augmented reality to teach the processes of
687 transcription and translation.
688

689 **Figure 2.** Impact of the Genetic Code Kit on student performance on content-based questions
690 involving baseline or transcription (Tx) and translation (Tl) questions. Student score distributions are
691 depicted as follow: solid lines indicate median, dotted lines indicate mean, boxes demarcate the 25th
692 and 75th percentiles, whiskers represent the 10th and 90th percentiles, and points represent outliers.
693 Control group scores represent a population of 15 students and intervention group scores represent a
694 population of 54 students. The content-based portion of the questionnaire contained 16 questions, 3
695 baseline and 13 transcription and translation. Specific questions can be found in Supplementary Data
696 Sheet 4. **(A)** Student score distributions for baseline questions. Pre- and post- scores for the control
697 group and intervention group were compared using a two-sided paired t-test (ns indicates p-value >
698 0.05) with a null hypothesis that pre- and post- scores will be equal. **(B)** Student score distributions
699 for transcription and translation questions. Pre- and post- scores for the control group and
700 intervention group were compared using a one-sided paired t-test (ns indicates p-value > 0.05, ***
701 indicates p-value < 0.001) with a null hypothesis that pre- and post- scores will be equal.
702

703 **Figure 3.** Magnitude of student learning gains on content-based questions upon implementing the
704 Genetic Code Kit. The control group represents a population of 15 students and the intervention
705 group represents a population of 54 students. The content-based portion of the questionnaire
706 contained 16 questions, 3 baseline and 13 transcription and translation. Specific questions can be
707 found in Supplementary Data Sheet 4. **(A)** Normalized gain by question category. Normalized gains
708 > 0.3 indicate a medium gain activity. **(B)** Effect size by question category. Effect sizes of 0.2
709 indicate small effects, 0.5 indicate medium effects, and 0.8 indicate large effects. **(C)** Question-based
710 normalized gain for each question. Question categories are indicated as follows: (B) baseline, (Tx)
711 transcription, (Tl) translation. Normalized gains > 0.3 indicate a medium gain activity.
712

713 **Figure 4.** Changes in intervention group student attitudes toward CFPS and conducting CFPS-based
714 experiments. Answer choices for **(A)** ranged from A – “I have no idea what this term means” to D –
715 “I know what this term means.” Answer choices for **(B)** – **(D)** ranged from A – “Strongly disagree”
716 to E – “Strongly agree.” Student answers were converted to a numerical value where A=1 and E=5,
717 in order to calculate p-values using a one-sided paired t-test with a null hypothesis that pre- and post-
718 scores would be equal. The intervention group contained 52 students. This is less than the number of
719 students in the content analysis, as some students did not complete the attitudinal section of the post-
720 questionnaire. All possible answer categories can be found in Supplementary Data Sheet 4.