TACKLING DISCIPLINARY SILOS: ENGAGING CHEMISTRY AND BIOLOGY STUDENTS IN MECHANISTIC REASONING THROUGH CAREFULLY DESIGNED ASSESSMENT TASKS

By

Clare Carlson

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemistry – Doctor of Philosophy

2024

ABSTRACT

In an increasingly complex world, it is imperative that undergraduate science courses prepare students to make confident, informed decisions and predictions. The research presented in this thesis focuses on supporting students to engage in mechanistic reasoning (MR) about phenomena in chemistry and biology (that is, thinking about how and why these phenomena occur). Using a resources perspective of student knowledge construction, I investigated and characterized explanations for two phenomena: (1) preferential protein-ligand binding, leveraging ideas about electrostatic forces and interactions; and (2) ATP-driven coupled reactions, leveraging ideas about energy, reactivity, and favorability.

We collected and analyzed over a thousand student responses for the protein-ligand binding task, a previously developed task that elicits productive student ideas and MR as a thinking strategy. Students across chemistry and biology courses engaged in MR to varying extents, but the responses that were fully mechanistic nearly always correlated with a correct prediction, suggesting the predictive power of this reasoning strategy. With a large selection of responses and access to demographic characteristics, I then calculated an ordinal regression model to determine if GPA, race/ethnicity, and/or binary gender predicted student engagement in MR. Binary gender identification did not contribute as a predictor; however, students with a higher GPA had significantly higher odds of engaging fully in MR, and being White (compared to Non-White), to a lesser extent than GPA, also resulted in higher odds of engaging fully in MR.

In subsequent studies, I investigated student and instructor understandings of the role of ATP, an integral biological molecule that is directly related to energy, a core idea in both chemistry and biology, suggesting another opportunity for interdisciplinary learning. Due to the

widespread misconception that breaking bonds releases energy, which is frequently associated with language about the "high-energy bond(s)" in ATP, we aimed to uncover a more coherent, mechanistic way of talking about ATP. To do this, we first interviewed a range of chemistry, biology, and biochemistry faculty to learn how they explain and teach the mechanism by which ATP provides energy. The findings from this study informed my final study, in which we designed a task that supports students' understanding of the mechanism by which ATP drives the unfavorable formation of glutamine from glutamate and ammonium. We focus specifically on the role of ATP in transferring a phosphoryl group to create a reactive intermediate, avoiding the common, yet irrelevant, ideas about ATP hydrolysis and bond energy. After working through this task, both molecular biology and organic chemistry students included more mechanistic resources in their final explanations about the role of ATP, suggesting the potential of formative tasks as learning opportunities for students to advance and refine their productive resources.

The findings from these four studies point to the challenge, but importance, of incorporating MR into undergraduate science courses. Leveraging interdisciplinary thinking and MR are ways to support making informed decisions as citizens; for example, reasoning about how and why vaccines prevent the spread of infectious diseases, among other reasoning strategies (e.g., social and historical), can support one to make an informed choice. Further, many of the students we serve are pre-health majors, meaning they will pursue careers which often require deep reasoning and understanding about how or why symptoms occur, potential side effects of treatments, and how to appropriately address challenging decisions. MR is an important tool for certain contexts to ultimately help them do this.

I dedicate this dissertation to my mother, Nancy Carlson. I am so proud to be your daughter, and who I am is because of you. Thank you for believing in every part of me, always.

ACKNOWLEDGEMENTS

The list of people that should be acknowledged is longer than that I will provide here. I have the most supportive team, family, and friends, all of whom contributed in some way – whether as listening ears to my outpours over a beer, as philosophical contributors to a mess of thoughts and perspectives, or as thoughtful editors and commentators to my written products – to the research communicated in this dissertation.

First and foremost, to Dr. Melanie Cooper, who trusted, supported, and challenged me as I took on these projects. I feel eternally grateful to have learned and worked with one of the greatest minds in Chemistry and in Science Education. Dr. Cooper, a walking body of wisdom, curiosity, and care, served as a supportive and inspiring advisor – her role in this work cannot be understated. Thank you, thank you!

Secondly, I must acknowledge (does that word capture my gratitude?) Dr. Keenan Noyes, who, as a third-year graduate student, pulled me under his wing immediately after I joined the Cooper group. Keenan provided the mentorship that every graduate student deserves, initially giving consistent support and encouragement, followed by gradual release as I took on my own projects. His interests inspired my own, and I am indebted to him. I deeply aspire to continue collaborating with Keenan on future projects.

Next, to my many collaborators across disciplines: Dr. Mike Klymkowsky, Dr. Jon Stoltzfus, Dr. Tammy Long, and Dr. Christina Schwarz, among others. Most graduate students attend group meetings within a single environment, but I had the rare and privileged opportunity to engage in frequent discussions with researchers outside of the chemistry education bubble. We frequently hear about the advantages of diversity of thought, and I truly

gained more than I can put into words from the many deep, thoughtful, difficult, inspiring, joyful, and wonderful conversations with these humans. Thank you!

To the Cooper group! To Olivia Crandell, Ryan Bowen, Samantha Houchlei, Sewwandi Abeywardana, Kriti Seth, Veeda Scammahorn, Jacob Starkie, Brit Eggly, and Dani Losinski – thank you for being the ones on which I leaned, the ones that made me smile amidst a mountain of work, the ones that never said no to The Dairy Store, or a Panera run, or a walk to Sparty's, the ones to which I could vent and who felt comfortable venting to me. Graduate school equates to an actual roller coaster, highs and lows of varying magnitude that make you want to look over your shoulder and cry/laugh with those on the same ride. I love you all.

To some specifics – Bill Weston, for cheering me on, always making me laugh, and helping raise me, thank you so deeply for taking up a part of my heart; and the Franovic's: Dajana, Marko, Selina, BJ, and Christine, thank you for your endless love and support. I am forever grateful for a second family, one in which I have felt so welcome and one that has lifted and respected me in all that I do. I love you all!

No part of this research could have been completed without my perfect, astounding, inspiring, supportive, beautiful mother and sisters. To my mom, Nancy Carlson, my number one fan – thank you for your genuine interest in what I do. Thank you for raising three girls on your own, inspiring me to chase my dreams, and believing without an ounce of doubt that I am capable of anything. You are superwoman, a selfless/caring nurse and instructor, a loquaciously curious mother and friend, and my forever hero. To Mary and Rachel – the absolute, undoubtedly, unequivocally best siblings. You both inspire me beyond what you might imagine, taking on the world with humility, joy, awe, and courage. Thank you for long voice memos, for

belly laughs, for sharing tears as we experienced loss, and for lifting me up always. My heart swells for the three most important women in my life.

Lastly, and most importantly, to my husband, Sreten. Thank you for being my rock, an unfailing, unwavering, sturdy, and comforting support throughout the past 4.5 years. These years have included the best and most difficult of our lives, your constancy in love and support serving as a gift for which I will forever be grateful. Thank you for being my shoulder, for listening to me, for adventuring and doing life with me, but above all, thank you for believing in me. I love you more than you know.

TABLE OF CONTENTS

Chapter I – Introduction

Undergraduate STEM courses are tasked with preparing students to become active citizens in an increasingly complex world. Thus, it is the responsibility of instructors and institutions to support these students in thinking and reasoning deeply about carefully chosen phenomena, so that they can make informed and productive decisions/predictions/explanations in their everyday lives.

Contributing productively to our complex world includes tackling local and global challenges, such as climate change and sustainability, which requires the integration of knowledge across disciplines. That is, an idea more relevant to one discipline might be integral to a decision being made in another discipline. For example, advances in polymer chemistry can have profound effects on improved environmental systems, or a deep understanding of electrostatic forces and interactions can inform predictions/decisions about the effectiveness of vaccinations. The skill of thinking across (often siloed) disciplines involves using and applying prior knowledge to novel situations, something that is implicitly expected for students enrolled in requisite courses (i.e., ideas learned in general chemistry, a pre- or co-requisite for many upper-level STEM courses, should be used and advanced in biology). In this thesis, I emphasize the importance of interdisciplinary learning in both research design and instructional implications. By interdisciplinary, which is often conflated with multidisciplinary or transdisciplinary, I mean the integration or overlap of ideas across disciplines (chemistry and biology) which can be used to explain or predict the occurrence of some phenomena.

Connecting across disciplines (i.e., interdisciplinary learning/thinking) can be supported by the thinking strategy of mechanistic reasoning, which requires connecting entities and their

activities to explain or predict a target phenomenon[1–5]. This thinking strategy, explained further in Chapter II, is essential to scientific disciplines and, when paired with interdisciplinary thinking, may result in powerful and informed decisions, predictions, and/or investigations about the world. While there are several productive scientific reasoning strategies (e.g., mathematical, probabilistic, systems thinking)[6], my thesis work focuses specifically on supporting students to reason mechanistically about interdisciplinary phenomena. By focusing on this specific reasoning strategy, I could more carefully design assessments to elicit this type of thinking and characterize student explanations based on an understanding of this type of thinking.

To support students as they develop and refine this thinking strategy (or epistemic heuristic), it is imperative that we provide opportunities for them to practice, without penalty, in their undergraduate courses. I have narrowed these broader goals by designing instructional materials, and corresponding rubrics, that engage students in mechanistic reasoning about phenomena that carry relevance in, and leverage ideas from, both chemistry and biology: (1) preferential protein-ligand binding and (2) ATP-driven coupled reactions. Through analysis of students' written explanations and instructor interviews, this work provides implications for the challenges of mechanistic reasoning as well as instructional tools and implications to mitigate these challenges.

Finally, there is limited work showing whether and how mechanistic reasoning tasks relate to equity in the classroom. Ralph et al. (2022) showed that exams emphasizing mechanistic reasoning assessment items predict more equitable student outcomes in general chemistry courses[7]. Further, curricula designed to emphasize mechanistic reasoning (through

three-dimensional learning) result in more students receiving a passing grade in "gateway" courses, which often inhibit historically marginalized students from pursuing STEM degrees[8]. While limited, this literature suggests that a focus on mechanistic reasoning may be an equitable practice; however, additional evidence should be uncovered to learn more about the interplay between mechanistic reasoning and equity. A step in this direction involves disaggregating student data to reflect on instructional practices/materials and their impact(s) on all students, which I have done in my second study.

The following sections provide a brief summary of the four studies that I conducted for my thesis work. Each of these studies contributes to an understanding of mechanistic reasoning and interdisciplinary learning in the context of undergraduate chemistry and biology courses. I use both quantitative and qualitative research methods to explore and characterize students' written explanations and instructors' discussions about phenomena that bridge between the disciplines of chemistry and biology.

Summary Of Study Goals And Research Questions

Study 1: Student Engagement In CMR About Protein-Ligand Binding

In the first study, we were interested in how students across a range of undergraduate chemistry and biology courses engaged in causal mechanistic reasoning (CMR) about preferential protein-ligand binding, an interdisciplinary phenomenon that can be explained using chemistry and biology ideas. We used a carefully designed coding scheme to characterize responses as non-casual mechanistic (CM), partially CM, or fully CM based on the presence or absence of three key ideas in students' explanations: the attraction of oppositely charged species, the negative/polar nature of atoms or amino acids in binding site(s), and the strength

of attraction as impacted by magnitude of charge. This study addressed the following three research questions:

1. How do students in chemistry and biology courses engage in causal mechanistic reasoning in the context of protein-ligand binding?
2. What conceptual resources do students enrolled in chemistry and biology courses use when explaining this phenomenon?
3. How does engagement in CMR relate to students' overall predictions?

Study 2: How Do Different Groups Of Students Engage In CMR About Protein-Ligand Binding?

In the second study, I used ordinal regression analysis to identify variables that predict, or do not predict, student engagement in CMR for the protein-ligand binding task. We coded an additional 800+ student responses to appropriately compare engagement in CMR based on (binary) gender identification and race/ethnicity, with GPA as a covariate. There is limited work showing whether mechanistic reasoning tasks (of which there are few) are equitable. Here, we compared student engagement in CMR based on two demographic variables (gender and race/ethnicity) to add to this literature base. The study addresses the following research questions:

1. How do males and females compare in their engagement in CMR about protein-ligand binding?
2. How do White students compare to Non-White students in their engagement in CMR about protein-ligand binding?
3. To what extent do cumulative GPA, (binary) gender, and race/ethnicity predict student engagement in CMR?

Study 3: Instructor Explanations For The Mechanism By Which ATP Drives Unfavorable Processes

In the third study, we interviewed 15 instructors across chemistry, biology, and biochemistry disciplines to explore how they think and teach about the mechanism(s) by which ATP provides energy. This phenomenon was chosen because of the well-documented misconception that breaking bonds releases energy (frequently associated with the role of ATP in biological systems), when, in fact, the opposite is true. In order to better understand the role of ATP and how to support instructors in teaching this role, we used a semi-structured interview protocol to discuss these ideas with a range of chemistry, biology, and biochemistry faculty. The analyses and results are driven by three research questions:

1. What ideas did instructors leverage when discussing ATP?
2. In what ways do instructors use these ideas to discuss how ATP drives unfavorable processes?
3. What teaching experiences did molecular biology (MB) instructors share regarding ATP in their course(s)?

Study 4: Investigating The Impact Of A Learning Task On Students' Use Of Mechanistic Resources When Explaining A Complex Phenomenon

In the final study, we leveraged themes from the instructor interviews to design and administer a learning task aimed to support students' mechanistic understanding of the role of ATP in driving unfavorable reactions. The task includes both an initial and final question prompting students to explain *how* ATP drives the formation of glutamine from glutamate and ammonium, allowing us to investigate the impact of the task on student use of mechanistic

resources. We administered the task to both Molecular Biology and Organic Chemistry students in order to support our interdisciplinary efforts. After developing and refining a coding scheme, we characterized the student explanations and answered the following research questions:

1. How do students explain the role of ATP in driving the unfavorable formation of glutamine from glutamate and ammonium?
2. What is the impact of the task on students' use of mechanistic resources when explaining this phenomenon?
3. How do Molecular Biology student explanations compare to Organic Chemistry student explanations for this phenomenon?

REFERENCES

(1) Krist, C.; Schwarz, C. V.; Reiser, B. J. Identifying Essential Epistemic Heuristics for Guiding Mechanistic Reasoning in Science Learning. *Journal of the Learning Sciences* **2019**, *28* (2), 160–205. https://doi.org/10.1080/10508406.2018.1510404.

(2) Russ, R. S.; Scherr, R. E.; Hammer, D.; Mikeska, J. Recognizing Mechanistic Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed from Philosophy of Science. *Science Education* **2008**, *92* (3), 499–525. https://doi.org/10.1002/sce.20264.

(3) Machamer, P.; Darden, L.; Craver, C. F. Thinking about Mechanisms. *Philosophy of Science* **2000**, *67* (1), 1–25. https://doi.org/10.1086/392759.

(4) van Mil, M. H. W.; Boerwinkel, D. J.; Waarlo, A. J. Modelling Molecular Mechanisms: A Framework of Scientific Reasoning to Construct Molecular-Level Explanations for Cellular Behaviour. *Sci & Educ* **2013**, *22* (1), 93–118. https://doi.org/10.1007/s11191-011-9379-7.

(5) Glennan, S. *The New Mechanical Philosophy*; Oxford University Press: Oxford, New York, 2017.

(6) Osborne, J.; Rafanelli, S.; Kind, P. Toward a More Coherent Model for Science Education than the Crosscutting Concepts of the next Generation Science Standards: The Affordances of Styles of Reasoning. *Journal of Research in Science Teaching* **2018**, *55* (7), 962–981. https://doi.org/10.1002/tea.21460.

(7) Ralph, V. R.; Scharlott, L. J.; Schafer, A. G. L.; Deshaye, M. Y.; Becker, N. M.; Stowe, R. L. Advancing Equity in STEM: The Impact Assessment Design Has on Who Succeeds in Undergraduate Introductory Chemistry. *JACS Au* **2022**. https://doi.org/10.1021/jacsau.2c00221.

(8) Matz, R. L.; Fata-Hartley, C. L.; Posey, L. A.; Laverty, J. T.; Underwood, S. M.; Carmel, J. H.; Herrington, D. G.; Stowe, R. L.; Caballero, M. D.; Ebert-May, D.; Cooper, M. M. Evaluating the Extent of a Large-Scale Transformation in Gateway Science Courses. *Science Advances* **2018**, *4* (10), eaau0554. https://doi.org/10.1126/sciadv.aau0554.

Chapter II – Theoretical Considerations

The culmination of my work is not only a result of the data itself, but also the theoretical frameworks which influenced how I chose to collect the data, how I analyzed the data, and the resulting synthesis and report of the data. My perception of how people learn and what they should learn stands firmly upon a few complementary theories, each of which I describe below.

How People Learn – Constructivism And Resources

Several works have been published regarding how people learn, notably the books *How People Learn I* and *II*[1,2], and in recent decades, education researchers have widely accepted a constructivist theory of knowledge, which builds on renown psychologist Jean Piaget's theory of cognitive development[3]. According to Piaget, humans construct knowledge (or make meaning) by integrating new information with their prior experiences and ideas. In 1986, George Bodner published an article that applied Piaget's theory to the chemistry classroom, initiating the constructivist perspective of learning in the chemistry education community[4]. This understanding of how people learn, which is based on learners *constructing* their knowledge frameworks by integrating new knowledge with prior knowledge, is in contrast to that of a coherence perspective, in which knowledge is perceived as more connected, coherent units which can be added or replaced by more expert-like ideas[4,5], or that the integration of new information is done in particular ways[6]. My work relies on the constructivist perspective and, more specifically, Hammer's description of "resources" in the context of this perspective.[7–10] According to Hammer (2000), resources represent fine-grained conceptual or epistemological knowledge elements that are connected or disconnected in a dynamic knowledge framework[8]. These resources, given their dynamic and context-dependent nature, are activated according to

different situations depending on (1) the connections that exist in a student's mind, and (2) the context in which the student or phenomenon is situated[8]. For example, the terms "donor" and "acceptor" in an organic chemistry class typically refer to Lewis acids and bases, with the "donor" being the molecule donating a pair of electrons and the "acceptor" accepting those electrons; however, in a biochemistry class "donors" and "acceptors" may be discussed in the context of hydrogen bond donors and acceptors, in which case the donor is the molecule with the hydrogen, while the "acceptor" is the molecule with a set of lone pairs participating in that H-bond. Therefore, the class (context) in which a student is situated would likely activate the "donor" definition that is most relevant to the class.

With this perspective, rather than knowledge being perceived as intact, stable conceptions which are either correct or incorrect one can think about conceptions as a network of resources which might be connected in productive or unproductive ways[7]. This theoretical framework allows for a mechanism of knowledge construction (refinement, development, conceptual change, etc.) which is in contrast with the idea that incorrect conceptions should be removed and replaced by correct or expert-like conceptions.

Hammer's resources perspective is also applicable in the realm of knowledge "transfer"[9], a term used less frequently as there is no evidence of full "transfer" of ideas to exist. As educators, we hope to provide students with knowledge and tools that they can *apply* to (or *activate* in) new contexts[9]. This is particularly true for requisite courses, in which an understanding of new contexts (i.e., systems or phenomena) relies on previous knowledge. For example, to enroll in molecular biology, students often need general chemistry as a pre- or co-requisite, suggesting and promoting the potential of knowledge transfer. Rather than

transferring stable conceptions, however, Hammer considers resources being activated in new contexts, allowing for further refinement and organization of students' knowledge structures. Developing canonical conceptions is not trivial, and it is the job of the educator to set students up so that resources can be activated in a variety of contexts, allowing for deep and meaningful connections of ideas.

What Should Students Know And Be Able To Do?

The constructivist/resources theory of cognition that I use provides information regarding how students learn; however, my work is also informed by frameworks considering *what* students should learn. That is, in science education, what do we want students to know and be able to do?

Three-Dimensional Learning

In 2012, the National Academies published a consensus report on a Framework for K-12 Education (hereon referred to as *the Framework*), which outlined three dimensions of learning: disciplinary core ideas, crosscutting concepts, and scientific and engineering practices[11]. Several works since 2012 have furthered our community's understanding of these three dimensions and how they can be put into practice, including works that have applied *the Framework* to post-secondary education. For example, Cooper et al. (2013) developed a general chemistry curriculum that leverages three-dimensional learning – Chemistry, Life, the Universe and Everything (CLUE)[12]. This curriculum has more recently been continued with organic chemistry (OCLUE)[13]. There is evidence that these curricula provide opportunities for students to engage in 3DL via course assignments, recitation activities, and assessments[14,15].

Prior to publication of *the Framework*, the American Association for the Advancement of Science (AAAS) published a call to action for undergraduate biology education called Vision and Change[16]. Similar to *the Framework* and CLUE, Vision and Change emphasizes "core concepts" that should be included across all undergraduate biology curricula. Unlike CLUE, however, Vision and Change does not specifically leverage 3DL. My work focuses on core ideas outlined by CLUE and Vision and Change (*structure-property relationships* and *energy*); two cross-cutting concepts (*energy* and *cause and effect: mechanism and explanation*); and the scientific practice *constructing explanations*. Note the overlap in some areas, for example *energy* is considered both a core idea and a cross-cutting concept.

Mechanistic Reasoning

Mechanistic reasoning (MR) about scientific phenomena, or reasoning about how and why phenomena occur, is a powerful explanatory and predictive tool. Engaging in this type of reasoning during instruction provides students the opportunity to engage in 3DL because it involves the crosscutting concept of cause and effect, the practice of constructing explanations, and typically leverages a core idea depending on the phenomenon under consideration. The education community clearly values MR; however, without a doubt, the science community also values this type of reasoning (and investigation). I searched the term "mechanism" in both *Science* and *Nature*, to find 10,000+ and 366,603 hits, respectively.

MR (or the thinking strategy involved in constructing a mechanistic explanation/prediction) has been defined by several researchers, all of whom agree on the importance of (1) entities, (2) "activities" of those entities, and (3) linking these ideas to the phenomenon under consideration[17–21]. For example, my initial research relied on an

understanding of Krist et al.'s framework for MR (2018) (a modified version of that outlined by Russ et al. (2008)), which they describe as an epistemic resource, or thinking strategy, involving three heuristics. According to this framework, the first step requires considering entities at the scalar level below that of the target phenomenon; the second step involves "unpacking" the behaviors/properties of entities at that lower scalar level; and the third step requires linking those lower-level interactions and behaviors to the target phenomenon. These steps are illustrated in Figure 2.1.

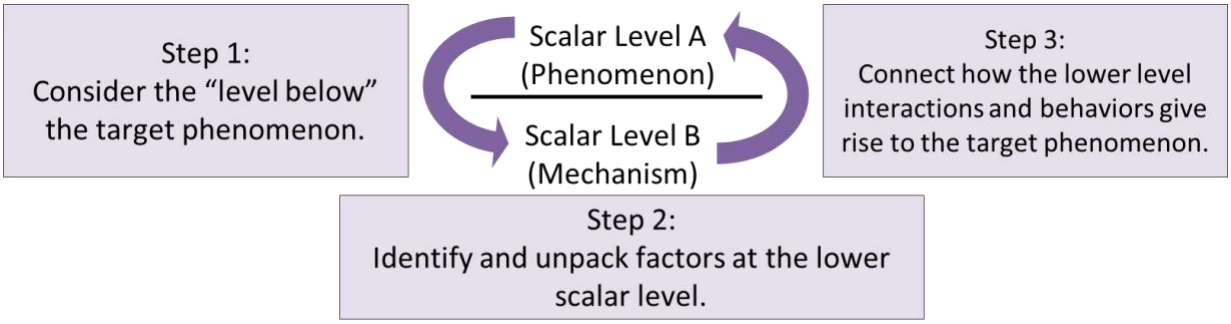


Figure 2.1. Krist et al.'s (2018) essential epistemic heuristics for mechanistic reasoning.[16]

While the authors developed this MR framework with K-12 education in mind, its utility transfers to undergraduate education as well. Consider a "mechanistic explanation" for what happens (and how/why it happens) when two neutral atoms approach each other. As the two atoms approach each other, one of them experiences an instantaneous dipole, that is, the electron density is randomly increased on one end of the atom. This dipole induces a dipole in the other atom, because the partially negative end of the first atom repels the electrons of the second atom, pushing its electron density to the opposite end. The opposite partial charges of each dipole attract each other, and the result is a weak, but stable, interaction (a London-

dispersion force). Noyes et al. (2019) examined general chemistry students' explanations of this phenomenon, characterizing the explanations using an understanding of the frameworks proposed by Russ et al. (2008) and Krist et al. (2018), each of which identifies the importance of the lower scalar level (in this case, the subatomic level) and the activities/changes that occur at this level[15,18,20].

These MR frameworks can be implemented for a variety of simple scientific phenomena; however, many biological phenomena are complex and multi-scalar, spanning from the microscopic/genetic levels to entire ecosystems. This poses a challenging question: if there are multiple scalar levels below that of the target phenomenon, how "deep" does one need to go in order to construct a fully mechanistic explanation? And in what context (e.g., to satisfy an exam question, to predict what might happen in a particular context)? A more recent approach to this question involves the recognition and use of explanatory black boxes, or "mechanisms within mechanisms". Haskel-Ittah (2023) called for increased attention on explanatory black boxes in science education, particularly to help students avoid an "illusion of explanatory depth, or the sense that one understands causally complex phenomena more deeply than one really does"[22]. My work has progressed into helping students construct explanations for more complex phenomena, thereby increasing the importance of explanatory black boxes.

Equity And Justice-Oriented Research

In 2022, The Boyer 2030 Commission published a report emphasizing "the equity-excellence imperative", calling on institutions to put both equity and excellence at the forefront of undergraduate education[23]. The commission argues that equity and excellence are

intertwined, that you cannot have one without the other, and it is imperative that evidence-based practices to support historically marginalized students should be implemented alongside practices that engage students in deep learning opportunities and experiences. While discipline-based education research (DBER) is defined as "an empirical approach to investigating learning and teaching that is informed by an expert understanding of [STEM] disciplinary knowledge and practice"[24], engaging in this research should also be informed by equity and justice-oriented approaches in order to best serve both teachers and learners. There is evidence that certain instructional practices/norms are more equitable than others – for example, grading on a curve disproportionately favors the majority, often "weeding out" historically marginalized students[25]. As educators, we must create spaces for all students to succeed, and the types of assessments we use can serve as a means by which to achieve more equitable outcomes. Further, the assessments we use send a strong message to students about what is valued in the course or discipline[26]. Mechanistic reasoning tasks require students to draw and/or explain phenomena and may serve as equitable tasks that provide reliable evidence of what students know and can do. It has been shown that the use of three-dimensional assessments (that is, those that provide the opportunity for students to engage in three-dimensional learning, such as mechanistic reasoning) result in increased student success regardless of background[15,27,28]. Ralph et al. (2022) showed that assessments with more mechanistic reasoning items than those that require calculations or rote memorization result in more equitable outcomes for all students, specifically Black and Latinx students[28]. Using these approaches, I have dedicated a chapter to disaggregating student data in order to critically

investigate whether the task and analysis I conducted are equitable across race/ethnicity and binary gender identification.

Summary Of Theoretical Framework And Application To My Work

Unfortunately, there is evidence that some students see chemistry and biology courses as disconnected[29]. That is, they know what they need to say in chemistry and what they need to say in biology, even if the information is seemingly contradictory. In my research, I have used Hammer's resources perspective as a tool for both developing assessments and analyzing student responses, ultimately with the goal of helping students build connections across these disciplines. Specifically, I focused on student engagement in mechanistic reasoning in the context of biochemical phenomena including (1) preferential protein-ligand binding and (2) reaction coupling via ATP in common biological processes. With an emphasis on a resources-perspective of learning, my work has shown that engaging in MR leads to more correct predictions in the context of protein-ligand binding. This finding, along with student difficulties in connecting ideas about energy between chemistry and biology, led to my investigation of the mechanism by which ATP is used as an energy source. In this pursuit, I (and a colleague) interviewed faculty from chemistry, biology, and biochemistry to develop a picture of instructors' understanding of the mechanism, resources they use in explaining the mechanism, and instructional strategies used when teaching about ATP as an energy source. The rich discussions that emerged from these interviews supported my continuation of assessment development, specifically related to ATP. The resources-perspective of learning becomes especially apparent when asking students to explain complex phenomena such as how and why ATP drives the formation of glutamine from glutamate and ammonium. This theory of learning

paired with a deep understanding of what it means to reason mechanistically about complex phenomena form the pillars of my research. These two themes have firmly supported my investigation and analysis of my work, serving as a tightly woven thread by which my discussions and conclusions are based. Using this theoretical framework, I have aimed to help students activate and develop their resources associated with “chemistry” (i.e., electrostatic forces and interactions, and energy), by proposing biological phenomena which can be explained mechanistically by leveraging those resources.

REFERENCES

(1) Bransford, J. D.; Brown, A. L.; Cocking, R. R. *How People Learn: Brain, Mind, Experience, and School*, Expanded.; National Academies Press, 2000.

(2) National Academies of Sciences, Engineering, and Medicine. *How People Learn II: Learners, Contexts, and Cultures*, 2nd ed.; National Academies Press, 2018.

(3) Piaget, J. Part I: Cognitive Development in Children: Piaget Development and Learning. *Journal of Research in Science Teaching* **1964**, *2* (3), 176–186. https://doi.org/10.1002/tea.3660020306.

(4) Bodner, G. M. Constructivism: A Theory of Knowledge. *J. Chem. Educ.* **1986**, *63* (10), 873. https://doi.org/10.1021/ed063p873.

(5) Kuhn, T. S. *International Encyclopedia of Unified Science. Foundations of the Unity of Science*, 2nd ed.; University of Chicago Press: Chicago, IL, 1970; Vol. 2.

(6) Chinn, C. A.; Brewer, W. F. The Role of Anomalous Data in Knowledge Acquisition: A Theoretical Framework and Implications for Science Instruction. *Review of Educational Research* **1993**, *63* (1), 1–49. https://doi.org/10.3102/00346543063001001.

(7) Smith, J. P. I.; diSessa, A. A.; Roschelle, J. Misconceptions Reconceived: A Constructivist Analysis of Knowledge in Transition. *Journal of the Learning Sciences* **1994**, *3* (2), 115–163. https://doi.org/10.1207/s15327809jls0302_1.

(8) Hammer, D. Student Resources for Learning Introductory Physics. *American Journal of Physics* **2000**, *68* (S1), S52–S59. https://doi.org/10.1119/1.19520.

(9) Hammer, D.; Elby, A.; Scherr, R. E.; Redish, E. F. Resources, Framing, and Transfer. In *Transfer of learning from a modern multidisciplinary perspective*; Greenwich, CT, 2005; pp 89–120.

(10) diSessa, A. A. Toward an Epistemology of Physics. *Cognition and Instruction* **1993**, *10* (2–3), 105–225. https://doi.org/10.1080/07370008.1985.9649008.

(11) National Research Council. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*; National Academies Press: Washington, DC, 2012.

(12) Cooper, M.; Klymkowsky, M. Chemistry, Life, the Universe, and Everything: A New Approach to General Chemistry, and a Model for Curriculum Reform. *J. Chem. Educ.* **2013**, *90* (9), 1116–1122. https://doi.org/10.1021/ed300456y.

(13) Cooper, M. M.; Stowe, R. L.; Crandell, O. M.; Klymkowsky, M. W. Organic Chemistry, Life, the Universe and Everything (OCLUE): A Transformed Organic Chemistry Curriculum. *J. Chem. Educ.* **2019**, *96* (9), 1858–1872. https://doi.org/10.1021/acs.jchemed.9b00401.

(14) Crandell, O. M.; Lockhart, M. A.; Cooper, M. M. Arrows on the Page Are Not a Good Gauge: Evidence for the Importance of Causal Mechanistic Explanations about Nucleophilic Substitution in Organic Chemistry. *J. Chem. Educ.* **2020**, *97* (2), 313–327. https://doi.org/10.1021/acs.jchemed.9b00815.

(15) Noyes, K.; Cooper, M. M. Investigating Student Understanding of London Dispersion Forces: A Longitudinal Study. *J. Chem. Educ.* **2019**, *96* (9), 1821–1832. https://doi.org/10.1021/acs.jchemed.9b00455.

(16) American Association for the Advancement of Science. *Vision and Change in Undergraduate Biology Education: A Call to Action*; Washington D.C., 2011. https://live-visionandchange.pantheonsite.io/wp-content/uploads/2011/03/Revised-Vision-and-Change-Final-Report.pdf (accessed 2020-04-17).

(17) Glennan, S. *The New Mechanical Philosophy*; Oxford University Press: Oxford, New York, 2017.

(18) Krist, C.; Schwarz, C. V.; Reiser, B. J. Identifying Essential Epistemic Heuristics for Guiding Mechanistic Reasoning in Science Learning. *Journal of the Learning Sciences* **2019**, *28* (2), 160–205. https://doi.org/10.1080/10508406.2018.1510404.

(19) Machamer, P.; Darden, L.; Craver, C. F. Thinking about Mechanisms. *Philosophy of Science* **2000**, *67* (1), 1–25. https://doi.org/10.1086/392759.

(20) Russ, R. S.; Scherr, R. E.; Hammer, D.; Mikeska, J. Recognizing Mechanistic Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed from Philosophy of Science. *Science Education* **2008**, *92* (3), 499–525. https://doi.org/10.1002/sce.20264.

(21) van Mil, M. H. W.; Boerwinkel, D. J.; Waarlo, A. J. Modelling Molecular Mechanisms: A Framework of Scientific Reasoning to Construct Molecular-Level Explanations for Cellular Behaviour. *Sci & Educ* **2013**, *22* (1), 93–118. https://doi.org/10.1007/s11191-011-9379-7.

(22) Haskel-Ittah, M. Explanatory Black Boxes and Mechanistic Reasoning. *Journal of Research in Science Teaching* **2023**, *60* (4). https://doi.org/10.1002/tea.21817.

(23) The Boyer 2030 Commission. *The Equity-Excellence Imperative*; The Association for Undergraduate Education at Research Universities, 2022; pp 3–59. https://ueru.org/boyer2030 (accessed 2023-10-17).

(24) National Research Council. *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*; National Academies Press: Washington, D.C., 2012.

(25) Bowen, R. S.; Cooper, M. M. Grading on a Curve as a Systemic Issue of Equity in Chemistry Education. *J. Chem. Educ.* **2022**, *99* (1), 185–194. https://doi.org/10.1021/acs.jchemed.1c00369.

(26) Stowe, R. L.; Scharlott, L. J.; Ralph, V. R.; Becker, N. M.; Cooper, M. M. You Are What You Assess: The Case for Emphasizing Chemistry on Chemistry Assessments. *J. Chem. Educ.* **2021**, *98* (8), 2490–2495. https://doi.org/10.1021/acs.jchemed.1c00532.

(27) Matz, R. L.; Fata-Hartley, C. L.; Posey, L. A.; Laverty, J. T.; Underwood, S. M.; Carmel, J. H.; Herrington, D. G.; Stowe, R. L.; Caballero, M. D.; Ebert-May, D.; Cooper, M. M. Evaluating the Extent of a Large-Scale Transformation in Gateway Science Courses. *Science Advances* **2018**, *4* (10), eaau0554. https://doi.org/10.1126/sciadv.aau0554.

(28) Ralph, V. R.; Scharlott, L. J.; Schafer, A. G. L.; Deshaye, M. Y.; Becker, N. M.; Stowe, R. L. Advancing Equity in STEM: The Impact Assessment Design Has on Who Succeeds in Undergraduate Introductory Chemistry. *JACS Au* **2022**. https://doi.org/10.1021/jacsau.2c00221.

(29) Kohn, K. P.; Underwood, S. M.; Cooper, M. M. Energy Connections and Misconnections across Chemistry and Biology. *LSE* **2018**, *17* (1), ar3. https://doi.org/10.1187/cbe.17-08-0169.

Chapter III – Literature Review

Three-Dimensional Learning In Chemistry And Biology

In the past decade, there has been a push in the science education community to establish courses that emphasize and value three-dimensional learning (3DL) – as outlined by the National Research Council's consensus report (2012) – which calls for the integration of disciplinary core ideas, crosscutting concepts, and scientific practices[1]. The disciplinary core ideas were established for K-12 life science, physical science, earth and space science, and engineering science[1]; however, undergraduate educators/researchers have defined core ideas for chemistry, biology, and physics as well. For example, Cooper et al. (2013) identified four core ideas for undergraduate chemistry including (1) electrostatic and bonding interactions, (2) atomic/molecular structure and properties, (3) energy, and (4) change and stability in chemical systems[2]. Each of these core ideas is related to the others and can be used to explain or predict various chemical phenomena[2]; for example, we can explain how and why the boiling point of ethanol is higher than that of acetone by leveraging *atomic/molecular structure and properties* to identify types of intermolecular interactions (*electrostatic and bonding interactions*) and, finally, relate the strength of these interactions to how much *energy* is required to overcome the attractive electrostatic forces between molecules. The other two dimensions, crosscutting concepts and scientific practices, do not depend on the discipline – that is, they span the sciences and should be incorporated in each. Figure 3.1 displays the three dimensions (with chemistry core ideas).

Similar to 3DL, but specific for undergraduate biology education, the American Association for the Advancement of Science (AAAS), published a call to action (*Vision and*

Change) for the use of five core concepts (for example, information flow and transformations of energy & matter) and six core competencies (for example, the ability to apply the process of science and use quantitative reasoning) to guide instruction of undergraduate biology courses[3].

Figure 3.1. The three dimensions of 3DL. Disciplinary core ideas are represented by those outlined for Undergraduate Chemistry[2].

While each of the three dimensions carries significance in science education, I have aimed to link specific core ideas and crosscutting concepts across chemistry and biology through the practice of *constructing explanations*. In particular, my work focuses on two chemistry core ideas: (1) electrostatic and bonding interactions and (2) energy; two biology core ideas: (1) structure & function and (2) energy; and three crosscutting concepts: (1) cause and effect: mechanism and explanation, (2) structure & function, and (3) energy (Figure 3.1). These core ideas and crosscutting concepts overlap significantly, meaning there should (theoretically) be several ways in which we can investigate student learning across chemistry and biology using these themes. For example, the crosscutting concept of *structure & function* can be used to explain the phenomenon of protein-ligand binding both in chemistry (when

thinking about the types of interactions that cause ligand binding) and in biology (when thinking about how and why different proteins bind different ligands)[4].

Previous work in curriculum design has leveraged three-dimensional learning to transform undergraduate chemistry and/or biology courses, such as those proposed by Cooper and Klymkowsky including (*Organic*) *Chemistry, Life, the Universe and Everything* ((O)CLUE) and *Biofundamentals*[5–7]. Ample evidence exists for the efficacy of CLUE in helping students to build a deeper understanding of chemistry (when compared to students in traditional general chemistry courses) by frequently asking them to construct explanations and model phenomena (i.e., engage in 3DL)[8–12]. My work draws largely from the impact of courses designed to implement 3DL, as these courses served as the context for much of my data collection.

Interdisciplinary Learning – Connecting Chemistry And Biology

In tandem with 3DL, there should also be an emphasis on interdisciplinary learning, or the coordination of ideas across disciplines in approaching scientific challenges. These challenges include global ones, e.g., mitigating climate change effects, as well as local or personal, e.g., making predictions or decisions about potential health risks. The term "interdisciplinary" differs from "multidisciplinary" and "transdisciplinary", which are often conflated or poorly defined. By "interdisciplinary", we mean the coordination or coherence between two or more disciplines – that is, how the ideas in the disciplines are connected or complementary. A multidisciplinary perspective involves using ideas that are distinct in different disciplines (for example, one might take either a historical or artistic approach when examining a painting[13]; transdisciplinary involves "transcending" (going beyond) disciplines or taking a more holistic approach to develop new perspectives[13]. While multiple disciplinarity can

include each of these dimensions, my research specifically aims to leverage interdisciplinary learning in supporting students as they build coherency between ideas in chemistry and biology.

Most undergraduate STEM curricula require introductory chemistry and/or biology courses as pre- or co-requisites for upper-level courses in these disciplines, implying and suggesting disciplinary connections; however, these courses typically remain siloed and disconnected to other disciplines (or even other courses in the same discipline), and there is limited work investigating how to help students build connections between chemistry and biology[4,14–18]. Work that has been done includes tasks developed to encourage students to use chemistry core ideas when explaining biological (or biochemical) phenomena. In one of these tasks, Martinez et al. (2021) used a scaffolded osmosis activity to help students use their knowledge of entropy (a critical chemistry idea) to explain osmosis across a cell membrane (a common biological phenomenon)[17]. Similarly, Green et al. (2021) developed an activity linking energy changes during bond formation and breaking to the phenomenon of ATP coupling as an opportunity for students to connect ideas across these disciplines, finding via interviews that students valued this opportunity[15]. There are several themes and ideas that carry relevance in both chemistry and biology, and my work focuses on leveraging two of these to support students' mechanistic reasoning about biochemical phenomena: (1) structure, property, function relationships, and (2) energy.

Structure-Property-Function Relationships

In a chemistry context, structure-property-function (SPF) relationships typically focus on the link between structure and properties. That is, chemical properties (such as melting point)

are predicted or explained by considering the chemical structure. In a biology context, SPF relationships more often involve connecting structure to function. For example, the three-dimensional structure of a protein will dictate its function by, for example, preferentially binding to certain ligands over others. These ideas, while related, emerge as a result of the aims of the discipline and may seem, to a student, unrelated if not explicitly connected. However, faculty in chemistry and biology agree that these ideas should be connected for students, even though there are challenges in doing so[19]. Further, Kohn et al. (2018) interviewed students who were co-enrolled in introductory chemistry and biology to investigate the big ideas in each course and the connections (or lack thereof) that students noticed between the courses. They found that many students built a critical connection between chemistry and biology by recognizing links between structure, property, and function relationships[20]. These students noted structure & properties as a big idea in chemistry, and structure & function as a focus in biology, and they proposed this as an opportunity for both courses to help students identify and use this connection for more meaningful, interdisciplinary learning. Thus, SPF relationships may serve as an excellent area of investigation aimed at helping students to build connections across disciplines.

Energy

While the students that Kohn et al. (2018) interviewed recognized connections between chemistry and biology regarding SPF relationships,[20] these same students noticed distinct differences, and contradictions, in regard to energy in their courses[21]. Energy, a core idea in both disciplines and a crosscutting concept[1,3,5], should, theoretically, be a potential bridge between disciplines; however, one student said, in relation to energy, “I know for biology what

[the instructor] wants us to say and then for chemistry what we have to say"[21]. Energy in chemistry is frequently discussed in terms of bonds and the energy changes that occur when bonds are formed or broken. Biology, on the other hand, is typically not concerned with this fine-grained level of bonds breaking/forming; rather, its range spans from cellular/metabolic processes to population dynamics and ecosystems – i.e., explaining how energy is "captured" or "transferred" in order for some outcome to occur. This disconnect poses a challenge in linking ideas about energy between the disciplines; however, little research has been done to help students do this[15,18,22–24].

I intend to expand opportunities for students to build connections across chemistry and biology as they traverse STEM curricula, which include pre- and co-requisite courses designed with the implicit idea that students will apply and refine knowledge from one to another. We posit that by providing these opportunities, specifically by helping students to engage in mechanistic reasoning consistently, thereby developing habitual productive reasoning strategies, students might be able to better use their knowledge and build these connections.

Students' Mechanistic Reasoning In Chemistry And Biology

While there is an extensive literature base on mechanistic reasoning in K-12, this section focuses solely on mechanistic reasoning in chemistry education and biology education literature. Chemistry education research is fairly rich with evidence on how students construct (causal) mechanistic explanations[10,25–31]. In this work, researchers characterize student engagement in mechanistic reasoning by analyzing and making sense of student explanations or models for chemical phenomena. For example, Becker et al. characterized student explanations for the origins and causes of London dispersion forces, finding that explanations ranged from

non-mechanistic and non-canonical causes to full mechanisms and canonical electrostatic causes.[25] Some mechanistic reasoning work involves supporting students to reason this way by carefully designing assessments and frequently asking students in transformed courses to engage in mechanistic reasoning (via formative and summative assessment tasks)[5,10,32]. That is, by frequently asking students to explain how and why phenomena occur, they are better able to engage in this thinking strategy when faced with unfamiliar contexts[33]. Graulich et al. developed activities that capture students' identification and use of implicit properties when constructing mechanistic explanations in organic chemistry courses[27,34]. Similarly, Shultz et al. established a "writing to learn" program, which uses formative writing assignments in organic chemistry courses that engage students in constructing explanations and peer review about reaction mechanisms[25]. More recently, advances have been made in characterizing student explanations via machine learning, which involves analyzing large sets of students' responses to characterize explanations, providing practical implications for instruction[35,36].

To engage in mechanistic reasoning in a chemistry context often requires identifying lower-level entities (subatomic level) and their properties/activities, and then causally linking these aspects to the target phenomenon (which may occur at the molecular or macroscopic levels). A biological context, on the other hand, does not usually require going as deep as the subatomic level. In fact, because of the range of scalar levels explored in biology, the requirement to "identify entities at a scalar level below the target phenomenon", according to Krist et al. (2018), can be quite challenging. What scalar level is "deep enough" for a mechanistic explanation? Further, some phenomena occur as a result of entities that might be considered the same scalar level, and it is their interactions/activities that give rise to some

occurrence. Despite the importance of and emphasis on mechanisms in biology, only a small body of literature focuses on thinking about mechanisms as a way of reasoning rather than a process to memorize[37–39]. Further, studies on student mechanistic reasoning in biology have largely focused on genetic mechanisms or information flow[38,40–43], and less so on metabolic or molecular mechanisms[44], which may serve as a conceptual link between chemistry and biology, thereby providing an opportunity for interdisciplinary learning.

[Some] Research seems to support the idea that mechanistic reasoning may be a supportive, equitable, and reasonable (achievable) practice for students, even in large-enrollment undergraduate courses[5,7,26,45]. The surge in development of machine learning (ML) and artificial intelligence (AI) provides a tool for instructors to characterize large numbers of student explanations, allowing for feedback that can guide instructional decisions. Work in both chemistry and biology education has resulted in efficient ML models for various open-ended prompts[35,36,46], ideally resulting in increased use of explanatory assessment items in large-enrollment undergraduate courses. Further, extensive use of formative assessment tasks (which can both challenge and support students to construct complex explanations), may be a more equitable (and effective) teaching practice when compared to high-stakes summative assessments. The challenge then becomes, how do we design assessments to support and elicit this type of thinking?

Assessment Design – Eliciting Students' Mechanistic Explanations

Historically, assessments have been used to sort students – that is, which students know which material and how well? However, the push for "knowledge-in-use" (and 3DL) in the 21st century, has led to an emphasis on assessments designed to help students use and apply their

knowledge (as practicing scientists do)[1] – it is not simply what you know, but what you can do with what you know. Thus, the task for instructors becomes designing tasks that provide enough information for the students to understand what is being asked, but not so much that they can complete the task without thoughtful effort. I have relied on Mislevy's evidence-centered design (ECD) approach[47] to build an assessment that both (1) supports students in leveraging appropriate resources to explain a complex phenomenon (reaction coupling) and (2) elicit valid and reliable evidence of what students know and can do. ECD is based on an iterative process in which the designer first identifies what it is they want students to know and be able to do. Then, they decide what they will accept as evidence that students know or can do what was intended (evidence statements). Finally, the task is iterated accordingly based on the evidence in students' responses. For example, Noyes et al. (2022) explicated this process in their design of a task that elicited students' mechanistic explanations about preferential protein-ligand binding[48]. This iterative process of designing tasks, analyzing responses, and modifying tasks such that they are optimized, is no trivial feat – there is a narrow range within which an assessment provides enough, but not too much, information for the student. Once achieved, however, a task can be used to both assess *and support* student learning.

Supporting student learning via carefully designed assessments requires an understanding of different types and amounts of scaffolding. Scaffolding dates back to work by Wood et al., which proposed strategies that a tutor could use to better support student learning (e.g., *reducing degrees of freedom*)[49]. This work builds off of Vygotsky's zone of proximal development (ZPD) – a zone between what a learner can do without assistance and that which they cannot do even with assistance (i.e., support from experts or scaffolds)[50]. That

is, the ZPD represents a unique learning space in which students, with guidance, can build on what they know to expand or refine that knowledge. Originally, work on scaffolding was limited to individual student-teacher situations; however, it has since been applied to assessment design in the form of scaffolded activities[34,48,51], or those that include additional structure, hints, or questions that better support the learner to build an explanation, model, prediction, etc. For example, Noyes et al. (2022) showed that minor iterations on a task can have large impacts on the ideas that students include in their explanations.

Deciding the appropriate amount and type of scaffolding, as noted, can greatly influence what students include (or not) in responses to assessments. Thus, assessment design is crucially important, as we use assessments as evidence of what students know and can do[52]. Inadequately designed assessments can misrepresent students' knowledge and abilities, leading us to make inappropriate conclusions about our course instruction or the students themselves. By using ECD, we can mitigate the uncertainty about evidence elicited from assessments, and trust that optimized tasks (that have gone through the cycle of ECD) appropriately reflect students' knowledge. Further, assessments send a strong message to students about what is important and valued in the course[53]; therefore, if we want students to be able to explain/make predictions about complex phenomena and apply ideas from one course to another, then we must give them the opportunity to do so on assessment tasks.

REFERENCES

(1) National Research Council. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*; National Academies Press: Washington, DC, 2012.

(2) Cooper, M. M.; Posey, L. A.; Underwood, S. M. Core Ideas and Topics: Building Up or Drilling Down? *J. Chem. Educ.* **2017**, *94* (5), 541–548. https://doi.org/10.1021/acs.jchemed.6b00900.

(3) American Association for the Advancement of Science. *Vision and Change in Undergraduate Biology Education: A Call to Action*; Washington D.C., 2011. https://live-visionandchange.pantheonsite.io/wp-content/uploads/2011/03/Revised-Vision-and-Change-Final-Report.pdf (accessed 2020-04-17).

(4) Shiroda, M.; Franovic, C. G.-C.; de Lima, J.; Noyes, K.; Babi, D.; Persson-Gordon, E.; Beltran-Flores, E.; Kesh, J.; McKay, R. L.; Cooper, M. M.; Long, T. M.; Schwarz, C. V.; Stoltzfus, J. R. Examining and Supporting Mechanistic Explanations across Chemistry and Biology Courses, under review.

(5) Cooper, M.; Klymkowsky, M. Chemistry, Life, the Universe, and Everything: A New Approach to General Chemistry, and a Model for Curriculum Reform. *J. Chem. Educ.* **2013**, *90* (9), 1116–1122. https://doi.org/10.1021/ed300456y.

(6) Klymkowsky, M. W.; Rentsch, J. D.; Begovic, E.; Cooper, M. M. The Design and Transformation of Biofundamentals: A Nonsurvey Introductory Evolutionary and Molecular Biology Course. *CBE Life Sci Educ* **2016**, *15* (4), ar70. https://doi.org/10.1187/cbe.16-03-0142.

(7) Cooper, M. M.; Stowe, R. L.; Crandell, O. M.; Klymkowsky, M. W. Organic Chemistry, Life, the Universe and Everything (OCLUE): A Transformed Organic Chemistry Curriculum. *J. Chem. Educ.* **2019**, *96* (9), 1858–1872. https://doi.org/10.1021/acs.jchemed.9b00401.

(8) Cooper, M. M.; Underwood, S. M.; Hilley, C. Z.; Klymkowsky, M. W. Development and Assessment of a Molecular Structure and Properties Learning Progression. *J. Chem. Educ.* **2012**, *89* (11), 1351–1357. https://doi.org/10.1021/ed300083a.

(9) Crandell, O. M.; Kouyoumdjian, H.; Underwood, S. M.; Cooper, M. M. Reasoning about Reactions in Organic Chemistry: Starting It in General Chemistry. *J. Chem. Educ.* **2019**, *96* (2), 213–226. https://doi.org/10.1021/acs.jchemed.8b00784.

(10) Crandell, O. M.; Lockhart, M. A.; Cooper, M. M. Arrows on the Page Are Not a Good Gauge: Evidence for the Importance of Causal Mechanistic Explanations about Nucleophilic Substitution in Organic Chemistry. *J. Chem. Educ.* **2020**, *97* (2), 313–327. https://doi.org/10.1021/acs.jchemed.9b00815.

(11) Houchlei, S. K.; Crandell, O. M.; Cooper, M. M. "What About the Students Who Switched Course Type?": An Investigation of Inconsistent Course Experience. *J. Chem. Educ.* **2023**. https://doi.org/10.1021/acs.jchemed.3c00345.

(12) Williams, L. C.; Underwood, S. M.; Klymkowsky, M. W.; Cooper, M. M. Are Noncovalent Interactions an Achilles Heel in Chemistry Education? A Comparison of Instructional Approaches. *J. Chem. Educ.* **2015**, *92* (12), 1979–1987. https://doi.org/10.1021/acs.jchemed.5b00619.

(13) Choi, B. C. K.; Pak, A. W. P. Multidisciplinarity, Interdisciplinarity and Transdis- Ciplinarity in Health Research, Services, Education and Policy: 1. Definitions, Objectives, and Evidence of Effectiveness. *Clin Invest Med* **2006**, *29* (6).

(14) Allred, Z. D. R.; Farias, A. J.; Kararo, A. T.; Parent, K. N.; Matz, R. L.; Underwood, S. M. Students' Use of Chemistry Core Ideas to Explain the Structure and Stability of DNA. *Biochemistry and Molecular Biology Education* **2021**, *49* (1), 55–68. https://doi.org/10.1002/bmb.21391.

(15) Green, A. I.; Parent, K. N.; Underwood, S. M.; Matz, R. L. Connecting Ideas across Courses: Relating Energy, Bonds & How ATP Hydrolysis Powers a Molecular Motor. *The American Biology Teacher* **2021**, *83* (5), 303–310. https://doi.org/10.1525/abt.2021.83.5.303.

(16) Luckie, D. B.; Bellon, R.; Sweeder, R. D. The BRAID: Experiments in Stitching Together Disciplines at a Big Ten University. *Journal of STEM Education: Innovations and Research* **2012**, *13* (2).

(17) Martinez, B. L.; Kararo, A. T.; Parent, K. N.; Underwood, S. M.; Matz, R. L. Creating and Testing an Activity with Interdisciplinary Connections: Entropy to Osmosis. *Chem. Educ. Res. Pract.* **2021**, *22* (3), 683–696. https://doi.org/10.1039/D0RP00353K.

(18) VandenPlas, J. R.; Herrington, D. G.; Shrode, A. D.; Sweeder, R. D. Use of Simulations and Screencasts to Increase Student Understanding of Energy Concepts in Bonding. *J. Chem. Educ.* **2021**, *98* (3), 730–744. https://doi.org/10.1021/acs.jchemed.0c00470.

(19) Yoho, R.; Foster, T.; Urban-Lurain, M.; Merrill, J.; Haudek, K. C. Interdisciplinary Insights from Instructor Interviews Reconciling "Structure and Function" in Biology, Biochemistry, and Chemistry through the Context of Enzyme Binding. *Disciplinary and Interdisciplinary Science Education Research* **2019**, *1* (1), 16. https://doi.org/10.1186/s43031-019-0016-7.

(20) Kohn, K. P.; Underwood, S. M.; Cooper, M. M. Connecting Structure–Property and Structure–Function Relationships across the Disciplines of Chemistry and Biology: Exploring Student Perceptions. *LSE* **2018**, *17* (2), ar33. https://doi.org/10.1187/cbe.18-01-0004.

(21) Kohn, K. P.; Underwood, S. M.; Cooper, M. M. Energy Connections and Misconnections across Chemistry and Biology. *LSE* **2018**, *17* (1), ar3. https://doi.org/10.1187/cbe.17-08-0169.

(22) Cooper, M. M.; Klymkowsky, M. W. The Trouble with Chemical Energy: Why Understanding Bond Energies Requires an Interdisciplinary Systems Approach. *CBE Life Sci Educ* **2013**, *12* (2), 306–312. https://doi.org/10.1187/cbe.12-10-0170.

(23) Dreyfus, B. W.; Gouvea, J.; Geller, B. D.; Sawtelle, V.; Turpen, C.; Redish, E. F. Chemical Energy in an Introductory Physics Course for the Life Sciences. *American Journal of Physics* **2014**, *82* (5), 403–411. https://doi.org/10.1119/1.4870391.

(24) Dreyfus, B. W.; Sawtelle, V.; Turpen, C.; Gouvea, J.; Redish, E. F. Students' Reasoning about ``high-Energy Bonds'' and ATP: A Vision of Interdisciplinary Education. *Phys. Rev. ST Phys. Educ. Res.* **2014**, *10* (1), 010115. https://doi.org/10.1103/PhysRevSTPER.10.010115.

(25) A. Schmidt-McCormack, J.; A. Judge, J.; Spahr, K.; Yang, E.; Pugh, R.; Karlin, A.; Sattar, A.; C. Thompson, B.; Ruggles Gere, A.; V. Shultz, G. Analysis of the Role of a Writing-to-Learn Assignment in Student Understanding of Organic Acid–Base Concepts. *Chemistry Education Research and Practice* **2019**, *20* (2), 383–398. https://doi.org/10.1039/C8RP00260F.

(26) Becker, N.; Noyes, K.; Cooper, M. Characterizing Students' Mechanistic Reasoning about London Dispersion Forces. *J. Chem. Educ.* **2016**, *93* (10), 1713–1724. https://doi.org/10.1021/acs.jchemed.6b00298.

(27) Caspari, I.; Kranz, D.; Graulich, N. Resolving the Complexity of Organic Chemistry Students' Reasoning through the Lens of a Mechanistic Framework. *Chemistry Education Research and Practice* **2018**, *19* (4), 1117–1141. https://doi.org/10.1039/C8RP00131F.

(28) Cooper, M. M.; Corley, L. M.; Underwood, S. M. An Investigation of College Chemistry Students' Understanding of Structure–Property Relationships. *Journal of Research in Science Teaching* **2013**, *50* (6), 699–721. https://doi.org/10.1002/tea.21093.

(29) Cooper, M. M.; Kouyoumdjian, H.; Underwood, S. M. Investigating Students' Reasoning about Acid–Base Reactions. *J. Chem. Educ.* **2016**, *93* (10), 1703–1712. https://doi.org/10.1021/acs.jchemed.6b00417.

(30) Graulich, N.; Caspari, I. Designing a Scaffold for Mechanistic Reasoning in Organic Chemistry. *Chemistry Teacher International* **2021**, *3* (1), 19–30. https://doi.org/10.1515/cti-2020-0001.

(31) M. Watts, F.; A. Schmidt-McCormack, J.; A. Wilhelm, C.; Karlin, A.; Sattar, A.; C. Thompson, B.; Ruggles Gere, A.; V. Shultz, G. What Students Write about When Students Write about Mechanisms: Analysis of Features Present in Students' Written Descriptions of an Organic

Reaction Mechanism. *Chemistry Education Research and Practice* **2020**, *21* (4), 1148–1172. https://doi.org/10.1039/C9RP00185A.

(32) Matz, R. L.; Fata-Hartley, C. L.; Posey, L. A.; Laverty, J. T.; Underwood, S. M.; Carmel, J. H.; Herrington, D. G.; Stowe, R. L.; Caballero, M. D.; Ebert-May, D.; Cooper, M. M. Evaluating the Extent of a Large-Scale Transformation in Gateway Science Courses. *Science Advances* **2018**, *4* (10), eaau0554. https://doi.org/10.1126/sciadv.aau0554.

(33) Houchlei, S. A Step into the Unknown: Exploring Students' Construction of Mechanistic Arrows for Both Familiar and Unfamiliar Reactions in Organic Chemistry. Doctoral dissertation, Michigan State University, East Lansing, MI, 2022. https://www.proquest.com/openview/88e9dd8894ca65fe39ac27072f056a67/1.pdf?pq-origsite=gscholar&cbl=18750&diss=y.

(34) Graulich, N.; Schween, M. Concept-Oriented Task Design: Making Purposeful Case Comparisons in Organic Chemistry. *J. Chem. Educ.* **2018**, *95* (3), 376–383. https://doi.org/10.1021/acs.jchemed.7b00672.

(35) Dood, A.; Das, K.; Qian, Z.; Finkenstaedt-Quinn, S.; Gere, A.; Shultz, G. A Dashboard to Provide Instructors with Automated Feedback on Students' Peer Review Comments. In *LAK23: 13th International Learning Analytics and Knowledge Conference*; LAK2023; Association for Computing Machinery: New York, NY, USA, 2023; pp 619–625. https://doi.org/10.1145/3576050.3576087.

(36) Jescovitch, L. N.; Scott, E. E.; Cerchiara, J. A.; Merrill, J.; Urban-Lurain, M.; Doherty, J. H.; Haudek, K. C. Comparison of Machine Learning Performance Using Analytic and Holistic Coding Approaches Across Constructed Response Assessments Aligned to a Science Learning Progression. *J Sci Educ Technol* **2021**, *30* (2), 150–167. https://doi.org/10.1007/s10956-020-09858-0.

(37) Doherty, J. H.; Scott, E. E.; Cerchiara, J. A.; Jescovitch, L. N.; McFarland, J. L.; Haudek, K. C.; Wenderoth, M. P. What a Difference in Pressure Makes! A Framework Describing Undergraduate Students' Reasoning about Bulk Flow Down Pressure Gradients. *LSE* **2023**, *22* (2), ar23. https://doi.org/10.1187/cbe.20-01-0003.

(38) Haskel-Ittah, M. How Can We Help Students Reason About the Mechanisms by Which Genes Affect Traits? In *Genetics Education: Current Challenges and Possible Solutions*; Haskel-Ittah, M., Yarden, A., Eds.; Contributions from Biology Education Research; Springer International Publishing: Cham, 2021; pp 71–86. https://doi.org/10.1007/978-3-030-86051-6_5.

(39) Haskel-Ittah, M. Explanatory Black Boxes and Mechanistic Reasoning. *Journal of Research in Science Teaching* **2023**, *60* (4). https://doi.org/10.1002/tea.21817.

(40) Haskel-Ittah, M.; Duncan, R. G.; Vázquez-Ben, L.; Yarden, A. Reasoning about Genetic Mechanisms: Affordances and Constraints for Learning. *Journal of Research in Science Teaching* **2020**, *57* (3), 342–367. https://doi.org/10.1002/tea.21595.

(41) Haskel-Ittah, M.; Yarden, A. Students' Conception of Genetic Phenomena and Its Effect on Their Ability to Understand the Underlying Mechanism. *LSE* **2018**, *17* (3), ar36. https://doi.org/10.1187/cbe.18-01-0014.

(42) Southard, K.; Wince, T.; Meddleton, S.; Bolger, M. S. Features of Knowledge Building in Biology: Understanding Undergraduate Students' Ideas about Molecular Mechanisms. *LSE* **2016**, *15* (1), ar7. https://doi.org/10.1187/cbe.15-05-0114.

(43) Uhl, J. D.; Shiroda, M.; Haudek, K. C. Developing Assessments to Elicit and Characterize Undergraduate Mechanistic Explanations about Information Flow in Biology. *Journal of Biological Education* **2022**, *0* (0), 1–20. https://doi.org/10.1080/00219266.2022.2041460.

(44) Trujillo, C. M.; Anderson, T. R.; Pelaez, N. J. A Model of How Different Biology Experts Explain Molecular and Cellular Mechanisms. *LSE* **2015**, *14* (2), ar20. https://doi.org/10.1187/cbe.14-12-0229.

(45) Ralph, V. R.; Scharlott, L. J.; Schafer, A. G. L.; Deshaye, M. Y.; Becker, N. M.; Stowe, R. L. Advancing Equity in STEM: The Impact Assessment Design Has on Who Succeeds in Undergraduate Introductory Chemistry. *JACS Au* **2022**. https://doi.org/10.1021/jacsau.2c00221.

(46) Noyes, K.; McKay, R. L.; Neumann, M.; Haudek, K. C.; Cooper, M. M. Developing Computer Resources to Automate Analysis of Students' Explanations of London Dispersion Forces. *J. Chem. Educ.* **2020**, *97* (11), 3923–3936. https://doi.org/10.1021/acs.jchemed.0c00445.

(47) Mislevy, R. J.; Almond, R. G.; Lukas, J. F. A Brief Introduction to Evidence-Centered Design. *ETS Research Report Series* **2003**, *2003* (1), i–29. https://doi.org/10.1002/j.2333-8504.2003.tb01908.x.

(48) Noyes, K.; Carlson, C. G.; Stoltzfus, J. R.; Schwarz, C. V.; Long, T. M.; Cooper, M. M. A Deep Look into Designing a Task and Coding Scheme through the Lens of Causal Mechanistic Reasoning. *J. Chem. Educ.* **2022**, *99* (2), 874–885. https://doi.org/10.1021/acs.jchemed.1c00959.

(49) Wood, D.; Bruner, J. S.; Ross, G. The Role of Tutoring in Problem Solving*. *Journal of Child Psychology and Psychiatry* **1976**, *17* (2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x.

(50) Vygotsky, L. S. *Mind in Society: The Development of Higher Psychological Processes*; Harvard University Press, 1980.

(51) Reiser, B. J. Scaffolding Complex Learning: The Mechanisms of Structuring and Problematizing Student Work. *Journal of the Learning Sciences* **2004**, *13* (3), 273–304. https://doi.org/10.1207/s15327809jls1303_2.

(52) National Research Council. *Knowing What Students Know: The Science and Design of Educational Assessment*; National Academies Press: Washington, DC, 2001. https://doi.org/10.17226/10019.

(53) Stowe, R. L.; Scharlott, L. J.; Ralph, V. R.; Becker, N. M.; Cooper, M. M. You Are What You Assess: The Case for Emphasizing Chemistry on Chemistry Assessments. *J. Chem. Educ.* **2021**, *98* (8), 2490–2495. https://doi.org/10.1021/acs.jchemed.1c00532.

Chapter IV – Undergraduate Chemistry And Biology Students' Use Of Causal Mechanistic Reasoning To Explain And Predict Preferential Protein-Ligand Binding Activity

Preface

In this study, we investigate how chemistry and biology students engage in causal mechanistic reasoning to explain and predict how a Mg^{2+} ion preferentially binds to a protein. This research has been previously published in the *Journal of Chemical Education* and is reprinted with permission from Franovic, C. G.-C.; Noyes, K.; Stoltzfus, J. R.; Schwarz, C. V.; Long, T. M.; Cooper, M. M. Undergraduate Chemistry and Biology Students' Use of Causal Mechanistic Reasoning to Explain and Predict Preferential Protein–Ligand Binding Activity. *J. Chem. Educ.* **2023**, *100* (5), 1716–1727. Copyright 2023 American Chemical Society.

A copy of permissions obtained is included in Appendix A. Supporting Information for this manuscript is included in the Appendix.

Background And Significance

Connecting Across Disciplines

Having both an interdisciplinary understanding of science and a mechanistic understanding of scientific phenomena are valued ways of thinking in science education.[1–5] In an increasingly complex world with problems and phenomena not bounded by disciplinary silos, it is important to think deeply about how and why they occur to make informed decisions and to recognize when and how to use and apply prior knowledge.

The importance of interdisciplinary learning at the undergraduate level has increased in science education discourse. These efforts include large projects (for example, the BRAID

project[6]), as well as smaller, single-activity research articles, such as those proposed by Underwood et al.[7–10] While interdisciplinarity may focus on the integration of science with social and economic studies (for example, addressing climate change or sustainability issues), students should also be able to integrate ideas between their science courses, such as chemistry and biology – particularly pre-health students who need to leverage both disciplines in their future courses and careers. Students in undergraduate STEM majors frequently enroll in both chemistry and biology courses to meet their degree requirements. In fact, general chemistry often serves as a pre- or co-requisite for introductory biology courses, implying interdisciplinary connections; however, these connections are not often obvious to students,[11–13] making this an intriguing area of study.

The National Academies, in their 2012 consensus report, further emphasized the importance of integrating ideas across science courses, and to make this actionable, introduced three-dimensional learning (3DL). 3DL aims to engage science students in three dimensions during their learning experience: disciplinary core ideas, scientific practices, and cross-cutting concepts.[3] Causal mechanistic reasoning (CMR) about phenomena can be used as a tool for instructors who are aiming to implement 3DL in coursework, because this reasoning strategy incorporates all three dimensions – a disciplinary core idea relevant to the phenomenon under consideration, the scientific practice of constructing explanations, and the crosscutting concept of cause and effect. While both chemistry and biology education research have implemented these ideas in their respective fields, there is limited work probing student engagement in CMR using chemical principles in biological contexts.[14]

Causal Mechanistic Reasoning

Mechanistic reasoning has been defined as a type of causal reasoning that involves explaining the sequential stages of the underlying causal events leading to a phenomenon, or how and why one or more factors behave to give rise to a phenomenon.[15,16] Krist et al. describe this as an epistemic resource, or thinking strategy, consisting of three steps: (1) considering the scalar level below the phenomenon of interest, (2) identifying and unpacking the properties and behaviors of entities at that lower scalar level, and (3) connecting how those interactions and behaviors give rise to the phenomenon of interest.[16] This understanding of CMR guided our work in both the activity development process (i.e., the type of thinking we aimed to elicit) and in the assessment process (i.e., our interpretation of student explanations and predictions).

When considering a phenomenon, we call upon conceptual resources, or pieces of knowledge/information, such as the idea that opposite charges attract each other, in order to identify relevant concepts for specific situations.[17] In addition to conceptual resources, we can leverage epistemic resources, such as CMR, which we use as a productive way of thinking about, explaining, or predicting phenomena in science and in our everyday lives.[16] Our world is made up of mechanisms (in science as well as social and economic studies among others) and models that explain mechanisms, supporting this type of reasoning as a crucial part of education. Therefore, as educators, we must provide opportunities for students to use prior knowledge and promote their ability to construct causal mechanistic explanations – something that is critical in science and beyond.

Several research findings from studies in general and organic chemistry courses emphasize the importance and utility of CMR.[18–24] CMR can help focus attention to the

important parts of the phenomena so that learners do not try to memorize every detail. This can help people learn "better" rather than "more" and understand phenomena more deeply. Such approaches can support all students and, in fact, a recent report provides evidence that mechanistic reasoning tasks are more equitable than tasks assessing rote knowledge or skills.[25] We posit that CMR is also useful for spanning the disciplinary boundary between chemistry and biology, because the two disciplines may emphasize the same phenomenon at different scalar levels, and the information at each scalar level is important for a deep understanding. In this work, we focus on the (macro)molecular level of proteins and define atoms or amino acids as the scalar level below that of proteins, since it is the behavior and interactions of relevant atoms and side chains which govern protein activity, such as the binding of ligands. Thus, it is the "chemical" ideas (electrostatic forces and interactions occurring at the atomic/electronic level) dictating how and why ligands bind to proteins at certain binding sites.

Structure, Properties, And Function

Structure, property, and function (SPF) relationships, in a chemical context, refer primarily to the structures of atoms or molecules, which govern their chemical properties, such as melting or boiling point. In biology, these ideas can be extended to macromolecules, such as proteins, which carry out specific functions as a result of their chemical properties, which, as noted, are dictated by their structures. While the connections regarding these ideas between chemistry and biology may be accepted and understood by experts, these connections may not always be so explicit for students.[11–13] For example, Kohn et al. interviewed students who were co-enrolled in second semester general chemistry and introductory molecular biology to investigate how they thought about the "big ideas" in each course. When discussing chemistry,

these students identified the relationship between the molecular structure of a substance and its properties as central to understanding the discipline; however, these same students referred to "structure-function" as the big idea in biology. Thus, they saw these ideas as being discussed differently in the two courses. Further, these same students talked about how they actively separate ideas related to "energy" in biology versus chemistry.[11] In this study, one student said, "I know for biology what [the instructor] wants us to say and then for chemistry what we have to say," suggesting that students may be constructing different responses for the same phenomena based on the course in which they are situated. While the students noted a disconnect between SPF relationships in these two disciplines, unlike with energy, they eventually recognized that these SPF ideas were quite similar and, even more encouraging, a number of students spontaneously voiced the idea that both chemistry and biology courses should attempt to explicitly help students make connections from structure to properties to function. That is, each course had part of the connection, but neither had the whole.

The fact that students suggested the idea that chemistry and biology courses should emphasize and clarify the connections between structure, properties, and function, and that it is critical to do so for a sustained understanding and use of the content,[26] makes this an important area of investigation. Research exploring students' understandings of ideas about SPF relationships in the contexts of specific phenomena relevant to chemistry and biology is limited;[10,14] however, in a recent publication, Yoho et al. found that faculty in chemistry, biology, and biochemistry all emphasized the importance of these relationships in each discipline but also the importance and challenges (via instructor opinions) of bridging the disciplines using these ideas.[27]

Purpose Of This Paper

Given (1) our interest in students building connections via CMR between chemistry and biology, (2) the importance of context in activating resources and how this might impact students' explanations, and (3) a gap in the literature studying student responses across these disciplines, we investigated student engagement in CMR using a task centered around protein-ligand binding, a biological structure-function phenomenon that can be explained using chemical principles associated with structure and properties. This task, however, serves as just one part of a three-part activity. Parts two and three also aim to elicit CMR, but the phenomena focus on how different proteins result in different functions (part 2) and how protein variation emerges in populations (part 3). Because of our large team of discipline-based education researchers, with varying interests involving student engagement in CMR across chemistry and biology undergraduate courses, we gathered responses from a range of chemistry, biology, and biochemistry courses. To our knowledge, a study of this magnitude and depth has not previously been done, fueling our interest to conduct an exploratory and descriptive investigation that can inform future work to support both CMR and interdisciplinary learning. In this manuscript, we focus solely on student responses to the protein-ligand binding task, which has deepened our understanding of the ideas students use to reason causal mechanistically about this phenomenon as well as the degree to which engagement in CMR influences overall predictions.

Research Questions

By collecting and analyzing responses from a wide array of courses, we could explore how students at different points in an undergraduate degree engaged in CMR about this phenomenon. In this analysis, we aimed to answer the following research questions:

1. How do students in chemistry and biology courses engage in causal mechanistic reasoning in the context of protein-ligand binding?
2. What conceptual resources do students enrolled in chemistry and biology courses use when explaining this phenomenon?
3. How does engagement in CMR relate to students' overall predictions?

Methods

Development Of The Task

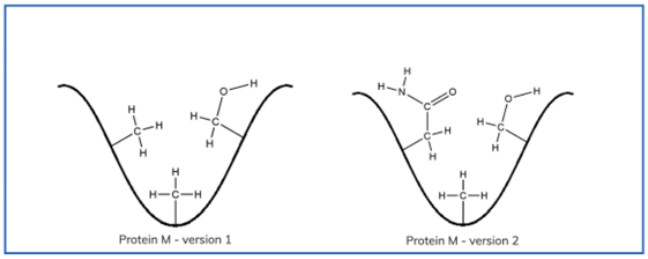
Noyes et al. developed the protein-ligand task (hereon referred to as "PL task") and the subsequent coding approach that we used to analyze the responses in this paper.[28] The task aimed to elicit both CMR (an epistemic resource, or way of thinking) and the appropriate knowledge pieces (conceptual resources) that would be required to construct an accurate prediction and explanation, while also providing a balanced amount of information so that the task would sufficiently activate these resources without giving away the answer. We hoped to activate conceptual resources related to electrostatics in this context, so that the students could then use and advance those ideas while reasoning about this relevant, biological phenomenon. To scaffold the writing of the task, we used a modified version of evidence-centered design,[29,30] a process outlined by Mislevy et al. in which design of the task is based on statements that the designers deem acceptable evidence of learning.[29] In our case, we defined

acceptable evidence as a response that appropriately leverages electrostatics and explicitly identifies lower-level entities, as well as how these entities and properties link to preferential binding. Based on the design process discussed by Noyes et al.,[28] this task seems to activate appropriate resources (conceptual and epistemic) for explaining this phenomenon and prompt students to include these ideas in their written explanations.

The PL Task

The PL task asks students to choose one of two protein binding sites to which a positive magnesium ion would most likely bind (Figure 4.1). The students are prompted to draw the Mg^{2+} in the site they chose, showing why it is binding there, then explain what causes the ion to bind to the protein, and, finally, explain why the version they chose has the better binding site and how structural differences cause the difference in binding.

The drawings below represent binding sites in two different versions of protein M showing only the atoms in relevant amino acid side chains. Consider a positively charged magnesium ion (Mg^{2+}). Pick the binding site you think is most likely to bind the magnesium ion and **draw** the ion in the binding site **showing why it is binding** in that site.



Explain what causes the magnesium ion to bind to the protein **making specific references to your drawing.**

Explain why the protein you chose has the better magnesium binding site and **how** the structural differences in the site cause this difference in binding.

Figure 4.1. The full PL task. Students are first prompted to draw Mg^{2+} in the version that would better bind the ion. Then they are prompted to explain what causes Mg^{2+} to bind and why the version they chose is the better binding site.[28]

Based on structure-property predictions and the information available in the diagrams, the correct protein version is version 2, because that binding site has two polar amino acids (as opposed to one in version 1) which have partial negative charges on the oxygen atoms. The two partially negative oxygen atoms attract the positive magnesium ion more strongly than the one partially negative oxygen in version 1. We recognize that this represents a simplified model of a complex system, one in which other factors like 3-dimensional protein folding, access to the binding site, potential solvation of Mg^{2+}, etc., may impact binding activity; however, here we only provided the students with structural images of two binding sites, so we did not expect them to consider anything outside of this given context.

Participants And Data Collection

During spring semester 2020, we administered the full activity containing the PL task to students in seven different courses at a large, public, research-intensive midwestern university. The courses in which we administered the task included general chemistry I (GCI), general chemistry II (GCII), molecular biology (MB), organismal biology (OB), organic chemistry II (OCII), biochemistry I (BCI), and biochemistry II (BCII). We chose to administer this activity to these courses so that we could create "co-enrolled", "longitudinal", and "cross-sectional" cohorts of students in order to address the broader interests of our research group. In this study, we are primarily interested in the "co-enrolled" and "cross-sectional" cohorts of students; however, we explain the selection process of all three cohorts in Appendix B as these processes influence one another. We provide additional information about the courses and data collection process in Figures 4.2 and 4.3 and Table 4.1.

Students completed the task on an online assessment platform called beSocratic, which allows students to provide both drawn and written responses.[31] Students were provided a small amount of course credit for completing the task, but could choose to exclude their responses from our study. All students in this study agreed and consented to our use of their responses according to our IRB protocol, and, after removing incomplete or un-codable responses, there was a total of 4,092 student explanations. We deidentified the responses by assigning each response a random (using a random number generator) identification value. Of the 4,092 total responses, we decided to begin coding with the two smaller, more feasibly sized, cohorts, that represent the focus of this study: the "cross-sectional cohort" and the "co-enrolled cohort".

The Cross-Sectional Cohort

In the cross-sectional cohort, we targeted responses from students in all courses. After removing students who did not provide a complete or codable response, we aimed to limit other confounding factors by selecting responses from students who had only seen the task once up until that point (i.e., some students were enrolled in more than one of the listed courses or completed both a pre and post response). For example, the MB and OB post responses were constructed by students who did not complete a pre response. Finally, we randomly selected a maximum of 50 responses from each course/timepoint (Figure 4.2), resulting in a total of 395 responses for the cross-sectional cohort.

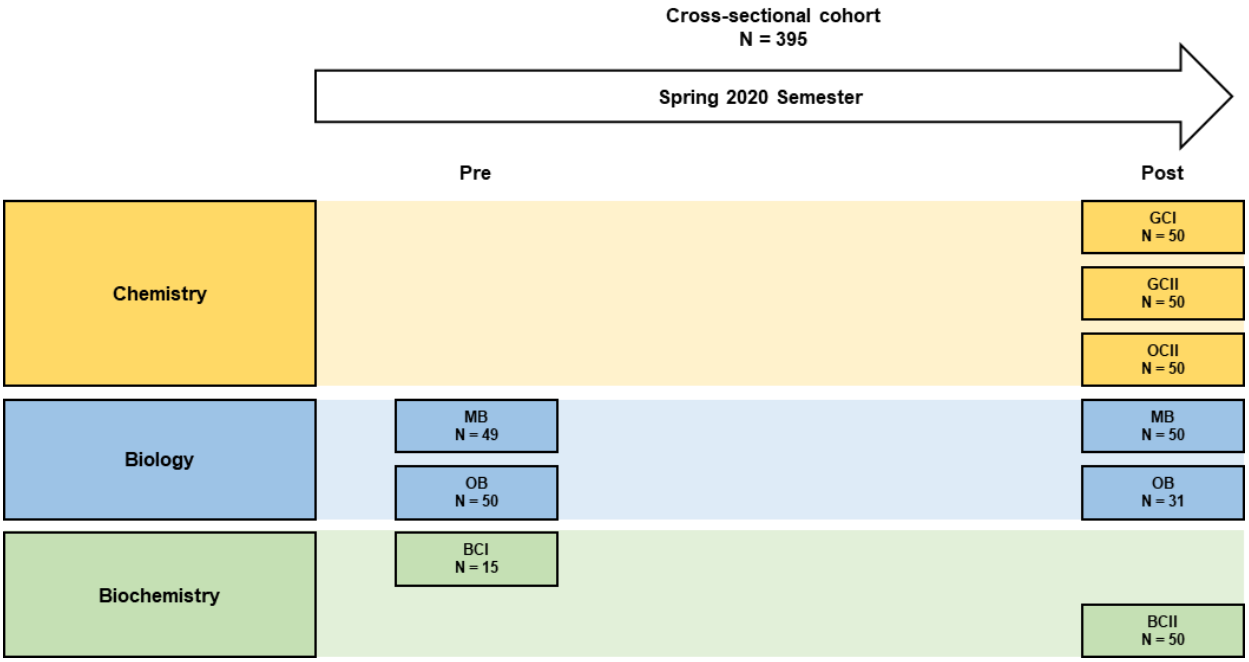


Figure 4.2. All students in the cross-sectional cohort took the PL task for the first time in the specified course and timepoint displayed.

The Co-Enrolled Cohort

To answer our broader questions about how students in chemistry and biology courses use CMR and conceptual/epistemic resources to explain this phenomenon, we also needed to see if the course in which the task is asked affects the students' responses. We included the co-enrolled cohort (students who were co-enrolled in GCII and MB) in this study, because (1) resources are contextually activated and (2) there is some evidence that students actively separate their chemistry and biology knowledge.[11] This cohort consists of 99 students, and while not the direct intent of selecting this cohort, because we administered the task at the beginning of the semester in MB, we could observe how the students' responses changed over the course of the semester. Each student in this cohort completed the task three times: once at the beginning of the semester (for their MB course) and twice at the end of the semester (for

both their MB and GCII courses) (Figure 4.3). Specifically, we chose the same 49 students who were coded for the "MB pre" point in the cross-sectional cohort, as well as an additional 50 students who fit this description. Thus, this set consisted of 297 total coded responses from these 99 students. We note that, while we administered the PL task in GCII prior to MB at the end of the semester, there was some slight overlap in the timing of the tasks and three of these students completed their post response in MB prior to GCII. Table 4.1 outlines the course names, the total number of responses collected, the total number of responses coded, and the time point of administration of the task.

We conducted independent samples t tests to evaluate the difference in grade point average (GPA) and SAT scores between the students who constructed responses that were coded and the remaining students (not coded) in each respective group. Because of the small sample size coded for each course, we did not compare other factors (e.g., age, race/ethnicity, course history, major, etc.). Each of the independent samples t test results were not significant ($p > 0.05$). Thus, according to GPA and SAT scores, the responses we coded were constructed by students that are representative of each course population – the p values for each group are shown in Appendix C.

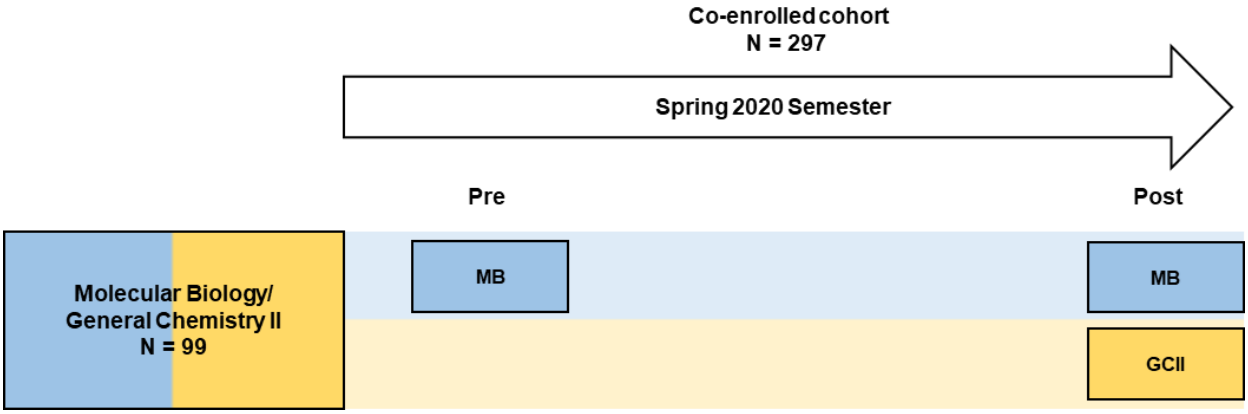


Figure 4.3. All students in the co-enrolled cohort completed the task either in MB or GCII at the three timepoints displayed. Ninety-nine students and N = 297 responses.

Table 4.1. Number of responses collected and analyzed in each course and timepoint.

Course Name	Time in semester	Number of responses collected	Responses coded for cross-sectional cohort	Responses coded for co-enrolled cohort
General Chemistry I	Post	1016	50	0
General Chemistry II	Post	837	50	99
Molecular Biology	Pre	584	49*	99*
Molecular Biology	Post	636	50	99
Organismal Biology	Pre	350	50	0
Organismal Biology	Post	368	31	0
Organic Chemistry II	Post	225	50	0
Biochemistry I	Pre	20	15	0
Biochemistry II	Post	56	50	0
Total		**4092**	**395**	**297**

*The 49 MB pre responses in the cross-sectional cohort were also included in the co-enrolled cohort.

COVID-19

During this semester, all courses at the university switched to an online format in response to the COVID-19 pandemic. Thus, pre responses (in BCI, MB, and OB) were

constructed prior to the pandemic, while post responses (in GCI, GCII, OCII, MB, OB, and BCII) were constructed during online learning and amid the pandemic. This unprecedented event had (and continues to have) wide-spread effects on students that should not be ignored when considering the results; however, it is important to continue research on teaching and assessing students so that we may understand how to provide the best support possible.

Coding Rubric

Noyes et al. described the coding approach that we used to characterize students' engagement in CMR from their constructed explanations to the PL task.[28] The engagement in CMR (non-CM, partially CM, or fully CM) depends on the presence or absence of three key ideas in the explanation: (1) the attraction of oppositely charged species, (2) the negative or polar nature of atoms and amino acids in the binding sites, and (3) the larger negative charge in one of the binding sites, causing the preferential binding of Mg^{2+}. If a response provided evidence of an understanding of all three key ideas, then we coded it as fully CM. If the response provided evidence of one or two of the ideas, but not all three, then we assigned the code partially CM. Lastly, if there was no evidence for any of these three key ideas, the response was deemed non-CM. For a more comprehensive description and student examples of the coding approach, we direct readers to our previous publication detailing its development and implementation.[28]

Initially, the first and second authors used the CMR coding approach to code the deidentified responses from the cross-sectional cohort. However, while coding this larger set (as compared to the smaller sets used for development of the coding approach),[28] we noticed an additional resource being leveraged by some students: space, or some physical attribute

causing Mg^{2+} to bind, instead of, or in tandem with, electrostatics. Because our goal is to characterize both student engagement in CMR (RQ1, which uses the holistic codes) as well as the different resources that students leveraged when constructing their explanations for this phenomenon (RQ2), we expanded the analytic rubric to include a coding bin for space and specified the codes for the original analytic rubric to include charge, polarity, both, or neither (to identify resource(s) rather than just "yes" or "no") (Table 4.2). For the holistic scheme, whether the students used charge or polarity to explain the binding of Mg^{2+} did not impact the code, since either resource is sufficient for explaining this phenomenon; however, we did identify which of the two (or both) was used in the responses. Finally, we added one last bin to see whether the students chose the correct binding site (version 2), allowing us to address RQ3. The expanded analytic rubric was used by the first and second authors to code the co-enrolled cohort as well.

Table 4.2. The expanded analytic rubric. The original analytic rubric bins are bolded.

Bin	Options	Criteria
Attraction of oppositely charged species	Polarity, charge, both, no	**Provides evidence of the attraction between oppositely charged species or polar species and ions**
Negative or polar nature of atoms and amino acids in the binding sites	Polarity, charge, both, no	**Identifies an atom or amino acid as polar or negative**
Magnitude of charge/polarity causing the preferential binding of Mg^{2+}	Polarity, charge, both, no	**Identifies the site they chose as more or less negative/polar than the other**
Correct binding site (version 2)	Yes, no	Clearly (through drawing and/or text) indicates that Version 2 is the better Mg^{2+} binding site.
Space, or some physical attribute causing Mg^{2+} to bind	Yes, no	Discusses the space or accessibility, or some physical aspect of the binding site to explain Mg^{2+} binding

Analysis Of Responses And Inter-Rater Reliability

During coding, the coders could not see the course or any other identifying student information. The first and second authors individually coded a set of 40 responses from the cross-sectional cohort, reaching percent agreement between 93% and 100% for each analytic bin. Then, after coding the remaining responses in the cross-sectional cohort individually, these authors discussed and came to consensus on any disagreements, which could then be used for the final analysis. We used Cohen's kappa values[32] to calculate inter-rater reliability (IRR), as the holistic codes are ordinal and mutually exclusive. Our calculated Cohen's kappa was 0.814, suggesting "almost perfect" agreement[33] between the two coders. After coding the cross-sectional cohort, the two authors coded the co-enrolled cohort in the same manner (Cohen's kappa = 0.785). The agreement between the iterations of coding for each holistic code was 98% for non-CM, 98% for partially CM, and 100% for fully CM.

We calculated quantitative statistics associated with Cohen's kappa, Pearson's $\chi 2$, independent t tests, ANOVA, and sign tests with IBM SPSS Statistics Version 27.[34] For these tests, we used a significance threshold of 0.05 and, when appropriate, conducted a Bonferroni adjustment for increased risk of type I error (i.e., false positives). For significant Pearson's χ^2 tests, we calculated Cramér's V, a modified version of ϕ for contingency tables with more than 2 rows or columns,[35] to characterize the effect size.

Results

In the following sections, we answer our research questions by leveraging both the holistic and analytic coding rubrics. We share the results for RQ1 using the holistic codes, which characterize student engagement in CMR. For RQ2, however, we use the analytic rubric to

characterize student use of conceptual resources in their explanations. Lastly, RQ3 focuses on how student engagement in CMR (holistic codes) related to their binding site predictions.

RQ1 – How Do Students In Chemistry And Biology Courses Engage In CMR In The Context Of Protein-Ligand Binding?

To address our first research question, we coded responses from nine different courses/timepoints (Table 4.1) characterizing engagement in CMR. We present raw results (i.e., breakdown of non-CM, partially CM, and fully CM responses) for each of these timepoints and courses in Appendix D, which shows an encouraging, stark increase in fully CM responses from GCI to BCII. In this section, we highlight comparisons between specific groups to illustrate how chemistry and biology students engage in CMR. Specifically, we focus on (1) the responses from the chemistry students in the cross-sectional cohort and (2) the responses from the co-enrolled cohort as representative of the biology students.

Chemistry Students' Engagement In CMR

Student Sample Population

To address this question, we analyzed data from students in the cross-sectional cohort who completed the PL task in one of three chemistry courses: GCI, GCII, and OCII. In order to determine whether students from each course were roughly equivalent, we used a one-way analysis of variance (ANOVA) to evaluate the difference in mean SAT scores between the students analyzed from each course. The results showed a significant difference (p = 0.012), with the GCI students having a lower mean SAT when compared to the GCII and OCII students (GC1 mean = 1165, GCII mean = 1236, OCII mean = 1239; for more information, see Appendix E). This is not surprising, given that a wider range of students enroll in GCI. However, we also

did a one-way ANOVA to evaluate the difference in mean cumulative GPA, which showed no difference between the students in these three courses (p = 0.888). Appendix E provides descriptive statistics for each of these performance measures.

Students In GCII And OCII Construct More Fully CM Responses Than Students In GCI

Chemistry course and engagement in CMR were found to be significantly related, with GCI students constructing the least amount of fully CM responses (Pearson's χ^2 = 26.450, p < 0.001, Cramér's V = 0.297). The percentage of students constructing fully CM responses in GCI, GCII, and OCII were 2%, 24%, and 30% respectively (Figure 4.4). We did follow-up pairwise comparisons to evaluate the difference among these proportions across courses (Table 4.3). While the GCI students constructed significantly different explanations from both the GCII and OCII students, the GCII and OCII students did not differ in their engagement in CMR (p = 0.342). Only one student in GCI constructed a fully CM response, while 24% and 30% of the GCII and OCII students did so. We posit that this is due to resource refinement – that is, students in GCI do not have complete access to the resources with which they need to construct a causal mechanistic explanation, while the GCII and OCII students have had both time and opportunity to (1) reason mechanistically about phenomena (i.e., develop this epistemic resource), and (2) refine their conceptual resources, so that they can be called on and used in appropriate contexts.

Table 4.3. Pearson's χ^2 test results exploring the relationship between chemistry courses and engagement in CMR.

Comparison	χ^2 Value	*p* Value[a]	Cramér's V[b]
GCI and GCII	14.904	$p < 0.001$	0.386
GCI and OCII	25.211	$p < 0.001$	0.502
GCII and OCII	2.146	$p = 0.342$	0.146

a. Bonferroni-adjusted $\alpha = 0.017$

b. Suggestions for interpreting Cramér's V[36]: small − 0.1, medium − 0.3, large − 0.5

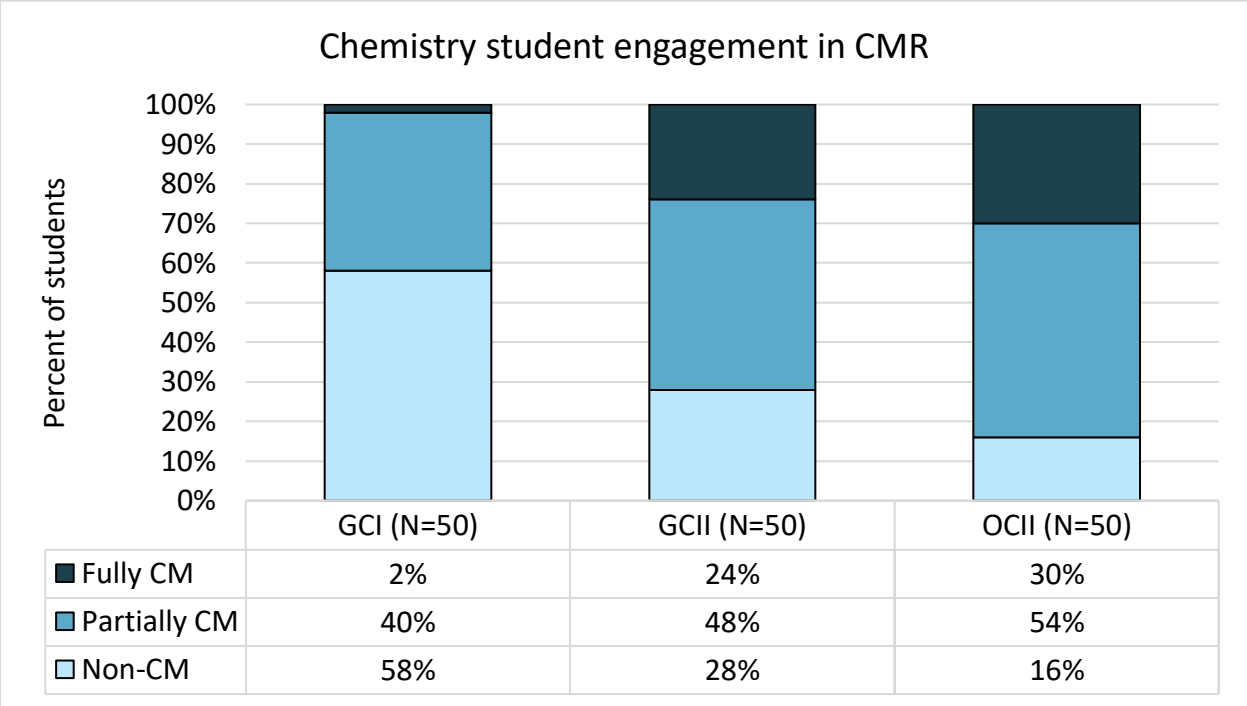


Figure 4.4. Chemistry student engagement in CMR (all cross-sectional cohort students who completed the PL task at the "post" timepoint).

Biology Students' Engagement In CMR

Student Sample Population

To understand how students from biology engaged in CMR, we focused on the responses from the co-enrolled cohort. We chose to focus on this cohort, because students in the cross-sectional cohort had not seen the PL task prior; therefore, the post OB and MB

students may not be representative of the rest of the group, since they had missed the pre administration of the PL task for some reason (e.g., transferred into the course late, poor engagement early in the semester, etc.).

MB Student Explanations Improved From Pre To Post

We coded responses from ninety-nine MB students who provided explanations both at the beginning and end of their MB course; however, we removed three of these students from this analysis, because they had not completed their GCII post response yet. Thus, to maintain consistency in the order in which each student completed their explanations, we only analyzed the responses constructed by these 96 students. We used the sign test for significance (a repeated-measures test) to compare the differences in engagement in CMR at these two timepoints (MB pre and post) for the related sample, showing 47 positive changes, 15 negative changes, and 34 ties ($p < 0.001$) (Figure 4.5). That is, nearly half (49%) of the students improved their response type from non-CM to partially or fully CM, or from partially CM to fully CM.

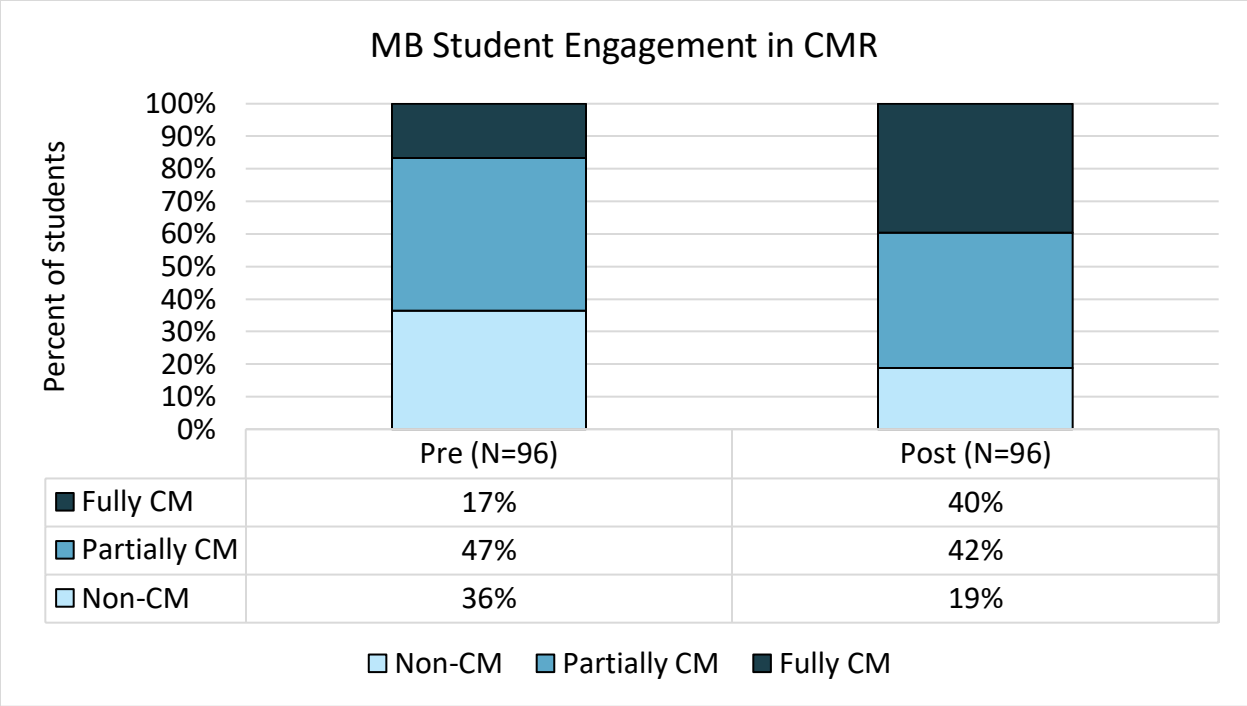


Figure 4.5. MB (N=96) pre and post engagement in CMR. Note that it is the same cohort of students represented in the pre and post categories.

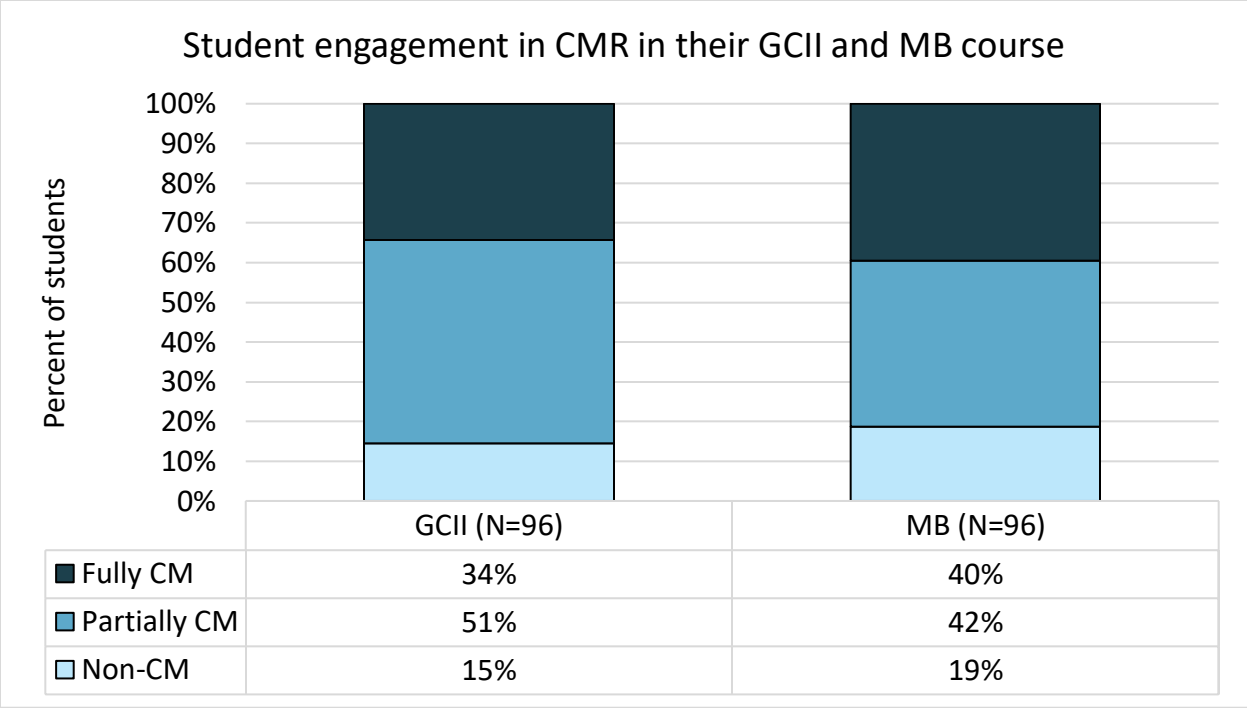
These findings indicate that student engagement in CMR improved throughout the duration of their MB course; however, like the chemistry students, these students were enrolled in other courses during this time which may have also contributed to their improvement. Additionally, having seen the prompt twice already could have impacted student explanations by allowing more time to wrestle with these ideas.

Comparing Student Engagement In CMR For Chemistry And Biology

According to the resources perspective, context matters when constructing explanations or thinking about certain ideas. Additionally, Kohn et al. showed that students think about SPF relationships in different ways when prompted about their biology course versus chemistry course, and some actively separate ideas about energy for these two courses, even when they are enrolled in them at the same time.[11,12] For these reasons, we expected that students would

construct different responses “for their chemistry course” versus “for their biology course”, even though both responses were completed on beSocratic. To test this, we examined the post responses from those students in the co-enrolled cohort who completed the PL task for both MB and GCII. All the students included in this analysis (N = 96 students, N = 192 responses) completed the GCII post task first, followed by completion of the MB post task within the next week.

We used a sign test to measure the difference between students’ post engagement in CMR in their chemistry course and their biology course. This sign test indicated no significant difference between these groups, with 13 negative changes (e.g., partially CM response in chemistry changing to non-CM in biology) and 17 positive changes ($z = -0.548$, $p = 0.584$). After examining the responses from students who changed their engagement in CMR, we did not discern any trends in the slight differences observed. Thus, even though some students had marginally altered responses, from a statistical standpoint, the responses were the same regardless of the course in which they were constructed. Figure 4.6 displays this nearly identical engagement in CMR between courses.



	GCII (N=96)	MB (N=96)
Fully CM	34%	40%
Partially CM	51%	42%
Non-CM	15%	19%

Figure 4.6. Frequency of response type constructed by students at the end (post) of their GCII course and end of their MB course. Recall that each student completed their response in GCII before completing it in MB.

We find this result encouraging, because it points to the interdisciplinary nature of the task, as it is equally effective at eliciting appropriate resources and CMR from students, regardless of the "course context". Further, the MB and GCII courses at this institution emphasize CMR in frequent formative and summative assessments, likely contributing to this result. Had one of these two courses put less emphasis on CMR, we may have seen different results.

Ideally these students, regardless of context, leveraged ideas from both courses, and therefore are building connections to refine their knowledge frameworks when explaining this phenomenon; however, the only evidence here is the nearly identical engagement in CMR in

both courses. An additional study involving think-aloud interviews would provide more support for this by asking students how and why they call upon certain ideas.

RQ2 – What Conceptual Resources Do Students Enrolled In Chemistry And Biology Use When Explaining This Phenomenon?

The similar engagement in CMR seen between the co-enrolled group of students suggests similar activation of electrostatic resources, such as opposite charges attracting each other. While most students in both cohorts leveraged these ideas, we also noted a handful of students discussing a related idea – polarity – and some students discussing an unrelated idea – space – which we could quantify using our analytic rubric. For example, student 1554 from OB leveraged charge (not polarity) in their response by writing, "The partial negative charges on the carbonyl group and on the OH group will attract the positive charge on the Mg." In contrast, student 1764 from MB leveraged polarity instead of charge by writing, "I chose version 2 for the better binding site because it has more polar side chains, thus making it easier for the magnesium to interact with the side chains within that protein." Another student (1267 from OCII) leveraged both space and charge by writing, "The partial negative charge on the oxygen draws the positive magnesium ion. Protein 1 has less steric hindrance. Protein 1 has less steric hindrance and thus has more space for the magnesium ion to bind." We included this research question to investigate what resources the cross-sectional cohort students used to explain this phenomenon and whether there were differences in the resources used between disciplines.

Polarity Versus Charge

We predicted that students in biology courses would more often invoke polarity when compared to their chemistry student peers, because polarity is the descriptor most used in

biology to discuss the uneven distribution of electron density in a molecule, whereas in chemistry the underlying cause of the uneven distribution is discussed in much greater depth. With our expanded analytic rubric, we were able to identify whether students used charge, polarity, or both in their explanations. For this analysis, we removed the GCI (N=50) students, since we saw from RQ1 that they did not have the content background needed to construct a fully CM explanation. Of the remaining 345 responses analyzed from the cross-sectional cohort, 52 of them invoked polarity in some way, with 48 of them coming from biology (MB or OB) or biochemistry (BCI or BCII) courses. Only one GCII student and three OCII students leveraged polarity. Because the number of responses per course from the cross-sectional cohort was not consistent, we grouped each discipline and calculated the percent of partially or fully CM responses that leveraged polarity (Table 4.4 – non-CM responses removed). We conducted a Pearson's χ^2 test which showed an overall significant difference ($\chi^2 = 14.820$, $p < 0.001$, Cramér's V = 0.236) in resource activation between chemistry, biology, and biochemistry. Over a quarter of biology students used polarity in their responses, 23% of biochemistry students used polarity, and only 5% of chemistry students used polarity (Table 4.4). However, even though polarity was more often invoked in biology than chemistry, a pair-wise Pearson's χ^2 test showed no significant difference in engagement in CMR between these disciplines ($\chi^2 = 1.267$, $p = 0.531$, Cramér's V = 0.067) (see Appendix D for additional statistics comparing engagement in CMR across disciplines).

Table 4.4. Number and percent of students leveraging polarity in their explanations.

Discipline	Total partially CM and fully CM responses	Number of responses using polarity	Percent of responses using polarity
Chemistry (GCII, and OCII, all post, N=100)	78	4	5%
Biology (MB and OB, pre and post, N=180)	136	36	26%
Biochemistry (BCI pre and BCII post, N=65)	53	12	23%

Space

While polarity could be used productively to explain the protein-ligand binding, some students in the cross-sectional cohort (N = 48) leveraged the idea of space to explain the preferential binding, an idea that probably emerges from the representation of the structure, and that is not a productive approach to generating a causal mechanistic explanation of the phenomenon. We consider "space" to be a surface feature, and thus, a resource that does not support a causal mechanistic explanation, because a student who only recognized the available space would not have provided any evidence of considering a lower scalar level (the first requirement for CMR). Further, while space may be a productive approach to explaining binding activity when the ligand is exceptionally large, it is a negligible factor for the binding activity of small, spherical ions like Mg^{2+} (and even in that case, identifying space (or lack thereof) can be thought of as an explanatory black box[37] for the repulsive interactions that emerge when a binding site is "too hindered").

In this set of responses, most frequently (67%) the idea of space showed up as a student noting "more room" or "less steric hindrance" in one of the protein versions (i.e., the spatial availability); however, explanations were also characterized by this bin if they discussed, for example, one version being a better "fit". Unlike the use of polarity, we did not find a significant

difference (Pearson's $\chi^2 = 4.490$, $p = 0.106$) in the proportion of students from each discipline (chemistry, biology, or biochemistry) using space to explain the protein-ligand binding.

Most (79%) of the students in this group selected version 1 (or were unclear in their selection) as the version which would better bind Mg^{2+}. For example, consider student 4000 from BCI who said, "…I think Version 1 has a better binding site simply because there is less steric hinderance in Version 1." Similarly, student 1818 from OB said, "…Protein 1 has a better binding site because there is less 'clutter' within the protein." Along with 58% of all explanations coded for space, this student also referenced electrostatics in their response by writing, "…I show the Mg^{2+} binding to an oxygen because while there are no negative charges on these proteins, the oxygens within the proteins at least have a partially negative charge, which the Mg^{2+} is attracted to, causing it to bind." Thus, even though these students recognized the role of electrostatics in causing Mg^{2+} to bind, they leveraged the idea of space (instead of magnitude of charge) to explain why Mg^{2+} binds preferentially to one site over another. A response such as this would be coded as partially CM, because they included the first two key ideas, but their invocation of space would not be enough for the "linking" code (third step in CMR), because according to Krist et al., this step requires linking the underlying entities/interactions (in this case, the negative oxygens, or polar amino acids) to the target phenomenon. Although the numbers are small, this is an important finding from a resource's perspective, because, while students leveraged appropriate resources (i.e., electrostatics), the format of the prompt may have also activated this alternative resource (i.e., space) for some students, indicating a need to further refine the instruction or activity. That being said, it is also important to note that we do not view the use of space as a misconception, since it can be a

useful idea in other situations, such as macroscopic contexts. Rather, the use of space likely emerged as an artifact of (1) how we represented the binding sites, and/or (2) the organization or framework of resources in students' minds. For example, students who chose to leverage space when making their prediction may not have refined epistemic heuristics that support predicting phenomena of this complexity (i.e., choosing between competing resources).

RQ3 – How Does Engagement In CMR Relate To Students' Overall Predictions?

The most compelling evidence from our analysis of this data is the significant correlation between constructing a fully CM explanation and making a correct prediction. We used all responses from the cross-sectional cohort, as well as the 50 additional MB pre responses from the co-enrolled cohort (N = 445), to address this research question. Results indicated that nearly all (97%) of the students who constructed fully CM explanations also selected the correct protein version (version 2) (Figure 4.7). An overall Pearson's χ^2 test of all three types of responses was significant; however, a pairwise comparison showed no significant difference between the non-CM and partially CM responses (Table 4.5). Thus, the fully CM group's predictions must be driving the significance, which we see in the χ^2 values comparing fully CM predictions to both non-CM ($\chi^2 = 65.689$) and partially CM ($\chi^2 = 91.579$). With large effect sizes (Table 4.5), these results indicate the utility and power of engaging completely in CMR because doing so nearly always leads students to a correct prediction.

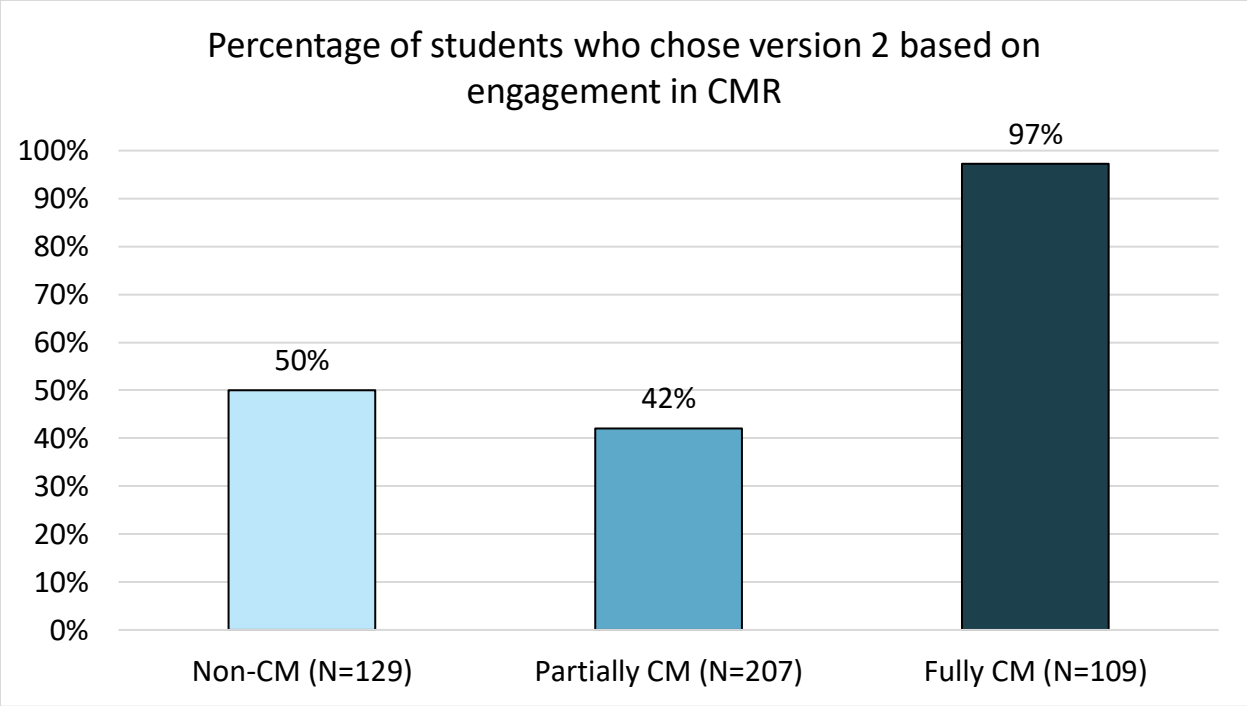


Figure 4.7. Relationship between engagement in CMR and the predictions that students made (i.e., which protein version they chose).

Table 4.5. Pearson's χ^2 test results exploring the relationship between engagement in CMR and binding site predictions.

Comparison	χ^2 Value	*P* Value	Cramér's V[b]
All	94.164	$p < 0.001$	0.460
Non-CM and Partially CM	1.847	$p^a = 0.174$	0.074
Non-CM and Fully CM	65.689	$p^a < 0.001$	0.525
Partially CM and Fully CM	91.579	$p^a < 0.001$	0.538

a. Bonferroni-adjusted $\alpha = 0.017$

b. Suggestions for interpreting Cramér's V[36]: small − 0.1, medium − 0.3, large − 0.5

This analysis clearly indicates that CMR is a powerful tool for accurately predicting and explaining phenomena. Specifically, for preferential protein-ligand binding, comparing the magnitude of charge/polarity between the two sites served as the critical idea for choosing the

appropriate version. While many students identified the role of electrostatics in causing Mg^{2+} to bind, fewer explained why Mg^{2+} binds better to version 2; that is, these students explained how, but not why this phenomenon occurs.

Discussion

In this work, we examined students' causal mechanistic reasoning (CMR) in an interdisciplinary context: preferential protein-ligand binding, a biologically relevant phenomenon that relies on understanding core chemistry ideas. While much of the work done regarding CMR is limited to students' responses from one discipline, we have shared explanations provided by students across three chemistry, two biology, and two biochemistry courses, for the same task. Research that engages students with tasks about interdisciplinary ideas is an important step towards addressing national calls for interdisciplinary education at all levels of science. Tasks such as this one are also important because they can engage students in and examine how students in different disciplines think about these ideas and in which areas they struggle or succeed.

While the results show varying types of student engagement in CMR across all courses, we provide evidence of nearly identical engagement in CMR between students co-enrolled in GCII and MB. These students maintained the level of engagement in CMR regardless of the discipline for which they completed the task, showing to our surprise that the course did not have a large impact on their explanations. This co-enrolled cohort may be leveraging ideas from both of their courses when constructing their explanations; however, additional work should be done to further support this, such as think-aloud interviews probing how and why students are including certain ideas in their explanations. Our results from RQ1 also show that students in

more advanced chemistry courses used CMR more frequently, and students in MB improved their explanations by the end of the semester, suggesting that students refined both their conceptual and epistemic resources over time.

Engagement in CMR did not differ significantly between students in chemistry (GCII and OCII) and biology (MB and OB); however, we did see notable differences in resource activation of charge or polarity based on discipline. The idea of polarity was leveraged far more frequently by students in biology than students in chemistry. This finding, while perhaps not surprising given the scale at which biology courses discuss molecular properties, highlights the importance of context in activating certain resources.[17] From an instructor perspective, it is imperative that we acknowledge the ideas students leverage, even if those ideas are not the intent of the assessment, because it may illuminate alternative appropriate explanations or, if incorrect, potential prompt revisions that better support students to use and apply relevant resources. For example, the handful of students who leveraged space might not have called upon that resource if a bulkier nonpolar amino acid was used in version 1 instead of alanine, or if a more accurate representation of the structure (e.g., a space filling model) were presented. However, an important part of CMR involves choosing appropriate and explanatory resources, even amidst other (perhaps more salient) resources that, when used, provide competing explanations. In this case, invocation of space correlated with an incorrect prediction; however, that is not to say that these students have misconceptions about how/why ligands bind to proteins. Rather, the representation of the prompt activated a resource that may have been more accessible or more strongly connected to binding in their knowledge frameworks.

While the number of students constructing fully CM explanations was not particularly high across courses, the number of students engaging in some aspect of causal mechanistic reasoning (i.e., providing either a fully CM or partially CM response) is promising (75%). All these students leveraged electrostatics in some productive way, providing evidence of an understanding of at least one of the three key ideas needed for a fully CM explanation. Thus, this task and coding approach may serve as a means by which instructors can engage students in this type of deep reasoning, while highlighting key ideas that carry significance in both chemistry and biology. Additionally, using this coding approach allows instructors to identify the ideas present or absent in student responses, which can then serve as a guide for feedback and instruction.

Finally, the most significant finding in this work is the compelling evidence for the predictive power of CMR. The strong correlation between fully CM explanations and correct predictions (97%) shows the value of CMR – that is, considering relevant, lower-level entities, their behaviors and interactions, and how these key ideas link to the target phenomenon. We addressed this research question because, while constructing a fully CM response did not depend on correctness, making correct predictions using CMR shows that these students have developed coherent and sophisticated networks of knowledge (i.e., they have refined their epistemic and conceptual resources so that they can use and apply them to novel situations). For the phenomenon of preferential protein-ligand binding, most students identified the role of electrostatics in causing the Mg^{2+} to bind, but far fewer students linked the electrostatics to the preferential binding. This key idea – that a larger charge results in a stronger attractive force – proved to be the key factor in choosing the correct site, but, because of the small number of

students recognizing this idea, there should be additional work to support students in linking underlying factors to make predictions and explain phenomena.

It would be reasonable to predict that the partially CM explanations would more often choose the correct protein version when compared to the non-CM explanations; however, we saw that this did not happen. While there are six combinations of the three key ideas which could result in a partially CM code, we most often saw that these responses consisted of either the first idea alone, identifying the attraction of oppositely charged species (N=44, 21.3%), or the first and second ideas, identifying this attraction and a lower-level negative entity (N=152, 73.4%). All other responses (N=11, 5.3%) included either the second idea alone, the second and third ideas, the first and third ideas, or the third idea alone. Thus, the partially CM responses rarely included the third idea: comparing the magnitude of charge between sites. That is, even though these students recognized the importance of electrostatics in causing Mg^{2+} to bind, they did not link these ideas to the phenomenon of preferential binding. In fact, a handful of students (N=42) who constructed partially CM explanations used the idea of space to explain the preferential binding; the vast majority (90%) of these students explained that Mg^{2+} would more easily bind to the incorrect binding site (version 1) because there was “more room” or “less steric hindrance”.

The relationship between fully CM responses and correct predictions points to the importance of linking underlying factors to the phenomenon under consideration when constructing a causal mechanistic explanation. This third step in CMR about our phenomenon aligns with the third key idea which involves comparing the magnitude of the charge between sites to explain why Mg^{2+} binds preferentially to version 2 – students who linked ideas about

electrostatics to the phenomenon in this way nearly always (97% of the time) chose the correct binding site.

Limitations

There are a few notable limitations that should be addressed. The number of responses coded in the cross-sectional cohort represents a small number of students in the courses; for example, there were more than a thousand GCI responses, but we only coded 50. It is our intention to code additional responses in the future so that we may have a more complete understanding of each course as a whole; however, the smaller selection that we discussed in this paper provides a starting point in identifying patterns and themes among the student responses. Also due to these small N values, we were not able to group students by previous coursework. For example, some of the OCII students did not complete GCII, and took different GCI courses (e.g., residential college vs large enrollment). Additionally, each student was co-enrolled in several other courses and in different academic years, which undoubtedly impacted their prior knowledge. Typical higher education studies such as these focus on a phenomenon specific to the discipline in which it is administered, but, in this case, the phenomenon is general and unfamiliar to students in most of the courses shown here, with perhaps the exception of the biochemistry courses, which discuss enzyme binding and active sites in more detail.

In addition to the sample selection and distribution, we note that the chemistry and biology courses in which this task was given were transformed. The GCI, GCII, and OCII courses use transformed curricula (CLUE and OCLUE),[2,38] which emphasize CMR in frequent homework activities, recitation activities, and other assessments. Thus, these students may be better

equipped than students in more traditional chemistry courses to explain this phenomenon. Likewise, the MB and OB courses at this institution encourage and provide opportunities for students to explain how and why phenomena occur. It is likely that students at a different institution would show different engagement in CMR.

Lastly, upon further analysis of student responses to parts 2 and 3 of this task, our colleagues found a small number (N=18) of responses that were identical to one or more other student responses (not all 18 were identical). Six of these responses were constructed by students in the co-enrolled cohort, meaning three students likely copied and pasted their response. Ten of these students seem to have worked with a classmate (or two) to construct identical responses, and the remaining two students had unique responses for the PL task, but their responses to parts 2 and/or 3 were identical to one or more other student responses. We did not re-analyze all of our data after discovering this, because the numbers were small in each course. With the activity being completed online, it is not surprising that a few students worked together or copied an explanation. Thus, while having negligible effects on our results, it is important to note that [such unobvious examples of] copied and pasted responses will likely play a role for future studies of open-ended responses in large-enrollment courses.

Conclusion And Future Directions

Causal mechanistic reasoning is applicable across a range of phenomena – and important for deeper and longer-lasting learning in science. This study emphasized CMR across the disciplinary boundary of chemistry and biology. Our task (addressing the phenomenon of preferential protein-ligand binding) requires an understanding of three key ideas: (1) the attraction of oppositely charged species, (2) the negative or polar nature of specific atoms or

amino acids in a binding site, and (3) a larger charge (or more polar species) resulting in a stronger attractive force. These ideas are, ideally, learned early in an undergraduate science curriculum and used in future contexts such as protein-ligand binding in biology. However, causal mechanisms in earlier courses like chemistry might not be tied together explicitly within and across courses. However, each idea, when connected appropriately, can be used to construct a causal mechanistic explanation and productive prediction for the phenomenon.

The results of analysis from students' explanations indicate that most explanations were partially mechanistic, meaning that these students see the relevance of electrostatics in the phenomenon of protein-ligand binding but need additional guidance in explaining the preferential binding. Thus, we plan to develop feedback statements that might better support these students in developing and refining their resources involving electrostatics and structure-property relationships. However, providing feedback in a timely manner to a large population of students poses a hurdle we are working to overcome; we may investigate whether peer feedback or automated feedback could be used to navigate this hurdle. Additionally, the use of space, as a resource that competes with magnitude of charge to explain preferential binding, resulted in several rich discussions among our group. A revised activity that better supports students in choosing appropriate resources (without making it obvious that space is irrelevant) would be an interesting study. For example, adding a third binding site consisting of three alanine amino acids (and therefore the most "space" available) might provide additional insight into student use of this idea as it compares to their use of electrostatics. We encourage others in the field to further this investigation using related ideas or phenomena.

Our ultimate goal is to support students' interdisciplinary learning and mechanistic reasoning by engaging them in tasks that elicit CMR in the context of phenomena that span disciplinary boundaries. CMR, as evidenced by our results, is a useful way of thinking that spans scalar levels and results in powerful predictions and explanations. The work published here is intended for both chemistry and biology instructors at the undergraduate level. We hope that these findings encourage our audience to engage students in CMR, because it will support them in their future lives as scientists and as citizens in an increasingly complex world.

REFERENCES

(1) American Association for the Advancement of Science. Vision and Change in Undergraduate Biology Education: A Call to Action; Washington D.C., 2011. https://live-visionandchange.pantheonsite.io/wp-content/uploads/2011/03/Revised-Vision-and-Change-Final-Report.pdf (accessed 2020-04-17).

(2) Cooper, M.; Klymkowsky, M. Chemistry, Life, the Universe, and Everything: A New Approach to General Chemistry, and a Model for Curriculum Reform. *J. Chem. Educ.* **2013**, *90* (9), 1116–1122. https://doi.org/10.1021/ed300456y.

(3) National Research Council. A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas; National Academies Press: Washington, DC, 2012.

(4) Ledford, H. How to Solve the World's Biggest Problems. *Nature* **2015**, *525* (7569), 308–311. https://doi.org/10.1038/525308a.

(5) Cooper, M. M.; Caballero, M. D.; Ebert-May, D.; Fata-Hartley, C. L.; Jardeleza, S. E.; Krajcik, J. S.; Laverty, J. T.; Matz, R. L.; Posey, L. A.; Underwood, S. M. Challenge Faculty to Transform STEM Learning. *Science* **2015**, *350* (6258), 281–282. https://doi.org/10.1126/science.aab0933.

(6) Luckie, D. B.; Bellon, R.; Sweeder, R. D. The BRAID: Experiments in Stitching Together Disciplines at a Big Ten University. *Journal of STEM Education: Innovations and Research* **2012**, *13* (2).

(7) Roche Allred, Z.; Shrode, A. D.; Gonzalez, J.; Rose, A.; Green, A. I.; Swamy, U.; Matz, R. L.; Underwood, S. M. Impact of Ocean Acidification on Shelled Organisms: Supporting Integration of Chemistry and Biology Knowledge through Multidisciplinary Activities. *J. Chem. Educ.* **2022**, *99* (5), 2182–2189. https://doi.org/10.1021/acs.jchemed.1c00981.

(8) Green, A. I.; Parent, K. N.; Underwood, S. M.; Matz, R. L. Connecting Ideas across Courses: Relating Energy, Bonds & How ATP Hydrolysis Powers a Molecular Motor. *The American Biology Teacher* **2021**, *83* (5), 303–310. https://doi.org/10.1525/abt.2021.83.5.303.

(9) Martinez, B. L.; Kararo, A. T.; Parent, K. N.; Underwood, S. M.; Matz, R. L. Creating and Testing an Activity with Interdisciplinary Connections: Entropy to Osmosis. *Chem. Educ. Res. Pract.* **2021**, *22* (3), 683–696. https://doi.org/10.1039/D0RP00353K.

(10) Allred, Z. D. R.; Farias, A. J.; Kararo, A. T.; Parent, K. N.; Matz, R. L.; Underwood, S. M. Students' Use of Chemistry Core Ideas to Explain the Structure and Stability of DNA. *Biochemistry and Molecular Biology Education* **2021**, *49* (1), 55–68. https://doi.org/10.1002/bmb.21391.

(11) Kohn, K. P.; Underwood, S. M.; Cooper, M. M. Energy Connections and Misconnections across Chemistry and Biology. *LSE* **2018**, *17* (1), ar3. https://doi.org/10.1187/cbe.17-08-0169.

(12) Kohn, K. P.; Underwood, S. M.; Cooper, M. M. Connecting Structure–Property and Structure–Function Relationships across the Disciplines of Chemistry and Biology: Exploring Student Perceptions. *LSE* **2018**, *17* (2), ar33. https://doi.org/10.1187/cbe.18-01-0004.

(13) Gayford, C. ATP: A Coherent View for School Advanced Level Studies in Biology. *Journal of Biological Education* **1986**, *20* (1), 27–32. https://doi.org/10.1080/00219266.1986.9654772.

(14) Halmo, S. M.; Sensibaugh, C. A.; Bhatia, K. S.; Howell, A.; Ferryanto, E. P.; Choe, B.; Kehoe, K.; Watson, M.; Lemons, P. P. Student Difficulties during Structure–Function Problem Solving. *Biochemistry and Molecular Biology Education* **2018**, *46* (5), 453–463. https://doi.org/10.1002/bmb.21166.

(15) Russ, R. S.; Scherr, R. E.; Hammer, D.; Mikeska, J. Recognizing Mechanistic Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed from Philosophy of Science. *Science Education* **2008**, *92* (3), 499–525. https://doi.org/10.1002/sce.20264.

(16) Krist, C.; Schwarz, C. V.; Reiser, B. J. Identifying Essential Epistemic Heuristics for Guiding Mechanistic Reasoning in Science Learning. *Journal of the Learning Sciences* **2019**, *28* (2), 160–205. https://doi.org/10.1080/10508406.2018.1510404.

(17) Hammer, D. Student Resources for Learning Introductory Physics. *American Journal of Physics* **2000**, *68* (S1), S52–S59. https://doi.org/10.1119/1.19520.

(18) Becker, N.; Noyes, K.; Cooper, M. Characterizing Students' Mechanistic Reasoning about London Dispersion Forces. *J. Chem. Educ.* **2016**, *93* (10), 1713–1724. https://doi.org/10.1021/acs.jchemed.6b00298.

(19) Cooper, M. M.; Williams, L. C.; Underwood, S. M. Student Understanding of Intermolecular Forces: A Multimodal Study. *J. Chem. Educ.* **2015**, *92* (8), 1288–1298. https://doi.org/10.1021/acs.jchemed.5b00169.

(20) Crandell, O. M.; Lockhart, M. A.; Cooper, M. M. Arrows on the Page Are Not a Good Gauge: Evidence for the Importance of Causal Mechanistic Explanations about Nucleophilic Substitution in Organic Chemistry. *J. Chem. Educ.* **2020**, *97* (2), 313–327. https://doi.org/10.1021/acs.jchemed.9b00815.

(21) Noyes, K.; Cooper, M. M. Investigating Student Understanding of London Dispersion Forces: A Longitudinal Study. *J. Chem. Educ.* **2019**, *96* (9), 1821–1832. https://doi.org/10.1021/acs.jchemed.9b00455.

(22) Crandell, O. M.; Kouyoumdjian, H.; Underwood, S. M.; Cooper, M. M. Reasoning about Reactions in Organic Chemistry: Starting It in General Chemistry. *J. Chem. Educ.* **2019**, *96* (2), 213–226. https://doi.org/10.1021/acs.jchemed.8b00784.

(23) Caspari, I.; Kranz, D.; Graulich, N. Resolving the Complexity of Organic Chemistry Students' Reasoning through the Lens of a Mechanistic Framework. *Chemistry Education Research and Practice* **2018**, *19* (4), 1117–1141. https://doi.org/10.1039/C8RP00131F.

(24) Graulich, N.; Caspari, I. Designing a Scaffold for Mechanistic Reasoning in Organic Chemistry. *Chemistry Teacher International* **2021**, *3* (1), 19–30. https://doi.org/10.1515/cti-2020-0001.

(25) Ralph, V. R.; Scharlott, L. J.; Schafer, A. G. L.; Deshaye, M. Y.; Becker, N. M.; Stowe, R. L. Advancing Equity in STEM: The Impact Assessment Design Has on Who Succeeds in Undergraduate Introductory Chemistry. *JACS Au* **2022**. https://doi.org/10.1021/jacsau.2c00221.

(26) Schwarz, C.; Cooper, M.; Long, T.; Trujillo, C.; Noyes, K.; de Lima, J.; Kesh, J.; Stolzfus, J. Mechanistic Explanations Across Undergraduate Chemistry and Biology Courses. In *ICLS 2020 Proceedings*; International Society of the Learning Sciences: Nashville, TN, USA, 2020; pp 625–628. https://doi.org/10.22318/icls2020.625.

(27) Yoho, R.; Foster, T.; Urban-Lurain, M.; Merrill, J.; Haudek, K. C. Interdisciplinary Insights from Instructor Interviews Reconciling "Structure and Function" in Biology, Biochemistry, and Chemistry through the Context of Enzyme Binding. *Disciplinary and Interdisciplinary Science Education Research* **2019**, *1* (1), 16. https://doi.org/10.1186/s43031-019-0016-7.

(28) Noyes, K.; Carlson, C. G.; Stoltzfus, J. R.; Schwarz, C. V.; Long, T. M.; Cooper, M. M. A Deep Look into Designing a Task and Coding Scheme through the Lens of Causal Mechanistic Reasoning. *J. Chem. Educ.* **2022**, *99* (2), 874–885. https://doi.org/10.1021/acs.jchemed.1c00959.

(29) Mislevy, R. J.; Almond, R. G.; Lukas, J. F. A Brief Introduction to Evidence-Centered Design. *ETS Research Report Series* **2003**, *2003* (1), i–29. https://doi.org/10.1002/j.2333-8504.2003.tb01908.x.

(30) Harris, C. J.; Krajcik, J. S.; Pellegrino, J. W.; DeBarger, A. H. Designing Knowledge-In-Use Assessments to Promote Deeper Learning. *Educational Measurement: Issues and Practice* **2019**, *38* (2), 53–67. https://doi.org/10.1111/emip.12253.

(31) Bryfczynski, S. BeSocratic: An Intelligent Tutoring System for the Recognition, Evaluation, and Analysis of Free-Form Student Input. *All Dissertations* **2012**.

(32) Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* **1960**, *20* (1), 37–46. https://doi.org/10.1177/001316446002000104.

(33) Landis, J. R.; Koch, G. G. The Measurement of Observer Agreement for Categorical Data. *Biometrics* **1977**, *33* (1), 159–174. https://doi.org/10.2307/2529310.

(34) IBM SPSS Statistics, 2020.

(35) Green, S. B.; Salkind, N. J. Two-Way Contingency Table Analysis Using Crosstabs. In *Using SPSS for Windows and Macintosh: Analyzing and Understanding Data*; Pearson Education Inc.: Upper Saddle River, NJ, 2011; pp 366–376.

(36) Cohen, J. A Power Primer. *Psychol Bull* **1992**, *112* (1), 155–159. https://doi.org/10.1037//0033-2909.112.1.155.

(37) Haskel-Ittah, M. Explanatory Black Boxes and Mechanistic Reasoning. *Journal of Research in Science Teaching* **2023**, *60* (4). https://doi.org/10.1002/tea.21817.

(38) Cooper, M. M.; Stowe, R. L.; Crandell, O. M.; Klymkowsky, M. W. Organic Chemistry, Life, the Universe and Everything (OCLUE): A Transformed Organic Chemistry Curriculum. *J. Chem. Educ.* **2019**, *96* (9), 1858–1872. https://doi.org/10.1021/acs.jchemed.9b00401.

(39) Morris, J.; Hartl, D.; Knoll, A.; Lue, R. *Biology: How Life Works*, 2nd ed.; W. H. Freeman: New York, NY, 2016.

(40) Nelson, D. L.; Cox, M. M. *Lehninger Principles of Biochemistry*, 7th ed.; W. H. Freeman: New York, NY, 2017.

APPENDIX A. PERMISSIONS

Figure 4.8. Permissions to reproduce manuscript in its entirety.

APPENDIX B. COURSE INFORMATION

Course And Timepoint Selection

To explore how co-enrolled students responded to our task for both their chemistry and biology courses, we administered the activity to both GCII and MB, two courses frequently taken in the same semester. To capture a longitudinal cohort, we administered the activity at the beginning and end of both MB and OB. Administering the activity in these three courses provided us with important timepoints for our cross-sectional cohort, but to capture a more complete picture of the chemistry and biology sequences, we also administered the activity in GCI, OCII, BCI, and BCII. We administered the activity to GCI and OCII to collect responses from students during both the first and last of the introductory chemistry course sequence. We did not collect "pre" responses from GC1, because this course is typically required for other chemistry and biology courses, meaning students (at the beginning of the semester) would not yet have the chemistry knowledge necessary to understand this mechanism. Finally, we also administered this activity to BCI and BCII students to see how students explained this phenomenon in the more senior-level courses to which these introductory chemistry and biology courses lead. We administered the activity to the BCI students at the beginning of the semester to capture students at the start of the biochemistry sequence. Table 4.6 characterizes each course with a brief description and required textbooks.

Table 4.6. Information for each course in which the PL task was administered.

Course	Description	Number of sections	Number of instructors	Textbook
GCI	This first semester general chemistry course uses a lecture-style format paired with weekly one-hour recitation sections (led by graduate teaching assistants), where students work in small groups on activities.	48	3	[a]*Chemistry, Life, the Universe, and Everything*, by Melanie M. Cooper and Michael W. Klymkowsky
GCII	This second semester general chemistry course uses a lecture-style format paired with weekly one-hour recitation sections (led by graduate teaching assistants), where students work in small groups on activities.	36	2	[a]*Chemistry, Life, the Universe, and Everything*, by Melanie M. Cooper and Michael W. Klymkowsky
OCII	This second semester organic chemistry course uses a lecture-style format paired with weekly one-hour recitation sections (led by graduate teaching assistants), where students work in small groups on activities.	12	1	[b]*Organic Chemistry, Life, the Universe, and Everything*, by Melanie M. Cooper and Michael W. Klymkowsky
MB	This molecular biology course uses a traditional lecture-style curriculum, with in-class activities engaging students in questions that probe how and why phenomena occur	6	8 (4 sections had a single faculty instructor, 2 sections were team taught by two faculty)	[c]*How Life Works*, 2nd edition by James Morris
OB	This organismal biology course uses a traditional lecture-style curriculum, with in-class activities engaging students in questions that probe how and why phenomena occur	4	5 (3 sections had a single faculty instructor, 1 section was team taught by two faculty)	[c]*How Life Works* 2nd edition by James Morris
BCI	This first-semester biochemistry course uses a traditional lecture-style curriculum, with in-class clicker questions about course content	1	4 (each instructor taught for a given unit or portion of the course)	[d]Nelson, D.L. and Cox, M.M. *Lehninger's Principles of Biochemistry*

Table 4.6 (cont'd)

BCII	This second-semester biochemistry course uses a traditional lecture-style curriculum, with in-class clicker questions about course content	1	4 (each instructor taught for a given unit or portion of the course)	[d]Nelson, D.L. and Cox, M.M. *Lehninger's Principles of Biochemistry*

a. CLUE[2]

b. OCLUE[38]

c. How Life Works[39]

d. Lehninger[40]

APPENDIX C. INDEPENDENT T TESTS FOR SAMPLE POPULATIONS

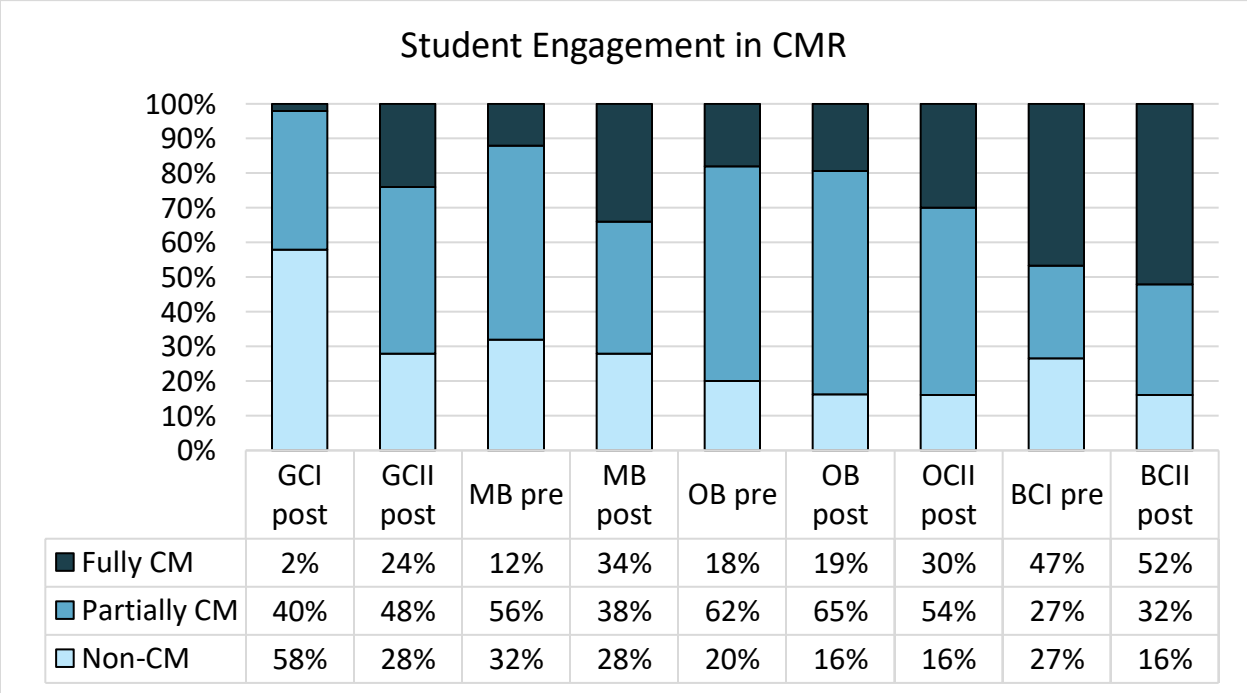
The responses selected for coding were constructed by students that are representative of their group. We conducted independent samples *t* tests for seven groups, and the results for each comparison are shown in Table 4.7. The GPAs for MB and OB represent cumulative GPA at the start of the term, while GPAs for the chemistry courses represent cumulative GPA for the course term. We did not include the BCI pre or BCII post groups, since the majority of the responses collected were also coded. Note: we were missing GPA data from 20 MB pre students, three OB pre students, six GCI students, one GCII student, and one OCII student. These 31 students were not included in the *t* tests, nor were they coded or used for analyses. Within the groups of students that were coded, we were missing SAT scores from five MB students, two OB students, five GCI students, four GCII students, and three OCII students. The results comparing both GPA and SAT metrics confirm that the students selected for coding are representative of their respective populations, since there were no significant differences.

Table 4.7. Independent samples *t* tests results evaluating the difference between the selected students for coding and the rest of the students in their representative group.

Group	GPA *P* value	SAT *P* value
GCI post	$p = 0.146$	$p = 0.525$
GCII post	$p = 0.374$	$p = 0.266$
OCII post	$p = 0.877$	$p = 0.806$
MB pre	$p = 0.186$	$p = 0.724$
MB post	$p = 0.066$	$p = 0.545$
OB pre	$p = 0.765$	$p = 0.465$
OB post	$p = 0.195$	$p = 0.179$

APPENDIX D. ENGAGEMENT IN CMR ACROSS DISCIPLINES

In Figure 4.9, we show student engagement in CMR for each course and timepoint collected for the cross-sectional cohort. The order in which we have placed these courses (left to right) represents the approximate order in which students might progress through this sequence; however, some courses are frequently taken at the same time (such as GCII and MB). In general, these findings show that students who have progressed through this sequence engage more fully in CMR, with the BCII post students constructing the most fully CM responses (and GCI students the least).



	GCI post	GCII post	MB pre	MB post	OB pre	OB post	OCII post	BCI pre	BCII post
Fully CM	2%	24%	12%	34%	18%	19%	30%	47%	52%
Partially CM	40%	48%	56%	38%	62%	65%	54%	27%	32%
Non-CM	58%	28%	32%	28%	20%	16%	16%	27%	16%

Figure 4.9. Proportion of non-CM, partially CM, and fully CM explanations from each course/time point in the cross-sectional cohort. See N values in Table 4.1.

We also show student engagement in CMR according to discipline (Table 4.8 and Figure 4.10), with the GCI responses removed (since they were shown to not have the appropriate

resources required to reason causal mechanistically about this phenomenon). There was a statistically significant difference when comparing all three disciplines ($p < 0.001$); however, a pair-wise comparison of only chemistry and biology showed no significant difference in engagement in CMR between these two disciplines ($p = 0.531$).

Table 4.8. Pearson's χ^2 test results exploring the relationship between chemistry, biology, and biochemistry student engagement in CMR.

Comparison	χ^2 Value	p Value[a]	Cramér's V[b]
Chemistry, Biology, and Biochemistry	21.246	$p < 0.001$	0.175
Chemistry and Biology	1.267	$p^{\alpha} = 0.531$	0.067
Chemistry and Biochemistry	10.107	$p^{\alpha} = 0.006$	0.247
Biology and Biochemistry	20.800	$p^{\alpha} < 0.001$	0.291

a. Bonferroni-adjusted $\alpha = 0.017$

b. Suggestions for interpreting Cramér's V[36]: small − 0.1, medium − 0.3, large − 0.5

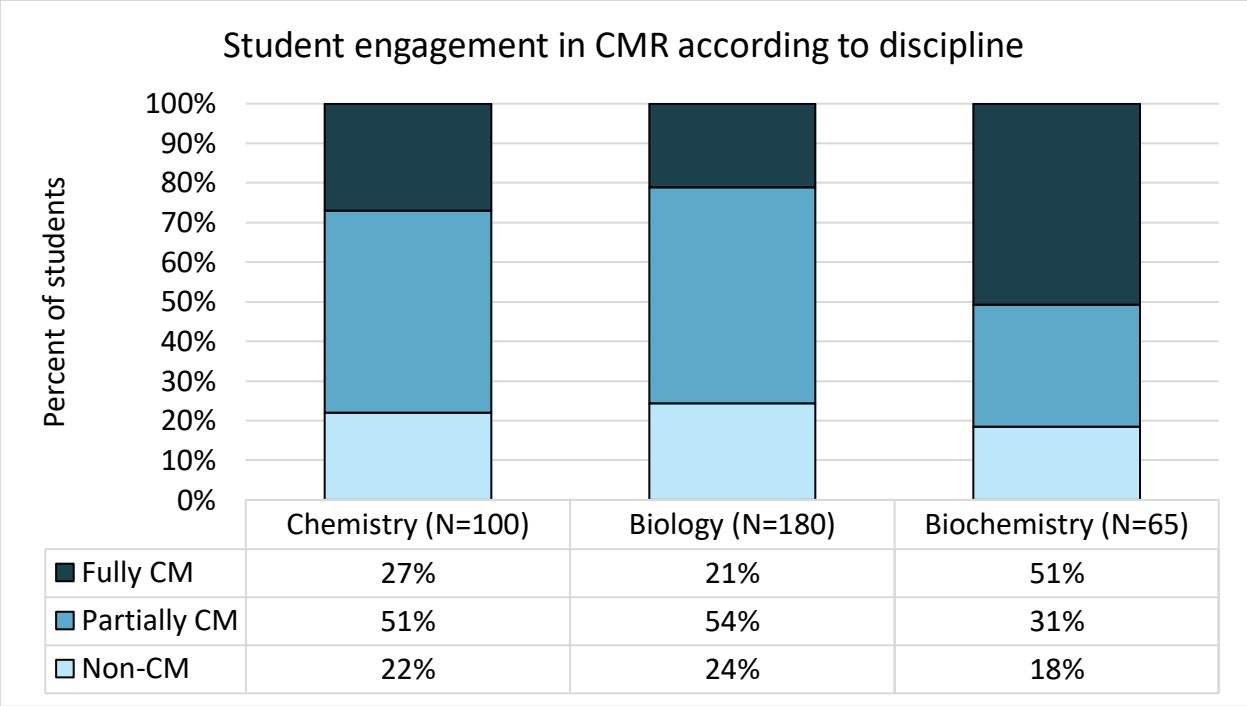


Figure 4.10. The percentage of non-CM, partially CM, and fully CM responses according to discipline (GCI responses removed).

APPENDIX E. CHEMISTRY GPA DESCRIPTIVE STATISTICS

Table 4.9. GCI, GCII, and OCII Mean GPA and SAT Scores.

	GPA		SAT	
Course	Mean	Std. Deviation	Mean	Std. Deviation
GCI	3.514	0.3964	1165	137.772
GCII	3.472	0.4338	1236	137.838
OCII	3.490	0.4639	1239	123.140
Total	3.492	0.4297	1213	136.478

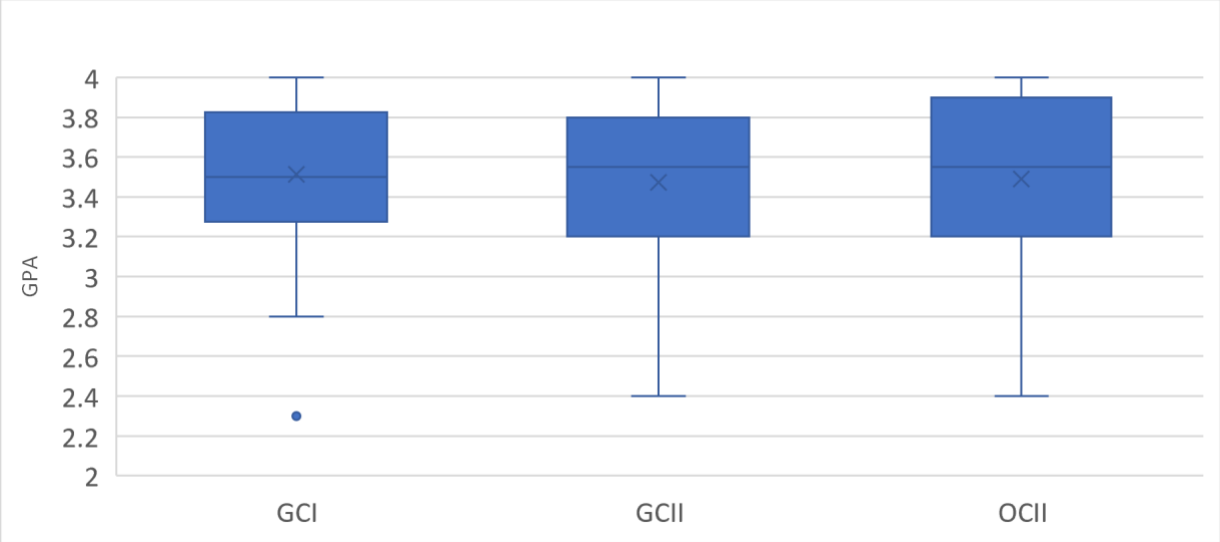


Figure 4.11. Box plots showing the distribution of GPA values in each of the three chemistry courses (N = 50 for each course).

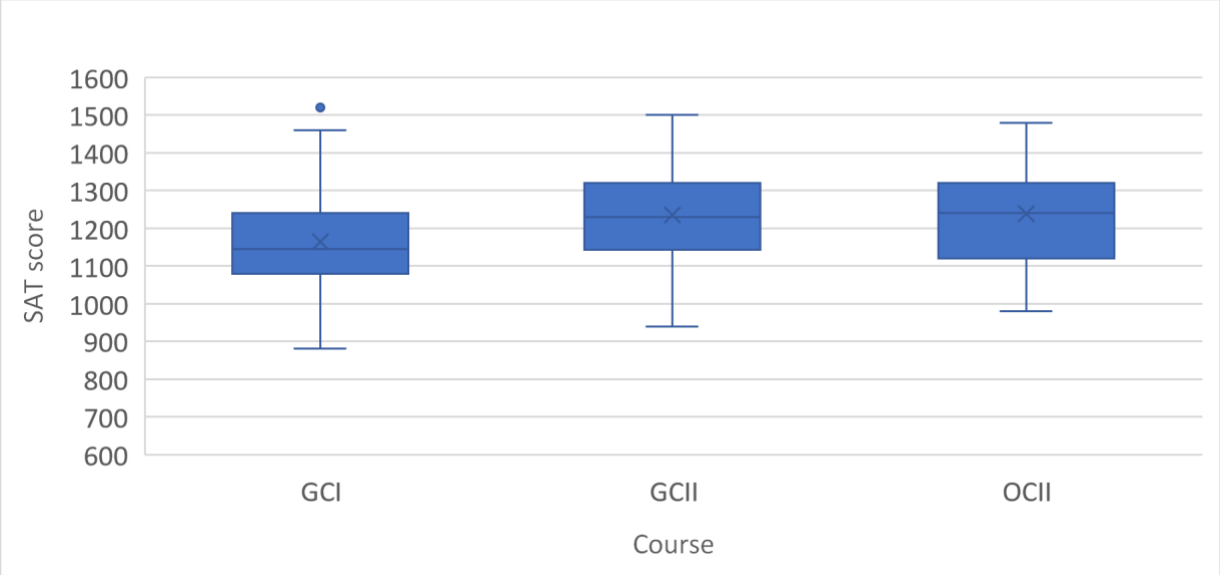


Figure 4.12. Box plots showing the distribution of SAT scores in each of the three chemistry courses.

Chapter V – How Do Different Groups Of Students Engage In CMR About Protein-Ligand Binding?

Introduction

Course culture emerges from an amalgamation of instructional practices, materials, social interactions, physical environment, etc., and can have a profound an impact on students (be it their motivation, experiences, success, etc.). Assessments are an important aspect of course culture, because they send a strong message to students about what is valued and important in a course or discipline[1]. Thus, as educators, we must think deeply about the types of assessments we use. There is strong evidence suggesting that some assessments, such as high-stakes multiple-choice exams, disproportionately and negatively affect marginalized students[2,3]. Further, these types of assessments (rote knowledge or calculations) provide poor evidence about what students know and can do. On the other hand, items that are carefully designed to elicit deep reasoning (such as constructing explanations or engaging in other scientific practices) give instructors a much more reliable idea of what their students know[4]. Using appropriately designed assessments not only provides instructors with trustworthy data, but it also gives the students the opportunity to engage in scientific practices (i.e., the ways in which scientists use their knowledge)[5]. Further, courses that frequently assess students' engagement in scientific practices by asking them to construct explanations or models, result in students being more likely to identify the goals of the course in alignment with instructor goals[6]. Bowen (2022) administered open-ended prompts, to students in both transformed and traditional Organic Chemistry (OC) courses, asking students how they perceived they were expected to think and how they were assessed, finding that students in the transformed course

were more likely to perceive the expectation of "Use of Knowledge", whereas students in the traditional course more often reported rote learning or memorization[6].

Knowing that the types of assessments we use have an impact on student success, it is important to pay close attention to the design and implementation of equitable assessments. That is, assessments that engage students in scientific practices and emphasize core ideas rather than those assessing rote knowledge or skills, which provide little evidence about student knowledge. Courses that include and emphasize these types of assessments begin to deconstruct the barriers that often exclude historically marginalized students. For example, the CLUE curriculum, a transformed general chemistry curriculum that frequently engages students in three-dimensional learning via homework, recitation, and other assessments[7], provides an environment in which more students than the previously implemented traditional curriculum (which assessed surface-level recall/calculations of a broad range of topics) are successful[2].

Our aims focus on developing and administering tasks that engage students in mechanistic reasoning, which, when emphasized on assessments, has been shown to be more equitable than assessments emphasizing calculations or rote knowledge[3]. Mechanistic reasoning (MR) serves as a thinking strategy that involves explaining how and why a target phenomenon occurs by identifying relevant (often lower-level) entities and their interactions/properties/activities and linking these ideas to the phenomenon[8–12]. This thinking strategy is useful, because, when practiced, it may better prepare students for a world in which they can make informed predictions and decisions. For example, reasoning about how and why vaccines are (or are not) effective results in a more confident, informed decision about whether to be vaccinated. With many data sources and media at our fingertips, it is imperative that we

help students develop useful thinking strategies (and metaknowledge – that is, thinking about knowledge; for example, what counts as knowledge or what makes someone a knower).

We use a modified version of evidence-centered design, as outlined by Mislevy[13], to design tasks that support students' mechanistic reasoning and elicit reliable evidence of their use of conceptual and epistemic resources when explaining/predicting phenomena. This process is outlined by Noyes et al.[14], explicating the design of a task (and coding scheme) about protein-ligand binding. We collected responses to this task from students in a range of chemistry and biology courses and characterized their engagement in causal mechanistic reasoning (CMR). We found that some students did not engage in CMR, while others did so either partially or fully, with the Fully CM responses correlating to more accurate predictions, pointing to the power and utility of this reasoning strategy[15].

Research Questions

In this study, we are interested in how students with different demographic characteristics engage in CMR about preferential protein-ligand binding. In our earlier work, we investigated how students in different courses engaged in CMR and what resources they used when completing this task[15]; however, we did not have a large enough sample of coded responses to disaggregate the data and compare across student demographics. We do so here to answer the following research questions:

1. How do males and females compare in their engagement in CMR about protein-ligand binding?
2. How do White students compare to Non-White students in their engagement in CMR about protein-ligand binding?

3. To what extent do cumulative GPA, (binary) gender, and race/ethnicity predict student engagement in CMR?

Methods

In this section, we describe the context and methods used to conduct this study. We used a previously developed mechanistic reasoning task and corresponding coding scheme to group students from a range of courses according to their engagement in CMR[14,15]. The interdisciplinary task focuses on protein-ligand binding, a phenomenon relevant in both chemistry and biology. It was developed with the goal of helping students to build connections between chemistry and biology via reasoning mechanistically about phenomena.

Task And Coding Scheme

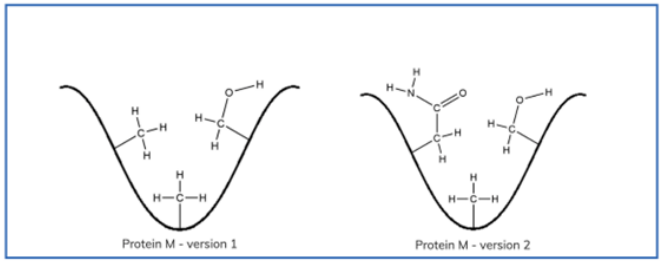
Noyes et al. (2022) described the development of the task and the coding scheme, which Franovic et al. (2023) used to show how students in a range of courses engage in CMR about protein-ligand binding. Here, we briefly describe these findings in order to provide the context for the analysis done in the present study. Should the reader need additional details about this prior work, we direct them to Noyes et al. (2022) and Franovic et al. (2023) (Chapter IV in this dissertation).[14,15]

PL Task

The protein-ligand binding task, hereon referred to as the *PL task*, asks students to predict which of two hypothetical protein binding sites a positive Mg^{2+} ion would most likely bind. The students are shown the two binding sites and asked to draw the Mg^{2+} ion in the site it would most likely bind; then, they are prompted to explain what *causes* the Mg^{2+} to bind and *why* the structural differences cause this difference in binding (Figure 5.1). A mechanistic

explanation for this phenomenon should leverage productive resources regarding the core idea *electrostatic bonding and interactions* as well as the implicit properties of atoms/molecules shown in the binding sites. For example, the student should recognize that the oxygens in Version 2 carry partial negative charges, which would attract the positive Mg^{2+} ion; further, Version 1 only has one oxygen, so Version 2 would have a larger negative charge (or more polar amino acids) and therefore attract the Mg^{2+} more strongly.

The drawings below represent binding sites in two different versions of protein M showing only the atoms in relevant amino acid side chains. Consider a positively charged magnesium ion (Mg^{2+}). Pick the binding site you think is most likely to bind the magnesium ion and **draw** the ion in the binding site **showing why it is binding** in that site.



Explain what causes the magnesium ion to bind to the protein **making specific references to your drawing.**

Explain why the protein you chose has the better magnesium binding site and **how** the structural differences in the site cause this difference in binding.

Figure 5.1. The full PL task (also shown in Figure 4.1 in Chapter IV of this dissertation).

We designed the holistic coding scheme through the lenses of Krist et al.'s framework for MR[9] and Hammer's resources perspective of knowledge construction[16]. With these perspectives and our interest in characterizing the degree to which students engage in CMR, we characterized student responses as (1) non-causal mechanistic (CM), (2) Partially CM, or (3) Fully CM. We characterized the responses in this way based on the presence or absence of three key ideas for this phenomenon: (1) opposite charges attract, or polar entities attract

charged entities, (2) identification of a negative/polar atom, amino acid, or functional group, and (3) a larger negative charge (or more polar binding site) more strongly attracts the Mg^{2+} (i.e., the magnitude of charge impacts the strength of attraction/binding). If all three of these ideas were present in a student's response, we coded it as Fully CM; if only one or two of the ideas were present, we coded it as Partially CM; and if the response did not provide evidence of any of the three ideas, we coded it as Non-CM. Table 5.1 shows each code, a description of the requirements to get that code, and an example student response. For a more detailed description of the coding scheme and less frequent themes that emerged in student responses, we refer the reader to Franovic et al. (2023). For the purposes of this manuscript, we are solely interested in how different students engaged in mechanistic reasoning.

Table 5.1. PL coding scheme.

Code	Description	Example student response
Non-CM	No evidence of any key (mechanistic) ideas. The response does not identify a negative/polar entity, does not indicate that opposite charges attract, and does not leverage magnitude of charge to predict preferential binding.	*"This protein has the better magnesium binding site because its easier to bind and the structural differences in the site causes a difference because it makes the structure longer. The extra carbon helps the magnesium ion to bind to the protein."*
Partially CM	Evidence of only one or two key (mechanistic) ideas. Most responses in this category identify oxygen as being (partially) negative and attracting the Mg^{2+} ion.	*"I picked version one to bind Mg^{2+} to O. This is because O is usually negatively charged, so there is a high chance that it will bind to the Mg which is positively charged. According to Coulomb's law, opposite charges attract. This is why I think Mg^{2+} will bind with O."*
Fully CM	Response provides evidence of understanding/leveraging all three key (mechanistic) ideas. The student identifies one or multiple lower-level negative/polar entities, explains the attraction between negative and positive charges, and explains that a larger charge will result in a stronger attraction.	*"The magnesium ion would bind to Protein M version 2 because the oxygens have lone pairs of electrons which would create a partially negative charge and draw in the positively charged magnesium ion. Protein M version 2 has a better magnesium binding site than the Protein M version 1 because there are more lone pairs of electrons available in Protein M version 2."*

Participants And Data Collection

For this study, we used the same student responses collected by Franovic et al. (2023), which we characterized according to the coding scheme in Table 5.1; however, in this work, we focus only on four mid-level courses: second-semester General Chemistry (GC2), Introductory Cell and Molecular Biology (MB), Organismal Biology (OB), and second-semester Organic Chemistry (OC2). All PL task responses analyzed in this study were collected (as an extra credit opportunity) at the end of the Spring 2020 semester via an online assessment platform called beSocratic, which is set up similar to PowerPoint in that students work through a series of slides where they can draw and construct explanations[17]. Students could choose to opt out of the study and still complete the task for extra credit; we removed these responses from our analysis. We gathered demographic information from the institution's registrar's office. We randomized (using a random number generator) and deidentified all responses by assigning an ID# (1001 – 5092) – this random ID# was the only identifier that coders (authors CF, KN, and DL) could see while coding. Franovic et al. (2023) analyzed 181 responses from these four courses[15]; however, in order to do a more thorough demographic analysis, we needed to code additional responses for a sufficient sample size (as this institution is predominantly White). After removing 7 responses that showed up as duplicates or blanks during coding, we ended with a total of 1,063 total coded responses for this analysis. We compared the coded responses from Franovic et al. (2023) (N = 181) to the additional responses we coded for this manuscript (N = 882), and there was no difference in the engagement in CMR for all courses except GC2, which showed a negligible difference (p = 0.008, Cramer's V = 0.119 (small[18])). Therefore, for this

manuscript, we combined the two sets of data (total N = 1,063) to maximize the number of responses representing each course. These numbers are listed in Table 5.2.

Table 5.2. Number of responses coded from each course.

Course	Number of responses
General Chemistry 2 (GC2)	533
Organic Chemistry 2 (OC2)	174
Cell & Molecular Biology (MB)	197
Organismal Biology (OB)	159

Student Demographics

We collected student demographic information in accordance with our IRB and the institution's registrar's office policies (and FERPA regulations). The data we requested include cumulative GPA, race/ethnicity, and gender (binary, m/f). Table 5.3 provides race/ethnicity, gender, and course information for the responses, as these are the variables we used to answer the RQs. The information here is limited, because it consists of previously determined selections. For example, while gender is not binary, the registrar information has it listed as such, limiting the ways in which students can identify. While these limitations exist, this work provides a glimpse into how our MR-focused tasks are completed by different groups of students. We used IBM SPSS Statistics Version 28[19] to calculate Pearson's Chi Square significance and Cramer's V effect sizes[20] for the following results, which compare (1) female students to male students and (2) White students to Non-White students. Then, we used IBM SPSS Statistics Version 28 to calculate an ordinal regression general linearized model that considers three independent variables: GPA, gender, and race/ethnicity, to identify which (if any) of these variables contributes to the model in predicting student engagement in CMR.

As noted, these students were enrolled in either GCII, MB, OB, or OCII, so they represent students in a range of curricular stages. We found there to be a negligible difference across courses in engagement in CMR ($\chi^2 = 17.565$, $p = 0.007$, Cramer's V = 0.091 (small)), which is consistent with findings from Franovic et al. (2023). A one-way analysis of variance showed that cumulative GPA was higher for the females (3.538) than males (3.440) ($p = 0.002$)), and higher for White students (3.555) than Non-White students (3.389) ($p < 0.001$).

Table 5.3. Race/ethnicity identification based on gender identification and course. All post responses.

	Gender		Course				Total
Race/ethnicity	Male	Female	GCII	MB	OB	OCII	
American Indian/Alaskan Native (non-Hispanic)	0	2	0	2	0	0	2
Asian (non-Hispanic)	39	54	54	15	14	10	93
Black	17	48	29	10	13	13	65
Hawaiian	0	1	1	0	0	0	1
Hispanic	13	39	23	11	12	6	52
International	33	23	37	5	7	7	56
Not reported	6	15	13	5	1	2	21
Two or more races (non-Hispanic)	12	27	20	8	3	8	39
White (non-Hispanic)	258	476	356	141	109	128	734
Totals	378	685	533	197	159	174	1,063

Results And Discussion

RQ1 – How Do Males And Females Compare In Their Engagement In CMR About Protein-Ligand Binding?

To answer this research question, we show here the number and percent of female and male students that constructed Non-CM, Partially CM or Fully CM responses. In this data set,

378 students identified as male, and 685 students identified as female (total = 1,063 students). When comparing male and female engagement in CMR for the PL task, we found there was no significant difference ($\chi^2 = 3.942$, $p = 0.139$). Figure 5.2 shows the distribution of male and female Non-CM, Partially CM, and Fully CM responses. This suggests that the PL task is equitable, in terms of difficulty, for males and females – that is, neither group constructed more Fully, Partially or Non-CM responses than the other. This is in alignment with previous work suggesting that mechanistic reasoning assessment items result in more equitable outcomes for students.[3]

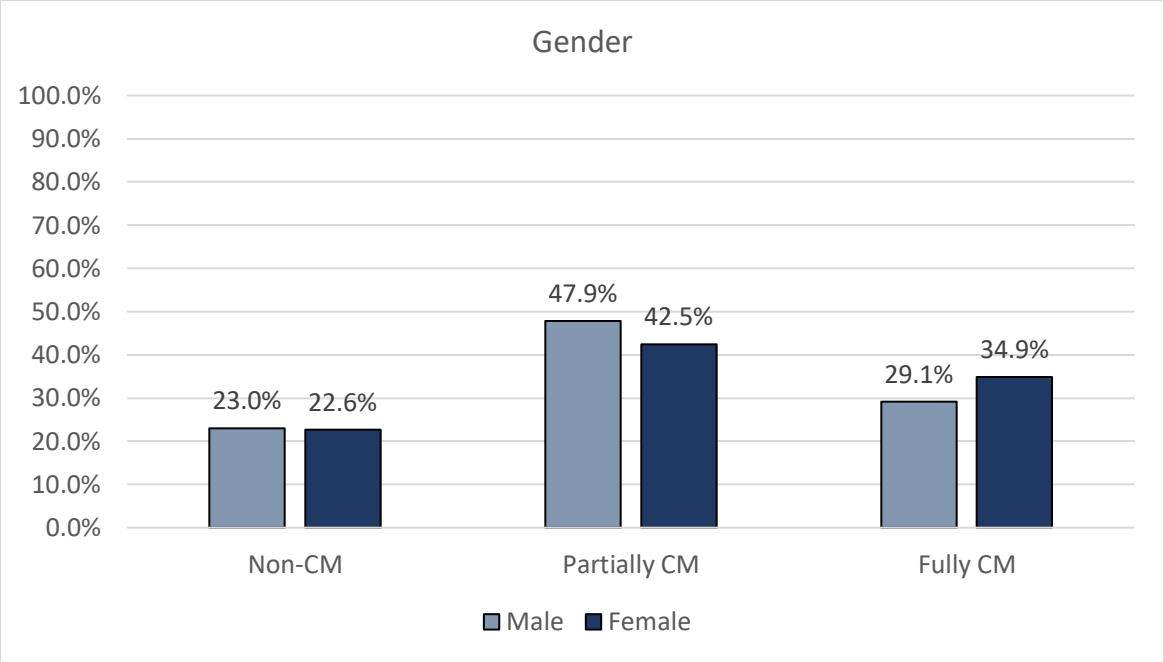


Figure 5.2. Male vs. Female engagement in CMR for the PL task.

RQ2 – How Do White Students Compare To Non-White Students In Their Engagement In CMR About Protein-Ligand Binding?

To answer RQ2, we compared the engagement in CMR between White students (N = 734) and Non-White students (N = 329) using a Pearson's χ^2 calculation, finding a negligible difference ($\chi^2 = 11.722$, $p = 0.003$, Cramér's V = 0.100 (small)), with a slightly greater percentage of White students constructing Fully CM responses. That is, while the statistical test was significant, the small effect size renders this difference negligible[20]. When running a pairwise comparison, comparing White students to Black, Asian, Hispanic, and students of Two Or More Races, there were no significant differences observed (Table 5.4). Figure 5.3 shows the distribution of Non-CM, Partially CM, and Fully CM for all race/ethnicity categories except American Indian and Hawaiian, as there were so few students (N = 3 and N= 1, respectively).

Table 5.4. Statistics.

Comparison groups	Chi-square	*p* value	Cramer's *V*
Non-White vs. White	11.722	0.003	0.100
Black vs. White	7.249	0.027	0.091
Asian vs. White	8.895	0.012	0.099
Hispanic vs. White	3.017	0.221	0.059
Two or more vs. White	1.93	0.381	0.048

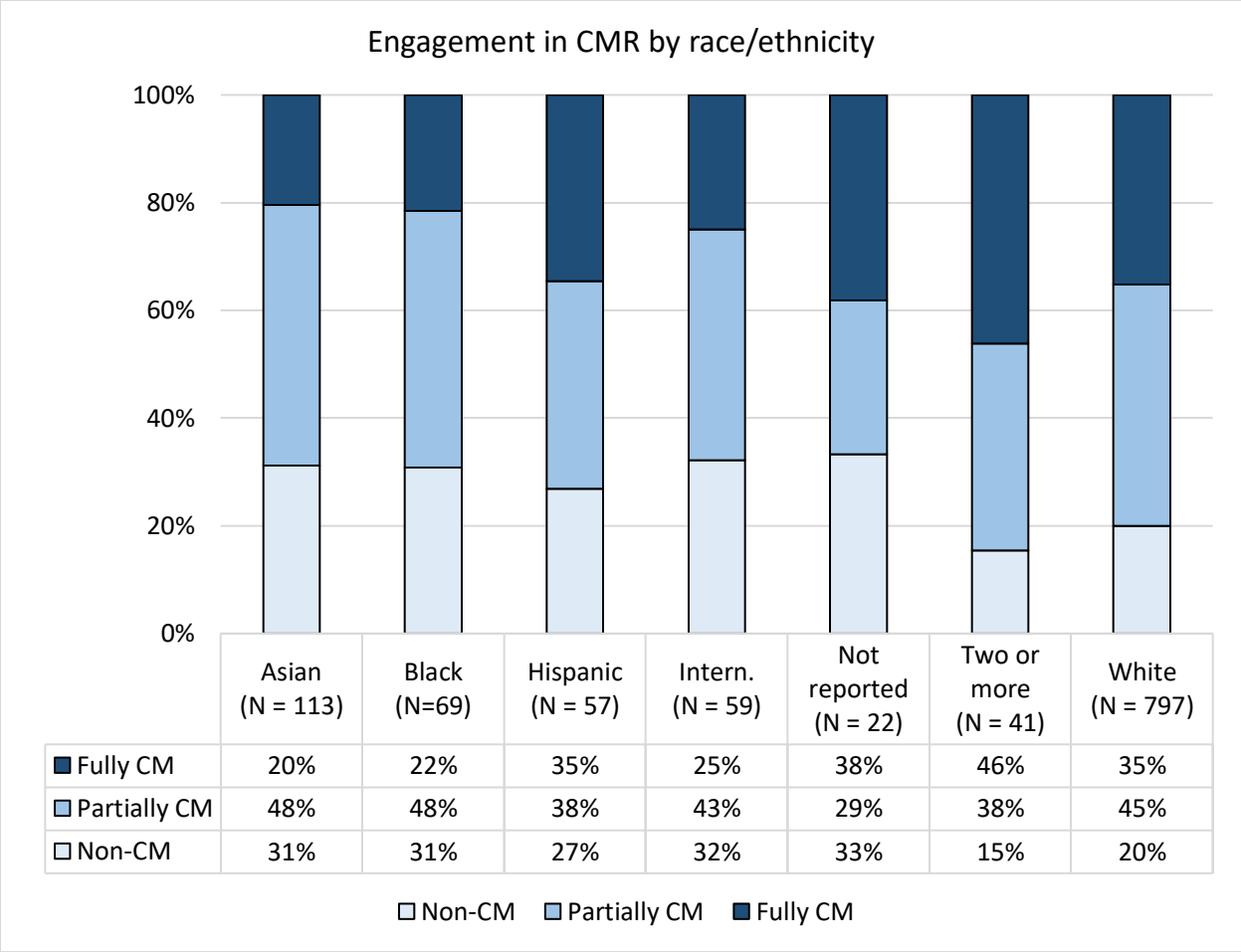


Figure 5.3. Distribution of Non-CM, Partially CM, and Fully CM responses across race/ethnicity groups.

We posit that these findings support the advancement of equity through use of MR tasks in courses that are either fully or partially transformed to emphasize this thinking strategy. Other courses at other institutions may show different results across gender and race/ethnicity.

RQ3 – To What Extent Do Cumulative GPA, (Binary) Gender, And Race/Ethnicity Predict Student Engagement In CMR?

For research questions 1 and 2, we were interested in providing a general picture comparing two demographic characteristics in the context of student engagement in CMR

about protein-ligand binding; however, to gain a more reliable understanding of the impact (if any) these characteristics have on engagement in CMR, we now answer research question 3. We calculated an ordinal regression (or proportional odds logistic regression) general linearized model to predict the outcome variable (engagement in CMR) based on three independent variables: cumulative GPA, binary gender identification, and race/ethnicity. Because of the small number of students in each Non-White category, we combined these races/ethnicities resulting in two categories (White and Non-White) for this variable. Cumulative GPA was treated as a continuous covariate variable.

The calculated model (with predictors) fits significantly better than the intercept-only model ($p < 0.001$), and all assumptions were met. Our results showed that both GPA and race/ethnicity are significant predictors of student engagement in CMR (code), while gender is not. GPA is the strongest predictor, as it had the highest estimate value (0.613). That is, students with a higher cumulative GPA have 0.613 higher odds of being in a higher category. On the other hand, there is a small negative association with race/ethnicity identification ($p = 0.014$). The estimate in the model for race is -0.308, which indicates that, compared to White students, being Non-White decreases the log odds of being in a higher category by 0.308. The estimate for race/ethnicity, however, was half as small as that for GPA, indicating that GPA is a stronger predictor of student engagement in CMR. Further, an ANOVA test comparing White and Non-White GPA showed that, for this group of students, White students had a significantly higher GPA than Non-White students. Thus, we posit that GPA is the most reliable predictor for determining student engagement in CMR. Gender identification had no effect (Table 5.5).

Table 5.5. Ordinal regression parameter estimates.

	Estimate	*p* value
GPA	0.613	$p < 0.001$
Gender	-0.102	$p = 0.397$
Race/ethnicity	-0.308	$p = 0.014$

Conclusions

In this study, we coded 800+ student responses to analyze alongside those that we coded and reported in Study 1. We conducted this study in order to advance equity efforts in education by disaggregating the results and reflecting on student outcomes. Mechanistic reasoning assessment items, when used extensively on exams that contribute to course grades, have been shown to result in more equitable outcomes for students (i.e., more students pass the course).[3] The PL task that we used was carefully designed to promote student use of productive ideas to construct mechanistic explanations; however, students in chemistry and biology courses engaged in CMR to varying extents. Some students did not include any key ideas to explain this phenomenon (Non-CM), while others included all key ideas, constructing Fully CM responses.

Using the coded responses, we grouped students based on their engagement in CMR and then compared whether there were differences between binary gender identification (male vs. female) or race/ethnicity (White vs. Non-White). While there was a small difference between White and Non-White student responses, there was no difference between male and female students. The most significant and reliable predictor of student engagement in CMR, according to an ordinal regression model, was cumulative GPA. Since this group of White students had a higher GPA than the Non-White students, it is likely that any differences seen

between White and Non-White responses are actually driven by the differences in GPA. This poses an additional challenge and implication, outside of the scope of this study. To address why Non-White students have lower cumulative GPAs than White students would involve larger systemic efforts. Another factor to consider, however, is how course grades are determined and the relative weighting of high-stakes exams (or summative assessments) and low-stakes assessments such as homework and recitation activities that encourage students to use and apply their knowledge without penalties based on correctness. We encourage instructors to consider using frequent, formative assessments that engage students in reasoning strategies such as mechanistic reasoning, providing credit based on a good-faith effort rather than correctness, as this likely promotes the free flow of ideas between peers and instructors, ideally advancing diversity, equity, and inclusion.

Limitations

The institution from which we collected responses has been engaged in the transformation of undergraduate chemistry and biology courses, with a focus on opportunities for students to use and apply their knowledge by engaging in scientific practices; thus, the student participants in this study may not be representative of students from other institutions.

These data were collected in Spring 2020, the semester in which the COVID-19 pandemic hit the United States. All courses moved to an online format in the middle of the semester. It is likely that many students had negative experiences during this time, which may have impacted their explanations and the data reported here.

REFERENCES

(1) Stowe, R. L.; Scharlott, L. J.; Ralph, V. R.; Becker, N. M.; Cooper, M. M. You Are What You Assess: The Case for Emphasizing Chemistry on Chemistry Assessments. *J. Chem. Educ.* **2021**, *98* (8), 2490–2495. https://doi.org/10.1021/acs.jchemed.1c00532.

(2) Matz, R. L.; Fata-Hartley, C. L.; Posey, L. A.; Laverty, J. T.; Underwood, S. M.; Carmel, J. H.; Herrington, D. G.; Stowe, R. L.; Caballero, M. D.; Ebert-May, D.; Cooper, M. M. Evaluating the Extent of a Large-Scale Transformation in Gateway Science Courses. *Science Advances* **2018**, *4* (10), eaau0554. https://doi.org/10.1126/sciadv.aau0554.

(3) Ralph, V. R.; Scharlott, L. J.; Schafer, A. G. L.; Deshaye, M. Y.; Becker, N. M.; Stowe, R. L. Advancing Equity in STEM: The Impact Assessment Design Has on Who Succeeds in Undergraduate Introductory Chemistry. *JACS Au* **2022**. https://doi.org/10.1021/jacsau.2c00221.

(4) Cooper, M. M. Why Ask Why? *J. Chem. Educ.* **2015**, *92* (8), 1273–1279. https://doi.org/10.1021/acs.jchemed.5b00203.

(5) Laverty, J. T.; Underwood, S. M.; Matz, R. L.; Posey, L. A.; Carmel, J. H.; Caballero, M. D.; Fata-Hartley, C. L.; Ebert-May, D.; Jardeleza, S. E.; Cooper, M. M. Characterizing College Science Assessments: The Three-Dimensional Learning Assessment Protocol. *PLOS ONE* **2016**, *11* (9). https://doi.org/10.1371/journal.pone.0162333.

(6) Bowen, R. S.; Flaherty, A. A.; Cooper, M. M. Investigating Student Perceptions of Transformational Intent and Classroom Culture in Organic Chemistry Courses. *Chem. Educ. Res. Pract.* **2022**, *23* (3), 560–581. https://doi.org/10.1039/D2RP00010E.

(7) Cooper, M.; Klymkowsky, M. Chemistry, Life, the Universe, and Everything: A New Approach to General Chemistry, and a Model for Curriculum Reform. *J. Chem. Educ.* **2013**, *90* (9), 1116–1122. https://doi.org/10.1021/ed300456y.

(8) Glennan, S. *The New Mechanical Philosophy*; Oxford University Press: Oxford, New York, 2017.

(9) Krist, C.; Schwarz, C. V.; Reiser, B. J. Identifying Essential Epistemic Heuristics for Guiding Mechanistic Reasoning in Science Learning. *Journal of the Learning Sciences* **2019**, *28* (2), 160–205. https://doi.org/10.1080/10508406.2018.1510404.

(10) Machamer, P.; Darden, L.; Craver, C. F. Thinking about Mechanisms. *Philosophy of Science* **2000**, *67* (1), 1–25. https://doi.org/10.1086/392759.

(11) Russ, R. S.; Scherr, R. E.; Hammer, D.; Mikeska, J. Recognizing Mechanistic Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed from Philosophy of Science. *Science Education* **2008**, *92* (3), 499–525. https://doi.org/10.1002/sce.20264.

(12) van Mil, M. H. W.; Boerwinkel, D. J.; Waarlo, A. J. Modelling Molecular Mechanisms: A Framework of Scientific Reasoning to Construct Molecular-Level Explanations for Cellular Behaviour. *Sci & Educ* **2013**, *22* (1), 93–118. https://doi.org/10.1007/s11191-011-9379-7.

(13) Mislevy, R. J.; Almond, R. G.; Lukas, J. F. A Brief Introduction to Evidence-Centered Design. *ETS Research Report Series* **2003**, *2003* (1), i–29. https://doi.org/10.1002/j.2333-8504.2003.tb01908.x.

(14) Noyes, K.; Carlson, C. G.; Stoltzfus, J. R.; Schwarz, C. V.; Long, T. M.; Cooper, M. M. A Deep Look into Designing a Task and Coding Scheme through the Lens of Causal Mechanistic Reasoning. *J. Chem. Educ.* **2022**, *99* (2), 874–885. https://doi.org/10.1021/acs.jchemed.1c00959.

(15) Franovic, C. G.-C.; Noyes, K.; Stoltzfus, J. R.; Schwarz, C. V.; Long, T. M.; Cooper, M. M. Undergraduate Chemistry and Biology Students' Use of Causal Mechanistic Reasoning to Explain and Predict Preferential Protein–Ligand Binding Activity. *J. Chem. Educ.* **2023**, *100* (5), 1716–1727. https://doi.org/10.1021/acs.jchemed.2c00737.

(16) Hammer, D. Student Resources for Learning Introductory Physics. *American Journal of Physics* **2000**, *68* (S1), S52–S59. https://doi.org/10.1119/1.19520.

(17) Bryfczynski, S. BeSocratic: An Intelligent Tutoring System for the Recognition, Evaluation, and Analysis of Free-Form Student Input. *All Dissertations* **2012**.

(18) Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* **1960**, *20* (1), 37–46. https://doi.org/10.1177/001316446002000104.

(19) IBM SPSS Statistics, 2020.

(20) Green, S. B.; Salkind, N. J. Two-Way Contingency Table Analysis Using Crosstabs. In *Using SPSS for Windows and Macintosh: Analyzing and Understanding Data*; Pearson Education Inc.: Upper Saddle River, NJ, 2011; pp 366–376.

Chapter VI – How Do Instructors Explain The Mechanism By Which ATP Drives Unfavorable Processes?

Preface

In this study, we investigate how chemistry, biology, and biochemistry instructors explain and teach the mechanism by which ATP drives unfavorable processes. This research has been previously published in the CBE Life Sciences Education Journal and is reprinted with permission from Franovic, C. G.-C.; Williams, N. R.; Noyes, K.; Klymkowsky, M. W.; Cooper, M. M. How Do Instructors Explain The Mechanism by Which ATP Drives Unfavorable Processes? *CBE Life Sci Educ* 2023, 22. https://doi.org/DOI:10.1187/cbe.23-05-0071.

A copy of the licensing agreement (with permissions) is included in Appendix A.

Supporting Information for this manuscript is included in the Appendix.

Introduction And Background

A History Of Educational Research Involving ATP

It could be argued that no scientific topic has been more widely used and less well understood than energy. Although it is difficult to define[1], energy plays a critical role within individual disciplines and as a crosscutting concept that transcends disciplines[2,3]. Given its central position, one might imagine that the concept of energy could provide a unifying framework through which to support connections between scientific disciplines. However, most disciplines have adopted siloed approaches to thinking about energy, contributing to the unfortunate fact that energy is a notoriously difficult topic for both students and instructors to grasp in a useful and productive way[2,4–6]. These difficulties are exacerbated when studying molecular-level phenomena, which involve unfamiliar symbols and is often based on non-

intuitive ideas, for example, "hydrophobic interactions" are attractive and not repulsive[2]. One of the most well-documented, problematic ideas about energy is the belief that breaking chemical bonds or non-covalent interactions releases "energy" into the surroundings when, in fact, the opposite is true. For nearly 50 years, researchers have documented the widespread presence of this misconception and tried to develop instructional approaches that support a more accurate understanding of energy[5,7–15]. According to this work, one molecule appears to contribute significantly to the issue: adenosine triphosphate or ATP, a key molecule in biological systems. ATP plays a central role in cellular metabolism. It is frequently described as the "energy currency" of the cell, storing energy so it may later be used to allow thermodynamically unfavorable processes to proceed[16–18]. In this paper, we refer to ATP as an energy "carrier" or "source" interchangeably. We also frequently discuss driving unfavorable *processes* as a more general term that includes driving unfavorable *reactions* (a more specific process that can be discussed at the molecular/chemical level).

The role of ATP as the link between energy storage and use in the cell was first proposed by Fritz Lipmann[19,20], which contributed to his Nobel Prize in 1953. Lipmann discussed the "energy-rich" bond between phosphate groups and noted that the hydrolysis of these groups could release a large amount of heat. Unintentionally, this led to the logical, but incorrect, simplification that it is the breaking of this energy-rich bond in ATP which releases energy. This idea directly contradicts the fact that energy is required to overcome (to break) the stabilizing attractive interactions of covalent bonds and non-covalent interactions.

Thus, in chemistry and biology courses, we may be offering contradictory descriptions of the same phenomena—a situation that can result in didaskalogenic (that is, instruction-

induced) misunderstandings that prevent students from developing a coherent and productive understanding of energy. Consider this quote regarding bonding and energy from Kohn et al., in which an undergraduate student concurrently enrolled in chemistry and biology courses said: "*I know for biology what [the instructor] wants us to say and then for chemistry what we have to say*"[21]. With all the confusion surrounding energy, bonds, and ATP, it is worth considering whether there is a more useful way to talk about ATP. More specifically, how can we talk about ATP in a way that is both productive in biology and in alignment with student understanding of chemical principles?

In the 1970's, Novick (1976) published one of the first studies documenting the presence of the "misconception" about bonds and energy among university students[13]. He found that when asked to provide a molecular interpretation of how "fats supply energy to the body", the majority of students suggested that energy was stored *within* chemical bonds that could be released when the bond was broken. Specifically, these students often referenced the bonds within ATP. For example, one student explained that "*the breakdown of fats produces a source of energy in the form of molecules (ATP), having special bonds which, when broken, release energy*." Since then, other studies have expanded upon this work, providing additional evidence for confusion among students regarding the energy associated with the functional roles of ATP in biological systems[9,11,12,14].

More recently, researchers have moved beyond documenting student misunderstandings about bond energies and ATP, to proposing instructional solutions to these issues[8,10,15]. The common thread across these efforts has been to explain how and why energy is released during ATP hydrolysis (that is, the reaction of ATP with water to form ADP and

inorganic phosphate), by leveraging bond energies more appropriately. Specifically, the focus of these efforts lies on how instructors can provide consistency in the way bond energies are discussed across disciplines or, at the very least, help students avoid developing misconceptions about how and why this reaction releases energy. The results show that, with a clear focus on these ideas, students can develop a more canonical understanding of why ATP hydrolysis is accompanied by release of energy (i.e., more energy is released upon bond formation, due to electrostatic attractive forces, than is required for bond breaking). However, focusing on the energy release of ATP hydrolysis may generate its own unproductive ideas, because in cellular systems, a negligible amount of energy is actually "released" to the surroundings. Rather, these systems have evolved such that energy is transferred (or used) both within and among systems and their components. In fact, many ATP-dependent processes involve phosphorylation of a molecule or protein and, thus, do not involve an ATP hydrolysis step, e.g., glutamine synthesis[22,23] or Ca^{2+} protein pumps[24–27]. For those reactions in which ATP hydrolysis is involved, for example the F0-ATPase proton pump, the hydrolysis step does not provide the major driving force for the process[28–33].

Perhaps even more importantly, if we invoke ATP hydrolysis by emphasizing energy release, this approach provides no mechanism for how the energy is transferred among components in biological systems. If ATP hydrolysis occurred in isolation, it would lead to an increase in the overall temperature of the system—an inefficient and non-specific way to speed up favorable reactions and to drive unfavorable processes (to say nothing of the effects on the organism). We are not the first to note this disconnect. As early as 1970, researchers outside of the educational field noted that "*the original concept put forward by Lipmann was ill-founded*

and that its effect is to divert attention from the genuine problem of the mechanism of events in which ATP takes part"[34]. Carusi (1992) noted that most textbooks provide no explanation for how the energy released from "high-energy" bonds is coupled to the unfavorable reactions they drive, lamenting that without this mechanism "*the student is left in the dark waiting to learn how the free energy available from these compounds actually is transferred and utilized*"[35]. While this paper is over 30 years old, and the mechanisms by which ATP drives these reactions has perhaps made it into some textbooks, current research shows the way ATP is discussed and represented (most often as the hydrolysis of ATP) in commonly-used textbooks is ripe with issues[36], emphasizing the need for additional work investigating instructional practices and materials associated with ATP.

Our work aims to address this educational gap by exploring common themes in the *mechanisms* by which ATP drives unfavorable processes. Thus, we leverage mechanistic reasoning, which can be used as not only an explanatory and predictive thinking strategy, but also an instructional emphasis intended to provide students with a productive understanding of the roles of ATP in biological systems.

Mechanistic Reasoning

There is emerging consensus in science education literature that mechanistic reasoning – that is, how and why phenomena occur – should be practiced by students and implemented in course design and instruction[3,37–39]. Not only does mechanistic reasoning provide students with ways to explain and predict phenomena, but also, as recently shown, assessments incorporating mechanistic reasoning are more equitable than more traditional test items that focus on procedural/mathematical skills or rote knowledge[40]. Thus, we argue that explicitly

discussing the mechanisms by which biological systems act would be useful for all students, and specifically introductory biology students[41,42]. Many molecular biologists and biochemists spend their time searching for and/or predicting cellular mechanisms, thus providing students the opportunity to engage in this practice in their courses is a useful endeavor. For these reasons, we posit that research intended to help students develop a mechanistic understanding of how ATP functions in cellular systems, is a valuable area of investigation.

In earlier publications[43,44], we used the mechanistic reasoning framework proposed by Russ et al. and elaborated on by Krist et al.[45,46]. According to Krist et al., mechanistic reasoning is an epistemic heuristic, or thinking strategy, that can be leveraged across all science disciplines and, when used appropriately, results in powerful (testable) predictions. Russ, Krist, and others[42,47,48] broadly define mechanistic reasoning as reasoning about how and why phenomena occur by identifying relevant (often lower scalar level) entities, the properties/activities of those entities, and how behaviors or interactions of the entities link together and give rise to the phenomenon under consideration. However, while this approach is successful for relatively simple phenomena[43,49,50], such as phase changes or chemical reactions, no published work to our knowledge has investigated its effectiveness for more complex biological phenomena. These complex phenomena often span several scalar levels and require ideas from more than one discipline, making it more difficult to identify the level of depth at which an explanation is appropriate[51]. Further, recent work in our group has shown that students struggle to reason mechanistically about complex phenomena, such as that of protein-ligand binding at both the biochemical and population levels[52].

One of the challenges learners encounter when crafting an explanation is determining the necessary and sufficient depth for their explanation. For example, if we consider the role of ATP in driving unfavorable processes, there are a number of valid approaches that can be used: (1) we might think macroscopically about the overall reaction and how the hydrolysis of ATP releases a larger amount of energy than that required to drive the unfavorable reaction – an approach that typically involves thermodynamic calculations using Gibbs free energy; (2) we might approach the problem at the molecular level, invoking the idea of ATP as a phosphorylating agent, leading to a more reactive common intermediate, which then reacts to produce the product; (3) we might use a systemic approach to explain how the cell generates and maintains a relatively high concentration of ATP, thereby increasing the rate of collisions between ATP and the reactant, pushing a sequence of reactions forward; (4) we might consider the sub-atomic level, i.e., the electron distributions of the participating entities and how a nucleophilic substitution reaction with ATP and a substrate occurs as a result of electrostatic forces and interactions (an approach that may be more suitable in organic chemistry and more advance biochemistry courses); or (5) we might consider reactions in which the three-dimensional structure of an enzyme is altered by the binding of ATP as opposed to the situation when adenosine diphosphate is bound. Deciding the appropriate level and type of explanatory depth can depend on several factors including, for example, the content knowledge available to the explainer, that of the student, the discipline and course level in which the question is posed, or simply the question being asked.

In undergraduate science courses, the weight of identifying what to explain mechanistically (i.e., deciding sufficient depth) and in turn, what to expect of students typically

falls on the instructors. How then do these instructors think about the mechanisms by which ATP drives unfavorable processes? What mechanisms do they emphasize during instruction when explaining how ATP "works"? These were the questions we were interested in exploring in this study.

Purpose And Research Questions

We interviewed instructors who had taught undergraduate courses in biology, chemistry, and/or biochemistry with the aim of gaining insight into instructor and disciplinary-expert understandings and explanations about energy, in particular how ATP provides the energy needed to drive unfavorable processes (i.e., how does ATP "work"). On its own, the statement "ATP drives unfavorable processes" is an explanatory black box, or a unit within an explanation that remains unexplained[51], which might be unpacked using different approaches (such as explaining the mechanism of energy release or energy transfer) during instruction. Black boxes have historically carried a negative connotation; however, in this work, we recognize the necessity and utility of explanatory black boxes, particularly for complex phenomena which consist of many "mechanisms within mechanisms", to explain all of which would be tedious and unproductive[51]. The findings we discuss here can support the development of alternative curricular materials (such as formative assessments) that better support students' mechanistic understanding of the role of ATP in driving unfavorable processes, as well as their understanding of energy.

While our aim was to investigate how instructors explain and think about the mechanism(s) by which ATP drives unfavorable processes, most participants also revealed personal experiences, providing affective insights about teaching ATP in their course(s). Thus,

we address three different research questions, the first two capturing our initial goal and the third as an artifact (bonus) that emerged from these rich discussions:

1. What ideas did instructors leverage when discussing ATP?
2. In what ways do instructors use these ideas to discuss how ATP drives unfavorable processes?
3. What teaching experiences did molecular biology instructors share regarding ATP in their course(s)?

In this analysis, we do not intend to portray any instructors as superior or inferior based on how they unpacked this phenomenon during their interviews. Rather, we share several excerpts which sparked deep discussions among us and potential avenues for future research in biology education.

Methods

Participants

Between 2018 and 2020, we conducted semi-structured interviews with 15 instructors of undergraduate courses in biology, chemistry, and/or biochemistry. We used a convenience sampling approach to select these instructors, soliciting participation from instructors with whom we were familiar. All participants consented to the interview in accordance with our IRB protocol. The instructors were not compensated for participating in these interviews. To protect the identity of each participant, we use pseudonyms and gender-neutral pronouns (they/them) throughout this manuscript. Table 6.1 provides the pseudonym, primary course taught, primary research (if applicable), interviewer, and the interview medium. All instructors

were from one large, public, research-intensive midwestern university, except Dracaena, who was from a public, mid-sized (~25,000 students), midwestern university.

Table 6.1. Participant information.

Pseudonym	Primary Course	Primary research	Interviewer	Interview medium
Eucalyptus[b]	Introductory Molecular Biology (MB)	Physiology; Education	KN	In-person
Ficus[b]	Introductory MB	Education	KN	In-person
Jade[b]	Introductory MB	Education	KN	In-person
Lily[b]	Introductory MB	n/a	CF	Zoom
Monstera[b]	Introductory MB	Evolutionary biology	KN	In-person
Myrtle[b]	Introductory Organismal Biology	Education	KN	In-person
Olive[b]	Introductory MB	Education	KN	In-person
Dracaena[bc]	Biochemistry	Computational biochemistry	CF	Zoom
Ivy[bc]	Biochemistry	Education	CF	Zoom
Philodendron[bc]	Advanced Biochemistry	Structural biology	KN	In-person
Basil[c]	General Chemistry	Education	KN	In-person
Fig[c]	General Chemistry	n/a	KN	In-person
Ginkgo[c]	Physical Chemistry	Physical chemistry	CF	Zoom
Orchid[c]	General Chemistry	Education	KN	In-person
Pothos[c]	Honors General Chemistry	Theoretical physical chemistry	KN	In-person

Instructors listed according to discipline (green[b] = biology, blue[bc] = biochemistry, grey[c] = chemistry)

Data Collection

Authors CF or KN conducted semi-structured interviews, which allowed for the inclusion of our interests as well as authentic ideas from participants, either in-person or via the online meeting software Zoom (Table 6.1). In these interviews, we explored how the instructors think about sources of energy in their course/discipline, and the role(s) of ATP in such systems. To start each interview, we asked them about the courses they teach, and to describe how they think about energy in chemical and biological systems. We began with these general questions to ease into the more complex discussions that emerged when we asked how they think about the mechanism(s) by which ATP is used as an energy source (Figure 6.1 shows this general interview protocol, and Appendix B provides the entire protocol).

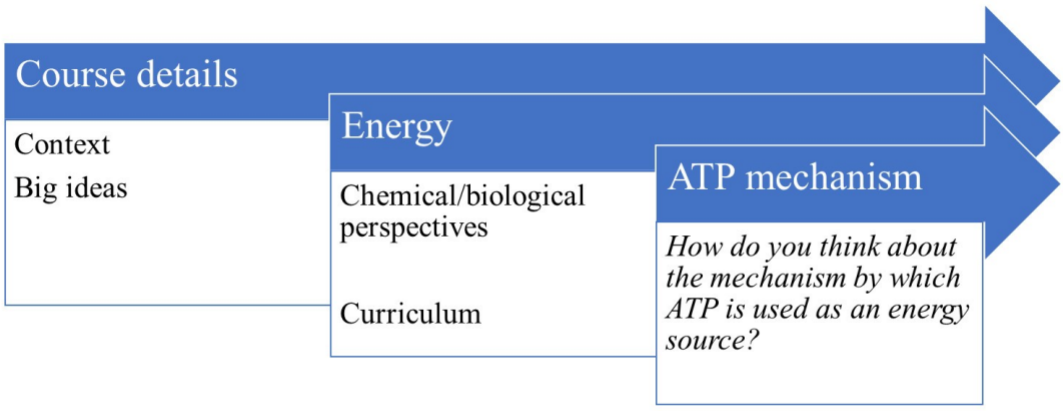


Figure 6.1. Outline of the semi-structured interview protocol.

In this study, we focused on their explanations of how ATP "works". Typically, this occurred in response to the question: *how do you think about the mechanism by which ATP is used as an energy source?* In addition to their thoughts about the mechanism, we asked if/how

they teach these ideas in their course(s). Some participants also discussed ATP and its role prior to our interview questions (unprompted).

We recorded the audio of each interview and transcribed them using either Express Scribe, software that facilitates human transcription of audio, or Otter, an automated transcription software. A second researcher read through each transcript while listening to the audio recordings to edit the transcript in order to confirm that it accurately captured the dialogue from each interview.

Positionality Statement

We intend to be as transparent as possible in presenting this work by stating our positions, which impacted both the data collection, analysis, and interpretation. Authors CF and KN were graduate students when conducting all interviews, placing them as "subordinate" to the faculty participants; however, as interviewers, we recognize our position of power in knowing the questions and topics to be discussed. To mitigate the effects of this position, we spent the first part of each interview establishing trust with the participants and engaging in casual conversation prior to recording. Second, all authors believe in the importance of using a mechanistic reasoning approach to teach and learn about scientific phenomena – it has been a driving force in various course redesign projects initiated by authors MWK and MMC[53–55]. Thus, we view mechanistic reasoning as a more productive way of thinking about science when compared to rote memorization or knowledge as a set of facts. In leading up to this study, we consulted the literature and thought deeply about the mechanisms of ATP in cellular systems. Thus, we had pre-existing ideas about what the participants might discuss when we asked them to explain the mechanism by which ATP drives unfavorable processes and the degree to which

they align with previously studied biological mechanisms. Specifically, the idea that ATP hydrolysis (i.e., the energy release mechanism) does not address the mechanism by which energy is transferred to the system to drive the unfavorable processes. These pre-existing ideas consciously and subconsciously shaped our analysis and, therefore, should be acknowledged. We did, however, share this analysis with scholars outside of our group for input and/or other interpretations. Lastly, our disciplinary identities include a range of chemistry and biology backgrounds.

Data Analysis

Research Question 1

To identify the different ideas instructors leveraged when discussing ATP, we analyzed the full interviews using MAXQDA 2020[56], a qualitative data analysis software. We began our analysis by first familiarizing ourselves with the data – that is, authors CF, KN, and MMC read through each interview multiple times, recording notes and questions, and discussed our findings[57]. These discussions informed the open coding conducted by authors CF and NW, in which they open-coded the interviews in sets of three, compared generated codes, and began coding scheme development. At this point in coding, we decided to narrow the scope of our analysis to focus solely on participant discussion of ATP; however, we continued to read through the entire interview of each participant to capture any unprompted discussions about ATP.

We continued refining the coding scheme using a constant comparative[58] approach until we felt that the codes represented the range of ideas expressed across all interviews (i.e., no additional codes were generated). The final coding scheme consisted of 19 codes (Appendix C,

Table 6.4). The length of coded segments varied based on what we deemed sufficient to capture the essence of how the instructor used that code. That is, we did not simply count the frequency of the term "bond energy"; rather, we read the transcripts in detail to identify sections in which the participants leveraged these ideas (and how). The majority of the codes are disciplinary content-related, as our protocol mainly focused on eliciting instructor explanations for how ATP functions (from both a chemical and biological perspective). However, some codes related to other aspects of the instructors' experiences (rather than content). For example, the code *teaching/learning* captured instructor comments related to an experience regarding ATP in the classroom. We provide descriptions and examples of some of the most frequently occurring codes in Table 6.2 (descriptions and examples of all other codes are included in Appendix C). We discuss the importance of our codes in the results section of RQ1.

Using this scheme, two coders (CF and NW) independently coded each interview. Our initial agreement was 85.67%; however, the two authors met and discussed any disagreements until we reached 100% consensus. The codes reported in this manuscript are the consensus codes. The final codes were non-mutually exclusive and could appear any number of times in each interview.

Research Questions 2 And 3

To address RQ2 and RQ3, we explored the participants' discussions of how ATP drives unfavorable processes (RQ2) and the experiences that molecular biology instructors shared when reflecting on their teaching of ATP (RQ3). The coding scheme we developed to answer RQ1 captured all ideas related to ATP that emerged throughout the entirety of each interview.

We used a smaller subset of these codes to define the larger themes which more directly answer RQ2 and RQ3. For RQ2, we used codes that specifically indicated instructor explanations of (1) the mechanism of energy release (*ATP hydrolysis* and *bond energy*) and/or (2) the mechanism of energy transfer (*phosphate transfer* and *common intermediate*) as a guide in our selection of excerpts (of varying length) highlighting these two broader themes. To answer RQ3, we used the *teaching/learning* code to identify the experiences instructors shared regarding their teaching practices about how ATP provides energy in a coupled reaction. Eliciting these experiences was not an intentional aim of our protocol; however, we would be remiss to exclude these experiences when they emerged through such little prompting. It is likely that these experiences inform instruction (and, subsequently, student learning), making this an important and meaningful research question. Due to the dependence of RQ2 and RQ3 on the codes discussed in RQ1 results, we have included additional details about the data analysis of RQ2 and RQ3 in the corresponding results sections.

Results

Research Question 1 – What Ideas Did Instructors Leverage When Talking About ATP?

We coded all the ideas that participants mentioned when discussing ATP in each interview, resulting in 19 total codes (italicized in the body text) (Table 6.2 and Appendix B). Our coding scheme captured the wide range of ideas mentioned when considering ATP; however, no instructor discussed every code, and most instructors only leveraged a smaller selection of these ideas. Additionally, given the nature of our coding scheme, some participants received the same code multiple times. Table 6.3 shows the presence and frequency of codes identified in each interview.

Table 6.2. Frequently occurring codes, descriptions of the codes, and example excerpts.

Code	Description	Example
Bond energy	Discussing energy changes associated with bond formation or bond breaking.	*"So those weak bonds are getting broken, and then stronger bonds are getting formed, and that the forming of those bonds is releasing energy" (*Olive*)*
Teaching/learning	When the participant discusses whether or not they understand a topic themselves or how to teach that topic; or when they discuss how students understand or learn ideas related to ATP and reaction coupling.	*"I've done it both ways. I've tried to show them the actual mechanism, and I feel like that's a little bit too much for them in second semester" (*Fig*)*
ATP hydrolysis	Discussing the hydrolysis of ATP, or how water should be included in an explanation of how ATP works.	*"we do go through a fairly detailed description of why hydrolysis of ATP is favorable" (*Ivy*)*
Coupled reactions	Discussing reaction coupling, vaguely or explicitly.	*"I talk about it in pairing the hydrolysis of ATP with unfavorable reactions." (*Jade*)*
Referencing disciplines	When the instructor explicitly discusses their own discipline or another discipline and how those disciplines teach/think about things.	*"what I want the students to do and the way that I talk about sort of the biology of it, is to highlight that there is a discrepancy between the way that biologists will often talk about it and the way chemists talk about it." (*Basil*)*
ATP Synthesis	Discussing the synthesis or formation of ATP	*"If you talk about ATP synthesis, right, you're using that protons all on the…the, inter membrane space coming back into the matrix." (*Dracaena*)*
Common intermediate	Discussing a "high-energy" or "phosphorylated" intermediate involved in reaction coupling.	*"instead of having the phosphate just go to free phosphate, let's donate it to another molecule and have it become a phosphorylated intermediate." (*Monstera*)*
Phosphate transfer	Discussing the transfer of a phosphoryl group from one entity to another (typically from ATP to some other molecule).	*"I think most of the time, we mean, it is phosphorylating the substrate, and activating it based on a phosphoryl transfer." (*Ivy*)*

While the majority of the codes relate to disciplinary concepts (e.g., *bond energy*, *gradients*), two of the most frequently occurring codes do not: *teaching/learning* and *referencing disciplines.* Rather, these codes address experiences regarding how ATP is taught or understood in the classroom (*teaching/learning*) and how their own or different disciplines might discuss ATP (*referencing disciplines*). While our interview protocol prompted instructors to consider their teaching practices, the prominence of these codes highlights the range of personal experiences instructors felt inclined to share about how ATP is discussed in their courses and the courses of other disciplines. This was particularly common among biologists reflecting on how their teaching aligns or differs with the concepts taught in chemistry. We discuss these experiences in more depth in RQ3.

Table 6.3. Frequency of codes in each instructor interview.

	Biology Instructors[b]							Biochemistry Instructors[bc]			Chemistry Instructors[c]						
	Eucalyptus	Ficus	Jade	Lily	Monstera	Myrtle	Olive	Dracaena	Ivy	Philodendron	Basil	Fig	Ginkgo	Orchid	Pothos	Number of instructors	Total occurrence
Bond energy	1	5	0	2	5	2	6	1	3	0	10	2	2	1	3	13	43
Teaching/learning	4	4	5	2	4	7	9	0	9	0	9	6	0	1	1	12	61
ATP Hydrolysis	0	5	1	8	4	0	1	2	6	3	4	4	0	0	2	11	40
Coupled reactions	0	1	2	9	4	0	1	8	4	1	0	1	6	0	1	11	38
Referencing disciplines	3	2	1	1	4	1	2	2	0	0	9	4	1	0	0	11	30
ATP Synthesis	2	1	0	2	7	0	3	2	2	0	0	0	1	0	3	9	23
Stability	0	1	1	1	0	0	1	0	1	1	2	0	1	3	0	9	12
Gradients	2	0	0	2	4	0	1	2	3	0	0	0	1	0	2	8	17
Phosphate transfer	1	0	0	1	2	0	1	8	4	2	0	0	0	0	0	7	19
Enzyme	1	0	0	1	0	0	2	8	3	1	0	0	1	0	0	7	17
Common intermediate	0	0	1	6	2	0	0	0	7	1	0	0	6	0	0	6	23
Explicit exclusion of a mechanism	0	2	3	0	4	0	0	0	2	0	0	0	0	1	0	5	12
Conformational change	1	0	0	5	0	0	0	3	1	1	0	0	0	0	0	5	11
Biological example	0	0	0	3	0	1	5	1	0	1	0	0	0	0	0	5	11
Equilibrium	0	2	0	7	0	0	0	3	0	0	0	0	5	0	0	4	17
Activation energy	0	0	0	1	0	0	0	1	1	0	0	0	0	0	1	4	4
Calculations	0	3	0	3	0	0	0	0	5	0	0	0	0	0	0	3	11
Phosphate repulsion	0	0	0	1	0	0	2	0	2	0	0	0	0	0	0	3	5
Suggested resolution	0	0	0	0	3	0	0	0	0	0	2	0	0	0	0	2	5

The codes are listed in order of occurrence, with codes that appeared in the most interviews at the top, and those that appeared in the fewest interviews at the bottom. The instructors are ordered based on their discipline (green[b] = biology, blue[bc] = biochemistry, grey[c] = chemistry). Darker shading represents a higher frequency of occurrence.

Among the codes related to disciplinary concepts, participants discussed ideas like the production of ATP (*ATP synthesis* and *gradients*) and the relevance of *equilibrium*. However, some of the most popular codes correspond to two important mechanisms associated with ATP as an energy source (the primary interest of this study): (1) how energy is released and (2) how energy is transferred. Codes related to energy release, *bond energy* and *ATP hydrolysis,* were

used extensively across disciplines, with all instructors discussing one or both of these ideas at some point in their interviews. On the other hand, the codes *phosphate transfer* (N = 7) and *common intermediate* (N = 6), related to the energy transfer mechanism, were present in many, but not all, of the interviews. Another frequently occurring code, *coupled reactions*, overlapped often with explanations for both mechanisms; this code captured more general statements about reaction coupling and did not necessarily contribute to explaining how or why either mechanism occurs.

Interestingly, there were some notable differences in the degree to which participants leveraged ideas related to energy transfer and energy release. For example, Eucalyptus[b] discussed *phosphate transfer* during their interview but not *ATP hydrolysis*; Ficus[b] talked about *ATP hydrolysis* but not *common intermediate*; Lily[b] discussed *ATP hydrolysis* and *common intermediate* (each with high frequencies). We were interested in the instructors' use of these codes (and therefore their adherence to mechanisms of energy transfer and/or energy release) in their explanations for how ATP drives unfavorable processes, which we address in RQ2.

Research Question 2 – In What Ways Do Instructors Use These Ideas To Discuss How ATP Drives Unfavorable Processes?

To explore how instructors think about the mechanism by which ATP drives unfavorable processes, we used our coding from RQ1 to analyze the interview sections which reflected the ideas instructors leveraged when "unpacking" (or opening up) the black box of ATP "providing" energy. Based on the results of RQ1, we found that the two most popular mechanisms to unpack were (1) energy release and (2) energy transfer. One instructor, Orchid[c], did not explain either mechanism in enough detail for us to confidently place them in one of the themes (even

though they did receive one instance of the code *bond energy*); they were removed from the analysis for RQ2.

The instructors we interviewed discussed (1) how energy is released by explaining how ATP is hydrolyzed (*ATP hydrolysis*) and that the bonds formed are stronger than the bonds broken (*bond energy*), and/or (2) how ATP transfers energy via the transfer of a phosphoryl group to some other entity (based on the presence of the codes *phosphate transfer* and/or *common intermediate).* While some instructors only explained one mechanism, several explained both at some point in the interview, which we discuss by providing example excerpts below. In the excerpts, "[Editorializing]" indicates rephrasing (of the participant or interviewer).

Energy Release

Excluding Orchid[c], all 14 instructors explained the mechanism of energy release at some point in their interviews. These explanations included ideas about *ATP hydrolysis* and/or *bond energy*, neither of which were included in our interview protocol (i.e., all the instructors brought up these ideas without us mentioning either of them explicitly), with a focus on how ATP hydrolysis releases energy and/or the idea that the bonds in ATP (and water) are less stable than those in ADP and inorganic phosphate. For example, Basil[c] said, "*when they often talk about it in biology, […] hey we break the […] ATP bond and energy comes out. And they completely ignore sort of the fact that there are a whole bunch of new products that get formed and that water is involved…*" (codes: *referencing disciplines* and *ATP hydrolysis*), suggesting that water should be included when explaining the mechanism by which ATP drives unfavorable processes. Using the ideas of *ATP hydrolysis* and *bond energy* also encompass the ways in which recent publications have suggested teaching about ATP as a source of energy[8,10,15]. Figure 6.2

provides excerpts of three instructor explanations. In the interest of space, we condensed quotes by using ellipses in brackets ([…]), which indicate repetitive dialogue or filler words/sentences that did not contribute significantly to an understanding of the excerpt.

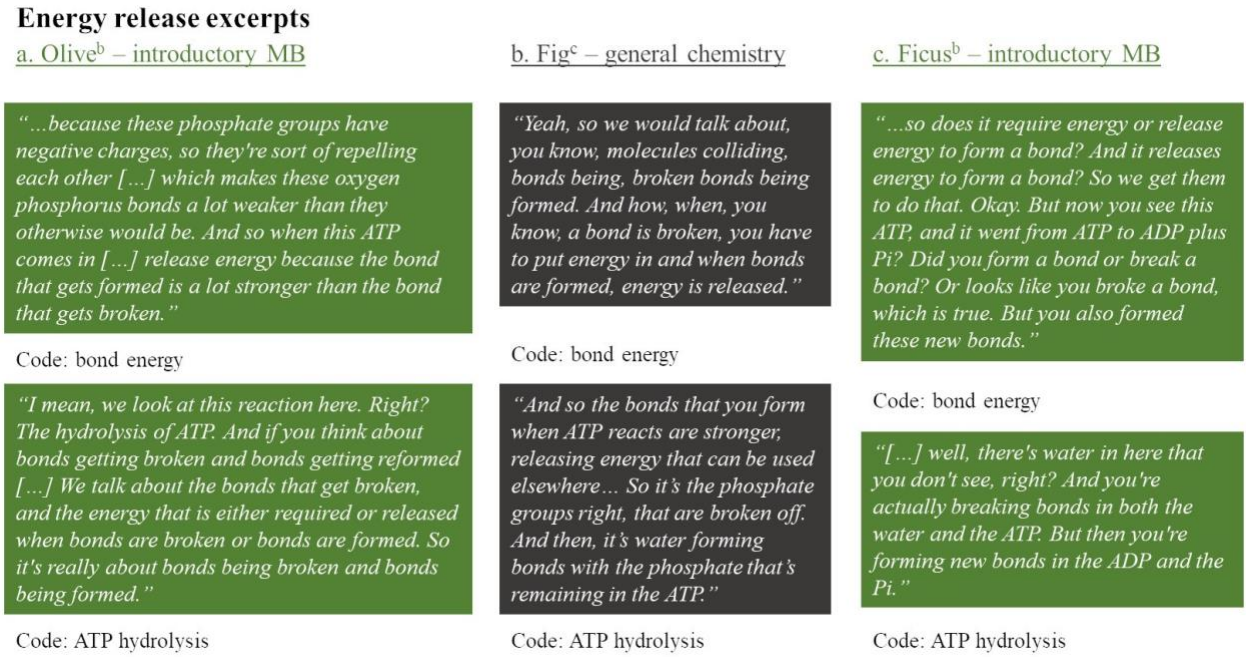


Figure 6.2. Example excerpts of explanations for energy release. [a, b, and c are color coded according to the discipline of each instructor. Green[b] = biology, blue[bc] = biochemistry, gray[c] = chemistry].

Energy Transfer

While all participants highlighted the mechanism of energy release by discussing *ATP hydrolysis* and/or *bond energy* at some point, only nine participants explained a mechanism of energy transfer, which we identified as those who included the ideas (codes) *phosphate transfer* and/or *common intermediate.* Figure 6.3 provides excerpts and codes from three participants (one from each discipline) who explained energy transfer using one or both of these ideas. For example, Ivy[bc] (Figure 6.3a) explained the mechanism by which ATP transfers

energy, by leveraging the idea of ATP as "*the universal phosphoryl group donor*" (code: *phosphate transfer*) and its role in forming a phosphorylated intermediate (code: *common intermediate*). Unlike Ivy[bc], Ginkgo[c] (physical chemistry) did not identify the role of ATP as a phosphorylating agent, but they emphasized the importance of energy being transferred mechanically or "*like a bicycle gear*", forming a common, high-energy intermediate to drive unfavorable reactions (code: *common intermediate*) (Figure 6.3b). Finally, Lily[b], an introductory MB instructor, discussed both *phosphate transfer* and *common intermediate*. However, Lily[b] also emphasized the importance of equilibrium and the concentrations of reaction sequence components in driving reactions forward (code: *equilibrium*) (Figure 6.3c).

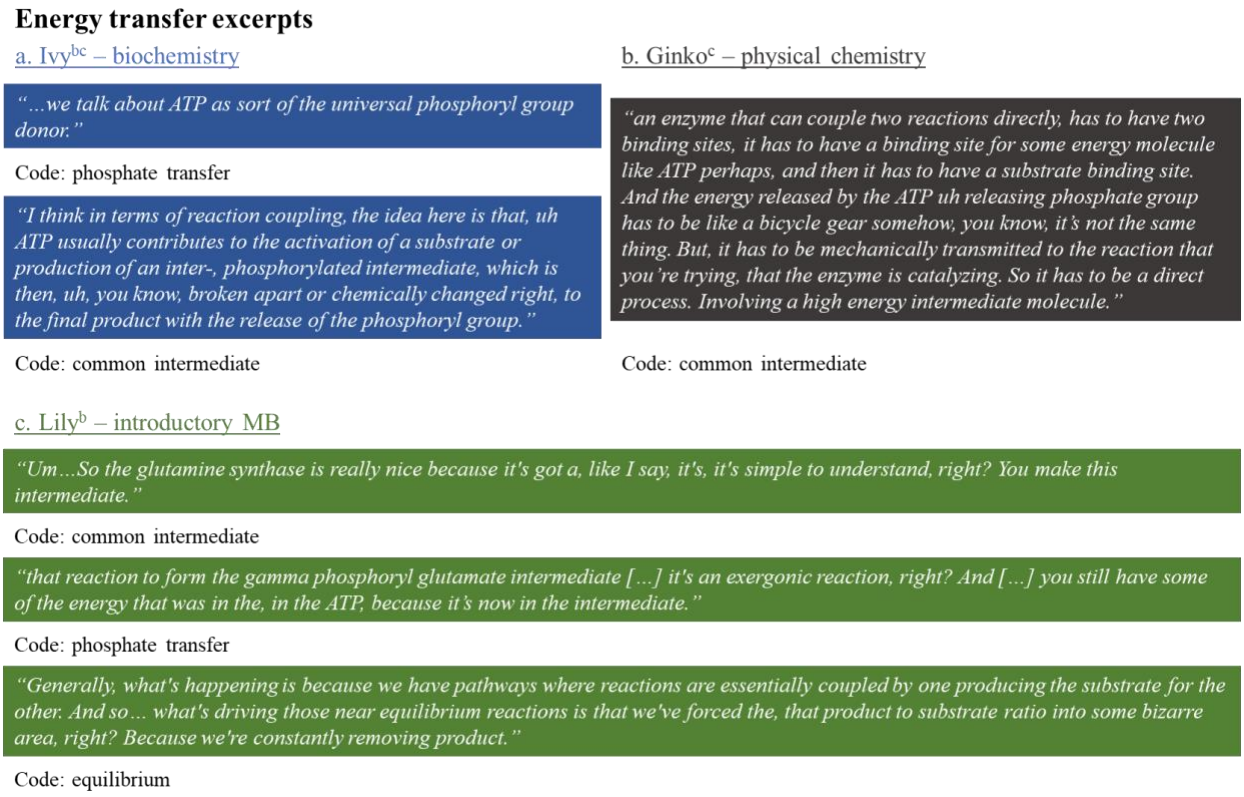


Figure 6.3. Example excerpts of explanations for energy transfer. [a, b, and c are color coded according to the discipline of each instructor. Green[b] = biology, blue[bc] = biochemistry, gray[c] = chemistry].

Instructors Who Explained Both Energy Release And Transfer

We were particularly interested in the nine instructors who leveraged both mechanisms and whether there was a difference in how or when they discussed each mechanism. In taking a closer look at the excerpts in which these codes emerged, we found that five of these instructors explained energy transfer in direct response to how ATP drives unfavorable processes, with their explanations for energy release occurring at different points or in different contexts during the interview. For example, Lily[b] explained how energy is transferred, but they also talked about how it is common for biologists to discuss the energy release mechanism, saying: *"… a lot of times […] biologists will just assume that well, okay. So there was this ATP,*

was broken, […] and usually they'll say it was hydrolyzed. And that has 30.5 kJ/mol, right? And then I can do this other reaction because I've already paid, but without realizing that there's no mechanistic coupling and that, that wouldn't work" (codes: *referencing disciplines* and *ATP hydrolysis*).

The other four instructors (all biologists) initially explained energy release in direct response to "how do you think about the mechanism by which ATP drives unfavorable processes?", but directly following this explanation, they leveraged the ideas of *phosphate transfer* and/or *common intermediate* to explain the mechanism of energy transfer. For example, Monstera[b], an introductory MB instructor, used *ATP hydrolysis* and *bond energy* when explaining how ATP drives unfavorable processes: *"[…] we talked about a high energy phosphate bond in the context of reactants and products. And so that's what we mean by when we say […] ATP plus water* (code: *ATP hydrolysis*) *is going to have more […] energy that is available"* (code: *bond energy*). They continued talking about ATP hydrolysis but eventually said, "*instead of having the phosphate just go to free phosphate, let's donate it* (code: *phosphate transfer*) *to another molecule and have it become a phosphorylated intermediate"* (code: *common intermediate*), thereby shifting their explanation to that of energy transfer.

Like Monstera[b], three other biology instructors who explained energy release via bond formation included ideas about *phosphate transfer* and/or *common intermediate* later during their interview, even though they initially used explanatory black boxes for how energy is transferred (Figure 6.4). While we recognize our small N value, we also found it interesting (and perhaps not surprising) that all three biochemists explained the energy transfer mechanism, with their ideas about energy release surfacing in other contexts (e.g., when referencing how

other disciplines discuss the role of ATP). Although the majority (N = 9) of instructors discussed both mechanisms at some point, five of the instructors (three chemists and two biologists) *only* explained energy release and, therefore continuously black boxed the mechanism of energy transfer (i.e., they had zero codes for both *phosphate transfer* and *common intermediate*) (Figure 6.4). We do not assume whether they recognize these ideas as components of the mechanism of energy transfer; however, as neither came up in their interviews, even if they hold a mechanistic understanding, they did not see it as relevant to our discussion. In other words, only ideas related to energy release were activated, and not any other conceptualizations that these instructors might hold regarding this phenomenon.

While this analysis provided insights into the mechanisms that instructors embrace (transfer, release, or both), it also surfaced the unprompted, negative experiences that instructors expressed regarding teaching ATP (RQ3).

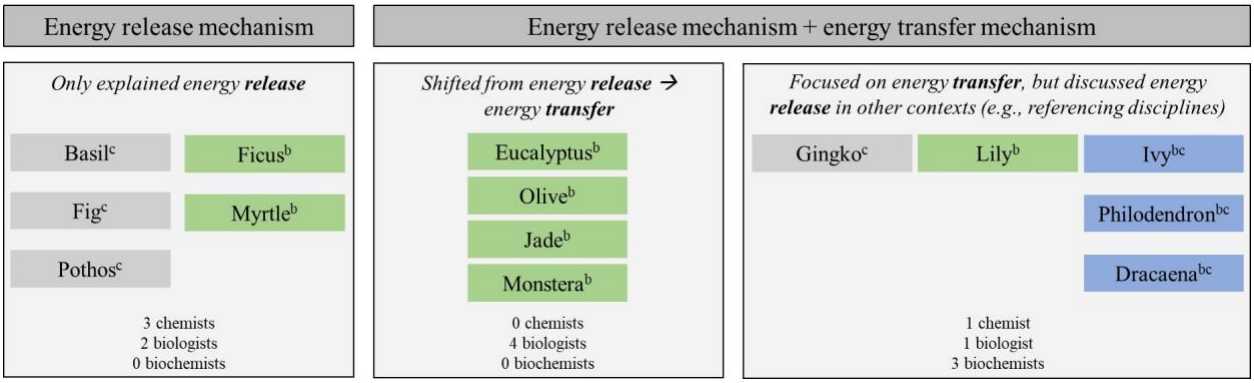


Figure 6.4. Instructor explanations for how ATP drives unfavorable processes using the mechanisms of energy release or energy transfer [Instructor pseudonyms are colored according to discipline. Green[b] = biology, blue[bc] = biochemistry, gray[c] = chemistry].

Research Question 3 – What Teaching Experiences Did Molecular Biology Instructors Share Regarding ATP In Their Course(s)?

As evidenced in our interview protocol, we did not intend to elicit themes related to affect; however, several instructors shared their thoughts regarding pedagogical experiences with ATP (as evidenced by the frequent occurrence (n=61) of the *teaching/learning* code). That is, rather than just explaining their current instructional practices, they also shared how they, as instructors, and/or their students, as learners, experience these instructional practices. Investigating this research question is warranted, (1) because it is likely that these experiences impact pedagogical/instructional approaches, and (2) because of its emergence despite minimal prompting. In this section, we highlight experiences shared by three introductory molecular biology (MB) instructors who shifted from an explanation of energy release to that of energy transfer (as seen in RQ2). We chose these three participants because they teach the introductory course that most often includes a discussion of the role of ATP, and they provided particularly rich discussions about their experiences with instruction on this topic. Because the *teaching/learning* code captured a broad range of experiences rather than more specified affective constructs, we share longer excerpts (Figures 6.5, 6.6, and 6.7) to exemplify the dissatisfaction that these instructors expressed while reflecting on their teaching of ATP.

Jade[b]

Consider the excerpt from Jade[b] in Figure 6.5, whose response to the initial question about ATP providing energy put them in the energy release theme, since they leveraged *ATP hydrolysis*. However, they showed concern and vulnerability at several points in this excerpt that should not be ignored. Jade[b] began by saying they're "*nervous*" to explain the mechanism,

because they know "*[they] don't get it all right*." While this was not further explored by the interviewer, there is a distinct possibility that this is a result of interactions with chemistry faculty and recent research publications[8,10,15] that have focused on the bond breaking/energy release misconception often associated with instruction about ATP. In fact, some MB instructors noted that they focus specifically on what their chemist colleagues tell them (see excerpts in Appendix D, captured by the code *referencing disciplines*).

Following this admission, interviewer KN attempted to mitigate Jade's[b] nerves/concerns by restating that this interview is not a test, and Jade[b] proceeded to note two reasons for why they do not explain anything beyond "*pairing the hydrolysis of ATP with unfavorable reactions*." First, they said, "*I'm not sure it matters… that students understand that next level in order to understand the bigger ideas*," and they went on to say that "*at some point, we had to assume that they had this knowledge when they get into cell and molecular biology, that they understand… that there's an intermediate and the intermediate is less stable than the final thing…*" (code: *common intermediate* and *teaching/learning*). After KN validated some of Jade's[b] thoughts (*[Editorializing]* in Figure 5), they continued to discuss a third, more personal, reason for not going to a deeper level; Jade[b] said, "*the course seems sort of overwhelming and exhausting to me. I think if I understood it better, I would be able to better explain why molecular reactions happen in general*" (code: *teaching/learning*). In just this excerpt, Jade[b] shared negative experiences (including nerves, lack of confidence in content, and concern that this content is irrelevant or repetitive), that we interpret as dissatisfaction regarding their current teaching practices about the role of ATP in their course and/or dissatisfaction in the

expectations surrounding teaching ATP. That is, regardless of the approach they take (energy release or energy transfer), there is some discomfort in the current pedagogy.

Interviewer: *[…] How would you explain the mechanism of how ATP is used as an energy source?*

Jade: *So the way I typically talk about it, and I feel nervous. Because I know that we don't get, I don't get it all right. So but I talk about it in pairing the hydrolysis of ATP with unfavorable reactions. That's, that's how I think, yeah, that's how I talk about it.*

Interviewer : *Okay. […] do you ever talk about, at the molecular level what's happening?*

Jade: *No.*

Interviewer: *Do you know what…? And this is not a test, this is, right, I just spent the last hour before this being, 'Where the heck does the energy come from?'*
Jade: *I, my, my knowledge pretty much stops there. So there's two things to it. One is I'm not sure it matters. [Identifying information], but I'm not sure it matters, that our… that, at least in our class, that students understand that next level in order to understand the bigger ideas that we're trying to get to. So that, I'm not convinced of that. And the second is, if I talked about that I might be, and I'd have to have some professional development in that area, but it may just end up being the same thing that's already been taught in the chemistry course. So at some point, we had to assume that they had this knowledge when they get into cell and molecular biology, that they understand what's happening during those chemical reactions and that there's an intermediate and the intermediate is less stable than the final thing, so… than the product. So I think, based on those two things. I do not go any further than that.*

Interviewer : *[Editorializing]*

Jade: *I think. And sometimes this is why the course seems sort of overwhelming and exhausting to me. I think if I understood it better, I would be able to better explain why molecular reactions happen in general. Right? So when I talk about DNA replication, the structure and the structure of a nucleotide is the way it's replicated, it's energetically favorable, but I don't really know all the details to that. And so I just say that, you know, this works well, because it's energetically favorable. Again, it's the combination of: Does it matter if I go the next level lower to explain that? And am I able to? And I'm saying the answer to both of those, at least right now, the last time I taught the class was no.*

Figure 6.5. Transcript excerpt from Jade[b].

Monstera[b]

Recall Monstera[b], an introductory MB instructor, who we saw in RQ2 discuss both *ATP hydrolysis* and how energy is released, as well as a follow-up explanation for energy transfer. In this subsequent discussion, Monstera[b] revealed (1) their understanding of the mechanism by which energy is transferred by leveraging both *phosphate transfer* and *common intermediate*

and (2) the challenges they experienced in teaching the role of ATP to students. Consider the excerpt in Figure 6.6.

Like Jade[b], Monstera[b] initially focused their discussion on ATP hydrolysis and the bond energies associated with the reaction components to explain the mechanism for energy release. In probing further, we learned that while those ideas were the focus of instruction, Monstera[b] specifically noted that they *"don't really get into how do we couple those reactions… I would hope that that would be something that we can incorporate into a chemistry curriculum…"* (code: *referencing disciplines*, and *explicit exclusion of a mechanism*). This quote provides evidence that Monstera[b] intentionally does not address the mechanism of energy transfer in their course. Following this admission, Monstera[b] specifically identified the mechanistic step of forming a phosphorylated intermediate (codes: *common intermediate* and *phosphate transfer*), suggesting that this be incorporated into a chemistry curriculum, because when they tried to do it in their course, they "*had neither the expertise nor the time to do it*". Here, similar to Jade[b], Monstera[b] expressed lack of confidence in the content and concern for course constraints (time). Finally, though less specific to ATP, Monstera[b] referenced the course textbook and how the chapter on bioenergetics is "*a sticky mess*" and "*it's hard*" for the students, further highlighting the negative experiences that they encounter as an instructor teaching these ideas in introductory MB.

Monstera: *[…] Now we can drive a reaction that has free energy change of less than 7.6 k calories per mole in the other direction. And so, we but we don't really get into how do we couple those reactions? How do we specifically go in there? And I would hope that that would be something that we can incorporate into a chemistry curriculum some place and say, well, look, here's a phosphoryl- instead of having the phosphate just go to free phosphate, let's donate it to another molecule and have it become a phosphorylated intermediate. And couple it to that reaction, and really kind of unpack what that looks like. I tried to do that at some level in my own course, maybe one or two times. But I had neither the expertise nor the time to do it.*

Interviewer: *[…] how would you explain it at the molecular level […] with regards to ATP or just energy transfer in general?*

Monstera: *Well, I mean, somehow that phosphate group gets stuck on another molecule. And so one would assume that that phosphorylated molecule now has higher energy content than it did prior to that point. And so what we don't do is really explore that idea.*

Interviewer: *In in chemistry or biology?*

Monstera: *In biology. Yeah. I mean, the, if you look at our if you look at our textbook, there's a formal chapter on bioenergetics. Which is, it's a sticky mess going in there. Because the students really, you know, it's hard. It's, it's hard.*

Figure 6.6. Transcript excerpt from Monstera[b].

Ficus[b]

Lastly, Ficus[b], who primarily taught introductory MB but also biochemistry, discussed the challenges of deciding what level to go to when teaching these ideas in their course(s). Ficus[b] only explained the mechanism of energy release, and their excerpts are shown in Figure 6.2 under RQ2. In addition to these excerpts, Ficus[b] shared the challenges of teaching these ideas to students. Consider the excerpt in Figure 6.7.

In this excerpt, Ficus[b] shared their concerns with teaching the mechanism by which ATP hydrolysis releases energy in an introductory MB course. At no point did they bring in ideas about *phosphate transfer* or *common intermediate*. When we first asked how they explain the mechanism of how ATP works, Ficus[b] said they tried teaching the mechanism (of how ATP hydrolysis releases energy, which they later note) in the past "*with very limited success… caused*

way more problems than it should have." Because of this, they actively avoid teaching these ideas, because they "*found it to be not productive*." Ficus[b], in our interpretation, expressed frustration and dissatisfaction when reflecting on their past experiences with teaching ATP by explaining the mechanism of energy release to introductory MB students. In their words, their efforts resulted in "*way more problems*", and were "*not productive*", and created "*mass confusion*" (code: *teaching/learning*).

***Interviewer**: So how do you explain the mechanism of how ATP works?*

***Ficus**: I try not to. I've tried in the past, with very limited success, it actually probably caused way more problems than it should have. I've done, so when I've tried to do it, and I go back and forth sometimes. And maybe next year, I'll bring it back. I don't know that. So what I've tried to do is get them to think about, when I do this, I recently I've tried to explicitly tie it, get them to tie it to what they know from chemistry. And so we start asking, so does it require energy or release energy to form a bond?*

[Ficus explaining the mechanism of energy release using ATP hydrolysis and bond energies]

***Ficus**: […] And I tried to very explicitly illustrate that. And so overall, where does, why do you get I mean it's a combination of entropy and enthalpy. But at this point, I've given up on using Delta G equals delta H minus T delta S, it's just, I've tried, tried really hard to use it in the intro class. And it, it I thought, it just wound up muddying a point in the student. The students trying to understand that equation didn't actually help them get to the bigger understanding that I was after. And so. So to really understand the mechanism of why ATP hydrolysis releases energy, I think you have to go there and to go there just was, I found it to be not productive.*

[Interviewer asks about biochemistry and Ficus explains how they teach ideas in biochemistry]

***Interviewer**: [Editorializing]*

***Ficus**: Yeah, but just try very hard* [in MB] *to avoid or directly address the idea that it's not breaking this bond that releases energy. But, and for students who are interested, I talked to one on one in small groups. I'll have that discussion in more detail, depending on where they are. But generally speaking, what I found is if you try to do that with the whole class, it just creates mass confusion.*

Figure 6.7. Transcript excerpt from Ficus[b].

Our use of a semi-structured interview protocol allowed for these rich discussions to emerge. While their emergence is perhaps not surprising given the purpose of the interviews,

the findings here show that minimal prompting activated personal and meaningful experiences that have shaped how these instructors think and teach about ATP. We provide shorter excerpts from additional instructors who shared negative experiences in Appendix E (Table 6.5).

Discussion

In this study, we set out to understand how a range of chemistry, biology, and biochemistry instructors think about the mechanism by which ATP drives unfavorable processes in relation to their discipline and course(s). The interviews revealed the range of ideas instructors think about when considering the roles of ATP (RQ1). From the analysis in RQ2, we found that the instructors primarily focused on two mechanisms when explaining how ATP drives energetically unfavorable processes: (1) **energy transfer** through the phosphorylation of a common intermediate and (2) **energy release** based on the breaking of weaker bonds and the forming of stronger bonds via ATP hydrolysis. Between the two, we found that most faculty leveraged the energy release mechanism (N = 14); however, nine of these instructors leveraged *both* mechanisms. In this case, some instructors described their use of the energy release mechanism when teaching about the roles of ATP but indicated understanding of the energy transfer mechanism as disciplinary experts (e.g., Monstera[b] and Jade[b]). Other instructors, like Lily[b] for example, focused on the mechanism of energy transfer in explaining the role of ATP, but explicitly brought up how certain disciplines might only explain the mechanism of energy release.

Our interviews showed that current instructional practices about ATP in biology courses appear to be dissatisfying and frustrating for instructors (RQ3). Within the three longer excerpts, we also noticed some comments which suggest that the instructors did not see

certain aspects as relevant to a general understanding of biological systems (e.g., Jade[b] said "*I'm not sure it matters…*"), perhaps speaking to the lack of empowerment of instructors. Further, it is likely that these instructors have conveyed this assumption to students. Some instructors shared the frustrations they experienced when teaching the energy release mechanism, especially Ficus[b] who said that teaching this mechanism "*caused way more problems than it should have*." These results highlight how current approaches are dissatisfying (evidenced by the negative experiences described by instructors). Additionally, as noted by Lily[b], this approach can be a misleading way to discuss the role of ATP: *"… usually they'll say it was hydrolyzed. And that has 30.5 kilojoules per mole, right? And then I can do this other reaction because I've already paid, but without realizing that there's no mechanistic coupling and that, that wouldn't work*".

Energy Release: A Misleading Mechanism

While both the energy transfer and energy release mechanisms were explained accurately by the instructors (that is, we did not note any misconceptions), the energy release mechanism explains a different phenomenon (how/why ATP hydrolysis releases energy in an isolated system) and is inconsistent with the actual molecular mechanistic events when ATP is involved in reaction coupling. Even a canonical discussion of ATP hydrolysis and bond energy does not address *how* the energy released from ATP hydrolysis drives the unfavorable process. While both energy release and energy transfer are important and related concepts, the two mechanisms often appear to be conflated with one another.

Lipmann (1941) used the ATP hydrolysis reactions to calculate bond energies, but his experimental design did not suggest that the hydrolysis of ATP was the biological mechanism

for the transfer of energy[20]. We are not the first to identify this issue in the discussion of the biological mechanisms of ATP – over half a century ago Banks and Vernon (1970) noted that "*simple thermodynamic parameters are irrelevant in discussing whole organisms: these must be understood in kinetic and mechanistic terms*."[34] Later in his career, Lipmann himself remarked that defining ATP as a phosphoryl group donor (that is, the energy transfer mechanism) "*could lead to a better understanding of such sequences of energy transfer where we are pretty much in the dark about events subsequent to an initial ATP-involving step*"[59]. The conflation of the energy release and energy transfer mechanisms in biology has led to much confusion about what is actually happening in biological systems and how the materials should be taught, so it is not surprising that this mechanism was discussed by nearly all of the instructors (n=14). However, no matter how well we improve the accuracy by which we discuss the energy release associated with ATP hydrolysis (by including an explicit discussion of the role of water and bond energies), this disconnect will not be addressed.

We do not believe anyone should be blamed for the conflation of these mechanisms, as our interviews provide evidence that biology instructors have a strong desire to align their courses with chemistry (see Appendix D excerpts from biologists) by communicating how ATP works in a way that mitigates the misconception that breaking bonds releases energy. Not only did Lipmann use bond energies to explain why ATP is a metabolically useful molecule[20], but bond energies are an important idea in chemistry and physics. Thus, it makes sense to focus on the importance of bond energies when discussing why ATP hydrolysis is a thermodynamically favorable reaction. However, if we focus solely on ATP hydrolysis, we miss out on the biological mechanisms by which ATP actually drives phenomena such as unfavorable chemical reactions,

the mechanism which all of the biochemists and most of the biologists explained at some point in their interviews (suggesting its importance in the discipline).

The Need For Additional Research On "Energy Transfer"

In our introduction, we noted the literature base in which researchers have explored the issue of ATP's role in biological systems; nearly all this work (based in physics and chemistry) has focused on ATP hydrolysis and the energy release mechanism. At the same time, no biology education research (to our knowledge) has focused on energy transfer mechanisms. That is, the ways in which ATP "provides" energy via transfer of (usually) its terminal phosphoryl group to the starting material, thereby activating a substrate or reactant (increasing its energy/reactivity), rather than through an adjacent reaction with water. Additionally, other mechanisms, such as the alteration in protein structure when ATP is bound, might also be more accessible if the current focus on ATP hydrolysis is removed. If we can understand how to incorporate these mechanisms into our undergraduate courses, specifically introductory biology, we may alleviate some of the negative experiences reported by instructors, while simultaneously helping students to be better equipped to make mechanistic predictions regarding the role of ATP in other biological processes. As a result, given the association between ATP and energy, this may help students develop and understand ideas about energy that are compatible across disciplines.

Research in this area is still in its infancy. We hope this study encourages the education research community to think deeply about the potential utility of the energy transfer mechanism (which we discuss below), as many important questions remain unanswered. For example, how do students consider both the energy release and energy transfer mechanism?

What level of explanatory depth is necessary and most appropriate for discussing these ideas in the disciplines (or different course levels)? How do students with an understanding of the energy transfer mechanism think about energy across disciplines? How does this mechanism contribute to other important ideas in biology like structure/function? The exploration of such questions is critical to the development of instructional materials that might better support the instructors who are tasked with teaching these ideas.

Suggestions For Instruction

Here we provide pedagogical suggestions that aim to better support instructors as they approach this topic in their course(s). Our findings emphasize that current approaches can lead to frustration for some instructors, and, if including discussions about energy release, misleading ideas for students. Further, most of the biology instructors (n=4), all of the biochemistry instructors (n=3), and only one chemistry instructor discussed phosphate transfer and/or common intermediates to explain the mechanism of energy transfer, suggesting its relevance in biology and biochemistry. Based on literature providing evidence for the usefulness and importance of mechanistic reasoning in education as outlined in the Introduction, we suggest the incorporation of an energy transfer mechanism and/or explicit recognition of explanatory black boxes, as proposed by Haskel-Ittah[51]. Recall that an explanatory black box is a "unit" within a mechanistic explanation that remains unexplained – that is, the entities and interactions which give rise to that unit are not identified or discussed, such that a gap is created between steps within the mechanistic explanation[51]. Explanatory black boxes, while often referred to as "hand-wavy" explanations that lack evidence of deeper knowledge, provide great utility in biology education because of the complexity of biological

systems and the diversity of mechanisms involved – it is unreasonable, impractical, and time-consuming to provide a fully mechanistic explanation for every phenomenon without any black boxes. Thus, deciding the relevant mechanism on which to focus becomes just as imperative as it is challenging. To add to the challenge, this decision must also be paired with explicit recognition of existing explanatory black boxes so as to avoid what has been termed an "illusion of explanatory depth"; that is, the sense that one understands complex phenomena deeply, when that may not be the case[51]. We believe this approach will alleviate the pressures felt by instructors to explain topics that they feel "nervous" about or that they find irrelevant to an understanding of biological systems (i.e., the energy release mechanism).

To illustrate how this could be done, we discuss the unfavorable, ATP-dependent process of the formation of glutamine (Figure 6.8). Typically, this process is shown in such a way that the mechanism by which ATP drives the reaction is presented as an explanatory black box, similar to how we presented it in words to participants in the interviews. Only the input and output are provided, and the mechanistic steps are missing ("phenomenon-based sketch"). Without a description/model of the entities that interact or link together to contribute to the outcome (i.e., the formation of glutamine), we can identify *that* ATP is involved, but not *how* it is involved. Further, without explicitly recognizing this as an explanatory black box, it is likely (and reasonable) that the learner will apply the mechanism that they know, the energy release mechanism, to explain how this unfavorable reaction occurs. However, such an application of the energy release mechanism does not provide a physical mechanism for *how* the energy is transferred from the hydrolysis to the reaction between glutamate and ammonium. We predict that by explicitly stating there is more going on at these reaction arrows (i.e., additional

relevant entities are not shown for simplification purposes), we both (1) better prepare the student for future learning about this topic and (2) mitigate unwanted Dunning-Kruger effects[60] associated with the over-estimation of one's understanding of a phenomenon (i.e., avoid the illusion of explanatory depth).

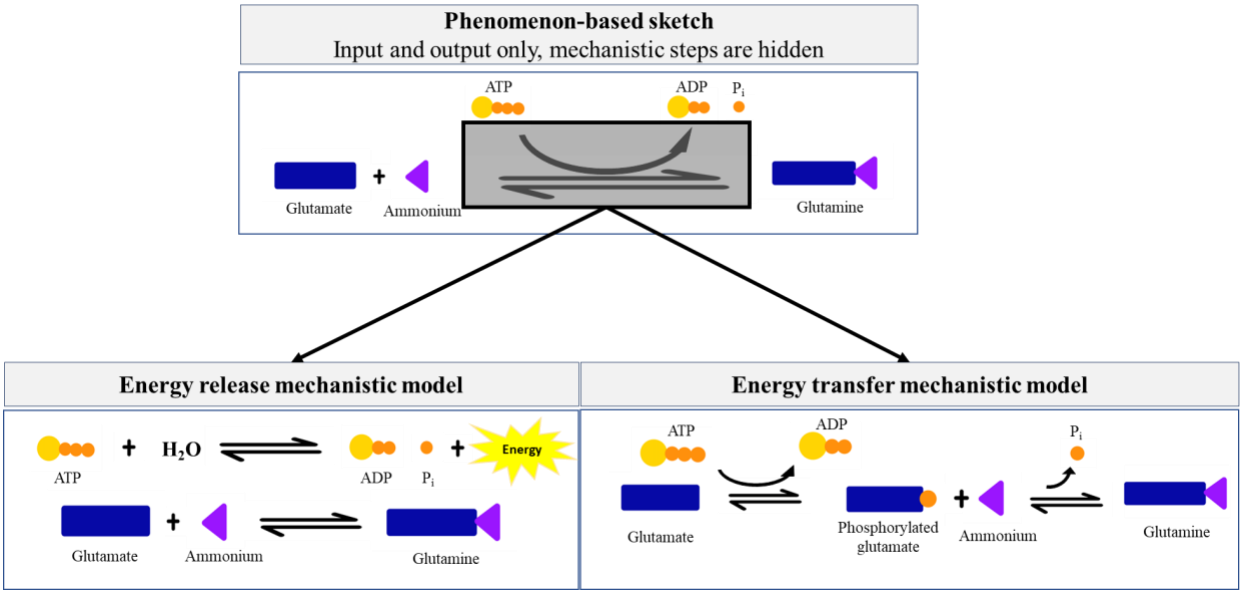


Figure 6.8. Using the energy release or energy transfer mechanistic models to unpack the black box for how ATP drives the formation of glutamine from glutamate and ammonium.

While explicitly recognizing this explanatory black box can prepare students for future learning, it still does not address the multiple mechanistic roles played by ATP, common processes that can (and perhaps should) be emphasized given the biological importance of this molecule. We posit that unpacking this black box by leveraging the energy transfer mechanism would support students' understanding of the mechanistic role of ATP, and thus, their ability to explain or predict other biological processes involving ATP or energy transfer in these complex systems.

Using the energy transfer mechanism outlined by both biology and biochemistry interviewees, we can describe the transfer of a phosphoryl group from ATP to glutamate, forming the phosphorylated glutamate. This more reactive common intermediate can then react with ammonium, releasing inorganic phosphate and producing glutamine. Broadly, a similar approach could be applied for phenomena in which phosphoryl transfer does not occur, like the non-covalent interactions formed with ATP binding to an enzyme. This *physical interaction,* through either non-covalent or covalent forces, is key to explaining the mechanism of energy transfer. This approach is also in better alignment with the actual mechanisms by which the biological phenomenon occurs, as glutamine formation does not involve the hydrolysis of ATP and instead occurs through the formation of a phosphorylated intermediate. Further, we propose investigating the role of prior knowledge in understanding and explaining these ideas, specifically by focusing on ideas about reactivity, which are often introduced in chemistry courses. Using this approach unpacks how ATP provides energy; however, additional explanatory black boxes which could be further unpacked still remain, for example, how and why does ATP react with glutamate? or how does the phosphoryl group change the reactivity of glutamate? These additional black boxes highlight the complexity of a relatively simple biological phenomenon, and therefore, the importance of making explanatory black boxes explicit, so as to best prepare students for future learning. This approach is our best recommendation based on the evidence uncovered in this study, our understanding of the roles of ATP, and the literature supporting mechanistic reasoning and its role in undergraduate science education; however, future work should investigate the effectiveness of these specific approaches. Based on this future work, we will refine these approaches to provide more

effective materials to support students' learning of the mechanism of a key role by which ATP works in cells.

Limitations

Our goal in sharing much of our qualitative data was to highlight the authentic thoughts of each participant; however, we recognize our biases, regardless of our attempts to mitigate these biases as outlined in our positionality statement. We urge our readers to think deeply about their own interpretations of these interview excerpts, as they may differ slightly from ours. We are grateful to the instructors for trusting us with their thoughts and have sincerely aimed to share this data appropriately and carefully.

Our focus with this work has been on chemical reactions, but there are other cellular events in which ATP acts as a regulator, chaperone, signaling molecule, hydrotrope[61], etc. These additional roles are critical in biology, as well as events leading to the synthesis of ATP and the maintenance of high (millimolar) and steady intracellular ATP concentrations. We plan to investigate these areas in future projects and encourage other experts to do the same. This work reflects a focus on chemical reactions in biological systems because of our concern with developing interdisciplinary teaching and learning techniques that help students link ideas between their undergraduate chemistry and biology courses.

Several of the instructors we interviewed are involved in education research, and perhaps a different group of instructors would include different ideas or different combinations of ideas regarding both their teaching and disciplinary understandings regarding ATP.

Lastly, we used the *teaching/learning* code to represent the experiences instructors shared with us during their interviews. This code captured a wide range of experiences, which

we did not categorize into sub-codes; however, additional investigation should be done to characterize how instructors feel about their current, past, or proposed teaching practices regarding this critical topic so as to best support them and, in turn, their students.

REFERENCES

(1) Feynman, R. P.; Leighton, R.; Sands, M. *The Feynman Lectures on Physics*, New Millennium Edition.; Basic Books: New York, NY, 2011; Vol. 1.

(2) Cooper, M. M.; Klymkowsky, M. W. The Trouble with Chemical Energy: Why Understanding Bond Energies Requires an Interdisciplinary Systems Approach. *CBE Life Sci Educ* **2013**, *12* (2), 306–312. https://doi.org/10.1187/cbe.12-10-0170.

(3) National Research Council. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*; National Academies Press: Washington, DC, 2012.

(4) Barak, J.; Gorodetsky, M.; Chipman, D. Understanding of Energy in Biology and Vitalistic Conceptions. *International Journal of Science Education* **1997**, *19* (1), 21–30. https://doi.org/10.1080/0950069970190102.

(5) Barker, V.; Millar, R. Students' Reasoning about Basic Chemical Thermodynamics and Chemical Bonding: What Changes Occur during a Context-Based Post-16 Chemistry Course? *International Journal of Science Education* **2000**, *22* (11), 1171–1200. https://doi.org/10.1080/09500690050166742.

(6) Goldring, H.; Osborne, J. Students' Difficulties with Energy and Related Concepts. *Physics Education* **1994**, *29*, 26–32. https://doi.org/10.1088/0031-9120/29/1/006.

(7) Boo, H. K. Students' Understandings of Chemical Bonds and the Energetics of Chemical Reactions. *Journal of Research in Science Teaching* **1998**, *35* (5), 569–581. https://doi.org/10.1002/(SICI)1098-2736(199805)35:5<569::AID-TEA6>3.0.CO;2-N.

(8) Dreyfus, B. W.; Sawtelle, V.; Turpen, C.; Gouvea, J.; Redish, E. F. Students' Reasoning about ``high-Energy Bonds'' and ATP: A Vision of Interdisciplinary Education. *Phys. Rev. ST Phys. Educ. Res.* **2014**, *10* (1), 010115. https://doi.org/10.1103/PhysRevSTPER.10.010115.

(9) Galley, W. C. Exothermic Bond Breaking: A Persistent Misconception. *J. Chem. Educ.* **2004**, *81* (4), 523. https://doi.org/10.1021/ed081p523.

(10) Green, A. I.; Parent, K. N.; Underwood, S. M.; Matz, R. L. Connecting Ideas across Courses: Relating Energy, Bonds & How ATP Hydrolysis Powers a Molecular Motor. *The American Biology Teacher* **2021**, *83* (5), 303–310. https://doi.org/10.1525/abt.2021.83.5.303.

(11) Hapkiewicz, A. Clarifying Chemical Bonding. Overcoming Our Misconceptions. *Science Teacher* **1991**, *58* (3), 24–27.

(12) Johnstone, A. H.; Mahmoud, N. A. Isolating Topics of High Perceived Difficulty School Biology. *Journal of Biological Education* **1980**, *14* (2), 163–166. https://doi.org/10.1080/00219266.1980.10668983.

(13) Novick, S. No Energy Storage in Chemical Bonds. *Journal of Biological Education* **1976**, *10* (3), 116–118. https://doi.org/10.1080/00219266.1976.9654072.

(14) Teichert, M. A.; Stacy, A. M. Promoting Understanding of Chemical Bonding and Spontaneity through Student Explanation and Integration of Ideas. *Journal of Research in Science Teaching* **2002**, *39* (6), 464–496. https://doi.org/10.1002/tea.10033.

(15) VandenPlas, J. R.; Herrington, D. G.; Shrode, A. D.; Sweeder, R. D. Use of Simulations and Screencasts to Increase Student Understanding of Energy Concepts in Bonding. *J. Chem. Educ.* **2021**, *98* (3), 730–744. https://doi.org/10.1021/acs.jchemed.0c00470.

(16) Freeman, S.; Quillin, K.; Allison, L.; Black, M.; Podgorski, G.; Taylor, E.; Carmichael, J. *Biological Science*, 7th ed.; Pearson: London, England, 2019.

(17) Storey, R. D. Textbook Errors & Misconceptions in Biology: Cell Energetics. *The American Biology Teacher* **1992**, *54* (3), 161–166. https://doi.org/10.2307/4449438.

(18) Urry, L. A.; Cain, M. L.; Wasserman, S. A.; Minorsky, P. V. *Campbell Biology*, 12th ed.; Pearson: London, England, 2020.

(19) *The Nobel Prize in Physiology or Medicine 1953*. NobelPrize.org. https://www.nobelprize.org/prizes/medicine/1953/lipmann/biographical/ (accessed 2023-04-07).

(20) Lipmann, F. Metabolic Generation and Utilization of Phosphate Bond Energy. In *Advances in Enzymology and Related Areas of Molecular Biology*; John Wiley & Sons, Ltd: Hoboken, New Jersey, 1941; pp 99–162. https://doi.org/10.1002/9780470122464.ch4.

(21) Kohn, K. P.; Underwood, S. M.; Cooper, M. M. Energy Connections and Misconnections across Chemistry and Biology. *LSE* **2018**, *17* (1), ar3. https://doi.org/10.1187/cbe.17-08-0169.

(22) Liaw, S. H.; Eisenberg, D. Structural Model for the Reaction Mechanism of Glutamine Synthetase, Based on Five Crystal Structures of Enzyme-Substrate Complexes. *Biochemistry* **1994**, *33* (3), 675–681. https://doi.org/10.1021/bi00169a007.

(23) Moreira, C.; Ramos, M. J.; Fernandes, P. A. Clarifying the Catalytic Mechanism of Human Glutamine Synthetase: A QM/MM Study. *J. Phys. Chem. B* **2017**, *121* (26), 6313–6320. https://doi.org/10.1021/acs.jpcb.7b02543.

(24) Das, A.; Rui, H.; Nakamoto, R.; Roux, B. Conformational Transitions and Alternating-Access Mechanism in the Sarcoplasmic Reticulum Calcium Pump. *Journal of Molecular Biology* **2017**, *429* (5), 647–666. https://doi.org/10.1016/j.jmb.2017.01.007.

(25) Inesi, G. Mechanism of Calcium Transport. *Annu Rev Physiol* **1985**, *47*, 573–601. https://doi.org/10.1146/annurev.ph.47.030185.003041.

(26) Møller, J. V.; Nissen, P.; Sørensen, T. L.-M.; Maire, M. le. Transport Mechanism of the Sarcoplasmic Reticulum Ca2+-ATPase Pump. *Current Opinion in Structural Biology* **2005**, *15* (4), 387–393. https://doi.org/10.1016/j.sbi.2005.06.005.

(27) Toyoshima, C.; Inesi, G. Structural Basis of Ion Pumping by Ca2+-ATPase of the Sarcoplasmic Reticulum. *Annual Review of Biochemistry* **2004**, *73* (1), 269–292. https://doi.org/10.1146/annurev.biochem.73.011303.073700.

(28) Feldman, R. I.; Sigman, D. S. The Synthesis of Enzyme-Bound ATP by Soluble Chloroplast Coupling Factor 1. *Journal of Biological Chemistry* **1982**, *257* (4), 1676–1683. https://doi.org/10.1016/S0021-9258(19)68090-7.

(29) Geeves, M. A. Review: The ATPase Mechanism of Myosin and Actomyosin. *Biopolymers* **2016**, *105* (8), 483–491. https://doi.org/10.1002/bip.22853.

(30) Johnson, K. A. Pathway of the Microtubule-Dynein ATPase and the Structure of Dynein: A Comparison with Actomyosin. *Annual Review of Biophysics and Biophysical Chemistry* **1985**, *14* (1), 161–188. https://doi.org/10.1146/annurev.bb.14.060185.001113.

(31) Kiani, F. A.; Fischer, S. Stabilization of the ADP/Metaphosphate Intermediate during ATP Hydrolysis in Pre-Power Stroke Myosin: QUANTITATIVE ANATOMY OF AN ENZYME *. *Journal of Biological Chemistry* **2013**, *288* (49), 35569–35580. https://doi.org/10.1074/jbc.M113.500298.

(32) Nakamoto, R. K.; Baylis Scanlon, J. A.; Al-Shawi, M. K. The Rotary Mechanism of the ATP Synthase. *Archives of Biochemistry and Biophysics* **2008**, *476* (1), 43–50. https://doi.org/10.1016/j.abb.2008.05.004.

(33) Prieß, M.; Göddeke, H.; Groenhof, G.; Schäfer, L. V. Molecular Mechanism of ATP Hydrolysis in an ABC Transporter. *ACS Cent. Sci.* **2018**, *4* (10), 1334–1343. https://doi.org/10.1021/acscentsci.8b00369.

(34) Banks, B. E. C.; Vernon, C. A. Reassessment of the Role of ATP in Vivo. *Journal of Theoretical Biology* **1970**, *29* (2), 301–326. https://doi.org/10.1016/0022-5193(70)90024-X.

(35) Carusi, E. A. It's Time We Replaced 'High-Energy Phosphate Group' with 'Phosphoryl Group.' *Biochemical Education* **1992**, *20* (3), 145–147. https://doi.org/10.1016/0307-4412(92)90055-Q.

(36) Yang, M. Conflicting Representations of Bond Breaking and Formation in Biology and Chemistry Textbooks: A Case Study of ATP Hydrolysis., 2023.

(37) Cooper, M. M. Why Ask Why? *J. Chem. Educ.* **2015**, *92* (8), 1273–1279. https://doi.org/10.1021/acs.jchemed.5b00203.

(38) Glennan, S. *The New Mechanical Philosophy*; Oxford University Press: Oxford, New York, 2017.

(39) Schwarz, C. V.; Passmore, C.; Reiser, B. J. *Helping Students Make Sense of the World Using Next Generation Science and Engineering Practices*; National Science Teachers Association: Arlington, VA, 2017.

(40) Ralph, V. R.; Scharlott, L. J.; Schafer, A. G. L.; Deshaye, M. Y.; Becker, N. M.; Stowe, R. L. Advancing Equity in STEM: The Impact Assessment Design Has on Who Succeeds in Undergraduate Introductory Chemistry. *JACS Au* **2022**. https://doi.org/10.1021/jacsau.2c00221.

(41) Klymkowsky, M. W. Making Mechanistic Sense: Are We Teaching Students What They Need to Know? *Dev Biol* **2021**, *476*, 308–313. https://doi.org/10.1016/j.ydbio.2021.04.004.

(42) van Mil, M. H. W.; Boerwinkel, D. J.; Waarlo, A. J. Modelling Molecular Mechanisms: A Framework of Scientific Reasoning to Construct Molecular-Level Explanations for Cellular Behaviour. *Sci & Educ* **2013**, *22* (1), 93–118. https://doi.org/10.1007/s11191-011-9379-7.

(43) Franovic, C. G.-C.; Noyes, K.; Stoltzfus, J. R.; Schwarz, C. V.; Long, T. M.; Cooper, M. M. Undergraduate Chemistry and Biology Students' Use of Causal Mechanistic Reasoning to Explain and Predict Preferential Protein–Ligand Binding Activity. *J. Chem. Educ.* **2023**, *100* (5), 1716–1727. https://doi.org/10.1021/acs.jchemed.2c00737.

(44) Noyes, K.; Carlson, C. G.; Stoltzfus, J. R.; Schwarz, C. V.; Long, T. M.; Cooper, M. M. A Deep Look into Designing a Task and Coding Scheme through the Lens of Causal Mechanistic Reasoning. *J. Chem. Educ.* **2022**, *99* (2), 874–885. https://doi.org/10.1021/acs.jchemed.1c00959.

(45) Krist, C.; Schwarz, C. V.; Reiser, B. J. Identifying Essential Epistemic Heuristics for Guiding Mechanistic Reasoning in Science Learning. *Journal of the Learning Sciences* **2019**, *28* (2), 160–205. https://doi.org/10.1080/10508406.2018.1510404.

(46) Russ, R. S.; Scherr, R. E.; Hammer, D.; Mikeska, J. Recognizing Mechanistic Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed from Philosophy of Science. *Science Education* **2008**, *92* (3), 499–525. https://doi.org/10.1002/sce.20264.

(47) Braaten, M.; Windschitl, M. Working toward a Stronger Conceptualization of Scientific Explanation for Science Education. *Science Education* **2011**, *95* (4), 639–669. https://doi.org/10.1002/sce.20449.

(48) Machamer, P.; Darden, L.; Craver, C. F. Thinking about Mechanisms. *Philosophy of Science* **2000**, *67* (1), 1–25. https://doi.org/10.1086/392759.

(49) Becker, N.; Noyes, K.; Cooper, M. Characterizing Students' Mechanistic Reasoning about London Dispersion Forces. *J. Chem. Educ.* **2016**, *93* (10), 1713–1724. https://doi.org/10.1021/acs.jchemed.6b00298.

(50) Crandell, O. M.; Lockhart, M. A.; Cooper, M. M. Arrows on the Page Are Not a Good Gauge: Evidence for the Importance of Causal Mechanistic Explanations about Nucleophilic Substitution in Organic Chemistry. *J. Chem. Educ.* **2020**, *97* (2), 313–327. https://doi.org/10.1021/acs.jchemed.9b00815.

(51) Haskel-Ittah, M. Explanatory Black Boxes and Mechanistic Reasoning. *Journal of Research in Science Teaching* **2023**, *60* (4). https://doi.org/10.1002/tea.21817.

(52) Shiroda, M.; Franovic, C. G.-C.; de Lima, J.; Noyes, K.; Babi, D.; Persson-Gordon, E.; Beltran-Flores, E.; Kesh, J.; McKay, R. L.; Cooper, M. M.; Long, T. M.; Schwarz, C. V.; Stoltzfus, J. R. Examining and Supporting Mechanistic Explanations across Chemistry and Biology Courses, under review.

(53) Cooper, M.; Klymkowsky, M. Chemistry, Life, the Universe, and Everything: A New Approach to General Chemistry, and a Model for Curriculum Reform. *J. Chem. Educ.* **2013**, *90* (9), 1116–1122. https://doi.org/10.1021/ed300456y.

(54) Cooper, M. M.; Stowe, R. L.; Crandell, O. M.; Klymkowsky, M. W. Organic Chemistry, Life, the Universe and Everything (OCLUE): A Transformed Organic Chemistry Curriculum. *J. Chem. Educ.* **2019**, *96* (9), 1858–1872. https://doi.org/10.1021/acs.jchemed.9b00401.

(55) Klymkowsky, M. W.; Rentsch, J. D.; Begovic, E.; Cooper, M. M. The Design and Transformation of Biofundamentals: A Nonsurvey Introductory Evolutionary and Molecular Biology Course. *CBE Life Sci Educ* **2016**, *15* (4), ar70. https://doi.org/10.1187/cbe.16-03-0142.

(56) VERBI Software. MAXQDA 2020 [Computer Software], 2019. maxqda.com.

(57) Creswell, J. W. Data Analysis and Representation. In *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*; SAGE Publications: Thousand Oaks, CA, US, 2013; pp 179–212.

(58) Corbin, J.; Strauss, A. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory, 3rd Ed*; Basics of qualitative research: Techniques and procedures for developing grounded theory, 3rd ed; Sage Publications, Inc: Thousand Oaks, CA, US, 2008; pp xv, 379. https://doi.org/10.4135/9781452230153.

(59) Lipmann, F. Attempts toward a Formulation of Biological Use of Energy in Terms of Chemical Potentials. In *Molecular biology: Elementary processes of nerve conduction and muscle contraction*; Academic Press: Cambridge, MA, 1960; pp 37–47.

(60) Dunning, D. Chapter Five - The Dunning–Kruger Effect: On Being Ignorant of One's Own Ignorance. In *Advances in Experimental Social Psychology*; Olson, J. M., Zanna, M. P., Eds.; Academic Press, 2011; Vol. 44, pp 247–296. https://doi.org/10.1016/B978-0-12-385522-0.00005-6.

(61) Patel, A.; Malinovska, L.; Saha, S.; Wang, J.; Alberti, S.; Krishnan, Y.; Hyman, A. A. ATP as a Biological Hydrotrope. *Science* **2017**, *356* (6339), 753–756. https://doi.org/10.1126/science.aaf6846.

APPENDIX A. PERMISSIONS

Figure 6.9 shows the licensing agreement that permits reproduction of this manuscript in its entirety.

CBE - Life Sciences Education

License and Publishing Agreement

MANUSCRIPT NO.: CBE-23-05-0071RR **DATE:** September 13, 2023
MANUSCRIPT TITLE: How do instructors explain the mechanism by which ATP drives unfavorable processes?
AUTHOR(S): Clare Franovic, Keenan Noyes, Nicholas Williams, Michael Klymkowsky, and Melanie Cooper

1. I (we) ("Author[s]"), on consideration of the acceptance of the work cited above along with all associated supplemental data, video files, and /or cover images ("Manuscript") for publication, do hereby grant to the American Society for Cell Biology ("ASCB") a license to publish the Manuscript in accordance with the following terms and conditions*:

(a) Retention of Rights. The Authors retain the copyright and the right to reprint the Manuscript in any publication of which Authors serve as an author or editor, subject to proper citation of the Manuscript in *CBE-Life Sciences Education (LSE)* and where feasible the presence of a link to the original publication of the Manuscript in *LSE*. Also, Authors are permitted to post the *LSE* PDF of their articles (and/or supplemental material) on their personal websites or in an online institutional repository provided there appears always the proper citation of the Manuscript in *LSE* and a link to the original publication of the Manuscript in *LSE*. Authors further retain the right to revise, adapt, prepare derivative works, present, or distribute the Manuscript provided that all such distribution is for noncommercial benefit and there appears always the proper citation of the Manuscript in *LSE* and where feasible a link to the original publication of the Manuscript in *LSE*.

(b) Licensing of Rights to ASCB. The Authors hereby grant to ASCB a perpetual, irrevocable, paid-up, world-wide license with the right to publish, distribute, reproduce, display, translate, sublicense for commercial purposes, and store the Manuscript in all forms now known or hereafter devised and to authorize others to do so.

(c) Licensing of Rights to the Public. The Authors hereby grant to the general public the nonexclusive right to copy, distribute, or display the Manuscript subject to the terms of the Creative Commons-Noncommercial-Share Alike 3.0 Unported license (http://creativecommons.org/licenses/by-nc-sa/3.0), by which the public:

(1) Is permitted to copy, distribute, transmit, and adapt the Manuscript;(2) Is **not** permitted to make commercial use of, or license or sublicense the Manuscript for commercial use; and,
(3) Is required:
a. To attribute the work to the author (in a way that does not suggest that the authors endorse the users or any user's use);
b. To include the terms of this license so that they apply to any subsequent use or distribution; and,
c. To respect the fair use rights, moral rights, and rights that the Authors and any others have in the Manuscript.

2. Authors warrant that the Manuscript is an original work of the Authors; that it does not infringe or violate any copyright, proprietary right, or any other right of any third party; and Authors agree to indemnify and hold the ASCB harmless against any claim to the contrary. Authors further warrant that they have the right to assign the copyright to the ASCB and that no portion of the copyright to the Manuscript has been previously assigned.

3. Authors have obtained written permission to use any quotation or excerpt from another work or from another's property not in the public domain or covered by fair use provisions of the U.S. Copyright law. Proper acknowledgment has been given in the article for the use of such material.

4. Authors authorize the editor to modify the Manuscript as necessary to prepare the article for publication-including changes in title, style, and format to conform to editorial usage, journal format, and ASCB editorial style. Minor stylistic changes may be made immediately before publication to meet ASCB editorial requirements. Authors may be asked to update material that appears in the Manuscript before publication.

5. Authors will receive a PDF page proof of the Manuscript that will show figures in position.

6. Authors agree to review and return the page proof within 48 hours of receipt of it and to mark on the page proof or an attached file or email message any corrections that need to be made to the Manuscript. Authors can be held responsible for payment of extra charges for changes that are not a result of a printer or editor error or query.

7. Authors confirm that the material presented in the Manuscript has not been published previously, is not currently being considered for publication elsewhere, that all authors listed on the Manuscript have seen and agreed to all information contained in the Manuscript, and that the information contained in the Manuscript is available for general dissemination.

8. Authors confirm that when reporting research that involves recombinant DNA, humans, and animals, they have carried out all of the experiments in accordance with the recommendations of the Declaration of Helsinki and the appropriate National Institutes of Health guidelines and that the research protocols have been approved where necessary by the appropriate institutional committees.

9. Authors understand that publication of the Manuscript in *LSE* implies that Authors agree to make available all propagative materials such as mutant organisms, cell lines, recombinant plasmids, vectors, viruses, and monoclonal antibodies that were used to obtain results presented in the article. Prior to obtaining these materials, interested scientists will provide Authors with a written statement that they will be used for noncommercial research purposes only. Authors are not required to make such material available to others when the terms under which the Authors originally obtained that material prohibit them from doing so.

Agent:* Clare Franovic
Date: 9/13/2023

Figure 6.9. CBE licensing and publishing agreement.

APPENDIX B. INTERVIEW PROTOCOL

Below is the entire interview protocol that KN and CF used when meeting with each participant. However, we wish to reiterate the semi-structured nature of these interviews. By no means did each interview ascribe directly to this protocol, as the instructors had several thoughts and ideas that we did not anticipate, or they brought up ideas early on in the interview that we had not intended to discuss until later. The protocol provided a general guide to maintain the discussion; however, most of the instructors provided far richer conversations than expected, resulting in altered trajectories well outside of the bounds to which this protocol is limited.

Provide Participants With A General Overview Of The Project

1. Things to mention:
    a. Goal of project – help students develop a more coherent understanding of energy in chemistry and biology.
    b. To achieve this – we are interested in understanding faculty perspectives on energy to see how you think about energy-related ideas as an instructor, and how you understand these ideas as an expert in your discipline.
2. Emphasize that **this is not a test**. We are genuinely interested in how you think about energy so we can get a better idea of how we might be able to help students understand these ideas as well.

Consent

3. Highlight important portions from the consent form
    a. Will record the interview and transcribe after

b. Will de-identify the data following this session, giving a pseudonym which would be used if this data is published.

4. Explain the materials (i.e., zoom or recording device).

Introductions

5. KN or CF introduce themselves and thank participant for meeting
6. Ask about courses they teach/have taught, and area of research (if applicable)

Energy

7. Depending on the discussion so far, some ideas to include are listed below, but not all are necessary – these may be dependent on the participant and what they have decided to bring up at that point:
    a. How would you define energy?
    b. Does energy come up in your course(s)? How so?
    c. How would you describe energy transfer? What is it?
    d. How would you explain the mechanism of energy transfer? How does it happen?
    e. Do you think about energy differently in a chemical or biological system?

The Role Of ATP In Driving Unfavorable Processes

8. If they have not brought up ATP on their own, ask them about its role (suggested questions below):
    a. How do you think about the mechanism by which ATP is used as an energy source?
    b. How does ATP drive unfavorable processes?

APPENDIX C. FULL CODEBOOK

Table 6.4 includes each code, a description of the code, an example segment, the number of interviews in which the code occurred ("instructors with code"), and the number of total occurrences across all transcripts ("segments with code").

Table 6.4. Full codebook including codes, descriptions, examples, number of interviews with code, and number of segments with code.

Code	Description	Example segment	Instructors with code	Segments with code
Bond energy	Discussing energy changes associated with bond formation or bond breaking.	"So those weak bonds are getting broken, and then stronger bonds are getting formed, and that the forming of those bonds is releasing energy" (Olive)	13	43
Teaching/learning	When the participant discusses whether or not they understand a topic themselves or how to teach that topic; or when they discuss how students understand or learn ideas related to ATP or reaction coupling.	"I've done it both ways. I've tried to show them the actual mechanism, and I feel like that's a little bit too much for them in second semester" (Fig)	12	61
ATP hydrolysis	Discussing the hydrolysis of ATP, or the reaction of ATP with water.	"we do go through a fairly detailed description of why hydrolysis of ATP is favorable" (Ivy)	11	40
Coupled reactions	Discussing reaction coupling, vaguely or explicitly.	"I talk about it in pairing the hydrolysis of ATP with unfavorable reactions." (Jade)	11	38

Table 6.4 (cont'd)

Referencing disciplines	When the instructor explicitly discusses their own discipline or another discipline and how those disciplines teach/think about ATP.	"what I want the students to do and the way that I talk about sort of the biology of it, is to highlight that there is a discrepancy between the way that biologists will often talk about it and the way chemists talk about it." (Basil)	11	30
ATP Synthesis	Discussing the synthesis or formation of ATP	"If you talk about ATP synthesis, right, you're using that protons all on the…the, inter membrane space coming back into the matrix." (Dracaena)	9	23
Stability	Discussing stability of systems or molecules.	"there's an intermediate and the intermediate is less stable than the final thing" (Jade)	9	12
Gradients	Discussing membrane potentials or chemical potential gradients in the context of coupling or ATP.	"that causes a change in the electrochemical potential energy across the membrane, which is stored there, it isn't stored in a particular molecule it is stored as a… gradient in a fi-, in a field, you know, and just a total potential energy thing." (Ginkgo)	8	17
Phosphate transfer	Discussing the transfer of a phosphoryl group from one entity to another (typically from ATP to some other molecule).	"I think most of the time, we mean, it is phosphorylating the substrate, and activating it based on a phosphoryl transfer." (Ivy)	7	19
Enzyme	Discussing enzymes in the context of ATP or reaction coupling (including mentions of kinase).	"And so when this ATP comes in, there's an enzyme that functions and breaks off two of those phosphates, and then there's, it will join with sort of your growing DNA chain." (Olive)	7	17

Table 6.4 (cont'd)

Common intermediate	Discussing a "high-energy" or "phosphorylated" intermediate involved in reaction coupling.	"instead of having the phosphate just go to free phosphate, let's donate it to another molecule and have it become a phosphorylated intermediate." (Monstera)	6	23
Explicit exclusion of a mechanism	If the participant explicitly mentions that they do NOT address a mechanism (how/why) of ATP being used as an energy source	"Now we've got 7.6. Now we can drive a reaction that has free energy change of less than 7.6 k calories per mole in the other direction. And so, we but we don't really get into how do we couple those reactions? How do we specifically go in there?" (Monstera)	5	12
Biological example	This code is used when the participant gives a biological example of ATP being used.	"the simplest example would be muscle contraction. Right. You need a lot of power to proceed for muscle cells to contract and relax." (Philodendron)	5	11
Equilibrium	Discussing equilibrium (or concentrations of reactants/products) in relation to reaction coupling or ATP.	"…then the second reaction will carry that in terms of the intermediate and it gets drawn along, drawn along, just an example of what le Chatelier's principle. You're, you're shifting equilibrium for the second reaction by loading up with react." (Gingko)	4	17

Table 6.4 (cont'd)

Conformational change	Discussing some conformation or shape change that is caused by ATP.	"No, so, the, the, the form of the enzyme, the conformation of the enzyme is different between the ATP bound and the ADP bound. So when you hydrolyze, and release, you know, so when you hydrolyzed, that ATP, you're going to usually release the phosphate but you get a conformational change of the protein related to that" (Lily)	5	11
Activation energy	Discussing a "barrier" or the activation energy required for a reaction.	"your molecules here and if you put enough energy, you kind of have an average energy that's up at this level. If you increase that energy up, now you have enough to actually overcome the activation energy that's associated with that" (Basil)	4	4
Calculations	Discussing calculations involved in ATP or reaction coupling from either a teaching perspective in the course, or content-related in understanding how ATP works.	"we would talk about the delta G equals delta H minus T delta S. And we will also then try to tie it into well, under cellular conditions, it's actually different than under standard conditions." (Ficus)	3	11

Table 6.4 (cont'd)

Phosphate repulsion	Discussing the charge repulsions of neighboring phosphate groups on ATP	"because these phosphate groups have negative charges, so they're sort of repelling each other. And which makes these oxygen phosphorus bonds a lot weaker than they otherwise would be." (Olive)	3	5
Suggested resolution	When the instructor mentions or suggests some way of talking about ATP that might be more productive or useful for students.	"we've always thought about how do we talk about energy in terms of ATP cleavage? Maybe we should be talking about energy in terms of ATP formation." (Monstera)	2	5

Note 1: Codes are listed in decreasing order of the number of interviews in which they occur.

Note 2: Some example segments include additional codes other than the one for which it is exemplifying.

APPENDIX D. EXAMPLE EXCERPTS FOR *REFERENCING DISCIPLINES*

Figure 6.10 provides example excerpts for the code *referencing disciplines*. These specific examples are from three biologists who mention a desire and/or effort to align content with chemistry.

Eucalyptus: *"so mostly I just copy what my chemist colleagues tell me. I mean, all of chemistry, according to my colleagues in the first year, both introductory courses is just all about bonds, bonds forming, what type of bonds and stuff and the movement of electrons, so I mean, that's the way I see energy also, then."*	**Monstera**: *"Yeah, maybe. But I but I had remembered walking past [chemist's name] class and having [them] go 'and the biologists won't explain it this way.' And I was like, what, hell we won't!"* **Interviewer**: *"… do you know what they were referring to or?* **Monstera**: *"Well, [they were] talking, I think [they were] talking about the energy within bonds and how you would determine differences between reactants and products?"*	**Olive**: *"I want to do things that are consistent with chemistry. I have lots of conversations with my chemistry, you know, peers to say, you know, am I doing this right? Does this make sense the way that I'm doing it? So we talk, we spend a reasonable amount of time talking about bond energy…"*

Figure 6.10. Example excerpts for referencing disciplines from three biology instructors.

APPENDIX E. EXAMPLE EXCERPTS FOR *TEACHING/LEARNING*

Table 6.5 includes additional excerpts for the *teaching/learning* code, showing that other instructors expressed experiences that were not positive.

Table 6.5. Example excerpts for the *teaching/learning* code.

Instructor	Excerpt
Myrtle[b]	*"…in terms of how we talked about the role of ATP, we've clearly, I've clearly been doing that wrong."*
Lily[b]	*"And we just had [inaudible] the difference between an intermediate and a transition state. So today-, I guess today caused some confusion."*
Olive[b]	*"You know, I have a, I have a, I would think a low-ish level of understanding, but I try to teach students in a way that I'm not undoing things that they're learning in chemistry."*
Eucalyptus[b]	*"ATP is an A and a T and a P. There's like, it's all highlighted in yellow with sparky little sunlight shapes around it like 'pchooooo!' that must be that right? So it must be in there? And then 'PCHOOOOO!' Yep. So I think it's, I think, I think the representations they see in the textbook become what they say on the test."*
Ivy[bc]	*"we tend to use a lot of shorthand in our language, we would get really sloppy in our language and we're expecting students to sort of see through that, um, you know, despite it being a high level biochemistry bond-, class, we still say ATP has high energy bonds."*
Basil[c]	*"it's really hard for students to carry this idea. Because when we talk- when they often talk about it in biology, they're sort of the, hey we break the, you know, ATP bond and energy comes out."*
Fig[c]	*"knowing how to teach that connection… It's hard. It's something I struggle with. And I kind of feel like I'm very fortunate, because chemistry is the first class."*

Instructors listed according to discipline (green[b] = biology, blue[bc] = biochemistry, grey[c] = chemistry)

Chapter VII: Investigating The Impact Of A Learning Task On Students' Use Of Mechanistic Resources When Explaining A Complex Phenomenon

Background And Introduction

There is consensus in the science education community that the abstract concept of energy should be included across grade levels and disciplines; however, energy is a notoriously difficult concept for students (and experts) to grasp.[1–9] Physics Nobel Prize winner Richard Feynman once said, "*It is important to realize that in physics today, we have no knowledge what energy is*."[10] We would argue that this statement remains true today, making it exceptionally difficult to productively discuss energy in science education contexts.

Despite its challenges, the central role of energy remains undisputed, making it an important area of investigation in science education research. Further, its utility across science disciplines (as both a crosscutting concept and core idea in chemistry, biology, and physics[11–13]) make it a theme by which we might support students' interdisciplinary learning. However, the ways in which energy is discussed in these disciplines can vary quite significantly. In physics, for example, energy is typically treated as quantifiable amounts that result from calculations, while in the life sciences, energy is often discussed as something that life "captures" or "uses" or "transforms" in order to survive.[2] That is, the nature of the discipline drives how energy is discussed in the classroom, potentially leaving students (and/or instructors) with siloed understandings that lack coherence. For example, Kohn et al. (2018) interviewed students who were co-enrolled in introductory chemistry and biology, finding that several students actively separated their ideas about energy (particularly in relation to changes during bond breaking

and forming) between their chemistry and biology courses. One student said, "*I know for biology what [the instructor] wants us to say, and then for chemistry what we have to say"*.[3]

Previous work has documented the common non-canonical idea that breaking bonds releases energy,[14–23] frequently attributing this idea (and/or its persistence) to the ways in which ATP is represented/discussed in biology courses. Research attempting to mitigate this idea (in chemistry and physics education research) has proposed activities or instructional materials that focus on the energy changes associated with ATP-mediated processes, by discussing a mechanism involving ATP hydrolysis (that is ATP + H_2O → ADP + Pi).[16,18,23] Indeed, ATP hydrolysis is highly exergonic, which is said to provide the energy to drive unfavorable reactions. The exergonic nature of the hydrolysis is explained using bond energies: the weak bond to phosphate in ATP is broken, and new stronger P-O bonds are formed in Pi and ADP. However, ATP rarely hydrolyzes under cellular conditions, and never does so when used in reaction coupling. This mechanism of energy release from hydrolysis begs the question of how the energy is used to drive the unfavorable reaction. In fact, most ATP-mediated processes involve a mechanism by which a phosphate group is transferred to a substrate or reactant, that then undergoes a conformational change, or further reaction. If we consider this in terms of energy, the energy is transferred through the reacting system rather than released.

In an earlier study, in which we interviewed a range of biology and chemistry instructors about the function of ATP, we found that the majority of instructors tend to discuss mechanisms involving energy release via ATP hydrolysis.[24] While some did discuss energy transfer (via phosphorylation by ATP), we also found that some instructors had dissatisfying experiences with teaching and learning about ATP. They were frustrated and uncomfortable

with approaches they were using. These results highlight the importance of finding an effective, appropriate, and satisfying way to talk about ATP in undergraduate introductory biology courses.

As a result of work highlighting disconnects between undergraduate chemistry and biology, course transformation efforts at our institution have aimed to better align these disciplines – that is, how can we better support students to build coherency between their chemistry and biology courses? One major aspect of these efforts involves engaging students in three-dimensional learning so as to support their construction and use of knowledge.[13] That is, rather than memorizing a set of facts for a given course, students are frequently prompted to explain, predict, and model phenomena using core ideas, thereby developing habitual engagement in these practices. Specifically, we focus on helping students explain how and why phenomena occur by activating productive conceptual and epistemological resources that will support powerful mechanistic explanations/predictions. In this study, we investigated the effectiveness of a learning task in building students' use of mechanistic resources as they explain the role of ATP in driving unfavorable reactions.

Theoretical Framework

We approach this work using Hammer's resources perspective of learning and knowledge construction.[25] According to Hammer (2000), resources represent fine-grained conceptual or epistemological knowledge elements that are connected or disconnected in a dynamic knowledge framework. These resources, given their dynamic and context-dependent nature, are activated according to different situations, one of which could be assessments in a classroom. Assessments have been shown to activate different resources based on the type and

amount of scaffolding included (Noyes, 2022).[26–28] For example, when shown a glucose molecule (with several O-H bonds), many students activated the idea of hydrogen bonding to explain how the molecule would bind to a protein binding site, without an explicit discussion of the electrostatic forces underpinning the interaction; however, when the authors swapped the glucose molecule for a positive magnesium ion (Mg^{2+}), far fewer students discussed hydrogen bonding and instead activated resources associated with electrostatic forces as the cause for Mg^{2+} binding.[26]

We use this resources perspective as a theory by which to support students' mechanistic reasoning in their chemistry and biology courses. Mechanistic reasoning, or explaining/predicting how and why phenomena occur, has been theorized as a productive thinking strategy by several scholars.[29–33] We have previously leveraged frameworks outlined by Krist et al. (2018) and Russ (2008); however, as we have expanded into explaining complex phenomena (i.e., phenomena that rely on connecting more than one or two core ideas/concepts or can be explained using different principles), we have found a more general framework better fitting. In general, all of the MR frameworks agree upon two components required for a mechanistic explanation: (1) identifying relevant entities, and (2) unpacking the activities of those entities which ultimately give rise to the target phenomenon.[29–33] We have found that some stipulations, such as identifying the "scalar level below",[30] become difficult to implement when explaining complex phenomena that can be explained using a number of different constructs. Amidst our struggles to apply a given MR framework to the explanation of complex phenomena, we came across the well-articulated article by Haskel-Ittah that

emphasizes the importance and utility of explanatory black boxes in science education (particularly biology).[34]

An explanatory black box is "*a 'unit' within a mechanistic explanation in which the entities, their activities and interactions are unknown to the person receiving the information or to the person constructing the explanation, such that a gap is created between one step in the process and the next*".[34] Haskel-Ittah points out the utility of explanatory black boxes when constructing mechanistic explanations for phenomena, particularly in education as each mechanistic/causal link may not be relevant to the topic of instruction; however, she also emphasizes the importance of explicitly recognizing these black boxes, so as to avoid the illusion of explanatory depth, or the belief that one fully understands causally complex phenomena, when that may not be the case.[34]

In this work, we focus on the ATP-driven formation of glutamine from glutamate and ammonium, a commonly occurring coupled reaction in biological systems, which can serve as a relatively simple model for many processes that occur via phosphate transfer. The mechanism is often omitted, or implicitly black boxed.[24] We chose to focus on reaction coupling, as the mechanism directly connects to ideas discussed in chemistry including reactivity, high-energy intermediate molecules, and predicted favorability of reactions, while carrying deep significance in biology courses (typically units on metabolism). We aimed to support students' use of mechanistic resources by using a modified version of evidence-centered design[35] to develop and refine a task that leverages the benefits of scaffolding in supporting students to integrate new ideas with their prior knowledge/resources.[27,36,37] In particular, the goal of the

task was to help students clarify and refine their prior ideas about ATP as an energy source by focusing on the mechanism of energy transfer (i.e., the role of ATP as a phosphorylating agent).

Conceptual Framework

Given our interests in mechanistic reasoning, resources, and connecting ideas across chemistry and biology, in this work we maintain a particular conceptual framework regarding the role of ATP in reaction coupling. While cells are complex systems and cellular phenomena result from several co-occurring events, we chose to focus on the role of ATP as a phosphorylating agent in coupled reactions. This critical event (phosphorylation) drives the direction of numerous processes in cells, one of which being the ATP-driven formation of glutamine from glutamate and ammonium; however, there are certainly other events that contribute to this outcome. For example, enzyme functionality, cellular concentrations of reaction components, and entropic effects all play a role. For the purposes of this research, we focus solely on ATP transferring a phosphoryl group and the resulting effect(s) of phosphorylation (specifically, energetic/reactivity changes).

Research Questions

Previous research has focused on ATP hydrolysis to support students' explanations for how energy is released;[16,18,23] however, explaining energy release leads to inexplicit use of explanatory black boxes for how energy is *transferred*, thereby leaving students in the dark about the mechanistic role of ATP in driving unfavorable reactions.[24] If the ATP hydrolysis approach is used, we believe that explicitly recognizing it as an explanatory black box is necessary in order to prepare students for future learning. However, there is no existing evidence as to whether introductory biology students can explain the mechanism of energy

transfer (i.e., the role of ATP as a phosphorylating agent and subsequent effects). Thus, we developed a task designed to support students' understanding of these ideas. After developing and implementing the task, we analyzed Molecular Biology and Organic Chemistry student responses to answer the following research questions:

1. How do students explain the role of ATP in driving the unfavorable formation of glutamine from glutamate and ammonium?
2. What is the impact of the task on students' use of mechanistic resources when explaining this phenomenon?
3. How do Molecular Biology student explanations compare to Organic Chemistry student explanations for this phenomenon?

Methods

Participants And Data Collection

We collected responses at the end of the Fall 2022 semester from students enrolled in Introductory Cell and Molecular Biology (MB) and Organic Chemistry I (OC). Both MB and OC are large-enrollment courses offered at a large, research-intensive midwestern university. Students were provided a small amount of extra credit for completing the task (described in the next section) on an online assessment platform called beSocratic, which allows students to work through a series of slides (similar to PowerPoint) and construct explanations.[38] Students could opt out of the research but still complete the task for extra credit if desired; we removed these students from our analysis. We also removed blank or incomplete responses (e.g., "*I don't know*"). Finally, we asked both an initial and final question about the role of ATP (to address RQ2), so we confirmed that each student had constructed both of these explanations,

removing those who only had one or the other. This left a total of 440 OC explanations (220 students) and 1,352 MB explanations (676 students).

Task And Coding Scheme

In this section, we describe key aspects of the ATP task, which aimed to support students' use of mechanistic resources when explaining how ATP drives the formation of glutamine from glutamate and ammonium. Ultimately, we leveraged scaffolding tactics to avoid the rather convincing, incorrect, and pervasive idea that ATP ***is*** energy so that students are less likely to state that ATP *provides* or *releases* energy that can be used, thereby black boxing (i.e., not unpacking or explaining) how energy is transferred and the mechanistic events contributing to the phenomenon. Then, we describe how we characterize responses using a carefully designed coding scheme, which includes both specific, conceptual resources that students used in their explanations as well as holistic themes we generated to inform broader research and practice implications.

The full final task is shown in Appendix A. To assess whether the task impacted student understanding of the role of ATP, we asked students both an initial and final free-response question about how ATP drives the formation of glutamine. The first few slides of the task contain general information to prepare the students for their consideration of this phenomenon. Then, in the initial prompt, we provide students with the image shown in Figure 7.1, followed by the prompt: *How do you think ATP drives this reaction? Specifically, what sequence of events occurs when ATP is added to the reactant system in order to produce glutamine?* We added the second question in this prompt to elicit the depth of students' understanding.

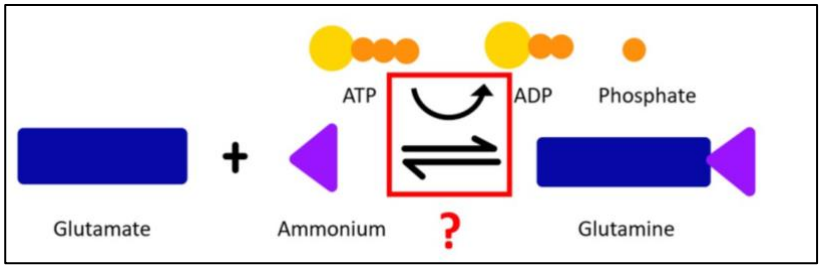


Figure 7.1. ATP driven formation of glutamine cartoon shown to students with initial prompt.

We will refer to responses to this question as students' "initial explanations". The subsequent slides introduce the common intermediate, phosphorylated glutamate, with questions probing ideas about reactivity and favorability. For example, we provide students with the general favorability of two reactions: (1) glutamate reacting with ammonium (unfavorable), and (2) phosphorylated glutamate reacting with ammonium (favorable) (Figure 7.2), followed by a question about relative reactivity.

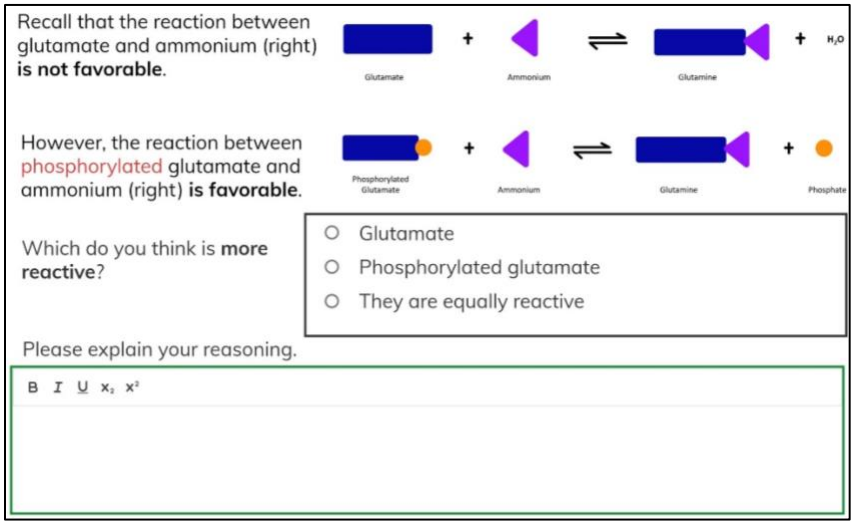


Figure 7.2. Questions prompting students to think about reactivity and favorability regarding the formation of glutamine.

The final prompt asks students the following question: *How was ATP used to "drive" the formation of glutamine? It is not enough to just say that ATP provides energy. Be sure to explain* **how** *ATP is involved in the formation of glutamine*. We noted that it is not enough to just say that ATP provides energy, in an attempt to elicit the mechanistic ideas that students held (from prior knowledge and/or from newly integrated ideas stemming from the task).

Characterizing Student Explanations

Authors CF and DL developed the coding scheme using responses (see example responses in Appendix B) from the second task version (Molecular Biology (N = 100) and Biochemistry (N = 79 students)), which slightly varied from the final version (Appendix A) in that the second version contained references to Gibbs free energy. Because of the variations in the task, we refined and finalized the coding scheme using 80 responses from the final task version. Our adherence to mechanistic reasoning and Hammer's resources perspective largely guided the development and refinement of the coding scheme. For example, we were interested in the mechanistic resources that students leveraged (such as linking ATP to phosphorylation or phosphorylation increasing the reactivity of the glutamate); however, we were also interested in other ideas that students leveraged. For example, many students included general ideas about ATP as an energy source, some mentioned activation energy, ATP hydrolysis, or a change in structure/conformation. These ideas were not explicitly addressed at any point in the task, and, thus, they likely represent pre-existing knowledge that the task activated in some way (this is not surprising as many, if not all, of the students had experienced learning about ATP in some capacity prior to completing this assignment).

The analytic coding scheme included 10 bins (Table 7.1), each of which is coded as "yes" or "no" depending on the evidence provided in the student response. The ideas students leveraged can be broken down into these conceptual resources; however, we also wanted to characterize their responses holistically. We have categorized some of the analytic bins as representative of "mechanistic" (shaded grey in Table 7.1) and others as "non-mechanistic" resources; however, we wanted the holistic characterizations to reflect students' conceptual frameworks. Further, more often than not, the "non-mechanistic" resources are not incorrect, and we do not intend for the category to be perceived as "worse" or "less sophisticated". Thus, we collapsed student responses, based on the presence or absence of specific analytic bins, into three overarching themes: "Phosphate transfer", "Energy release", and "Other". The essence of these themes are discussed in the Results section for RQ1.

Table 7.1. Analytic coding bins and corresponding student examples.

Coding Bin	Example response
Link ATP to phosphorylation	"*ATP provides glutamate with the phosphate it needs…"* – 1_1027 (OC)
Phosphorylation increases energy/reactivity	*"the phosphorylation of glutamate raises the free energy of the reactants in the first reaction in the sequence"* – B1048 (MB)
Phosphorylation increases favorability	*"ATP donates a phosphate group to the glutamate and that phosphate group makes the reaction favorable."* – 1_1021 (OC)
Change structure	*"ATP changes the structure of glutamate by giving it a phosphate. This new structure allows the reaction to become favorable and continue until glutamine is formed."* – 2_1019 (OC)
Making a reaction favorable via ATP	*"ATP makes this unfavorable reaction favorable as the dephosphorization [sic] of atp is favorable"* – 1_1089 (OC)
ATP as a source of energy	*"The ATP is what supplies the energy for the reaction"* – 1_1020 (OC)
ATP hydrolysis	*"…when this reaction is coupled to the favorable hydrolysis of ATP to ADP the unfavorable reaction turns into a favorable reaction."* – B1436 (MB)
Activation energy	*"ATP gives it a Phosphate, which allows the activation energy of the final reaction to be much lower."* – 2_1229 (OC)
Equilibrium	*"When there is a high cellular concentration of ATP, equilibrium shifts to produce more of the product (phosphorylated glutamate) …"* – B1511 (MB)
ATP "other"	*"It is the creation of ATP from ADP using energy from sunlight, and occurs during photosynthesis. ATP is also formed from the process of cellular respiration in the mitochondria of a cell…"* – B1375 (MB)

Note: Grey highlighted cells represented mechanistic resources.

Authors CF and DL used the analytic coding scheme shown in Table 7.1 to code 95 responses independently in order to determine acceptable inter-rater reliability. We used

Cohen's Kappa (κ) to determine our agreement for each analytic category, as the codes are ordinal and non-mutually exclusive.[39] The κ values for these categories ranged from 0.732 (substantial agreement) to 1.000 (perfect agreement).[40] We also calculated κ values for the three holistic categories (Phosphate transfer (κ = 1.000), Energy release (κ = 0.963), and Other (κ = 0.951)). For these 95 responses, CF and DL discussed disagreements to determine 100% consensus. Author DL coded the remaining responses independently, discussing edge cases with CF to determine the final codes used in our analysis. When coding the OC responses, we were aware whether it was an initial or final explanation; however, we were not aware of initial versus final for the MB responses. To test the significance of our findings, we used Pearson's Chi-Square (χ^2) tests, Cramér's *V*, and McNemar tests, each of which was calculated using SPSS Statistics Version 28.[41]

Results

We have organized the results to start by providing a detailed account of student responses. To answer RQ1, we focus on the different conceptual ideas that students included in their explanations, followed by a description of three holistic categories that we generated based on common themes. In these results, we show a general distribution of the ideas students included in their explanations to give an overall picture of what we saw in the data. Then, to answer RQ2, we discuss how students' responses changed from initial to final using the holistic characterizations that are described in the RQ1 results. In this section, we share results from individual students, showing examples from the most common change that we observed between initial and final explanations, thus identifying the impact that the task has on students understanding this phenomenon. Lastly, because we sampled from both MB and OC,

we compare the responses from students in each course to answer RQ3 in support of our interdisciplinary efforts.

Research Question 1: How Do Students Explain The Role Of ATP In Driving The Unfavorable Formation Of Glutamine From Glutamate And Ammonium?

Capturing Students' Use Of Conceptual Resources

We asked students both at the beginning and end of the activity how they think ATP "drives" the formation of glutamine. In analyzing responses to both questions, we found that students leveraged a range of ideas, which we captured using our analytic coding scheme (Table 7.1). These codes included ideas such as the role of ATP as a phosphorylating agent (*link ATP to phosphorylation*), ATP generally "providing" energy (*ATP as a source of energy*), the effects of phosphorylation (*phosphorylation increases energy/reactivity*, or *phosphorylation increases favorability*), etc., with the number of codes for each explanation varying as students could include any number of these ideas in their responses.

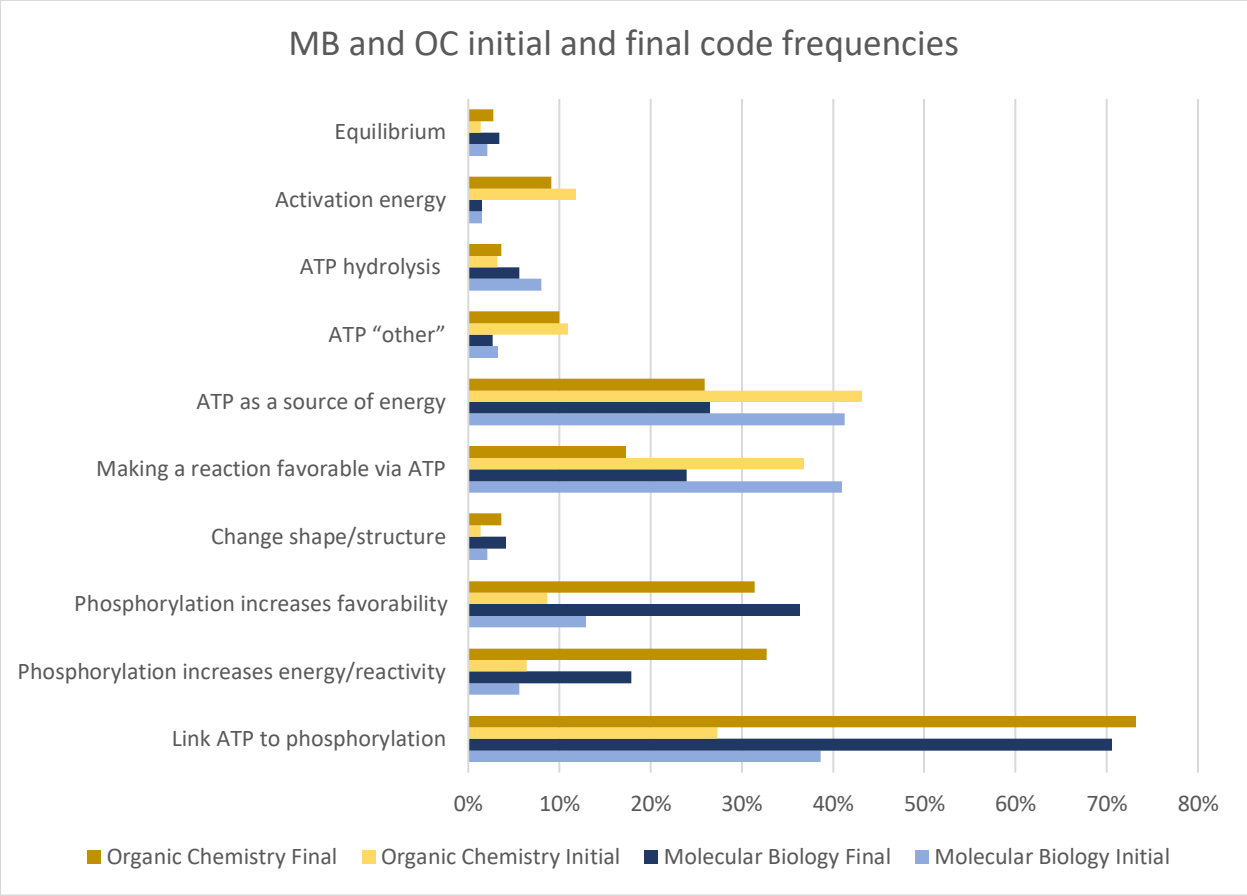


Figure 7.3. Frequency of codes for initial and final explanations in Organic Chemistry and Molecular Biology.

This analytic coding scheme allowed us to capture the fine-grained conceptual resources that students used to construct their explanations. For example, the idea that ATP is a phosphorylating agent is a mechanistic idea that many students used in their reasoning (particularly in their final responses as indicated by the dark yellow and dark blue bars in Figure 7.3). We consider this resource as mechanistic, because it involves an "activity" of an important "entity" (ATP) in this phenomenon or a relationship between two entities: the ATP and the reactant (glutamate). A smaller number of students included additional mechanistic resources, such as the effect that phosphorylation has on glutamate. By recognizing that phosphorylation

increases the energy of the reactant, 33% of OC students and 18% of MB students (in their final explanations) identified an additional entity (the phosphorylated glutamate) and an important property of that entity (its relative potential energy). These ideas support a fully mechanistic explanation for how ATP drives an unfavorable reaction.

Other frequent resources that students leveraged include generally linking ATP to energy or favorability. For example, many students in both courses identified ATP as a source of energy (43% OC and 41% MB initial responses), a productive idea for a macroscopic prediction (i.e., by "providing" energy, the reaction will favor products), but less productive for a mechanistic explanation (i.e., *how* ATP is involved). We also noticed some (<10%) students invoking ATP hydrolysis, which was intentionally excluded from the task design. Thus, we believe this is due to prior experiences students have had with learning about ATP (as an energy source).[24]

Capturing these ideas provide a detailed picture of what students leveraged in their reasoning; however, it is difficult to use this information practically (i.e., as instructors making holistic conclusions about students' conceptual understandings). The holistic characterizations, which we discuss next, provide a more general depiction of students' conceptual frameworks regarding the role of ATP, allowing us to draw more meaningful implications for research and practice.

Holistically Characterizing Student Responses

Mechanistic Resources: "Phosphate Transfer"

The responses characterized as "phosphate transfer" were those that incorporated the mechanistic resource of ATP acting as a phosphorylating agent (code: *link ATP to*

phosphorylation). These responses may have also referred to ATP as an energy source or included non-mechanistic resources. For example, consider student B1154 (MB) who wrote, "*ATP transfers a phosphate group to the reactant side, through the process called phosphorylation. The release of energy from ATP goes to the reaction and provides the energy for the unfavorable reaction to happen. Without the added energy from ATP the reaction would not happen*." Their first sentence clearly indicates an understanding of ATP transferring a phosphoryl group (code: *link ATP to phosphorylation*), but the following two sentences discuss ATP providing energy in a more general, non-mechanistic way (code: *ATP as a source of energy*).

Because of the large number of students who linked ATP to phosphorylation (over 70% of final responses from both OC and MB), we further divided this group into sub-categories (Figure 7.4) using our analytic scheme. The first group linked ATP to phosphorylation but did not discuss the effect(s) that phosphorylation has on the reactant or reaction (25% of all final responses). A second group of students linked ATP to phosphorylation and explained that the phosphorylation causes a change in structure (an idea not explicitly included in the task, thus indicating it as prior knowledge that some students activated). Only 34 students (< 4%) used this in their final responses; thus, with so few responses, it is unclear how or why students might be leveraging this idea, and whether it is being used productively. Some students do seem to productively leverage the core concept of structure-property-function to build their explanation, such as student 1046 (OC) who wrote, "*ATP is used to drive the formation of glutamine by providing energy. ATP does this by phosphorylating the glutamate which <u>changes the shape</u> of the molecule allowing it to be in a more reactive and favorable state*." Other

students discussed structural aspects more implicitly, such as student 1094 (OC) who wrote, "*Once the glutamate is phosphorylated the phosphorous can act as a leaving group and the ammonia can attach*," bringing in explicit structural resources from organic chemistry. Lastly, some students might be leveraging electrostatic forces when considering the effects of phosphorylation. For example student B1881 (MB) wrote, "*The phosphate in ATP is transferred to other molecules to make new unstable bonds. The unstable bonds are weak and easy to break, while the final bonds are strong and stable*."

The responses including ideas about structure, while there are few, contain interesting ideas that represent students' prior knowledge, as the task primarily aimed to activate resources related to energy/reactivity. Students that productively linked phosphorylation to energy or reactivity changes, in addition to linking ATP to phosphorylation, make up the largest Phosphate transfer sub-category (42% of all final responses). These responses reflect the ideas that our task aimed to activate so that students might productively link them to the phenomenon of ATP driving an unfavorable reaction. For example, some students correctly indicated that phosphorylating the glutamate increases its reactivity, such as student 1466 (MB) who wrote, "*The ATP creates a intermediate product by adding a phosphor [sic] group to the glutamate and the phosphorylated glutamate is at a higher energy state and can react with the ammonium…*" Students in this group may have also indicated that phosphorylation makes the overall reaction more favorable. For example, student 1051 (OC final) wrote "*ATP drives the formation of glutamine by donating a phosphorous, thereby phosphorylating the glutamate and making it more reactive, which in turn makes the reaction favorable*."

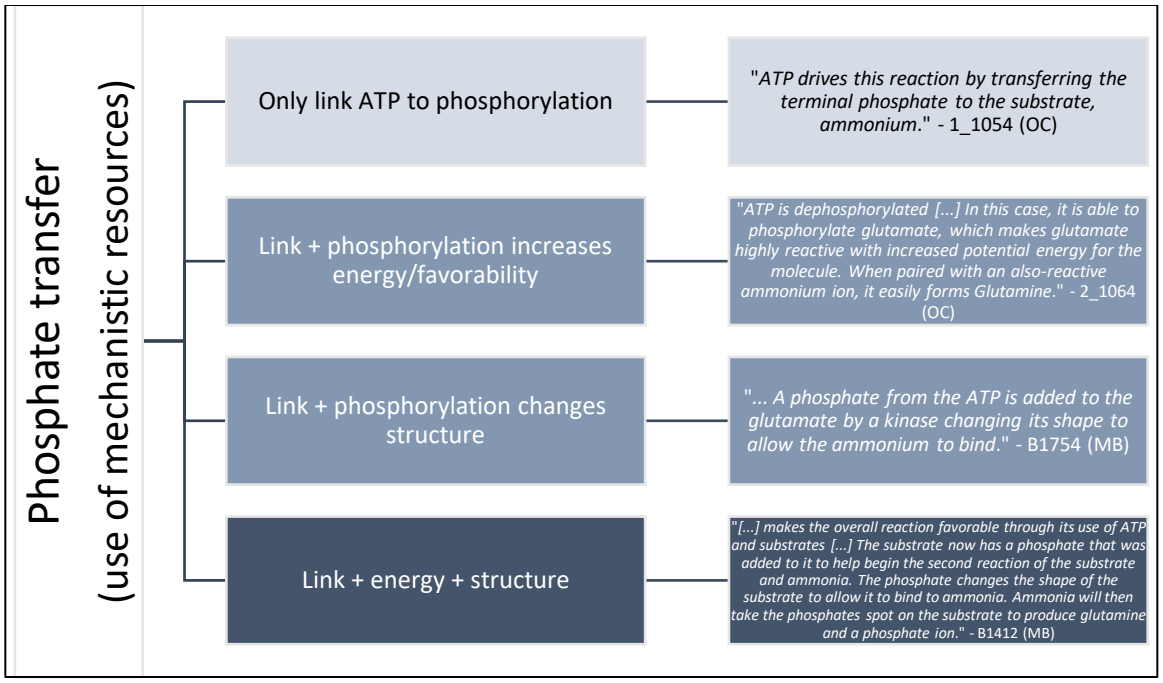


Figure 7.4. Subcategories of the Phosphate transfer group – darker shaded categories indicate inclusion of more mechanistic resources.

Non-Mechanistic Resources: "Energy Release"

Student responses that we characterized as "Energy release" were those that linked ATP to energy or favorability (analytic bins: *making a reaction favorable via ATP*, and/or *ATP as an energy source*), and did *not* link ATP to phosphorylation (nor did they explain the effects of phosphorylation). For example, student 1161 (OC) wrote, "*The breakdown of ATP to ADP releases energy which can then be used to assist the reaction*." Responses in this theme (20% of final responses) black-boxed the mechanism of energy transfer – however, we cannot be sure whether some of these students (1) intentionally chose to leave the mechanism black-boxed (perhaps believing this was sufficient for a complete answer) or (2) were unaware of this mechanism (initial responses) or did not follow the task to sufficiently understand the mechanism (final responses) and thus, their response reflected the extent of their

understanding. We believe it is a result of the latter explanation, as our previous work shows that students tend to give their best effort on formative assessment tasks and extra credit tasks, so we are inclined to believe that their written responses are an appropriate reflection of the depth of their understanding.[42]

Non-Mechanistic Resources: "Other"

We labeled the last category of responses as "Other", which included non-mechanistic ideas – that is, students leveraged resources that were not directly related to the task or phenomenon, or constructed explanations with vague/inconclusive ideas that we did not feel confident placing in the other categories. For example, consider student B1255 from MB who wrote, "*The sequence of events that occur when ATP is added to the reactant system to produce glutamine is some sort of synthesis, for example it could be when one molecule of glutamine enters the mitochondria*." Similarly, student 1056 from OC wrote, "*ATP is involved in the formation of glutamine as the ATP is the important part for the start of the formation of glutamine. If ATP was absent in the process, the process would stop and not proceed*." Neither of these explanations provides mechanistic information regarding how ATP is involved in the formation of glutamine. The first student seemed to be incorporating prior knowledge, but they did not link their prior knowledge to any mechanistic ideas related to reaction coupling; the second response gives a vague description of the involvement of ATP without providing ideas for *how* or *why* it is involved.

Figure 7.5 shows the overall percentage of final responses (both OC and MB) in each category, with the phosphate transfer group (blue shades) broken into its four sub-groups (see N values for each course in Table 7.2). It is not surprising that such a small number of students

leveraged structural changes in their explanations, as we did not scaffold for the activation of these ideas; however, the fact that students included it without being prompted to do so, suggests its potential utility in developing a mechanistic understanding of ATP, something that we plan to explore in future work.

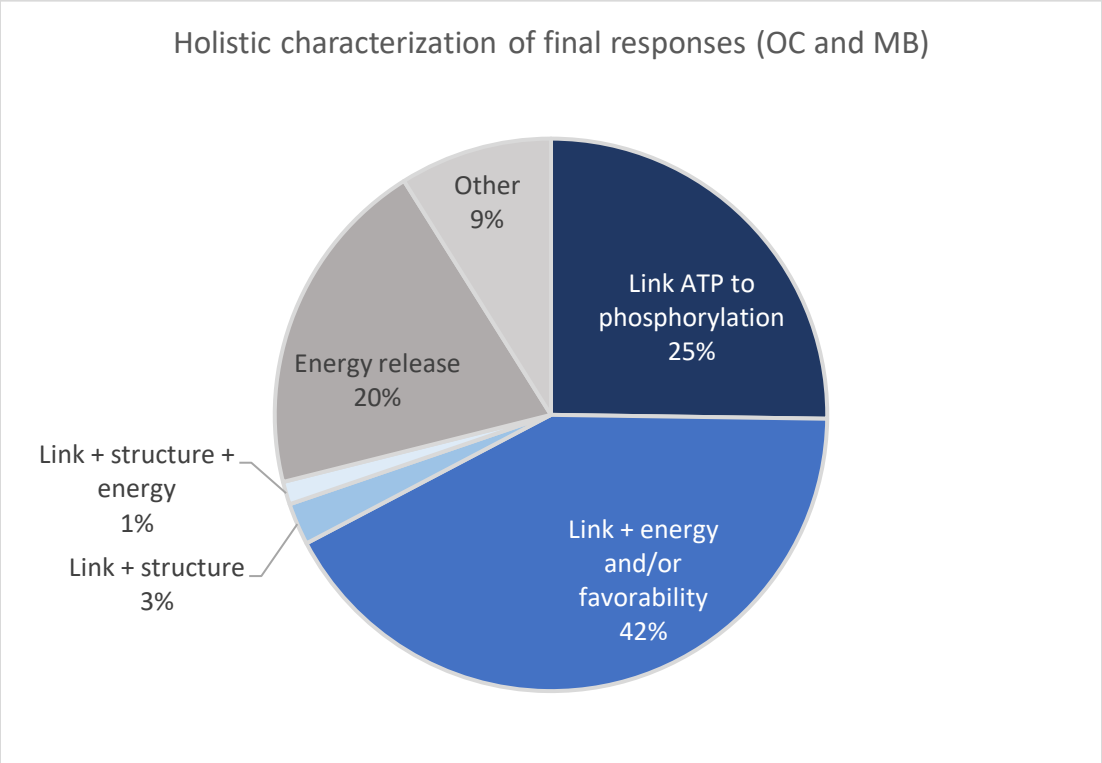


Figure 7.5. Distribution of final responses (both OC and MB; N = 896) according to holistic characterization, with the Phosphate transfer group broken into its four subcategories.

Table 7.2. Number (and %) of final responses (OC and MB) in each category.

Category	Number (%) of OC final responses	Number (%) of MB final responses	Total number of responses
Link ATP to phosphorylation (Link)	51 (23.2%)	175 (25.9%)	226
Link + energy and/or favorability	102 (46.4%)	275 (40.7%)	377
Link + structure	5 (2.3%)	17 (2.5%)	22
Link + structure + energy	3 (1.4%)	9 (1.3%)	12
Energy release	31 (14.1%)	148 (21.9%)	179
Other	28 (12.7%)	52 (7.7%)	80

Research Question 2: What Is The Impact Of The Task On Students' Use Of Mechanistic Resources When Explaining This Phenomenon?

Because we asked students to explain this phenomenon both at the beginning and end of the task, we could investigate whether working through this task supported students in building a mechanistic explanation for how ATP drives the unfavorable formation of glutamine from glutamate and ammonium. Using McNemar's test, we found that both the MB students and the OC students significantly (MB: $p < 0.001$, OC: $p < 0.001$) changed their explanations so that they included more mechanistic resources (Phosphate transfer group) in their final explanations. For this statistical test, we combined the Energy release and Other categories, since neither category includes mechanistic resources (for this phenomenon). Appendix B (Tables 7.5 and 7.6) provides the number of students in each course whose responses did not change or changed to include (or not include) mechanistic resources. We also created Sankey diagrams (Figures 7.6 and 7.7) to show the individual OC and MB student changes from initial to final according to each of the three categories (Other, Energy release, and Phosphate transfer).

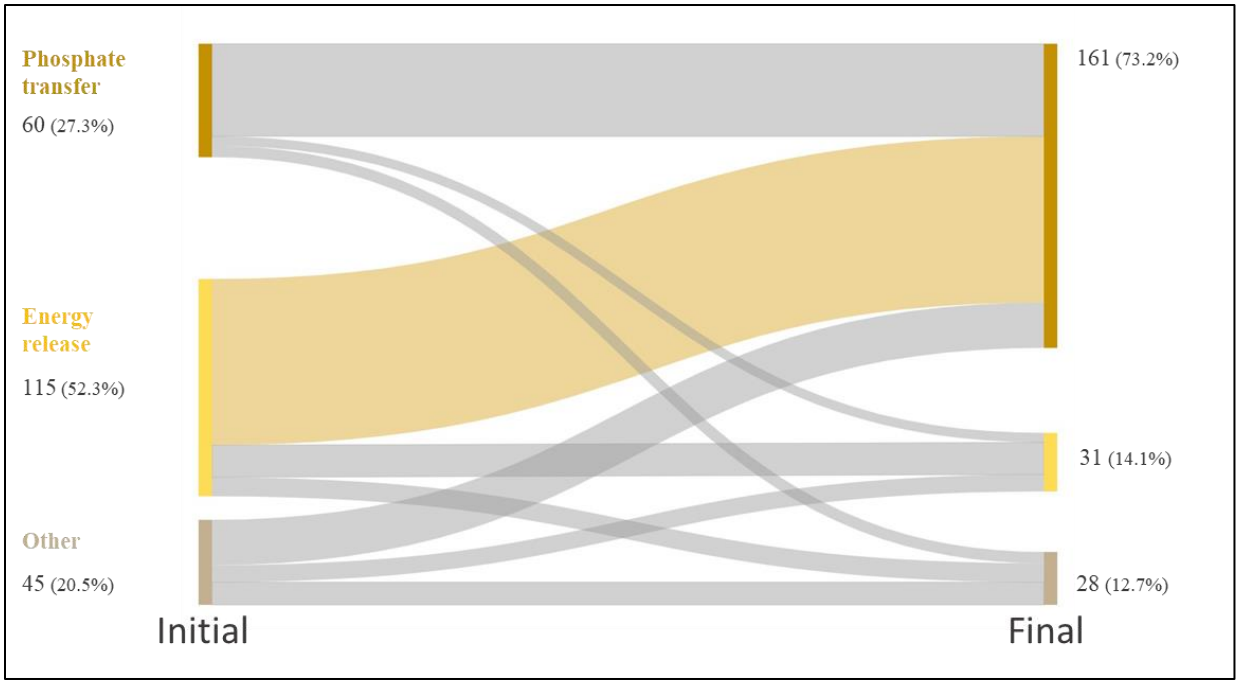


Figure 7.6. Individual changes from Organic Chemistry I students' initial and final explanations.

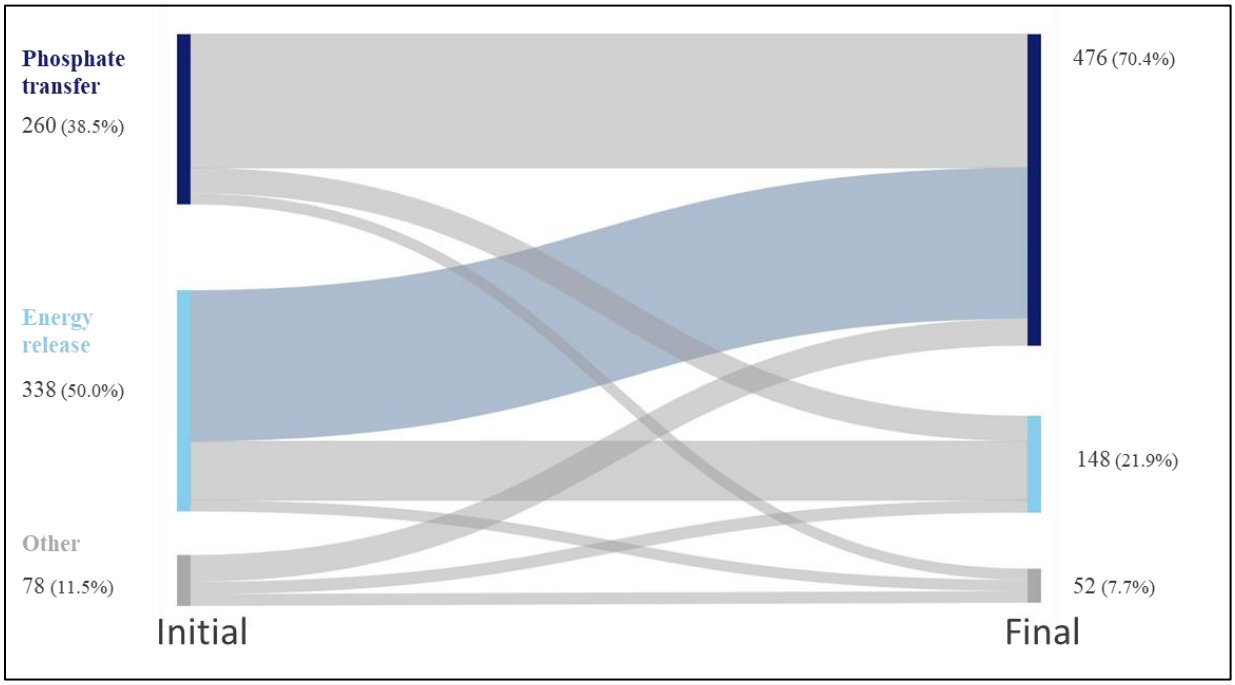


Figure 7.7. Individual changes from Molecular Biology students' initial and final explanations.

The most common change (colored flows in Figures 7.6 and 7.7) in students' explanations, was to shift from an Energy release explanation to a Phosphate transfer explanation (MB: N = 231, 34%; OC: N = 88, 40%; total: 35.6%). For example, consider the OC and MB student explanations shown in Figure 7.8. Both students' initial explanations placed them in the Energy release category. However, after working through the task, both students identified the role of ATP in transferring a phosphate to glutamate. The OC student recognized the role of ATP in phosphorylation, but they did not explain the effect that phosphorylation has on the reactant or the reaction overall (i.e., changing the glutamate's structure, energy, or reactivity, or changing the favorability of the reaction). The MB student did build this connection, and their response included several productive mechanistic ideas and connections (Link + energy + structure).

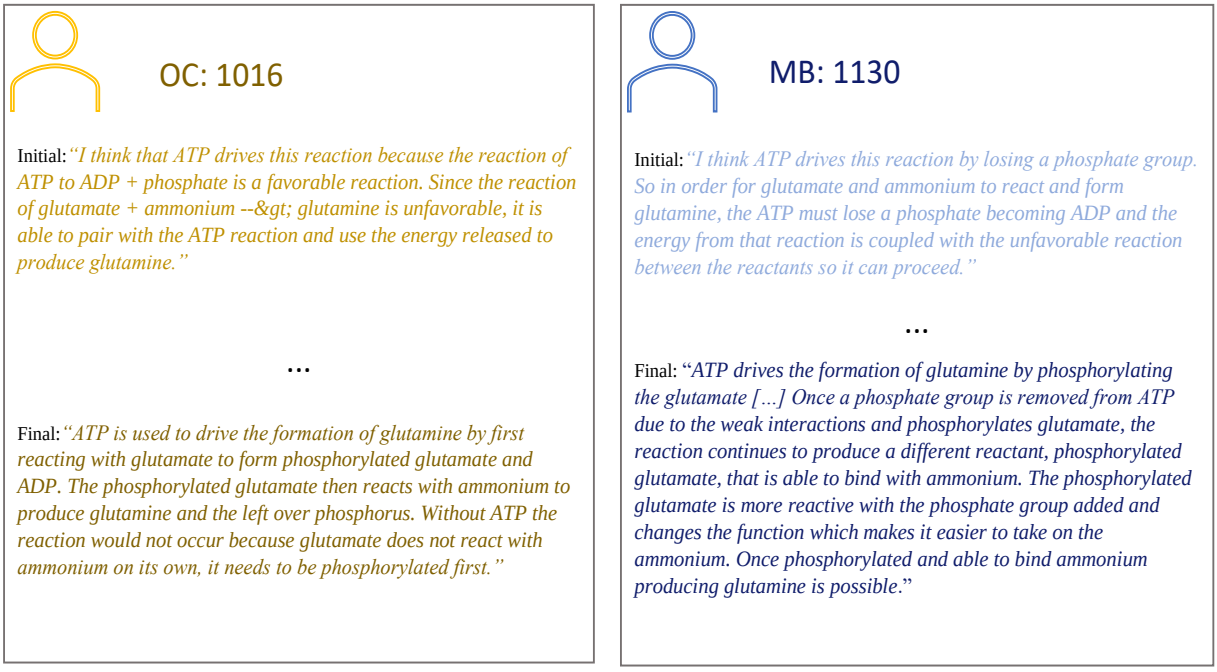


Figure 7.8. Example students whose explanations changed from Energy release (initial) to Phosphate transfer (final).

We posit that these students integrated their prior knowledge about ATP providing energy (an explanatory black box) with new ideas (presumably from the task) about the mechanism by which ATP does this, i.e., by transferring a phosphoryl group. This coherent integration of knowledge is ultimately our goal; however, not all responses were as well-constructed as these. For example, some students used both prior ideas and ideas from the task in their final explanations without linking these ideas coherently. We provide example responses in Table 7.3 to explicate this important finding.

Table 7.3 Examples of incoherent "Phosphate transfer" explanations.

Student ID	Final Explanation
B1198	*"A phosphate is removed from ATP and ADP and a phosphate is produced which is favorable because ATP is breaking down which is couples to this unfavorable reaction to produce glutamine in the forward reaction. The breaking down of ATP into ADP and phosphate is what is needed in order for the forward reaction to occur. The glutamate will be phosphorylated and then used to produce glutamine when the phosphorylated glutamate reacts with ammonium"*
2_1123	*"ATP drives the formation of glutamine by undergoing an exergonic reaction. When ATP is reduced to ADP, it is dephosphorylated which causes energy to exist the system in the form of phosphate. The phosphate in the cellular environment attaches to the glutamate, providing the energy needed for it to react favorable with ammonium. This allows for the reaction to proceed in the forward direction, driving the formation of glutamine."*
B1302	*"ATP drives the formation of glutamine by providing the needed phosphate to phosphorylate the glutamate so it is able to bond to ammonium and form glutamine. additionally, ATP [→] ADP + Pi is a favorable process therefore it can pair with the formation of glutamine to let that reaction occur"*
B1917	*"ATP drives this reaction by giving the system energy that escapes when the bond between ADP and the phosphate group breaks. This event is called hydrolysis and in this event the ATP acts like a charged battery, releasing free energy to the reaction allowing for the glutamate and ammonium to bond and form Glutamine. The free phosphate group bonds to the Glutamate which leads to the Ammonium eventually taking its place"*
2_1143	*"he [sic] first reaction involves cleaving off a phosphate from ATP to make ADP + P. This action releases an immense amount of energy that can be used to drive the initial reaction to make phosphorylated glutamate. This intermediate is extremely reactive because of the added phosphate and will then react with ammonium to give us a final product of glutamine and a phosphate."*
B1720	*"ATP is used to drive the formation of glutamine in a couple of ways. Firstly, the overall reaction that produces glutamine is unfavorable by itself, so it must be coupled to a reaction that is favorable. In this example, that favorable reaction would be the dephosphorylation of ATP. When a phosphate group is removed from ATP, that releases energy. That energy is then used to drive the reaction forward. Lastly, the phosphate group from the ATP is then attathed [sic] to glutamate, phosphorylizes [sic] it, and is further used as a reactant."*

Note: Orange highlighted text indicates what we assume to be prior knowledge (as these ideas were not included in the task), and the blue highlighted text indicate mechanistic ideas (either prior knowledge activated by the task, or newly integrated ideas from the task).

We found that this task resulted in an increase in the number of students explaining how ATP drives unfavorable reactions by leveraging its mechanistic role of phosphorylating (or "activating") a reactant (in this case, glutamate). However, these explanations represent a

range of coherent and incoherent integration of resources. We discuss the implications of these findings further in the discussion section.

Research Question 3: How Do MB Student Explanations Compare To OC Student Explanations For This Phenomenon?

Our final research question asks whether the responses from biology students (MB) differed from organic chemistry students (OC). There is some evidence that students see their chemistry and biology courses as disconnected.[3] However, since this publication, transformation efforts to better align chemistry and biology at this institution have been in effect, with introductory courses in both disciplines aiming to support students' mechanistic reasoning and use of prior knowledge (as opposed to siloed courses in which ideas in one are disconnected from those in another). Our recent work has shown that students co-enrolled in general chemistry and molecular biology do not construct different responses when situated in each course for a task about preferential protein-ligand binding,[43] supporting the effectiveness of the transformation efforts. To continue these efforts, we administered the current task to both biology and chemistry students (in this case, organic chemistry students, as this group of students had most likely already completed the MB course).

While a Pearson's Chi Square test showed there is a difference in final explanations between the MB and OC student explanations (χ^2 = 9.950; p = 0.007), the effect size is quite small (Cramer's V = .105), suggesting that this difference is not meaningful. We also measured whether there is a difference in MB and OC initial explanations. The results showed a significant difference (χ^2 = 15.600, p < 0.001), but, again, the effect size is small (Cramer's V = 0.132). The initial and final percentages for both courses are shown in Table 7.4, and the final percentages

of both courses are shown graphically in Figure 7.9. These findings suggest that there could be some minor context-driven differences (e.g., student 1094, who referred to the added phosphate as a good leaving group, a commonly discussed idea in OC); however, the overall distribution is similar – thus, the task has activated similar ideas for both groups of students, regardless of course context.

Table 7.4. Percent distribution of responses in MB versus OC (for initial and final explanations).

	MB initial	OC initial	MB final	OC final
Phosphate transfer	38.5%	27.3%	70.4%	73.2%
Energy release	50%	52.3%	21.9%	14.1%
Other	11.5%	20.4%	7.7%	12.7%

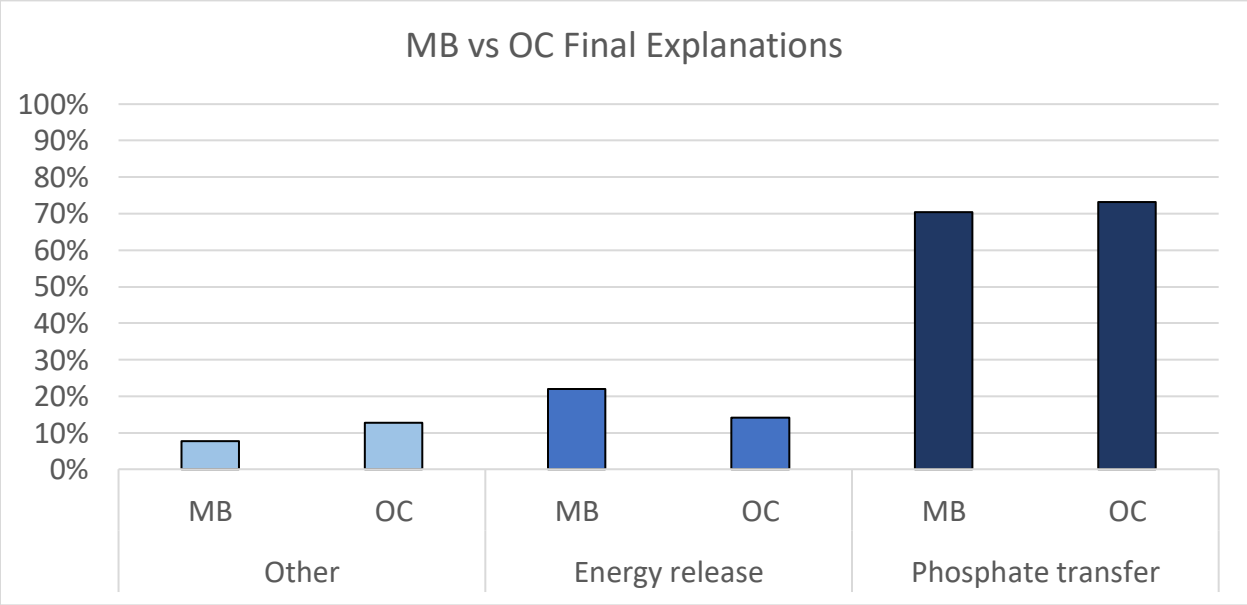


Figure 7.9. MB versus OC final explanations.

Discussion

We began this work as an effort to support students' mechanistic understanding of how ATP "provides" energy in common biological processes, like reaction coupling. The integral role

of ATP in biological systems is undisputed; however, our previous work (and a gap in the literature) suggests that this role may not often be explored using a mechanistic lens in introductory molecular biology courses.[24] It is our understanding that, traditionally, the role of ATP as a phosphorylating agent is inexplicitly black boxed – that is, representations and instruction leave students in the dark about *how* ATP (or the hydrolysis of ATP) provides energy to drive unfavorable processes by neglecting to explicitly recognize that there are additional mechanistic events and entities involved (such as a phosphorylated intermediate).[24,44] Explaining this phenomenon using a mechanistic reasoning lens provides an opportunity for students to leverage core concepts such as "energy transformations" or "structure & function"[11], which overlap significantly with the crosscutting concepts "energy" and "structure function" as outlined by the NRC (2012).[13] These themes are central to the scientific discipline and should be used as tools to support students' reasoning about complex phenomena; thus, it is imperative that, as educators, we provide the opportunity for them to do so. Our work aims to specifically help students build connections both within and between their undergraduate chemistry and biology courses, and these themes carry salient significance in both disciplines.

With all of this in mind, we were first interested in the ideas that students leveraged when asked to explain how ATP drives the unfavorable formation of glutamine from glutamate and ammonium. Using a resources perspective of learning to construct an analytic coding scheme, we captured the range of conceptual resources students used in their explanations. Some of these ideas emerged as prior knowledge – that is, the ideas were not explicitly included in the task, but were activated in the minds of some students. For example, some students leveraged the idea that phosphorylation would result in a structure, shape, or function

change, ultimately resulting in the formation of glutamine. Additionally, many students (particularly in their initial responses) discussed the role of ATP as an energy source (or as an entity which can make an unfavorable reaction more favorable). Alone, this latter idea, while productive and appropriate in some contexts, implicitly black boxes the mechanism of energy transfer, which, as noted, is a core concept in undergraduate biology courses. It was our goal to help students coordinate these general ideas about ATP and energy with the mechanistic events that result in their outcome. We found that most students, in their final explanations, generally relied on some combination of four mechanistic resources: (1) linking ATP to phosphorylation, (2) phosphorylation increasing the energy of the reactant, (3) phosphorylation increasing the favorability of the reaction, and/or (4) phosphorylation changing the structure of the reactant. While we expected to observe student use of the first three ideas, it surprised us to also see some students using ideas about structure-property-function (SPF) relationships in their explanations. The few students who used these ideas did so in varying ways, so we refrain from making any general claims here. However, if SPF relationships (and or forces/interactions) are leveraged productively, they may serve as a link between the role of ATP as a phosphorylating agent and the resulting energetic changes associated with coupled reactions.

To address the goal of helping students explain the mechanistic role of ATP (and therefore, integrate their prior knowledge with mechanistic resources), we asked students both at the beginning and end of the task to explain how ATP drives the formation of glutamine. In both courses, students' explanations significantly changed, with more Phosphate transfer explanations (i.e., use of mechanistic resources) in the final responses compared to initial responses. That is, by working through this learning task, students seem to have developed,

refined, and/or advanced their mechanistic ideas about the role of ATP. It is important to note that some students in the Phosphate transfer group had trouble coherently connecting their prior knowledge about ATP as an energy source to its role as a phosphorylating agent. For example, student B1917 wrote "*ATP drives this reaction by giving the system energy that escapes when the bond between ADP and the phosphate group breaks. This event is called hydrolysis and in this event the ATP acts like a charged battery, releasing free energy to the reaction allowing for the glutamate and ammonium to bond and form Glutamine. The free phosphate group bonds to the Glutamate which leads to the Ammonium eventually taking its place."* While the student has identified a phosphorylation event, they seem to be thinking about it as two isolated reactions – that is, the ATP first releases a phosphoryl group (prior knowledge about ATP hydrolysis), and *then* that group reacts with glutamate to form phosphorylated glutamate (rather than the more canonical explanation that ATP directly reacts (collides) with glutamate resulting in the phosphate transfer). Thus, while it is encouraging that a number of students recognize the importance of phosphorylation, integrating this idea with prior knowledge about ATP warrants additional investigation. Further, while the idea that breaking apart ATP results in energy release is often considered a misconception, we are aiming to activate resources that can be integrated with mechanistic ideas so as to help students build a robust framework of knowledge about the role of ATP. The goal is not to "replace" this misconception, but rather activate appropriate and productive resources that support students in their understanding of biological processes and energy transformations.

Lastly, we administered this task to students in both molecular biology (MB) and organic chemistry (OC), both of which have incorporated transformative efforts aimed at helping

students to (1) build connections between chemistry and biology and (2) engage in mechanistic reasoning about phenomena. The OC course was completely transformed for life science students – that is, it explicitly connects chemistry ideas to different biological phenomena – so as to incorporate three-dimensional learning throughout the curriculum.[45] In 2018, prior to MB course transformation efforts, Kohn et al. found that students actively separated their ideas about energy between introductory biology and chemistry;[3] thus, we hoped to see in this work that, regardless of course, there would be little difference in how students responded to the task (i.e., we have aimed to support students in using their knowledge to construct robust and thoughtful explanations about complex phenomena, regardless of course context). We found there to be a negligible difference in responses constructed by OC and MB students (the small effect size renders this "difference" rather meaningless). However, we predict that students in traditional curricula would respond quite differently to this task, as previous work has shown that students in transformed curricula (which emphasize mechanistic reasoning) are better able to construct mechanistic explanations.[42,46–49]

Conclusions And Future Directions

This work has shown that a carefully designed task can support learners to connect mechanistic resources with prior knowledge in order to engage in mechanistic reasoning about the role of ATP in reaction coupling. We have provided the full task in Appendix A with the hope that instructors across chemistry and biology courses will use and modify it to support students in refining these ideas.

Our previous work has shown that engaging fully in mechanistic reasoning strongly correlates with making correct predictions.[43] Thus, in future work, we plan to develop question

prompts that require activation, use, and refinement of these mechanistic ideas while making predictions about related phenomena. For example, how do students predict ATP will facilitate the transfer of ions across a membrane pump? We also recognize the complexity of phenomena involving ATP and, therefore, the utility of explanatory black boxes. Given the vast amount of information that introductory MB instructors are often tasked with covering, it may not always be reasonable to dive deep into a mechanism; however, in such cases, we urge instructors to make explanatory black boxes explicit, so as to best prepare students for future learning and avoid instilling an illusion of explanatory depth.[34] Many students adhered to the ideas associated with Energy release (thereby black boxing energy transfer), even after working through our task. ATP hydrolysis is pervasive in biology textbooks,[44] so it is not surprising that the general non-mechanistic descriptions of the role of ATP showed up in our data. It is possible that these students lost the thread at some point while working through the task, and then resorted to prior knowledge about ATP once posed the final question; however, we would need to conduct think-aloud-interviews to know this for sure.

Lastly, we are currently investigating whether this task (and other mechanistic reasoning tasks) are equitable. That is, does this task support/hinder students with different experiences/backgrounds? According to Ralph et al. (2022), exams with more mechanistic reasoning assessment items result in more equitable outcomes for students when compared to items assessing calculations or rote memorization.[50] Our future work will contribute to an understanding of the impact of MR tasks on all students. We also think it would be fruitful to interview students to better understand whether these tasks are meaningful for their lives/education, and, if so, how?

Limitations

The task we developed, while shown to increase student use of mechanistic resources, is perhaps not optimized. We included a slide that asked students how the concentration of ATP would impact the direction of the coupled reaction to activate ideas about Le Chetalier's principle. We expected that more final responses would include ideas about equilibrium; however, a very small percentage of students did so, and we will likely remove this from the task.

As noted, both courses from which we gathered student responses have been transformed to support student engagement in mechanistic reasoning. Thus, students in different courses or at a different university would likely construct different responses.

Upon coding the large set of data, author DL noticed that there may have been some copied or AI-generated responses. We did not remove these responses from our analysis, as it was not suspected until late in the coding process and was a negligible proportion (>1%).

REFERENCES

(1) Eckhard, J.; Rodemer, M.; Bernholt, S.; Graulich, N. What Do University Students Truly Learn When Watching Tutorial Videos in Organic Chemistry? An Exploratory Study Focusing on Mechanistic Reasoning. *J. Chem. Educ.* **2022**, *99* (6), 2231–2244. https://doi.org/10.1021/acs.jchemed.2c00076.

(2) Cooper, M. M.; Klymkowsky, M. W. The Trouble with Chemical Energy: Why Understanding Bond Energies Requires an Interdisciplinary Systems Approach. *CBE Life Sci Educ* **2013**, *12* (2), 306–312. https://doi.org/10.1187/cbe.12-10-0170.

(3) Kohn, K. P.; Underwood, S. M.; Cooper, M. M. Energy Connections and Misconnections across Chemistry and Biology. *LSE* **2018**, *17* (1), ar3. https://doi.org/10.1187/cbe.17-08-0169.

(4) Goldring, H.; Osborne, J. Students' Difficulties with Energy and Related Concepts. *Physics Education* **1994**, *29*, 26–32. https://doi.org/10.1088/0031-9120/29/1/006.

(5) Storey, R. D. Textbook Errors & Misconceptions in Biology: Cell Energetics. *The American Biology Teacher* **1992**, *54* (3), 161–166. https://doi.org/10.2307/4449438.

(6) Jewett, J. W. Energy and the Confused Student I: Work. *The Physics Teacher* **2008**, *46* (1), 38–43. https://doi.org/10.1119/1.2823999.

(7) Jewett, J. W. Energy and the Confused Student III: Language. *The Physics Teacher* **2008**, *46* (3), 149–153. https://doi.org/10.1119/1.2840978.

(8) Nordine, J. *Teaching Energy Across the Sciences, K-12*; National Science Teachers Association: Arlington, VA, 2015.

(9) Becker, N. M.; Cooper, M. M. College Chemistry Students' Understanding of Potential Energy in the Context of Atomic–Molecular Interactions. *Journal of Research in Science Teaching* **2014**, *51* (6), 789–808. https://doi.org/10.1002/tea.21159.

(10) Feynman, R. P.; Leighton, R.; Sands, M. *The Feynman Lectures on Physics*, New Millennium Edition.; Basic Books: New York, NY, 2011; Vol. 1.

(11) American Association for the Advancement of Science. *Vision and Change in Undergraduate Biology Education: A Call to Action*; Washington D.C., 2011. https://live-visionandchange.pantheonsite.io/wp-content/uploads/2011/03/Revised-Vision-and-Change-Final-Report.pdf (accessed 2020-04-17).

(12) Cooper, M.; Klymkowsky, M. Chemistry, Life, the Universe, and Everything: A New Approach to General Chemistry, and a Model for Curriculum Reform. *J. Chem. Educ.* **2013**, *90* (9), 1116–1122. https://doi.org/10.1021/ed300456y.

(13) National Research Council. A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas; National Academies Press: Washington, DC, 2012.

(14) Barker, V.; Millar, R. Students' Reasoning about Basic Chemical Thermodynamics and Chemical Bonding: What Changes Occur during a Context-Based Post-16 Chemistry Course? *International Journal of Science Education* **2000**, *22* (11), 1171–1200. https://doi.org/10.1080/09500690050166742.

(15) Boo, H. K. Students' Understandings of Chemical Bonds and the Energetics of Chemical Reactions. *Journal of Research in Science Teaching* **1998**, *35* (5), 569–581. https://doi.org/10.1002/(SICI)1098-2736(199805)35:5<569::AID-TEA6>3.0.CO;2-N.

(16) Dreyfus, B. W.; Sawtelle, V.; Turpen, C.; Gouvea, J.; Redish, E. F. Students' Reasoning about ``high-Energy Bonds'' and ATP: A Vision of Interdisciplinary Education. *Phys. Rev. ST Phys. Educ. Res.* **2014**, *10* (1), 010115. https://doi.org/10.1103/PhysRevSTPER.10.010115.

(17) Galley, W. C. Exothermic Bond Breaking: A Persistent Misconception. *J. Chem. Educ.* **2004**, *81* (4), 523. https://doi.org/10.1021/ed081p523.

(18) Green, A. I.; Parent, K. N.; Underwood, S. M.; Matz, R. L. Connecting Ideas across Courses: Relating Energy, Bonds & How ATP Hydrolysis Powers a Molecular Motor. *The American Biology Teacher* **2021**, *83* (5), 303–310. https://doi.org/10.1525/abt.2021.83.5.303.

(19) Hapkiewicz, A. Clarifying Chemical Bonding. Overcoming Our Misconceptions. *Science Teacher* **1991**, *58* (3), 24–27.

(20) Johnstone, A. H.; Mahmoud, N. A. Isolating Topics of High Perceived Difficulty School Biology. *Journal of Biological Education* **1980**, *14* (2), 163–166. https://doi.org/10.1080/00219266.1980.10668983.

(21) Novick, S. No Energy Storage in Chemical Bonds. *Journal of Biological Education* **1976**, *10* (3), 116–118. https://doi.org/10.1080/00219266.1976.9654072.

(22) Teichert, M. A.; Stacy, A. M. Promoting Understanding of Chemical Bonding and Spontaneity through Student Explanation and Integration of Ideas. *Journal of Research in Science Teaching* **2002**, *39* (6), 464–496. https://doi.org/10.1002/tea.10033.

(23) VandenPlas, J. R.; Herrington, D. G.; Shrode, A. D.; Sweeder, R. D. Use of Simulations and Screencasts to Increase Student Understanding of Energy Concepts in Bonding. *J. Chem. Educ.* **2021**, *98* (3), 730–744. https://doi.org/10.1021/acs.jchemed.0c00470.

(24) Franovic, C. G.-C.; Williams, N. R.; Noyes, K.; Klymkowsky, M. W.; Cooper, M. M. How Do Instructors Explain The Mechanism by Which ATP Drives Unfavorable Processes? *CBE Life Sci Educ* **2023**, *22*. https://doi.org/DOI:10.1187/cbe.23-05-0071.

(25) Hammer, D. Student Resources for Learning Introductory Physics. *American Journal of Physics* **2000**, *68* (S1), S52–S59. https://doi.org/10.1119/1.19520.

(26) Noyes, K.; Carlson, C. G.; Stoltzfus, J. R.; Schwarz, C. V.; Long, T. M.; Cooper, M. M. A Deep Look into Designing a Task and Coding Scheme through the Lens of Causal Mechanistic Reasoning. *J. Chem. Educ.* **2022**, *99* (2), 874–885. https://doi.org/10.1021/acs.jchemed.1c00959.

(27) Kang, H.; Thompson, J.; Windschitl, M. Creating Opportunities for Students to Show What They Know: The Role of Scaffolding in Assessment Tasks. *Science Education* **2014**, *98* (4), 674–704. https://doi.org/10.1002/sce.21123.

(28) Graulich, N.; Schween, M. Concept-Oriented Task Design: Making Purposeful Case Comparisons in Organic Chemistry. *J. Chem. Educ.* **2018**, *95* (3), 376–383. https://doi.org/10.1021/acs.jchemed.7b00672.

(29) Machamer, P.; Darden, L.; Craver, C. F. Thinking about Mechanisms. *Philosophy of Science* **2000**, *67* (1), 1–25. https://doi.org/10.1086/392759.

(30) Krist, C.; Schwarz, C. V.; Reiser, B. J. Identifying Essential Epistemic Heuristics for Guiding Mechanistic Reasoning in Science Learning. *Journal of the Learning Sciences* **2019**, *28* (2), 160–205. https://doi.org/10.1080/10508406.2018.1510404.

(31) Russ, R. S.; Scherr, R. E.; Hammer, D.; Mikeska, J. Recognizing Mechanistic Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed from Philosophy of Science. *Science Education* **2008**, *92* (3), 499–525. https://doi.org/10.1002/sce.20264.

(32) van Mil, M. H. W.; Boerwinkel, D. J.; Waarlo, A. J. Modelling Molecular Mechanisms: A Framework of Scientific Reasoning to Construct Molecular-Level Explanations for Cellular Behaviour. *Sci & Educ* **2013**, *22* (1), 93–118. https://doi.org/10.1007/s11191-011-9379-7.

(33) Glennan, S. *The New Mechanical Philosophy*; Oxford University Press: Oxford, New York, 2017.

(34) Haskel-Ittah, M. Explanatory Black Boxes and Mechanistic Reasoning. *Journal of Research in Science Teaching* **2023**, *60* (4). https://doi.org/10.1002/tea.21817.

(35) Mislevy, R. J.; Almond, R. G.; Lukas, J. F. A Brief Introduction to Evidence-Centered Design. *ETS Research Report Series* **2003**, *2003* (1), i–29. https://doi.org/10.1002/j.2333-8504.2003.tb01908.x.

(36) Wood, D.; Bruner, J. S.; Ross, G. The Role of Tutoring in Problem Solving*. *Journal of Child Psychology and Psychiatry* **1976**, *17* (2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x.

(37) Hammer, D.; Elby, A.; Scherr, R. E.; Redish, E. F. Resources, Framing, and Transfer. In *Transfer of learning from a modern multidisciplinary perspective*; Greenwich, CT, 2005; pp 89–120.

(38) Bryfczynski, S. BeSocratic: An Intelligent Tutoring System for the Recognition, Evaluation, and Analysis of Free-Form Student Input. *All Dissertations* **2012**.

(39) Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* **1960**, *20* (1), 37–46. https://doi.org/10.1177/001316446002000104.

(40) Landis, J. R.; Koch, G. G. The Measurement of Observer Agreement for Categorical Data. *Biometrics* **1977**, *33* (1), 159–174. https://doi.org/10.2307/2529310.

(41) IBM SPSS Statistics, 2020.

(42) Noyes, K.; Cooper, M. M. Investigating Student Understanding of London Dispersion Forces: A Longitudinal Study. *J. Chem. Educ.* **2019**, *96* (9), 1821–1832. https://doi.org/10.1021/acs.jchemed.9b00455.

(43) Franovic, C. G.-C.; Noyes, K.; Stoltzfus, J. R.; Schwarz, C. V.; Long, T. M.; Cooper, M. M. Undergraduate Chemistry and Biology Students' Use of Causal Mechanistic Reasoning to Explain and Predict Preferential Protein–Ligand Binding Activity. *J. Chem. Educ.* **2023**, *100* (5), 1716–1727. https://doi.org/10.1021/acs.jchemed.2c00737.

(44) Yang, M. Conflicting Representations of Bond Breaking and Formation in Biology and Chemistry Textbooks: A Case Study of ATP Hydrolysis., 2023.

(45) Cooper, M. M.; Stowe, R. L.; Crandell, O. M.; Klymkowsky, M. W. Organic Chemistry, Life, the Universe and Everything (OCLUE): A Transformed Organic Chemistry Curriculum. *J. Chem. Educ.* **2019**, *96* (9), 1858–1872. https://doi.org/10.1021/acs.jchemed.9b00401.

(46) Houchlei, S. K.; Crandell, O. M.; Cooper, M. M. "What About the Students Who Switched Course Type?": An Investigation of Inconsistent Course Experience. *J. Chem. Educ.* **2023**. https://doi.org/10.1021/acs.jchemed.3c00345.

(47) Houchlei, S. A Step into the Unknown: Exploring Students' Construction of Mechanistic Arrows for Both Familiar and Unfamiliar Reactions in Organic Chemistry. Doctoral dissertation, Michigan State University, East Lansing, MI, 2022. https://www.proquest.com/openview/88e9dd8894ca65fe39ac27072f056a67/1.pdf?pq-origsite=gscholar&cbl=18750&diss=y.

(48) Becker, N.; Noyes, K.; Cooper, M. Characterizing Students' Mechanistic Reasoning about London Dispersion Forces. *J. Chem. Educ.* **2016**, *93* (10), 1713–1724. https://doi.org/10.1021/acs.jchemed.6b00298.

(49) Crandell, O. M.; Lockhart, M. A.; Cooper, M. M. Arrows on the Page Are Not a Good Gauge: Evidence for the Importance of Causal Mechanistic Explanations about Nucleophilic Substitution in Organic Chemistry. *J. Chem. Educ.* **2020**, *97* (2), 313–327. https://doi.org/10.1021/acs.jchemed.9b00815.

(50) Ralph, V. R.; Scharlott, L. J.; Schafer, A. G. L.; Deshaye, M. Y.; Becker, N. M.; Stowe, R. L. Advancing Equity in STEM: The Impact Assessment Design Has on Who Succeeds in Undergraduate Introductory Chemistry. *JACS Au* **2022**. https://doi.org/10.1021/jacsau.2c00221.

APPENDIX A. THE FULL FINAL TASK

Figures 7.8 – 7.20 show screenshots from beSocratic of the full activity from a student's perspective.

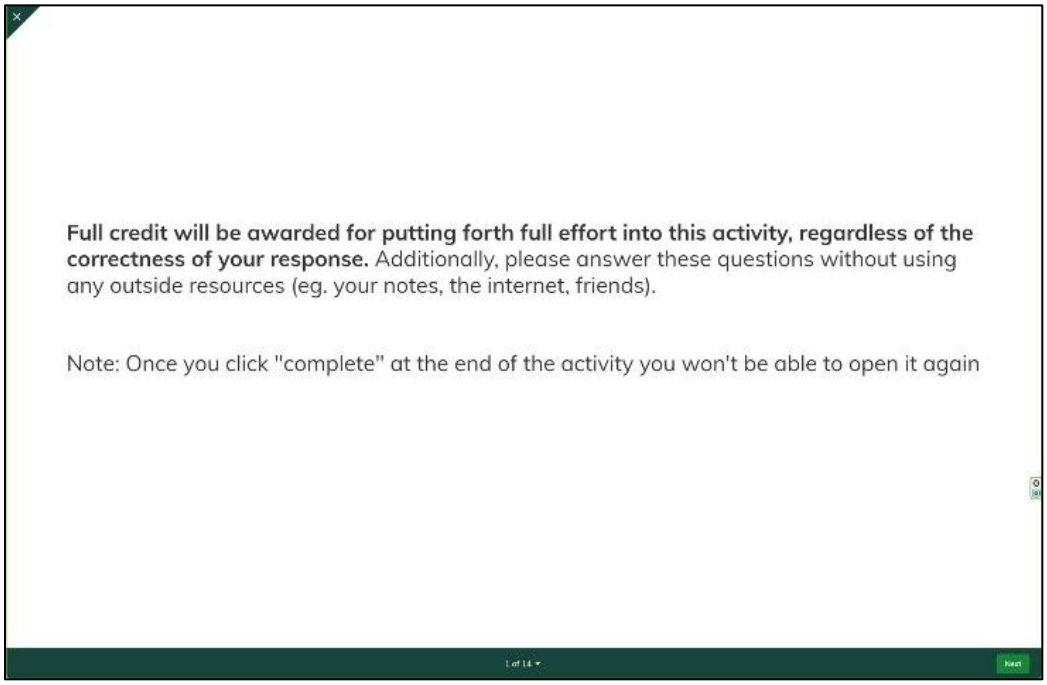


Figure 7.8. Slide 1 of the final task.

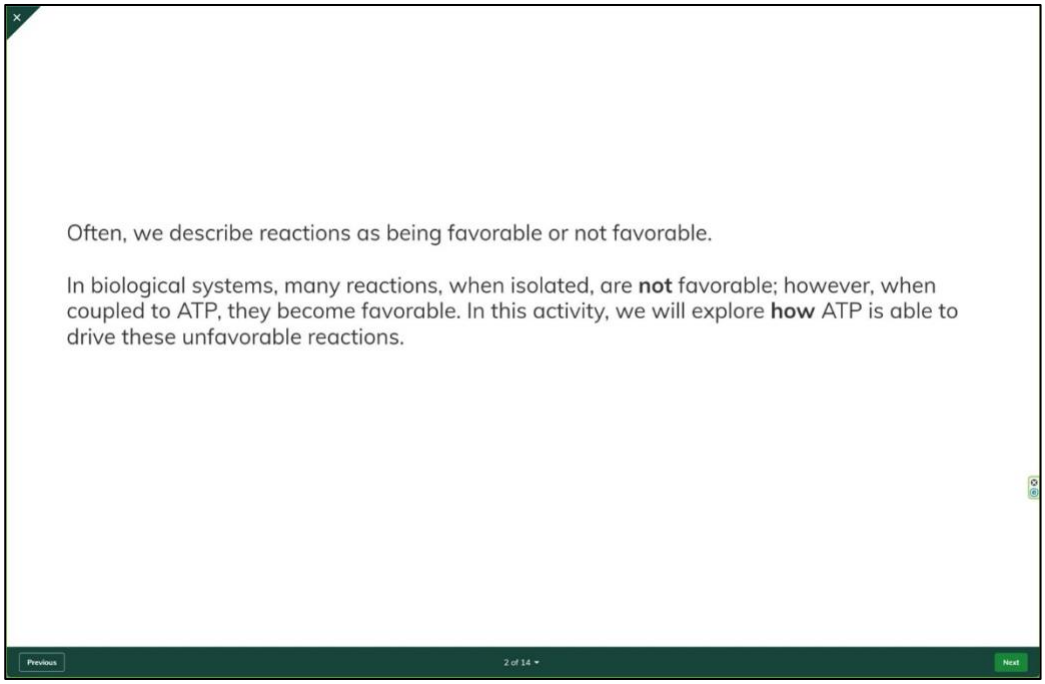


Figure 7.9. Slide 2 of the final task.

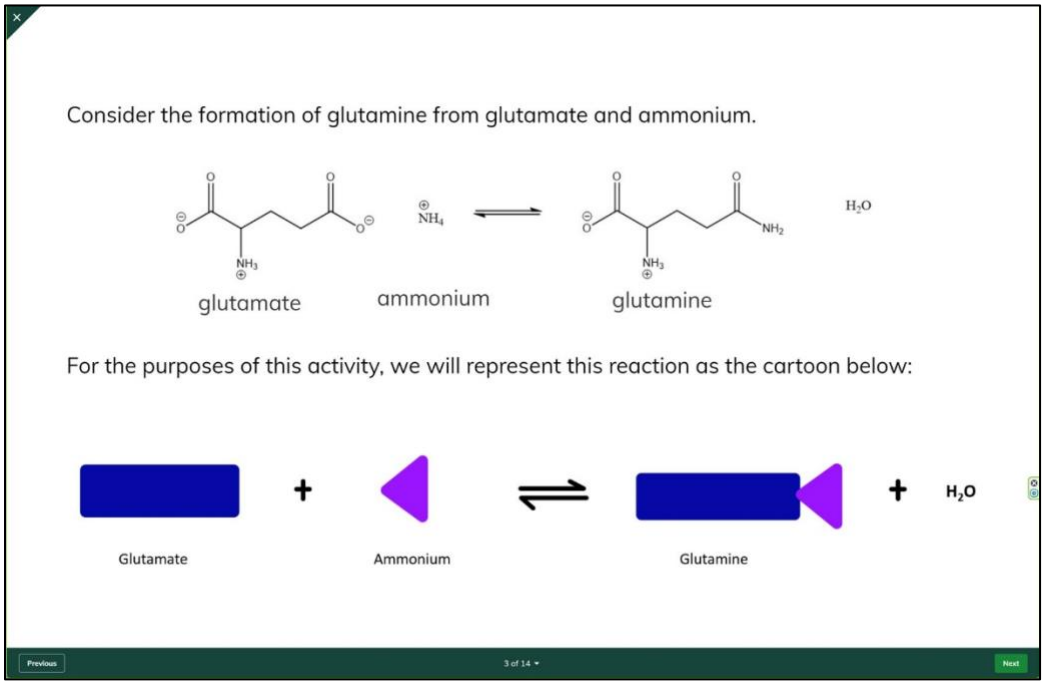


Figure 7.10. Slide 3 of the final task.

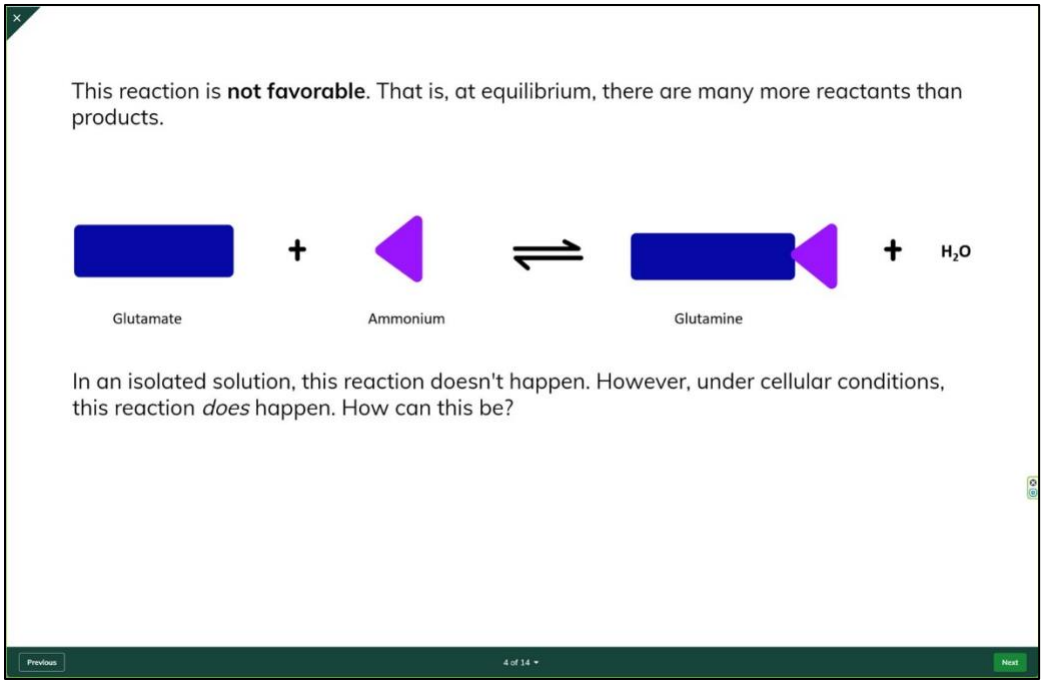


Figure 7.11. Slide 4 of the final task.

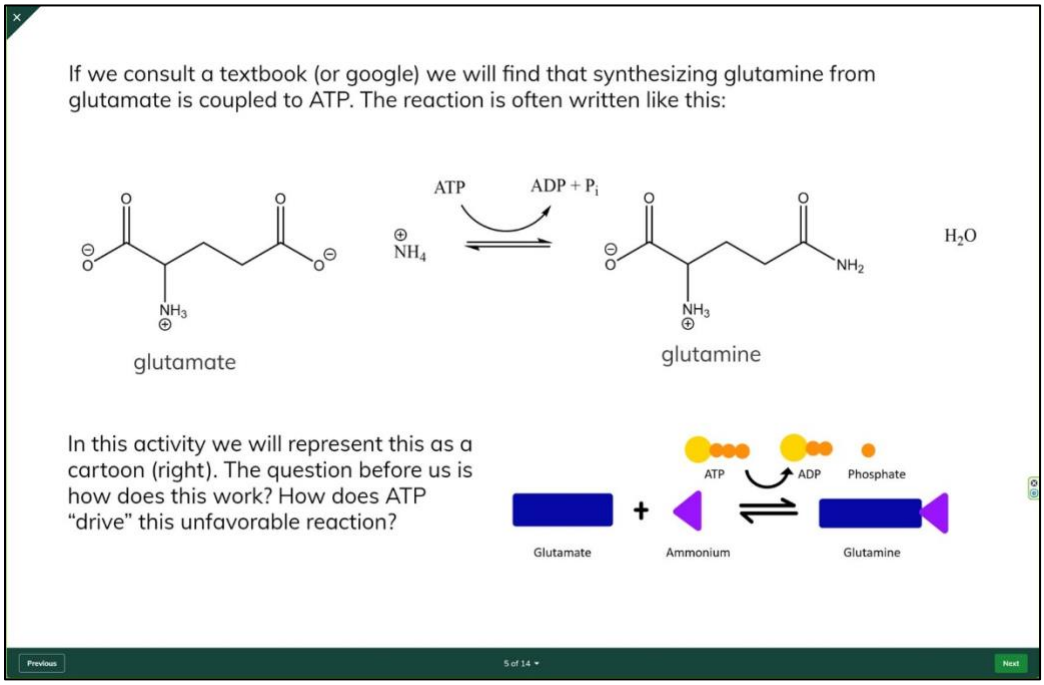


Figure 7.12. Slide 5 of the final task.

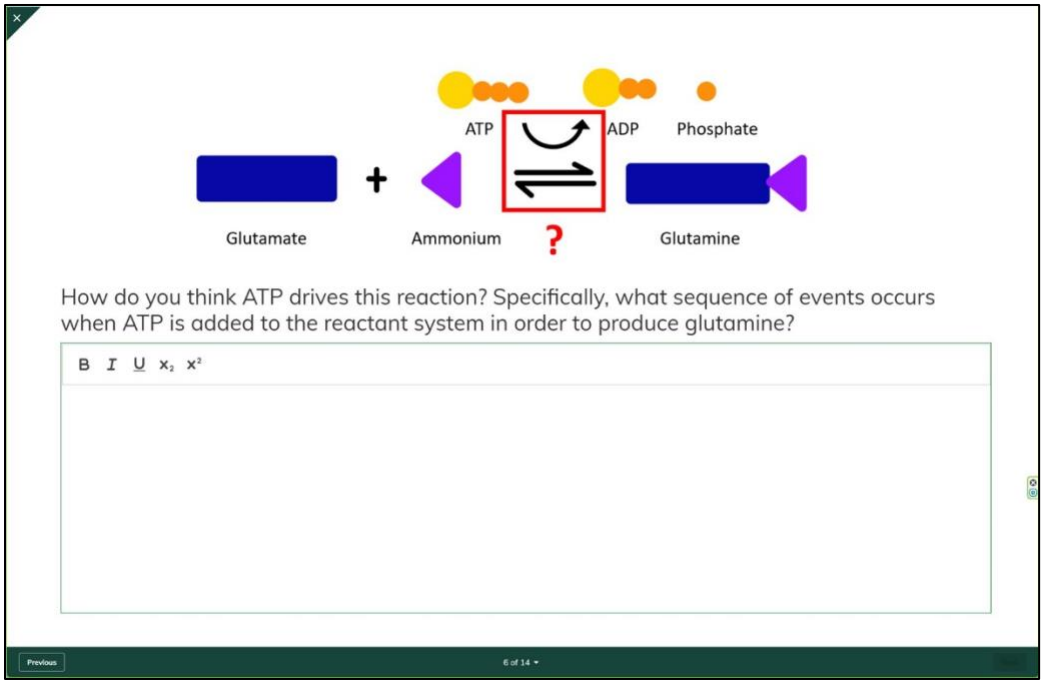


Figure 7.13. Slide 6 of the final task.

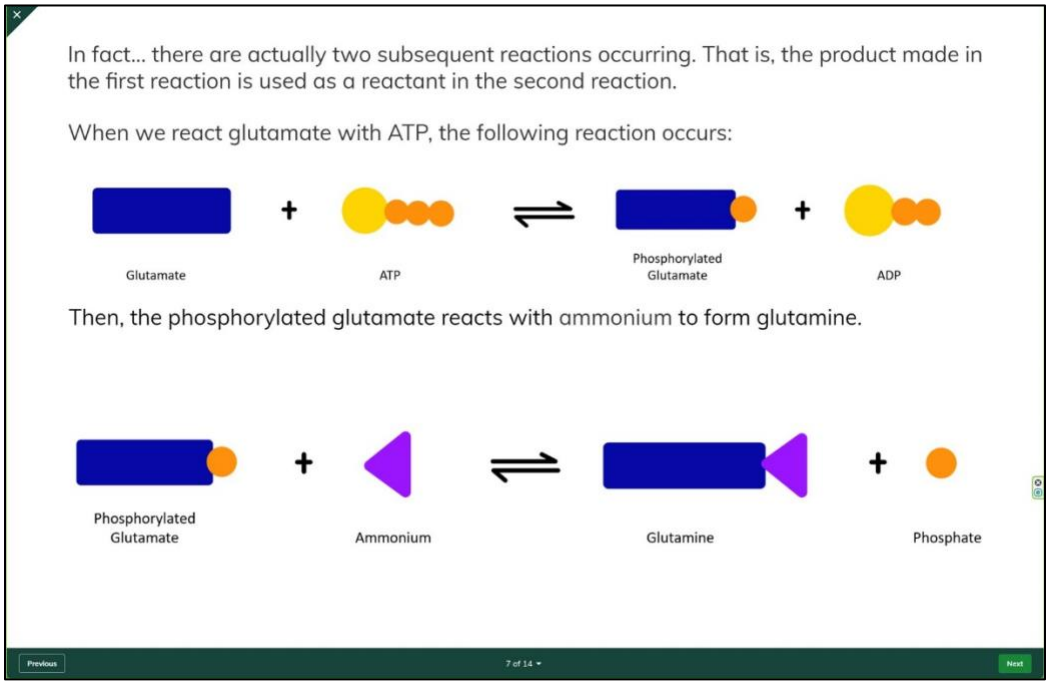


Figure 7.14. Slide 7 of the final task.

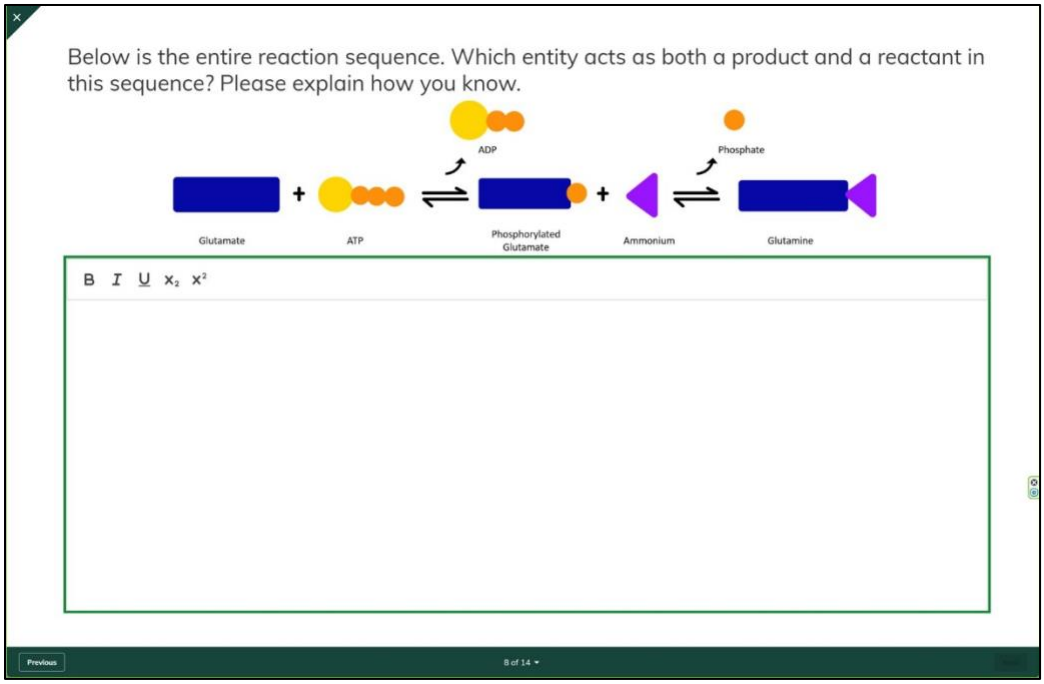


Figure 7.15. Slide 8 of the final task.

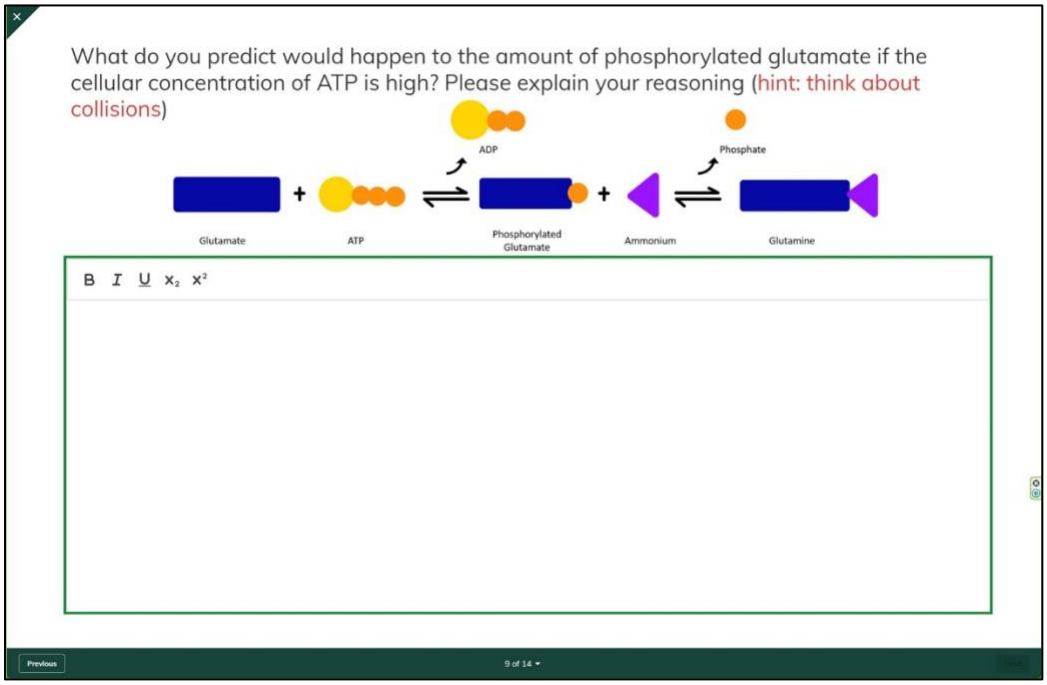


Figure 7.16. Slide 9 of the final task.

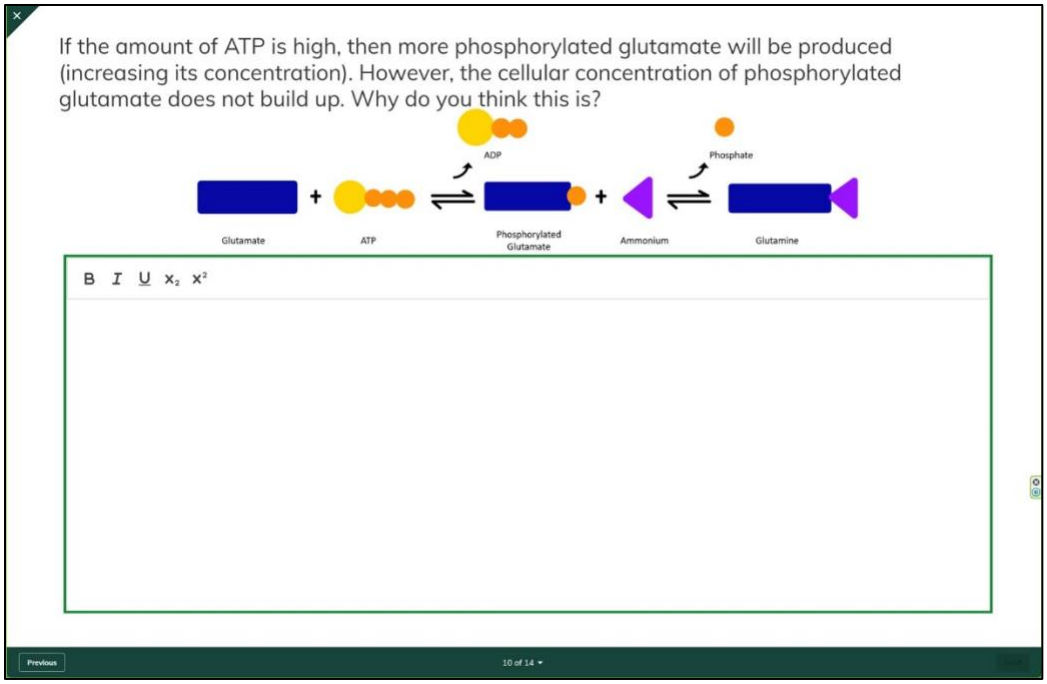


Figure 7.17. Slide 10 of the final task.

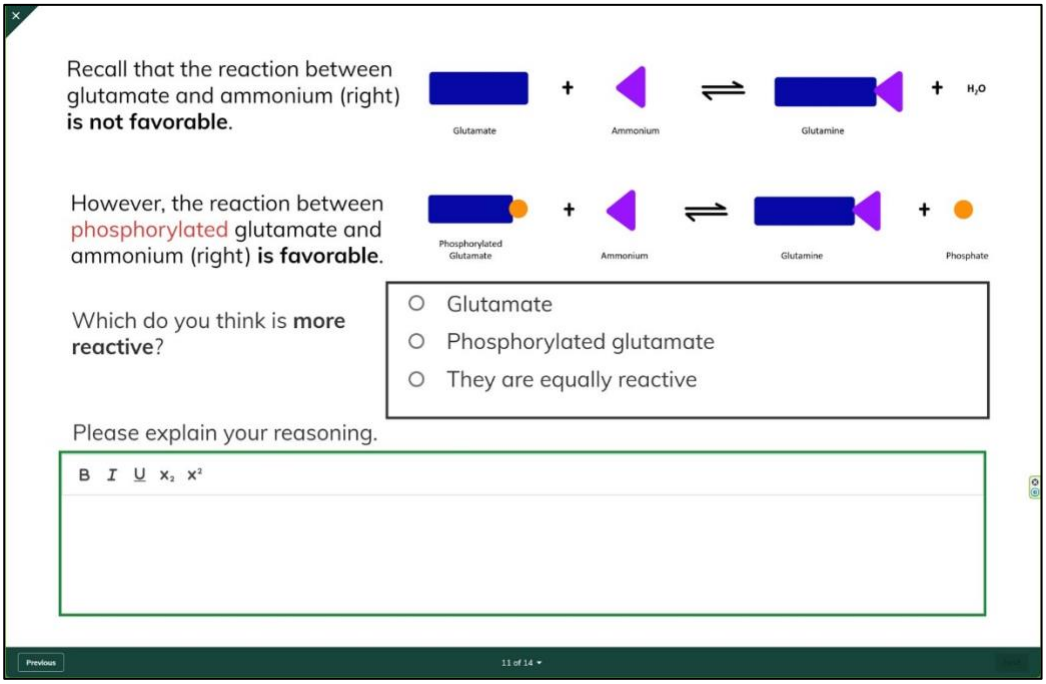


Figure 7.18. Slide 11 of the final task.

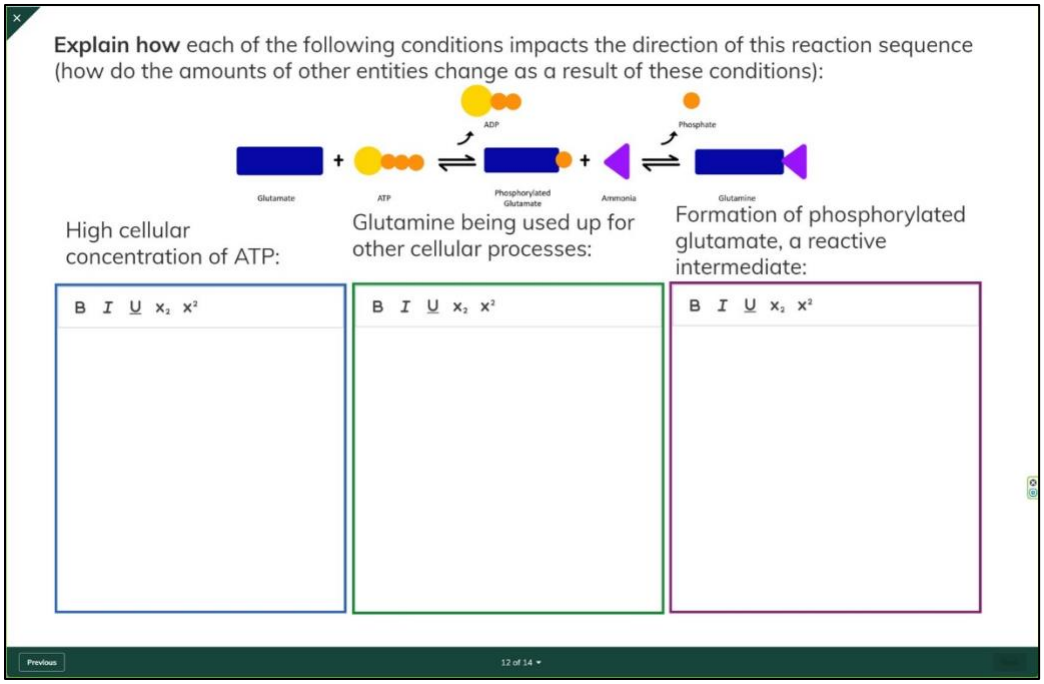


Figure 7.19. Slide 12 of the final task.

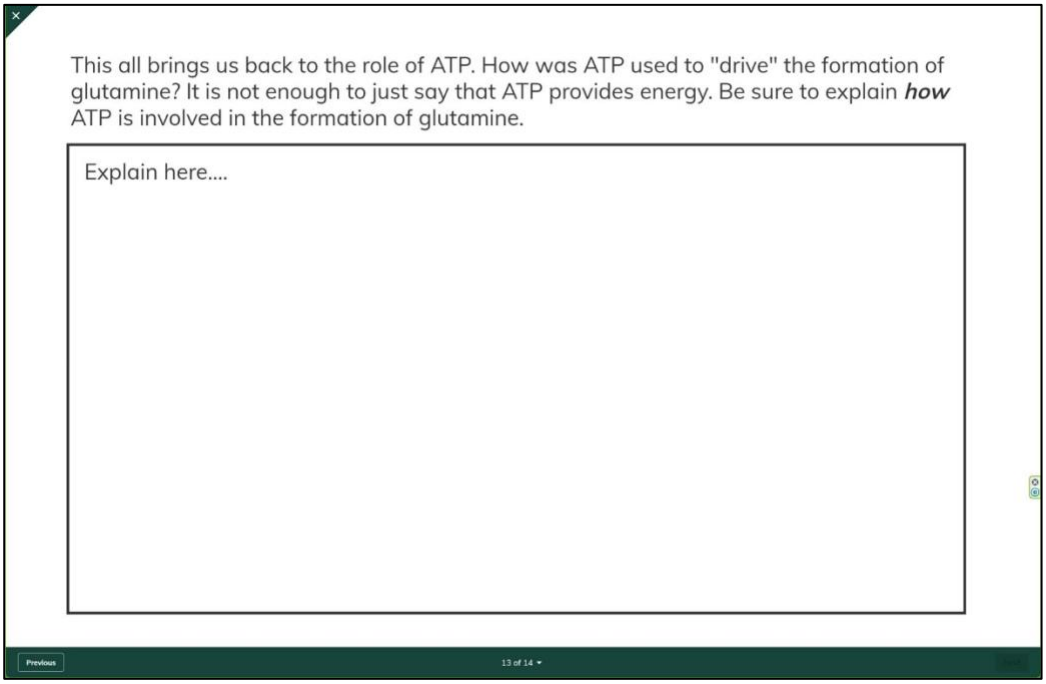


Figure 7.20. Slide 13 of the final task.

APPENDIX B. STUDENTS' INITIAL AND FINAL RESPONSES

Tables 7.5 and 7.6 represent how students' responses changed (or didn't change) between their initial and final explanations. There are four possible "changes": (1) non-mechanistic resources → non-mechanistic resources, (2) non-mechanistic resources → mechanistic resources, (3) mechanistic resources → non-mechanistic resources, or (4) mechanistic resources → mechanistic resources. For example, looking at Table 7.5, we see 272 MB students' initial responses relied on non-mechanistic resources, but their final responses leveraged mechanistic resources. This result was the most frequent for students in both courses.

Table 7.5. Two by two table showing how MB students' final responses changed (or not) from their initial responses.

MB initial	MB final	
	non-mechanistic resources	mechanistic resources
non-mechanistic resources	144	272
mechanistic resources	56	204

Table 7.6. Two by two table showing how OC students' final responses changed (or not) from their initial responses.

OC initial	OC final	
	non-mechanistic resources	mechanistic resources
non-mechanistic resources	48	112
mechanistic resources	11	49

Chapter VIII – Conclusions, Implications, And Future Directions

My dissertation research focused on mechanistic reasoning in the contexts of undergraduate chemistry and biology courses. Through informed design of tasks, characterization of students' explanations, and analysis of instructor interviews, I have drawn conclusions that provide implications for both research and instruction. Future work based on these conclusions and implications is proposed in the final section of this chapter.

Conclusions

Causal Mechanistic Reasoning Is A Predictive Tool That Supports Equitable Assessment

In Chapters IV and V, we used a previously designed task about protein-ligand binding to capture student engagement in causal mechanistic reasoning (CMR) across undergraduate chemistry and biology courses. With an initial snapshot of students across courses, we found that students engaged in CMR to varying extents (non-CM, partially CM, or fully CM). While there were differences across courses, students who were co-enrolled in introductory chemistry and biology did not change their explanations when situated in each course, which surprised us given the context-dependent nature of resource activation.[1,2] This suggests the interdisciplinary nature of this task, one that can be used in a variety of courses to engage students in this thinking strategy. In this study, we also found that engaging fully in CMR strongly correlated with accurate predictions – that is, 97% of the time, students who constructed a fully CM response also correctly chose the protein binding site to which Mg^{2+} would more likely bind, suggesting the predictive power of CMR.

In Chapter V, we coded and analyzed additional student responses to this task with the goal of disaggregating the data according to demographic characteristics. Previous work has

suggested that mechanistic reasoning (MR) assessment items, when used frequently on exams and in coursework, result in more equitable outcomes for students.[3] To advance our knowledge in the field, we calculated an ordinal regression model to determine whether cumulative GPA, race/ethnicity, and/or binary gender identification were predictors for student engagement in CMR for this task. We found that both GPA and race/ethnicity were significant predictors of student engagement in CMR ($p < 0.001$ and $p = 0.014$, respectively) – that is, students with a higher cumulative GPA have higher odds of being in a higher category, and compared to White students, being Non-White decreases the odds of being in a higher category. However, the GPA estimate (0.613) was twice as impactful as race/ethnicity (0.308), indicating it as a stronger predictor of student engagement in CMR for this task. Binary gender identification was not a significant predictor ($p = 0.397$). Thus, while GPA was the strongest predictor, race/ethnicity also contributed to the model, meaning additional work should be done to support historically marginalized students.

Instructors' Explanations And Experiences Suggest The Need For A Shift In Teaching Practices Regarding Adenosine Triphosphate (ATP)

In Chapter VI, we explored how instructors across chemistry and biology disciplines explain and teach the mechanism by which ATP provides energy or drives unfavorable processes. There is strong evidence for students' non-canonical (or unproductive activation of) ideas about the energy changes associated with bonds breaking or forming,[4–13] which may be exacerbated by discussions in biology classrooms about the "high-energy bond" of ATP.[11–16] While chemistry and physics education research has focused explicitly on the mechanism by which ATP hydrolysis releases energy,[11–13] this mechanism is irrelevant when considering

coupled reactions, which proceed via a mechanism of phosphate transfer (that is, ATP does not hydrolyze, rather it (often) transfers a phosphoryl group to raise the energy of a reactant). In this study, we found that instructors leveraged one or both of these mechanisms (energy release and/or energy transfer) to explain the role of ATP. Most intriguing, however, were the negative experiences that Molecular Biology instructors shared related to their teaching practices about ATP and mechanistic reasoning. These instructors shared vulnerable experiences including lack of confidence and frustration, suggesting the need for a new, more productive way of talking about ATP in these courses.

A Learning Task Supports Students' Use Of Mechanistic Resources When Explaining The Role Of ATP In A Complex Phenomenon

Using what we learned from interviews with instructors, as well as a desire to explore important topics that could benefit from MR in undergraduate molecular biology (MB) courses, we designed a scaffolded formative task to support students' understanding of the mechanism by which ATP drives the formation of glutamine from glutamate and ammonium. We administered the task to both MB students and Organic Chemistry (OC) students, finding that students in both courses used more mechanistic resources in their final explanations (end of the task) than their initial explanations (beginning of the task). This finding supports the utility of formative tasks in coursework, as they can serve as learning opportunities for students to integrate their prior knowledge with new, mechanistic information about complex phenomena such as this one. While many students successfully integrated prior knowledge about ATP as a source of energy with its mechanistic role as a phosphorylating agent, not all responses provided evidence of this idea. Some of the explanations included both ideas without

coherently integrating them, thus suggesting the co-existence of conceptions regarding the role of ATP.

Implications For Research And Practice

The evidence gathered from my studies has added to discipline-based education research on mechanistic reasoning and interdisciplinary learning, primarily through formative assessment tasks. The science education community values mechanistic reasoning (that is, thinking about how/why phenomena occur[17–21]) as an important practice, thereby warranting the studies outlined in this dissertation. Using the major conclusions, I now provide implications for both research and practice.

Designing And Implementing Tasks To Assess Student Engagement In Mechanistic Reasoning

This dissertation has contributed to previous work that highlighted the intricacies and challenges of effective task design in undergraduate science education.[22] In this work, we found that to design a task that provides enough information for students to understand what is being asked, but not so much that they can answer without thoughtful effort, is no trivial feat. The protein-ligand binding task (PL task) took several iterations in order to effectively focus students' attention on relevant, mechanistic components, without giving away the answer. Once the final task was designed, we used it to collect the evidence outlined in Chapter IV. This evidence highlighted the interdisciplinary nature of the task, as well as the predictive power of MR. That is, fully engaging in MR about protein-ligand binding led to more accurate predictions, providing evidence for the utility of this thinking strategy. As noted, however, designing the task to elicit this type of thinking was not trivial.[22] With full access to the activity, we hope that instructors in chemistry and biology disciplines will use (or modify) it in their courses so as to

provide the opportunity for students to leverage core ideas and think deeply about mechanisms spanning the disciplines of chemistry and (molecular) biology. We also propose that additional work be done to help students leverage productive resources. In this study, we found that a handful of students unproductively used the idea of "space" to incorrectly predict which protein the Mg^{2+} would bind. This was an important finding, because we did not expect to activate this resource – as instructors, it is critical that we pay close attention to the resources students use in order to help them refine their knowledge frameworks and choose between, potentially, competing resources.

Supporting Instructors To Incorporate Mechanistic Reasoning About ATP In Molecular Biology Courses

While the PL task leveraged core ideas related to electrostatic and bonding interactions (and structure-property relationships), my subsequent research investigated the core idea of energy across chemistry and biology. Chapter VI highlighted the mechanisms that instructors used and experiences they had when thinking about the role of ATP as an energy source, an important biological phenomenon. With extensive literature documenting students' non-canonical ideas about the energy changes associated with bonds breaking and forming, as well as the overlap of this idea with language surrounding the role of ATP as a source of energy, we set out to understand how instructors think about the mechanism by why which ATP drives unfavorable processes. We found that instructors explained the role of ATP by either discussing how ATP hydrolysis *releases* energy, or how ATP *transfers* energy (via phosphorylation). We also found that most instructors, but especially biology instructors, shared negative experiences (i.e., dissatisfaction) regarding their current teaching practices about ATP. While the two

mechanisms (energy release and energy transfer) are related, they are distinct and serve different purposes. Based on our conversations with these faculty, we urge MB instructors to think deeply about the mechanism(s) they want students to know. We believe the mechanism of energy transfer is more appropriate for biological contexts, as this mechanism was leveraged by all three biochemistry instructors and most of the biology instructors (indicating its relevance to the discipline), and because energy is rarely (and negligibly) released in these complex systems. Further, while it is simple to provide this suggestion, we recognize the challenges that instructors face in deciding the level of appropriate explanatory depth. Thus, we point instructors to the utility of explanatory black boxes, or unexplained mechanisms within mechanisms.[23] According to Haskel-Ittah, the explicit use of explanatory black boxes in education helps to avoid the illusion of explanatory depth, or the sense that one understands causally complex phenomena deeply when that may not be the case, while also preparing students for future learning (i.e., helping them to identify what they *don't* know).

Carefully Designed Tasks Can Serve As Learning Opportunities To Engage Students In MR

The rich discussions that emerged from our interviews with instructors led to the development of the activity discussed in Chapter VII. This activity, which, like the PL task, was designed via several iterations, aims to support students' understanding of the mechanism by which ATP drives (i.e., transfers energy to) the unfavorable formation of glutamine from glutamate and ammonium. This task differs from the PL task in that it does not ask students to make a prediction based on their prior knowledge; rather, it serves as a learning opportunity by providing information to students while drawing on their prior knowledge about reactivity and favorability. By collecting both initial and final explanations from students, we could assess

whether this task supported student use of mechanistic resources when explaining the role of ATP. After working through the task, both MB and Organic Chemistry (OC) students used more mechanistic resources, most notably the idea of ATP as a phosphorylating agent. Therefore, instructors and researchers can use carefully designed tasks to support students as they integrate prior knowledge (in this case, ATP as a source of energy) with (potentially) new information (ATP as a phosphorylating agent). However, we also found that not all students coherently integrated these ideas. More work should attend to this either through in-class discussions/activities (instructional implications) or via think-aloud interviews (research implications) to identify (1) productive avenues for knowledge integration and/or (2) the ideas that are more challenging for students and how best to support them in using and advancing those ideas.

Using And Improving MR Tasks To Support Equitable Practices

Chapters IV, VI, and VII were each driven by our efforts to incorporate MR into undergraduate chemistry and biology contexts, since this thinking strategy has been shown to be a productive and powerful tool; however, it is similarly important to learn whether these tasks (or this thinking strategy) are equitable. In Chapter V, we found that both GPA (reflecting prior academic achievement) and race/ethnicity significantly contribute, while gender does not contribute, to a model predicting student engagement in CMR for the PL task. That is, students with a higher GPA had higher odds of engaging fully in CMR; White students, compared to Non-White students, had higher odds of engaging fully in CMR; and being male or female did not impact engagement in CMR. While it is not surprising, nor concerning, that GPA is a strong predictor of student engagement in CMR, the small, but significant, contribution of

race/ethnicity as a predictor suggests that additional work should be done to support historically marginalized students. However, this is simply one task, a task that does not necessarily predict student course outcomes. Ralph et al. (2022) found that assessments which more frequently emphasize MR when compared to those assessing rote calculations result in more equitable outcomes (passing grades) for students.[3] MR tasks are important, because they engage students in deep thinking about phenomena; however, they are one component of the complex learning environment(s) in which students are situated. We urge instructors and researchers to disaggregate data according to demographic characteristics when available to learn whether assessments cater inequitably to different groups, as this will help future efforts to best support all students as they learn and inquire about science.

Future Directions

My dissertation research has contributed to building an understanding of MR in chemistry and biology undergraduate courses, highlighting challenges and gaps that future work can investigate. While other reasoning strategies (e.g., systems thinking) are important and relevant to include in undergraduate science courses, there is limited work on how to incorporate MR in Molecular Biology courses – a discipline that strongly relies on predicting and explaining mechanisms. Therefore, one area of future work should involve investigating how current MB instructional practices leverage MR and how to develop additional practices (i.e., such as incorporating intentionally designed formative assessment tasks) that support students and instructors to use this thinking strategy. This work should also leverage explanatory black boxes,[23] as cellular and molecular mechanisms are complex and often emerge as a result of

several mechanisms (for these reasons, systems thinking is also a highly relevant thinking strategy in biological contexts).

Further, there are calls to support equity- and justice-oriented education efforts.[24] Since MR is a powerful thinking strategy, it would be worthwhile to learn if and how it contributes to course culture or if tasks can be developed using both MR and equity lenses. As noted in the introduction, engaging in MR, and developing it as a productive epistemic heuristic can serve as a useful tool for all students' lives. For example, thinking about how and why things happen can support productive decision-making, as well as satisfaction and appreciation for curious minds. Through both an equity and MR lens, future work should aim to support students and instructors as they engage in meaningful science teaching and learning.

REFERENCES

(1) Hammer, D. Student Resources for Learning Introductory Physics. *American Journal of Physics* **2000**, *68* (S1), S52–S59. https://doi.org/10.1119/1.19520.

(2) Hammer, D.; Elby, A.; Scherr, R. E.; Redish, E. F. Resources, Framing, and Transfer. In *Transfer of learning from a modern multidisciplinary perspective*; Greenwich, CT, 2005; pp 89–120.

(3) Ralph, V. R.; Scharlott, L. J.; Schafer, A. G. L.; Deshaye, M. Y.; Becker, N. M.; Stowe, R. L. Advancing Equity in STEM: The Impact Assessment Design Has on Who Succeeds in Undergraduate Introductory Chemistry. *JACS Au* **2022**. https://doi.org/10.1021/jacsau.2c00221.

(4) Novick, S. No Energy Storage in Chemical Bonds. *Journal of Biological Education* **1976**, *10* (3), 116–118. https://doi.org/10.1080/00219266.1976.9654072.

(5) Johnstone, A. H.; Mahmoud, N. A. Isolating Topics of High Perceived Difficulty School Biology. *Journal of Biological Education* **1980**, *14* (2), 163–166. https://doi.org/10.1080/00219266.1980.10668983.

(6) Hapkiewicz, A. Clarifying Chemical Bonding. Overcoming Our Misconceptions. *Science Teacher* **1991**, *58* (3), 24–27.

(7) Boo, H. K. Students' Understandings of Chemical Bonds and the Energetics of Chemical Reactions. *Journal of Research in Science Teaching* **1998**, *35* (5), 569–581. https://doi.org/10.1002/(SICI)1098-2736(199805)35:5<569::AID-TEA6>3.0.CO;2-N.

(8) Barker, V.; Millar, R. Students' Reasoning about Basic Chemical Thermodynamics and Chemical Bonding: What Changes Occur during a Context-Based Post-16 Chemistry Course? *International Journal of Science Education* **2000**, *22* (11), 1171–1200. https://doi.org/10.1080/09500690050166742.

(9) Teichert, M. A.; Stacy, A. M. Promoting Understanding of Chemical Bonding and Spontaneity through Student Explanation and Integration of Ideas. *Journal of Research in Science Teaching* **2002**, *39* (6), 464–496. https://doi.org/10.1002/tea.10033.

(10) Galley, W. C. Exothermic Bond Breaking: A Persistent Misconception. *J. Chem. Educ.* **2004**, *81* (4), 523. https://doi.org/10.1021/ed081p523.

(11) Dreyfus, B. W.; Sawtelle, V.; Turpen, C.; Gouvea, J.; Redish, E. F. Students' Reasoning about ``high-Energy Bonds'' and ATP: A Vision of Interdisciplinary Education. *Phys. Rev. ST Phys. Educ. Res.* **2014**, *10* (1), 010115. https://doi.org/10.1103/PhysRevSTPER.10.010115.

(12) Green, A. I.; Parent, K. N.; Underwood, S. M.; Matz, R. L. Connecting Ideas across Courses: Relating Energy, Bonds & How ATP Hydrolysis Powers a Molecular Motor. *The American Biology Teacher* **2021**, *83* (5), 303–310. https://doi.org/10.1525/abt.2021.83.5.303.

(13) VandenPlas, J. R.; Herrington, D. G.; Shrode, A. D.; Sweeder, R. D. Use of Simulations and Screencasts to Increase Student Understanding of Energy Concepts in Bonding. *J. Chem. Educ.* **2021**, *98* (3), 730–744. https://doi.org/10.1021/acs.jchemed.0c00470.

(14) Lipmann, F. Attempts toward a Formulation of Biological Use of Energy in Terms of Chemical Potentials. In *Molecular biology: Elementary processes of nerve conduction and muscle contraction*; Academic Press: Cambridge, MA, 1960; pp 37–47.

(15) Cooper, M. M.; Klymkowsky, M. W. The Trouble with Chemical Energy: Why Understanding Bond Energies Requires an Interdisciplinary Systems Approach. *CBE Life Sci Educ* **2013**, *12* (2), 306–312. https://doi.org/10.1187/cbe.12-10-0170.

(16) Kohn, K. P.; Underwood, S. M.; Cooper, M. M. Energy Connections and Misconnections across Chemistry and Biology. *LSE* **2018**, *17* (1), ar3. https://doi.org/10.1187/cbe.17-08-0169.

(17) Russ, R. S.; Scherr, R. E.; Hammer, D.; Mikeska, J. Recognizing Mechanistic Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed from Philosophy of Science. *Science Education* **2008**, *92* (3), 499–525. https://doi.org/10.1002/sce.20264.

(18) Krist, C.; Schwarz, C. V.; Reiser, B. J. Identifying Essential Epistemic Heuristics for Guiding Mechanistic Reasoning in Science Learning. *Journal of the Learning Sciences* **2019**, *28* (2), 160–205. https://doi.org/10.1080/10508406.2018.1510404.

(19) Machamer, P.; Darden, L.; Craver, C. F. Thinking about Mechanisms. *Philosophy of Science* **2000**, *67* (1), 1–25. https://doi.org/10.1086/392759.

(20) van Mil, M. H. W.; Boerwinkel, D. J.; Waarlo, A. J. Modelling Molecular Mechanisms: A Framework of Scientific Reasoning to Construct Molecular-Level Explanations for Cellular Behaviour. *Sci & Educ* **2013**, *22* (1), 93–118. https://doi.org/10.1007/s11191-011-9379-7.

(21) Glennan, S. *The New Mechanical Philosophy*; Oxford University Press: Oxford, New York, 2017.

(22) Noyes, K.; Carlson, C. G.; Stoltzfus, J. R.; Schwarz, C. V.; Long, T. M.; Cooper, M. M. A Deep Look into Designing a Task and Coding Scheme through the Lens of Causal Mechanistic Reasoning. *J. Chem. Educ.* **2022**, *99* (2), 874–885. https://doi.org/10.1021/acs.jchemed.1c00959.

(23) Haskel-Ittah, M. Explanatory Black Boxes and Mechanistic Reasoning. *Journal of Research in Science Teaching* **2023**, *60* (4). https://doi.org/10.1002/tea.21817.

(24) The Boyer 2030 Commission. *The Equity-Excellence Imperative*; The Association for Undergraduate Education at Research Universities, 2022; pp 3–59. https://ueru.org/boyer2030 (accessed 2023-10-17).