Modeling Biochemistry Arguments

Nancy L. Green [1], Derek T. McLachlin [2] and Robert E. Mercer [3]

[1] *University of North Carolina Greensboro, Greensboro, North Carolina, USA*
[2] *University of Western Ontario, London, Ontario, CA*
[3] *University of Western Ontario, London, Ontario, CA*

Abstract

This paper describes our current research cataloging argument types in the Results and Discussions sections of biochemistry research articles. The arguments have a Toulmin-model structure and are instances of general argumentation schemes and Mill's "methods" of scientific reasoning. It may be possible to implement generalizations of the arguments as rules for argument mining.

Keywords

Argumentation schemes, Toulmin model, Mill's methods, argument mining, scientific arguments, experimental life sciences, biochemistry

1. Introduction

A number of challenges for argument mining in scientific documents have been noted by Al-Khatib et al. [1]. First, they claim that existing approaches such as Toulmin's model and argumentation schemes are "not a good fit" for modeling scientific arguments. Furthermore, they question whether a unified approach to argument mining is suitable for all domains of science and all genres of scientific documents. Another challenge, which we too have noted in our past and current research, is the common occurrence of arguments with implicit components and arguments whose claims and support are not adjacent in a text.

We agree that a single approach to argument mining may not be suitable for all domains and genres of scientific argumentation. However, based upon our past and current research on argumentation in the experimental life sciences, we disagree that Toulmin's model and argumentation schemes are problematic for modeling scientific arguments in this domain. Currently, we are cataloging argument types in biochemistry research articles for possible use in argument mining. We have found arguments in this domain that have a Toulmin-model data-warrant-claim structure [25] and are instances of general argumentation schemes [26] and Mill's "methods" of scientific reasoning [14].

In this paper we first describe related work on computational modeling of arguments in scientific documents, including our previous work on modeling argumentation in the experimental life sciences. Then we describe some results of our current research on argument types in the Results and Discussions sections of biochemistry research articles on cross-linked protein structures and signal transduction pathways. We suggest how these types may be used for argument mining. Finally, we discuss remaining challenges and possible future work.

2. Related work on computational modeling of scientific arguments

Our work is novel in modeling argumentation in research articles from the experimental life sciences. Previous research on argument mining of scientific research focused on areas outside of the natural sciences -- education [23] and computer graphics [12] -- or on abstracts of multidisciplinary

Computational Models of Natural Argument (CMNA2024), September 17, 2024, Hagen, Germany
EMAIL: nlgreen@uncg.edu (A. 1); derek.mclachlin@schulich.uwo.ca (A. 2); mercer@csd.uwo.ca (A. 3)
ORCID: 0000-0002-7810-1771 (A. 2); 0000-0002-0080-715X (A. 3)

articles on sustainable development [3]. Some research has modeled other aspects of argumentative discourse in the experimental life sciences: "argumentative zones" [18,24], "conceptualization zones" [13], and "rhetorical moves" [11] in the Methods section of full-text biochemistry articles [2]. Some of Green's earlier research focused on natural language generation of arguments in genetic counseling patient information [10].

The work reported in this paper carries forward our work on characterizing scientific argumentation in research articles in two experimental life sciences: biochemistry and genetics. Based upon analysis of five biochemistry articles, whose claims were annotated by a biochemist (McLachlin, who also is a coauthor of this paper), Moser and Mercer [19] proposed the Claim Graph model for visualizing the interrelationships between all claims and evidence in an article. Evidence could come from data reported in the text, data in the figures, claims made earlier in the article, cited works, or the reader's assumed knowledge. The arguments in the article were characterized by the Toulmin model [25] combined with argumentation schemes proposed by Green [4, 5] for describing arguments in a small corpus of genetics[2] research articles. For example, one argument in the biochemistry article was described by Moser and Mercer as:

> Premise (claim 19): The highest level of disulfide formation was found with the … combinations.
> Premise [missing warrant]: Proximity is necessary for disulfide binding between residues.
> Claim 21A (Green's Argumentation Scheme: Effect to Cause (5)): The result suggests that residue 65 of one subunit is close to residues 60 and 61of the other.

(The claims were numbered for depiction in the Claim Graph, i.e., an earlier claim (19) was a premise of the argument for claim 21A. Also, the warrant of this argument was implicit and assumed to be known by the intended reader.)

Subsequently, as an aid to corpus annotation of genetics research articles, an updated catalogue of argumentation schemes defined using terms of genetics was released [7]. The genetics schemes are specializations of general argumentation schemes [26] and Mill's "methods" of scientific reasoning [14]. For example, a specialization of Mill's Method of Difference for genetics was described in [7] as:

> Premise: A group of individuals I have atypical phenotype P
> Premise: All of the individuals in I have atypical genotype M
> Premise: A group of individuals C do not have P.
> Premise: None of the individuals in C have M.
> Conclusion: M may be the cause of P (in I)

To demonstrate how the genetics argumentation schemes could be used for argument mining, Green [8] implemented several as rules in a logic programming language. The rules contained semantic entities and relations that could be obtained by information extraction (IE); then arguments could be extracted by applying the rules to the output of IE. In later work [9], a method for acquisition of such rules by inductive logic programming was proposed.

Another challenge for argument mining in scientific documents noted in [1] is that the "line of reasoning" in experimental papers may make sense to a scientist but not to someone outside of the field, e.g., the reason that a specific experiment was done following another experiment may not be explicitly stated. Although not modeled in our current research, in [9] a discovery dialogue game was proposed to formally model the sequence of experimental goals in a genetics research article.

3. Arguments in biochemistry research articles

This section describes some preliminary results of our current investigation. We have analyzed certain arguments in the full-text Results and Discussion sections in research journal articles on two biochemistry topics: cross-linked protein structure [15,16,17], coauthored by a biochemist who also is a co-author of this paper (McLachlin); and signal transduction pathways [20]. McLachlin has identified

[2] Although some of Green's research on natural language generation also related to genetics [10], it involved arguments for patient communication. The work in [4] and later models arguments in genetics research articles. For a survey of that work, see [8].

the main scientific arguments of each article, paraphrasing the main *Claim* of each argument and its support. Supporting *Observations* are based on data reported in the text, figures, unpublished results, or a cited article. Supporting *Background* is information that is assumed to be known by the intended audience and which often functions as a warrant. Then for each argument, we have classified the argument type, which is related to an argumentation scheme [26] or one of Mill's "methods" of scientific reasoning [14]. Our goal is to compile a set of argument instances exemplifying different types of arguments in the biochemistry articles, in order to study their correspondences to the argumentation schemes we have previously defined for genetics [7] as well as to general models of argument [26, 14]. Also we would like to develop argument mining rules from the instances.

Figures 1 to 6 present a sample of the arguments, including a related excerpt from the text, and the paraphrased *Observations*, *Background* (if any) and *Claim* of each argument. In some cases the examples involve intermediate conclusions, described as *Inferences*. *Inferences* were not necessarily stated in the text, but are needed to show argument structure. See the example in Fig. 1, from an article on cross-linked protein structure [15]. Note that the analysis refers to entities such as Full XL, b158C, F1F0, Delta and BPM as a form of shorthand reference for the sake of readability. *Observation 1* describes how Full XL was created by linking the b158C subunit of F1F0 to another subunit of F1F0. The rest of the example attempts to answer the question, to what subunit was the b158C subunit linked in Full XL? Fig. 1 contains two subarguments: *Inference 1* follows from *Background 1* and *Observation 2* by *Inference to the best explanation*; and *Inference 2* from *Background 2* and *Observation 3* also by *Inference to the best explanation*. (Note that *Background 1* and *2* function as warrants.) The main claim is independently supported by each of *Inference 1* and *2*.

The two *Inference to the best explanation* arguments of Fig. 1 are instances of the general Abductive Argumentation Scheme for Argument from Effect to Cause [26, p. 172], in which "F is a finding or given set of facts in the form of some event that has occurred. E is a satisfactory causal explanation of F. No alternative causal explanation E′ given so far is as satisfactory as E. Therefore, E is plausible, as a hypothesis for the cause of [F]." In Fig. 1, each *Observation* presents F, a finding, and each *Inference* is a plausible causal explanation E of the finding. Although the *Claim* is independently supported by each *Inference*, the recognition of Full XL by Delta antibodies (*Observation* 3) would be very good evidence that the other protein in Full XL is Delta (since antibodies are normally highly specific in their binding interactions, such that they generally bind well only to one protein).

Individual cysteine residues were introduced into full-length b at positions 150, 151, and 155. A fourth construct was made that encoded a protein, referred to as b_{158C}, having two residues, glycine and cysteine, attached to the C terminus of b. … Membrane preparations bearing F_1F_0 complexes containing the mutated b subunits were incubated with the photoreactive cross-linker benzophenone-4-maleimide (BPM) and then exposed to ultraviolet light. The samples were analyzed by Western blotting, using ^{125}I-radiolabeled monoclonal antibodies raised against b as probes. No cross-linking was observed with either bD150C or bK151C (data not shown). However, cross-linked products of about the same size were observed with both bE155C and b158C (Fig.3). The new bands were approximately the size expected for a b-δ cross-link and showed reactivity with anti-δ polyclonal antibodies (Fig. 3), indicating that cross-links had been formed between b and δ [15].

Observation 1: Full XL [a *cross-linked protein*] arises when b158C [*modified b subunit of F1F0*], F1F0 [*F_1F_0 ATP synthase*], and activated BPM [*a cross-linker, which makes b158C highly reactive*] are mixed**.**
Background 1: If b158C were cross-linked to Delta in Full XL, then Full XL would have a certain size.
Observation 2: Full XL is the size expected if BPM joined b158C to the Delta subunit of F1F0.
Inference 1 (from Bk1, Ob2 **Inference to the best explanation**): b158C and Delta are cross-linked in Full XL.
Background 2: If b158C were cross-linked to Delta in Full XL, then Full XL would be recognized by antibodies specific to Delta
Observation 3: Full XL is recognized by antibodies specific to Delta.
Inference 2 (from Bk2, Ob3 **Inference to the best explanation**): b158C and Delta are cross-linked in Full XL.
Claim (independently supported by Inf1 and Inf2)**:** b158C and Delta are cross-linked in Full XL.

Figure 1: Inference to the best explanation

Fig. 2 illustrates arguments in another article on cross-linked protein structure [17]. There are three arguments: *Inference 1* is derived by *Default inference* from *Background 1* and *Observations 1 and 2*; *Inference 2* by *Default inference* from *Inference 1*, *Observation 3*, and *Background 2*; and the *Claim*

derived by *Inference to the best explanation* from *Background 3*, *Inference 2*, and *Observation 4*. Note that the *Background* premises function as warrants. *Default Inference* is plausible but defeasible deductive inference. As discussed with Fig. 1, *Inference to the best explanation* is related to the Abductive Argumentation Scheme for Argument from Effect to Cause [26], where the *Observations* present F and Claim is the plausible causal explanation E.

Treatment with 10 μM $CuCl_2$ for 20 min resulted in almost complete conversion of δ to b−δ in b158C/δM158C membranes, as estimated from the disappearance of the b and δ bands and the appearance of the cross-linked band (Figure 1) [17].

Observation 1: Treatment of F1F0 [*F_1F_0 ATP synthase*] with $CuCl_2$ [*a catalyst*] causes disappearance of b158C [*modified b subunit of F1F0*].
Observation 2: Treatment of the F1F0 with $CuCl_2$ causes disappearance of DeltaM158C [*modified δ subunit of F1F0*].
Background 1: If b158C and DeltaM158C disappear after treatment with $CuCl_2$, then the F1F0 no longer contains them in their original (un cross-linked) form.
Inference 1 (from Ob1, Ob2, Bk1 **Default inference**): The F1F0 no longer contains b158C and DeltaM158C in their original (un cross-linked) form.
Observation 3: Treatment of the F1F0 with $CuCl_2$ causes appearance of Full XL.
Background 2: If the F1F0 no longer contains b158C and DeltaM158C in their original (un cross-linked) form (Inf 1) and treatment of the F1F0 with $CuCl_2$ causes appearance of Full XL (Ob 3), then b158C may be cross-linked to DeltaM158C in the Full XL.
Inference 2 (from Inf1, Ob3, Bk2 **Default inference**): b158C may be cross-linked to DeltaM158C in the Full XL.
Observation 4: The Full XL is recognized by antibodies against the b and Delta subunits.
Background 3: If treatment of F1F0 with $CuCl_2$ causes b158C to be cross-linked to DeltaM158C in Full XL then treatment of the F1F0 with $CuCl_2$ causes appearance of Full XL (Ob3), and b158C may be cross-linked to DeltaM158C in Full XL (Inf 2), and the Full XL is recognized by antibodies against the b and Delta subunits (Ob 4).
Claim (from Inf2, Ob3, Ob4, Bk3 **Inference to the best explanation**): Treatment of F1F0 with $CuCl_2$ causes b158C to be cross-linked to DeltaM158C in Full XL.

Figure 2: Inference to the best explanation

The argument described in Fig. 3, from another article on cross-linked protein structure [16], refutes the proposal that a certain protein has a "coiled-coil" structure. *Background* 1 conflicts with *Observations 1* and *2*, thereby supporting the *Claim*. The argument type has been dubbed *Inconsistent with expectation*. This is an instance of the general, defeasible Argument from Falsification [26, p. 331]: "*Major premise*: If *A* (a hypothesis) is true, then *B* (… an event) will be observed to be true. *Minor Premise*: *B* has been observed to be false, in a given instance. *Conclusion*: Therefore, *A* is false."

A similar experiment was performed with bsyn proteins containing mutations at positions 59–65 or 68 in the heptad repeat region… (Fig. 3). Of these positions, the A59C (heptad b position) and S60C (heptad c position) proteins showed the highest tendency to form disulfides…. Cysteines in the other heptad positions had poor propensities to form disulfides (Fig. 3). In a coiled-coil domain, cysteines at the d positions of the heptad repeat (residues 61 and 68) would be expected to show the highest tendency to form disulfide bonds … The result that the b and c positions in this region of bsyn showed the highest propensities to form disulfide bonds is inconsistent with the presence of a coiled-coil structure, in which predominantly hydrophobic residues in the a and d positions form an interface [16].

Background 1 – In a coiled-coil domain, cysteines at the d positions [*e.g. 61 and 68 of bsyn,* a *truncated version of b subunit of F_1F_0 ATP synthase*] of the heptad repeat [*a repeating pattern of amino acid residues consisting of 7 residues per repeat that is characteristic of proteins interacting through a coiled-coiled domain*] are expected to show the highest tendency to form disulfide bonds .
Observation 1: Of positions 59-65 and 68 of bsyn, cysteines at positions 59 and 60 showed the highest tendency to form disulfides.
Observation 2: Cysteines at the other positions (61-65, 68) had poor propensity to form disulfides.
Claim (Inconsistent with expectation): Observations 1 and 2 are inconsistent with the presence of a coiled-coil domain in region 59-68 of bsyn.

Figure 3. Inconsistent with expectation

Fig. 4, based upon another article on cross-linked protein structure [17], concerns two experiments designed to answer: Does the formation of the cross-link between b158C and DeltaM158C subunits allow protons to move across the membrane ("leak") through ATP synthase in the absence of catalytic activity by ATP synthase? The experiments used NADH (nicotinamide adenine dinucleotide), a source of electrons that allows creation of a proton gradient across the membrane. The purpose of the first experiment, comparing wild type to the mutated ATP synthase, was to rule out that the mutations alone did not allow protons to "leak". After having shown that they did not, the second experiment was performed to show that the cross-link of the mutated subunits did not allow protons to "leak". (Basically, with respect to the property of leakiness to protons, the system behaves the same whether the cross-link is there or not. Therefore, the cross-link has no effect on this property of the enzyme.)

Observations 1 and *2* have been decomposed into two parts, *Observations 1a-b* and *2a-b*, respectively, to show that they each make an argument whose type is referred to as *Difference has no effect*. This type of argument is related to *Mill's Method of Difference* [14] (see the next example), but in this variant the observed effect of two situations differing in one respect is the same, thus the difference between the situations is not significant in producing that type of effect. This example also illustrates the composition of multiple arguments; the claim of the second *Difference has no effect* argument, *Inference 2*, is a premise of a *Default inference* argument, i.e. a plausible but defeasible deductive inference. Note that the premise of this argument annotated *Previous Claim* is a conclusion of an argument made earlier in the article.

The ability of the membranes to maintain a proton gradient when supplied with NADH was also tested. When disulfide bond formation was not induced, the b158C/δM158C membranes were able to maintain a proton gradient as effectively as the wild-type membranes (Figure 2C,D). Treatment of the b158C/δM158C membranes with 10 μM $CuCl_2$ resulted in no significant change in the proton gradient in either wild type or mutant membranes (Figure 2C,D) … These results show that the ability of the b158C/δM158C membranes to maintain a stable proton gradient was not affected by disulfide bond formation between b and δ [17].

Observation 1: Upon treatment with NADH, membranes that include ATP synthase with b158C and DeltaM158C mutations create a proton gradient just the same as membranes containing wild type ATP synthase [*i.e., not containing b158C and DeltaM158C mutations*].
1a: Membranes containing wild type ATP synthase are able to maintain an existing proton gradient.
1b: Membranes containing ATP synthase with the b and delta mutations are able to maintain a proton gradient the same as in 1a.
Inference 1 (Difference has no effect): Therefore, when supplied with NADH, the presence of mutations in b and Delta does not change ATP synthase function with respect to the ability to maintain a proton gradient.

Observation 2: Treatment with $CuCl_2$ does not impair the ability of wild type or mutant membranes to generate a proton gradient when supplied with NADH.
2a: Membranes containing wild type ATP synthase AND treated with $CuCl_2$ are able to maintain an existing proton gradient.
2b: Membranes containing ATP synthase with the b and Delta mutations AND treated with $CuCl_2$ are able to maintain a proton gradient equivalent to the one seen in 2a.
Inference 2 (Difference has no effect): Therefore, after treatment with $CuCl_2$ when supplied with NADH, the presence of mutations in b and Delta does not change ATP synthase function with respect to the ability to maintain a proton gradient

Previous Claim: Treatment of ATP synthase containing b158C and DeltaM158C mutations with $CuCL_2$ causes b158C and DeltaM158C to be cross-linked.

Claim (Default inference from Inference 2 & Previous claim**)**: Cross-linking of b158C to DeltaM158C in ATP synthase does not affect the ability of membranes to maintain a stable proton gradient.

Figure 4. Difference has no effect

Fig. 5, on an article [20] in a different subdomain of biochemistry, signal transduction pathways, involves multiple arguments, where the conclusions of the first two *Method of difference* arguments support the claim of the third argument, illustrating *Argument from Correlation*. (As in the previous example, *Observations 1* and *2* have been decomposed into *1a-b* and *2a-b*, respectively.) The namesake of our *Method of difference* argument type is Mill's Method of Difference: "where the only distinguishing feature marking situations in which phenomenon *a* occurs or does not occur is the presence or absence of phenomenon *A*, there is reason to think that *A* is an indispensable part of the cause of *a"* [14]. In the argument for *Inference 1*, the distinguishing feature marking situations in which cell cycle arrest occurs is treatment or stopping treatment with alpha factor; for *Inference 2* the distinguishing feature marking situations in which phosphorylation of FAR1 occurs is treatment or stopping treatment with alpha factor. Thus, alpha factor may have a causal role in cell cycle arrest and phosphorylation, i.e., they are correlated. In general, an argument of correlation makes a claim that there is an association between two events; when one occurs so does the other. (However, a correlation between two events does not imply a causal relationship between them.)

FAR1 protein is very rapidly phosphorylated when haploid a cells are exposed to α factor (Figure 1A; Chang and Herskowitz, 1992) and is dephosphorylated within minutes after cells are released from a factor arrest (M. P., unpublished data). These results show a correlation between phosphorylation of FAR1 in response to a factor and its ability to arrest the cell cycle [20].

Observation 1: When yeast cells are exposed to alpha factor *[a pheromone that induces cell cycle arrest in yeast*] and cell cycle arrest is induced, FAR1 [*a protein which is required for alpha-factor-induced cell cycle arrest in yeast*] is rapidly phosphorylated [*phosphorylation is the addition of a phosphate group to a protein molecule*].
1a: Treatment with alpha factor of yeast cells containing FAR1 results in cell cycle arrest.
1b: Treatment with alpha factor of yeast cells containing FAR1 results in phosphorylation of FAR1.
Observation 2: When cells are released from alpha-factor-induced arrest, FAR1 is rapidly dephosphorylated.
2a: Stopping treatment with alpha factor of cells containing FAR1 stops cell cycle arrest.
2b: Stopping treatment with alpha factor of cells containing FAR1 results in dephosphorylation.
Inference 1 (from Ob1a, Ob2a **Method of Difference**): Alpha factor has a causal role in cell cycle arrest.
Inference 2 (from Ob1b, Ob2b **Method of Difference):** Alpha factor has a causal role in phosphorylation of FAR1.
Claim (**Argument of Correlation**, from Inf 1-2)**:** Phosphorylation of FAR1 is correlated with alpha-factor-induced cell cycle arrest.

Figure 5: Method of Difference and Argument of Correlation

In another example (Fig. 6) from the same article, a similar chain of arguments from *Method of Difference* to *Argument of Correlation* ends in an *Argument from Correlation to Cause*. First, *Inference 1* follows from *Observation 1* and *Previous Observation 1* by *Method of Difference*. Likewise, *Inference 2* follows from *Observation 2* and *Previous Observation 2* by *Method of Difference*. *Inference 2.1* is a default inference from *Inference 2* and *Background 1*. The *Argument of Correlation*, from Inference 1 (i.e., the inability of the truncated FAR1 to arrest cell cycle in response to alpha factor) and Inference 2.1 (the inability of the truncated FAR1 to form the complex with (bind to) CDC28-CLN2 in response to alpha factor), "demonstrates a correlation between the ability of FAR1 to bind to CDC28-CLN2 and its ability to arrest the cell cycle."

The *Claim* of the main argument follows from *Background 2* and *Inference 3*. The main argument is an instance of the general, defeasible Argument from Correlation to Cause described as follows [26, p. 174]: "There is a positive correlation between *A* and *B*. Therefore, *A* causes *B*" (p. 174). Note that the causal conclusion is not strongly asserted; the authors hedge that they "interpret these results to indicate that formation of this complex is required for cell cycle arrest in response to external signals." However, some evidence in favor of the causal relationship is provided in *Background 2*.

To determine whether formation of a complex between FAR1 and the CDC28-CLN2 kinase is functionally important for cell cycle arrest in response to α factor, we have carried out a deletion analysis of the FAR1 protein. The truncated FAR1 proteins were expressed in a FAR1 deletion mutant and tested for their ability to arrest the cell cycle in response to α factor using halo assays (Figure 5A) and growth in liquid culture (M. P., unpublished data). Representative data are shown in Figure 5, demonstrating that deletion of residues 150-235 had no effect on arrest (Figure 5A) whereas deletion of residues 235-340 or 150-340 (Figures 5B and 5C) inactivated FAR1. The constructs were also tested for their ability to bind CDC28 and CLN2 using coimmunoprecipitation experiments (Figures 5B and 6). Figure 5B shows, for example, that FAR1 deleted for residues 150-235 was still able to bind CDC28 (lane 9) whereas other deletions, such as those removing residues 235-340 or 150-340 (lanes 10 and 11), did not bind to CDC28. ,,, The analysis of FAR1 mutants demonstrates a correlation between the ability of FAR1 to bind to CDC26-CLN2 and its ability to arrest the cell cycle. We interpret these results to indicate that formation of this complex is required for cell cycle arrest in response to external signals [20].

Observation 1: Cells in which the only version of FAR1 [*a protein required for alpha-factor-induced cell cycle arrest in yeast*] is FAR1Δ235-340 [*portion of FAR1 lacking residues 235-340*] do not arrest their cell cycle in response to alpha factor.
[**Previous Observation 1**: Cells expressing FAR1 arrest their cell cycle in response to alpha factor.]
Inference 1 (Method of Difference): The deleted part of FAR1 is necessary for alpha-factor-induced cell cycle arrest.
Observation 2: FAR1Δ235-340 from cells treated with alpha factor does not immunoprecipitate [*form a complex with*] CDC28.
[Previous Observation 2: FAR1 from cells treated with alpha factor does immunoprecipitate [*form a complex with*] CDC28.]
Inference 2 (Method of Difference): The deleted part of FAR1 is necessary for formation of a complex with CDC28.
Background 1: CDC28 requires CLN2 to act; if CDC28 is active, it must be as the CDC28-CLN2 complex.
Inference 2.1 (Default inference from Inf2 & Bk 1**):** The deleted part of FAR1 is necessary for formation of a complex with CDC28-CLN2.
Inference 3 (Argument of Correlation from Inf 1 & 2.1**):** Therefore, the (in)ability of FAR1 to form a complex with CDC28-CLN2 is correlated with the cell's (in)ability to arrest the cell cycle in response to alpha factor.
Background 2: Formation of a complex resulting in cell cycle arrest is typical of signaling systems.
Claim (from Inf 3 and Bk 2 **Argument from Correlation to Cause):** Binding of FAR1 to CDC28-CLN2 is required for cell cycle arrest in response to alpha factor.

Figure 6. Method of Difference, Argument of Correlation, and Argument from Correlation to Cause

To sum up what these examples illustrate, firstly, we could not have understood the arguments without the interpretation of the articles by a domain expert (McLachlin). The expert identified the main claims and their supporting data, a task made especially challenging since some arguments relied on implicit background knowledge, inferences from data in the figures, and/or observations or claims given earlier in the article. Secondly, the arguments have warrants (the *Background* premises) and are instances of general argumentation schemes and Mill's "methods" of science. In other words, the general models are indeed a "good fit" for certain arguments in this scientific domain and genre.

It is an interesting question why some argument types were found only in certain subdomains, i.e., Inference to the best explanation, Inconsistent with expectation, Difference has no effect in articles about cross-linked proteins, and Method of Difference, Argument of Correlation, Argument from Correlation to Cause in articles about signal pathways. This could be a coincidence (due to the small sample size) or happen to depend on the scientific goals of a particular article, e.g., to challenge a claim that a certain cross-linked protein has a coiled-coil structure.

4. **Discussion and Conclusions**

Signal transduction pathways and protein structure are currently active areas of biochemistry research with important practical applications, e.g., in cancer treatment and pharmacology. The ability to mine arguments in these domains would be very beneficial to researchers. In our previous work [5,6], we proposed how certain argument schemes implemented as rules could be used to mine arguments from

genetics research articles after extraction of entities and relations. As a step towards developing rules for mining arguments in the biochemistry literature, we are formulating rules based on the arguments that we have catalogued. In the rules, typed variables replace biochemical terms to enable the rule to apply to more cases than the example on which it is based. For example, from the argument in Figure 1 (repeated below with *Inferences* omitted), the following two rules were abstracted:

Argument:
Observation 1: Full XL arises when b158C, F1F0, and activated BPM are mixed**.**
Background 1: If BPM joined b158C to the Delta subunit of F1F0 in Full XL, then Full XL would have a certain size.
Observation 2: Full XL is the size expected if BPM joined b158C to the Delta subunit of F1F0.
Background 2: If b158C were cross-linked to Delta in Full XL, then Full XL would be recognized by antibodies specific to Delta.
Observation 3: Full XL is recognized by antibodies specific to Delta.
Claim (**Inference to the best explanation**): Full XL is b158C cross-linked to Delta.

Rule 1:
Premise: Mixing cross-linker C with synthase X containing mutated subunit B results in cross-linked protein XL
Premise: Size of XL is same as expected size of protein in which B is cross linked to subunit D of X
Conclusion: XL is a protein in which D and B subunits of X are cross-linked.

Rule 2:
Premise: Mixing cross-linker C with synthase X containing mutated subunit B results in cross-linked protein XL
Premise: D antibodies recognize XL.
Conclusion: XL is a protein in which D and B subunits of X are cross-linked.

Whether or not such an approach is feasible depends upon certain considerations. First, is it feasible to automatically extract from a text the biochemical entities and relations that are referred to in the rules? The Pathway Curation (PC) shared task competition [22] showed that "existing event extraction technology can generalize to meet the novel challenges represented by the … PC task settings, suggesting that extraction methods are capable of supporting the construction of knowledge bases on … biomolecular pathway models." It appears that the entities and relations targeted in that competition could be used to express the claims of the signal transduction pathway article we examined. However, we found that it would be necessary to extend the set of entities and relations that were defined for the PC task in order to describe the experimental observations referred to in the argument mining rules for signal transduction pathways. Similarly, current research on information extraction of protein-protein interactions [21] may provide a starting point for extracting entities and relations referred to in the protein structure rules.

Another feasibility issue to consider is that not all of the evidence used in the arguments is presented in the text, i.e., it may be presented only in figures. A related issue is the role of implicit and explicit background knowledge in the arguments. Should it be encoded in the rules? Is it feasible to create a knowledge base of such background knowledge for use in extracting arguments?

Our research on argumentation in biochemistry research articles is ongoing. So far, we have seen that the argument types are instantiations of argumentation schemes listed in [26] and "methods" of scientific reasoning [14]. In our previous research on genetics arguments [7], we found a number of instantiations of argumentation schemes and "methods" that we have not yet found in biochemistry articles but would not be surprised to find in the future, e.g., arguments related to Mill's Method of Agreement. In fact, if we were to revisit the genetics literature, no doubt we would find some of the argument types we have identified in biochemistry but not in genetics, e.g., argument of correlation and argument from correlation to cause. It remains to be seen whether argument schemes for biochemistry can be formulated at a useful level of abstraction comparable to that developed in [7] for genetics.

We are continuing to analyze argumentation in biochemistry research articles, and hope to create a small open-access catalog of the arguments for future argumentation studies.

5. Acknowledgements

We thank the CMNA24 reviewers for their comments and suggestions.

6. References

[1] K. Al-Khatib et al., Argument mining for scholarly document processing: Taking stock and looking ahead, in Proceedings of the ACL Workshop on Scholarly Document Processing, 2021.

[2] M. Alliheedi, R. E. Mercer, R. Cohen, Annotation of rhetorical moves in biochemistry articles, in Proceedings of the 6th Workshop on Argument Mining, Florence, Italy, August 1, 2019, pp. 113–123.

[3] A. Fergadis et al., Argumentation mining in scientific literature for sustainable development, in Proceedings of the 8th Workshop on Argument Mining, 2021, 100-111.

[4] N. L. Green, Identifying argumentation schemes in genetics research articles, in Proceedings of the 2nd Workshop on Argument Mining, 2015.

[5] N. L. Green, Towards mining scientific discourse using argumentation schemes, Argument and Computation 9 (2018) 121–135. DOI 10.3233/AAC-180038

[6] N. L. Green, Proposed method for annotation of scientific arguments in terms of semantic relations and argument schemes, in Proceedings of the 5th Workshop on Argument Mining, 2018.

[7] N. L. Green, An open-access catalog of argumentation schemes in biomedical genetics research articles, February 26, 2018. Available at https://github.com/greennl/BIO-arg.

[8] N. L. Green, Argumentation schemes for Computational Models of Natural Argument: From Genetics to International Relations to Science Policy to AI Ethics. *Argument and Computation*, 2021, 12(3):397-416.

[9] N. L. Green, L. J. Crotts, A first experiment using ILP for argument mining, in Proceedings of Computational Models of Natural Argument, CMNA 21, 2021. CEUR-WS.org/Vol-2937/paper7.pdf.

[10] N. L. Green, R. Dwight, K. Navoraphan, and B. Stadler. Natural Language Generation of Transparent Arguments for Lay Audiences. *Argument and Computation.* Volume 2, Issue 1 (2011), Pages 23-50.

[11] B. Kanoksilapatham, Rhetorical structure of biochemistry research articles, English for Specific Purposes 24 (2005) 269-292.

[12] A. Lauscher, G. Glavas, and S. P. Ponzetto. An argument-annotated corpus of scientific publications, in Proceedings of the 5th Workshop on Argument Mining, 2018.

[13] M. Liakata, S. Saha, S. Dobnik, C. Batchelor, and D. Rebholz-Schuhmann. 2012. Automatic recognition of conceptualization zones in scientific articles and two life science applications. Bioinformatics, 28(7):991-1000.

[14] C. Macleod, "John Stuart Mill", *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2020/entries/mill/>.

[15] D. T. McLachlin, J. A. Bestard, and S. D. Sunn. The b and δ subunits of the *Escherichia coli* ATP synthase interact via residues in their C-terminal regions, Journal of Biological Chemistry 273(24) (1998) 15162-8. (Available at https://www.jbc.org/article/S0021-9258(19)65404-9/fulltext)

[16] D. T. McLachlin and S. D. Dunn. 1997. Dimerization interactions of the b subunit of the Escherichia coli F_1F_0-ATPase. *Journal of Biological Chemistry*, 272(34):21233-21239.

[17] D. T. McLachlin and S. D. Dunn, Disulfide linkage of the b and δ subunits does not affect the function of the *Escherichia coli* ATP Synthase, Biochemistry 39(12) (2000)3486-90. (Available at https://pubs.acs.org/doi/10.1021/bi992586b)

[18] Y. Mizuta, A. Korhonen, T. Mullen, and N. Collier. 2006. Zone analysis in biology articles as a basis for information extraction. *International Journal of Medical Informatics*, 75(6):468-487.

[19] E. Moser and R.E. Mercer. Use of claim graphing and argumentation schemes in biomedical literature: A manual approach to analysis, in *Proceedings of the 7th Workshop on Argument Mining*, pp. 88-99, 2020.

[20] M. Peter, A. Gartner, J. Horecka, G. Ammerer, and I. Herskowitz, FAR1 links the signal transduction pathway to the cell cycle machinery in yeast, *Cell* (4) (1993) 747-760. (Available at https://doi.org/10.1016/0092-8674(93)90254-N)

[21] S. Pyysalo et al. Comparative analysis of five protein-protein interaction corpora, BMC Bioinformatics 2008, 9(3) 1-11.

[22] S. Pyysalo et al. Overview of the cancer genetics and pathway curation tasks of BioNLP shared task 2013, BMC Bioinformatics 2015, 16(Suppl 10):S2.

[23] C. Staab et al. 2014. Argumentation mining in persuasive essays and scientific articles from the discourse structure perspective. In Proceedings of the Workshop on Frontiers and Connections between Argumentation Theory and Natural Language Processing, Forlì-Cesena, Italy, July 21-25, 2014, volume 1341 of CEUR Workshop Proceedings. CEUR-WS.org.

[24] S. Teufel. 2010. *The Structure of Scientific Articles: Applications to Citation Indexing and Summarization*. CSLI Studies in Computational Linguistics. Center for the Study of Language and Information.

[25] S.E. Toulmin. *The uses of argument*. Cambridge University Press, Cambridge, 2003.

[26] D. Walton, C. Reed, and F. Macagno. *Argumentation schemes*. Cambridge University Press, Cambridge, 2008.