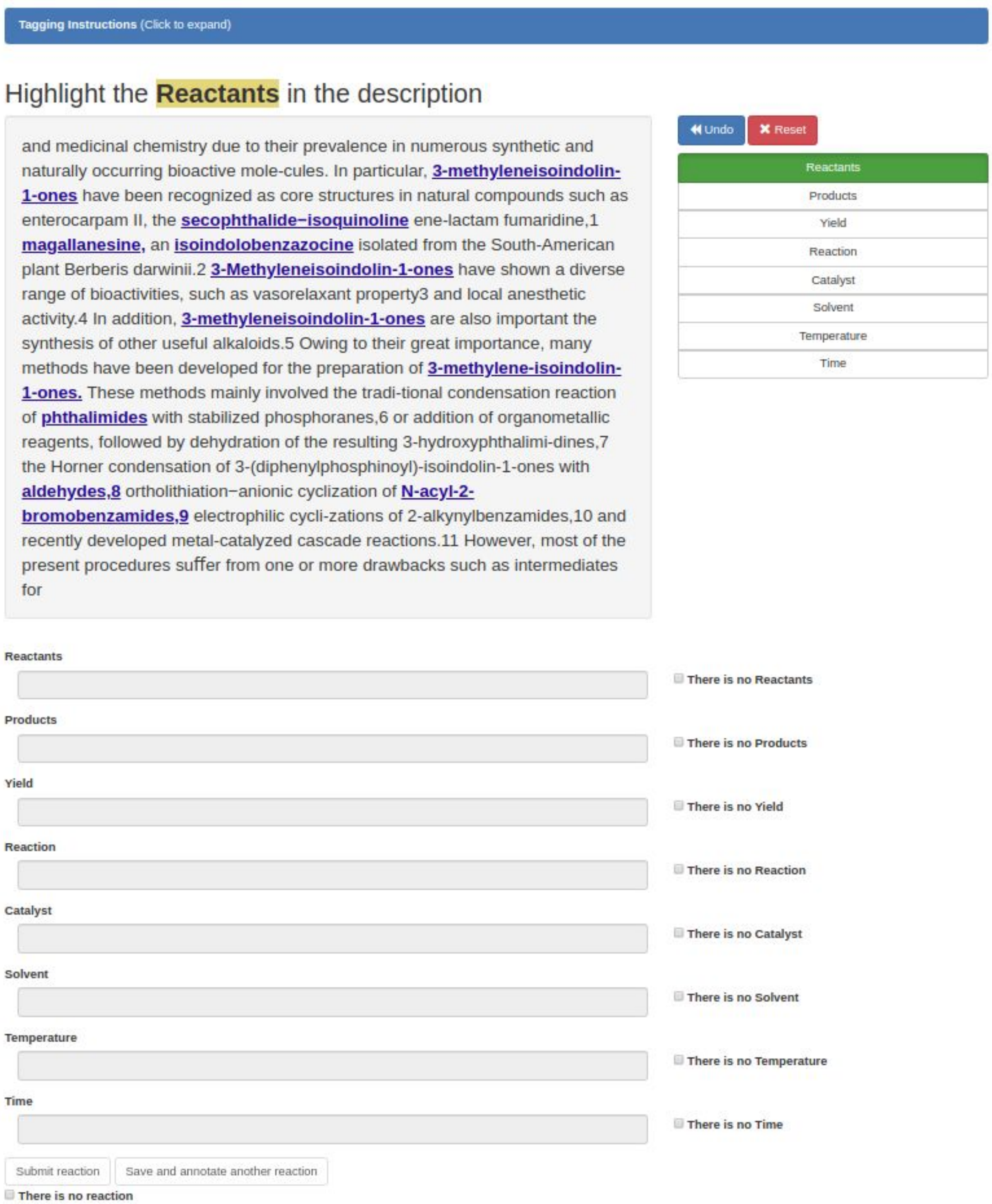
Annotation Instructions

A. INTRODUCTION

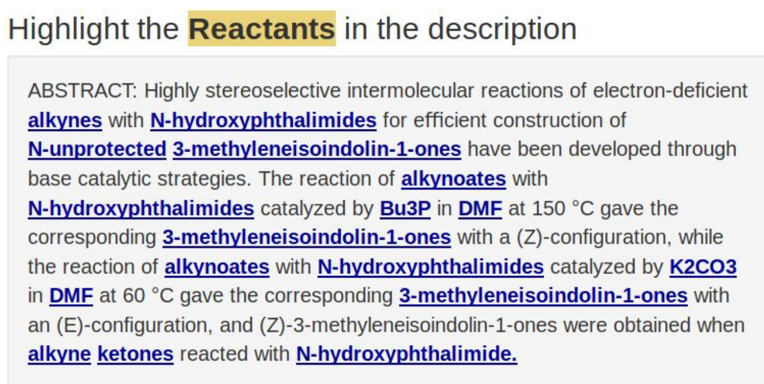
The goal of the Chemical Literature Annotation is the identification (annotation) of the following entities and their roles, related to specific chemical reactions in scientific literature:

- Reactants
- Products
- Yield
- Reaction type
- Catalyst(s)
- Solvent(s)
- Temperature
- Time (duration)



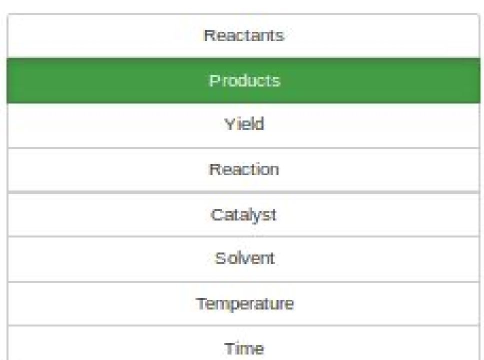
B. ANNOTATION PROCESS

1. The text paragraph to be annotated will be shown in the text window on the left side of the screen.
    - The paragraph shown may include <u>from none to multiple</u> chemical reactions.
    - <u>Only one</u> reaction at a time must be annotated.
    - Most chemical entities are automatically highlighted in order to facilitate annotation. Nevertheless, some relevant entities might not be identified automatically so that they should be manually identified by the user.

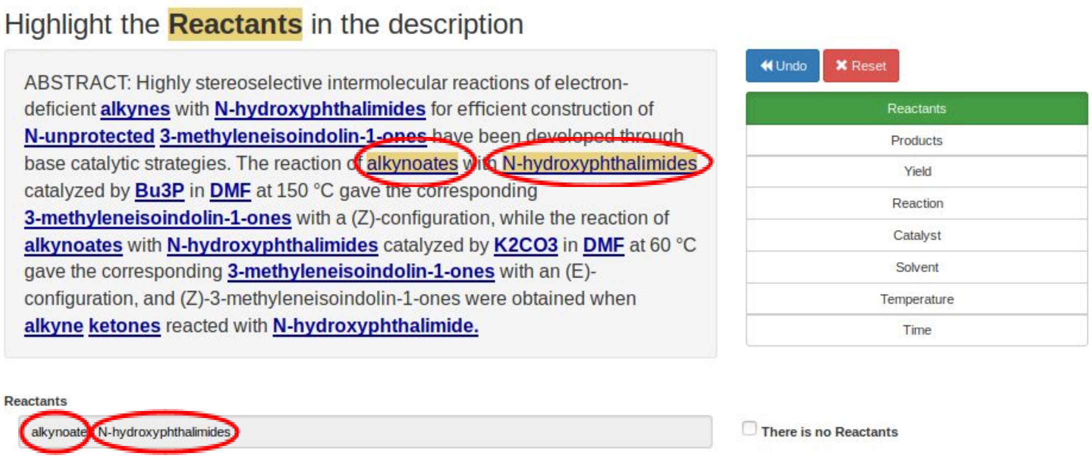


FOR EACH REACTION IN THE PARAGRAPH REPEAT SETPS 2 – 6 BELOW :

2. Select (highlight in green) in the window on the right the entity class to be annotated (Reactants, Products, Yield, Reaction type, Catalyst, Solvent, Temperature, Time).



3. Select the relevant word or span in the paragraph so that it is included in the entity field.

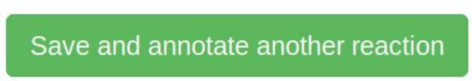


4. In case there is more than one entity belonging to an entity class (e.g. more than one reactant), repeat step 3, so that all of them are included in the entity class field.

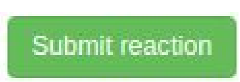
5. In case there are no entities related to the entity class being annotated, check the '**There is no [Item]**' box next to the entity class field.

6. Once the current reaction has been annotated:

- If the current paragraph includes additional reactions:
    - Click on **'Save and annotate another reaction'**



    - Repeat steps 2 - 6
- If the current paragraph does not include additional reactions:
    - Click on **'Submit Reaction'**



    - If there are more paragraphs to annotate: Repeat steps 1 - 7 for the subsequent paragraphs

7. Once there are no more paragraphs to annotate, the tool will show a 'DONE' message and no more paragraphs will be shown.

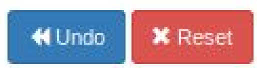
C. MISSING REACTIONS

- The annotation tool extracts those paragraphs which may be describing chemical reactions most likely. Nevertheless, some paragraphs may not include any reaction. In this case, the **'There is no reaction'** box should be checked, followed by the **'Submit Reaction'** button in order to start the annotation of the next paragraph.

D. CORRECTIONS

The **'Undo'** and **'Reset'** buttons allow to make corrections during the annotation process:

- Undo: Undoes the last action
- Reset: Clears the active entity field



E. GENERAL ANNOTATION RULES

This section describes the general annotation rules for the different types of entities.

Legend:

- 'Annotate'
- '~~Do not annotate~~'

1. General

- In case a paragraph is describing more than one reaction:
  - Annotate all entities explicitly linked to the reaction being annotated
  - In case an entity is not explicitly mentioned in the reaction being annotated but it is implicitly linked to it by being mentioned in other reaction within the paragraph, annotate it.

    Example:

    "On the basis of conventional Knoevenagel reactions, we initially investigated reaction conditions using benzaldehyde (1a) and dimethyl malonate (2a) as a model substrate (Table 1). When a reaction was performed with 10 mol % of InBr3 in toluene at 60 °C for

8 h, only 3% of Knoevenagel product 3aa was detected (entry 1) [Reaction 1]. Thus, to promote the initial abstraction of the activated proton, the addition of 1 equiv of several bases to the reaction mixture was examined. Consequently, when the primary amine 2-aminoethanol was added, the yield was remarkably increased to 61% (entry 2) [Reaction 2]."

Annotate [Reaction 2] as follows:

- **Reactants**: benzaldehyde, (1a), dimethyl malonate, (2a) [Reaction 1]
- **Products**: Knoevenagel, 3aa [Reaction 1]
- **Yield:** 61% [Reaction 2]
- **Reaction:** Knoevenagel [Reaction 1]
- **Catalyst:** InBr3 [Reaction 1], primary amine 2-aminoethanol [Reaction 2]
- **Solvent:** toluene [Reaction 1]
- **Temperature:** 60 [Reaction 1]
- **Time:** 8h [Reaction 1]

- Do not annotate expected results but actual results. Example:
  - 'Since CuCl(PPh 3) is a widely studied and common copper salt, the active catalytic species was expected to be ~~CuCN-(PPh 3)~~ in our optimized reaction condition'
- Annotate not only the reactions obtained in the actual experiments but also past reactions described. Example:
  - 'Both the Ullmann and Buchwald-Hartwig aminations are well-known copper-catalyzed crosscoupling reactions between an aryl halide and an amine'

2. **Conditions**

- Do annotate compound expressions when they form a single word. Examples:
  - 'Copper-catalyzed crosscoupling reactions'
  - 'Phosphine-catalyzed [3 + 2] annulation'
- Do not annotate references to reaction conditions. Examples:
  - '~~condition A~~ shows the highest reactivity'
- Do annotate generic reaction conditions. Example:
  - 'Elevated temperature'

3. **Chemical entities (reactants, products, catalysts, and solvents)**

- Annotate both the entity and its corresponding reference when they are shown together. Example:

- 'We started our investigation by examining the Bu3P catalyzed reaction of 2-hydroxyisoindoline-1,3-dione (1a) and ethyl propiolate (2a)'

- Annotate references (to Schemes, Tables, Figures, etc.) to entities. Example:
  - 'The reaction of 1a with 2a in the presence of Bu3P in DMF…'
- Annotate all references to an entity when shown together. Example:
  - pivalonitrile (2,2,-dimethylpropanenitrile, bp 105 °C)
- Do not annotate references to entity types but to specific entities. Examples:
  - 'the initial ~~hexameric PPh3-bounded copper cyanide~~ species'
  - 'the expected ~~alcohols~~'
  - 'the corresponding ~~alkene 3aa~~'
  - 'however, only ~~colorless crystals~~ were obtained'
- Do not annotate generic references to entities. Examples:
  - '~~the product~~', '~~final product~~', '~~the reactant~~'
  - '~~the corresponding product~~'
  - '~~same starting materials~~'
- Do not annotate entity classes but only specific entities. Examples:
  - 'While the ~~C-2 arylated product~~ (~~4aa~~) was the major product'
  - '~~Benzoxazoles bearing substituents with diverse electronic properties such as methyl (1b-1d), phenyl (1e), and chloro (1f) groups'~~
  - 'effectively promotes the Knoevenagel condensation of ~~aromatic / aliphatic / heteroaromatic aldehydes~~ with a ~~variety of activated methylene compounds~~'

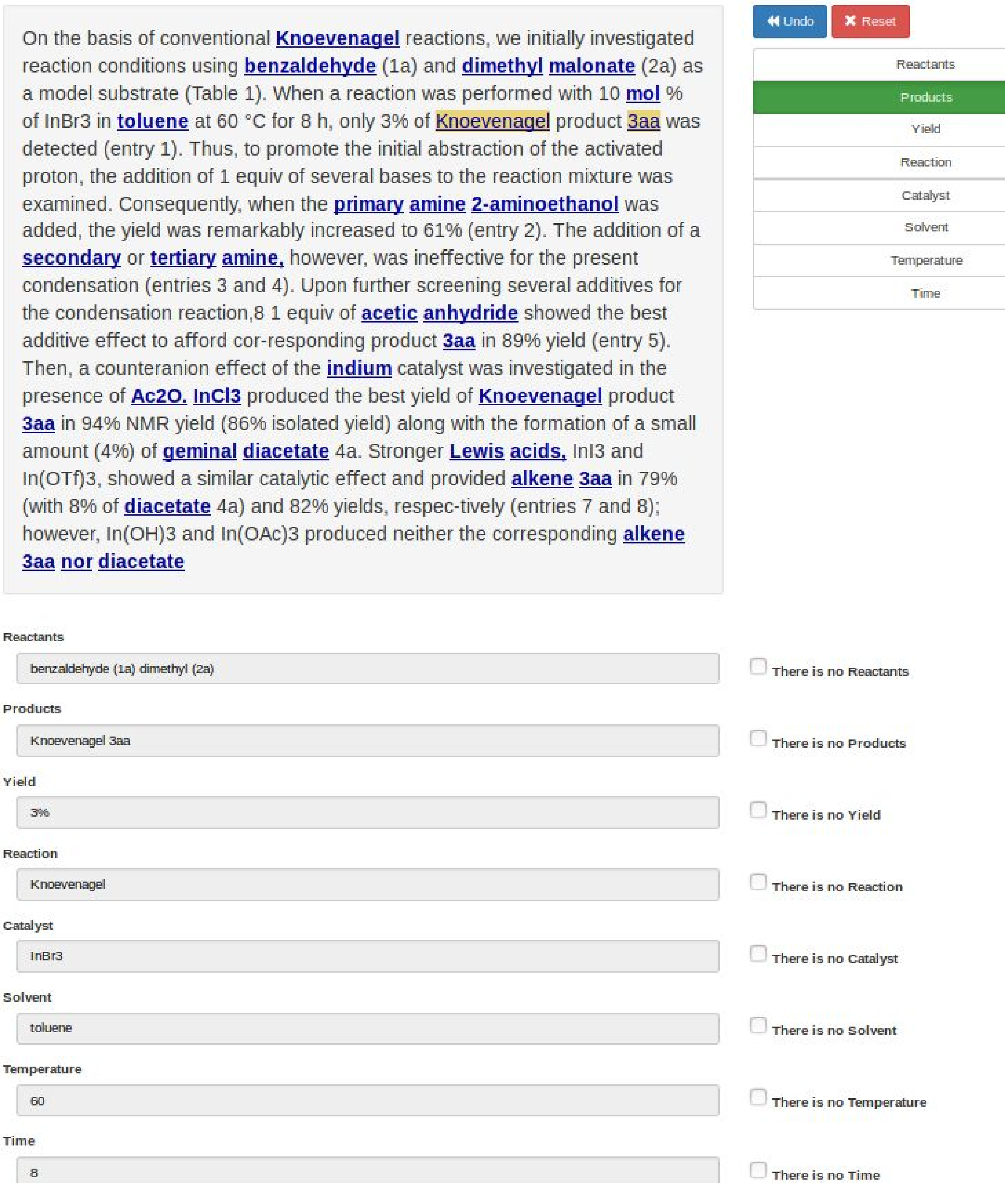
4. **Yields**

- Annotate generic yield expressions. Examples:
  - 'similar yield', 'lower yield', 'higher yield', 'diminished yield', 'improved reaction yield'
- Annotate all yield measures in case more than one is provided. Examples:
  - 'Higher yield (62%)'
- Alternative expressions of yield: **TBD – Ask Thomas**
  - 'increased catalytic efficiency'
  - 'improved reactivity'
  - 'perfect selectivity'
  - 'significant effect'
  - 'highest conversion'
  - 'no reactivity'
  - 'no further reaction observed'

F. EXAMPLE 1

The following example shows the annotation of two different reactions present in a single paragraph and with no missing entity classes.

Reaction #1:



Reaction #2:

On the basis of conventional **Knoevenagel** reactions, we initially investigated reaction conditions using benzaldehyde (1a) and dimethyl malonate (2a) as a model substrate (Table 1). When a reaction was performed with 10 **mol** % of InBr3 in **toluene** at 60 °C for 8 h, only 3% of **Knoevenagel** product **3aa** was detected (entry 1). Thus, to promote the initial abstraction of the activated proton, the addition of 1 equiv of several bases to the reaction mixture was examined. Consequently, when the **primary amine 2-aminoethanol** was added, the yield was remarkably increased to 61% (entry 2). The addition of a **secondary** or **tertiary amine,** however, was ineffective for the present condensation (entries 3 and 4). Upon further screening several additives for the condensation reaction,8 1 equiv of **acetic anhydride** showed the best additive effect to afford cor-responding product **3aa** in 89% yield (entry 5). Then, a counteranion effect of the **indium** catalyst was investigated in the presence of **Ac2O. InCl3** produced the best yield of **Knoevenagel** product **3aa** in 94% NMR yield (86% isolated yield) along with the formation of a small amount (4%) of **geminal diacetate** 4a. Stronger **Lewis acids,** InI3 and In(OTf)3, showed a similar catalytic effect and provided **alkene 3aa** in 79% (with 8% of **diacetate** 4a) and 82% yields, respec-tively (entries 7 and 8); however, In(OH)3 and In(OAc)3 produced neither the corresponding **alkene 3aa nor diacetate**

Undo | Reset

- Reactants
- Products
- Yield
- Reaction
- Catalyst
- Solvent
- Temperature
- Time

Reactants

benzaldehyde (1a) dimethyl malonate (2a)

☐ There is no Reactants

Products

Knoevenagel 3aa

☐ There is no Products

Yield

89%

☐ There is no Yield

Reaction

Knoevenagel

☐ There is no Reaction

Catalyst

InBr3 primary amine 2-aminoethanol

☐ There is no Catalyst

Solvent

toluene

☐ There is no Solvent

Temperature

60

☐ There is no Temperature

Time

8

☐ There is no Time

G. EXAMPLE 2

The following example shows the annotation of a reactions with missing entity classes (yield, reaction type and time).