

ChemScanner(-DB): an initiative to collect chemistry data by a community effort

Dr. Nicole Jung (Chemotion Coordinator)

Patrick Hodapp – Stefan Bräse Group

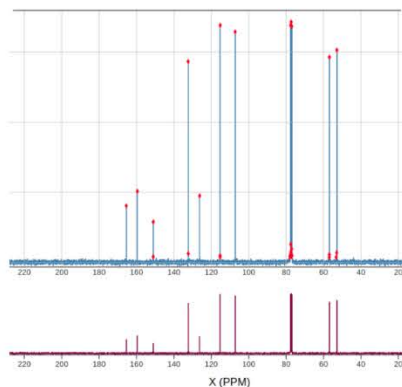
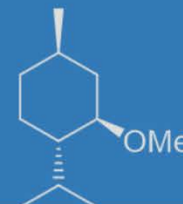
INSTITUTE OF ORGANIC CHEMISTRY



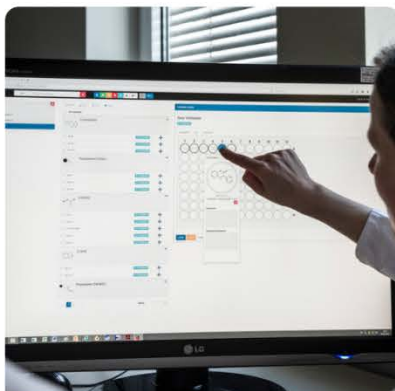
Chemotion
funded by
DFG



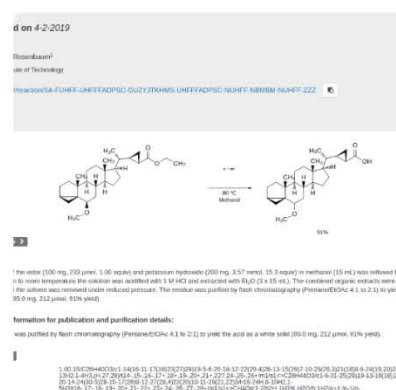
Electronic Laboratory Notebook & Repository for Research Data



Spectra Viewer



Electronic Lab Notebook



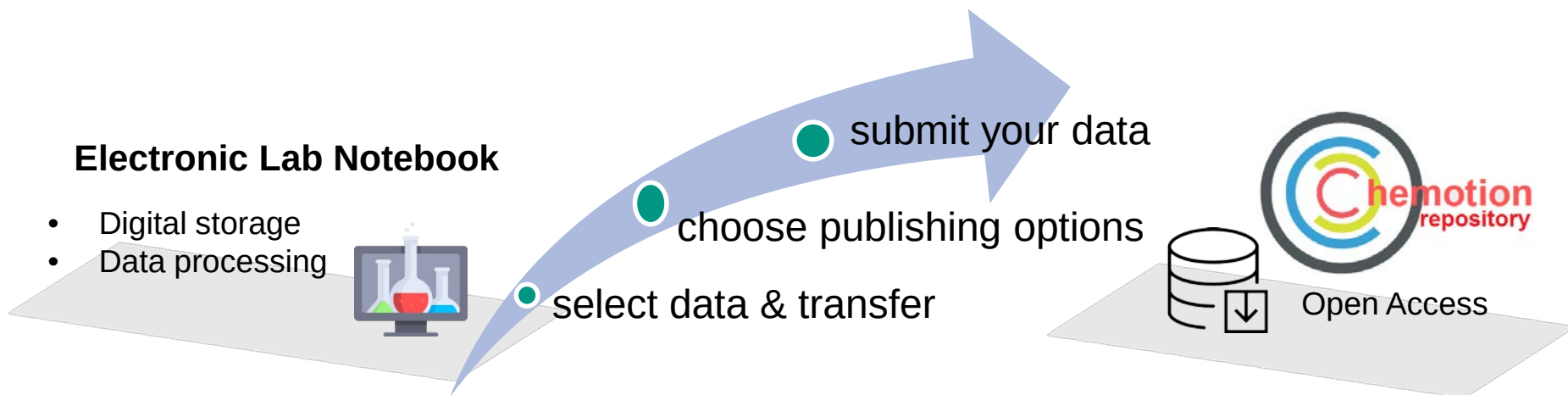
Chemotion Repository



ChemScanner

Visit us on YouTube

Chemotion Repository: Benefits and Workflow



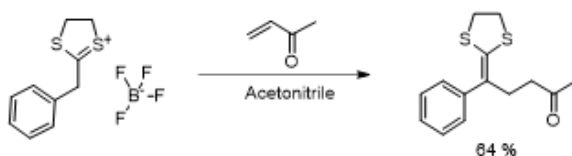
Repository for molecules and reactions including related research data
(descriptions, procedures, properties, analytics)

FAIR data concepts/infrastructure start now, will work in future, but what about the past?

- ELNs and repositories are basic instruments for our current and future work
- Knowledge of the past decades is reserved to commercial databases
- Scientific work depends on all data/community knowledge

How to re-use existing literature information?

5-(1,3-dithiolan-2-ylidene)-5-phenylpentan-2-one (4a)



According to General Procedure 2a: {A|1a} 2-benzyl-4,5-dihydro-1,3-dithiane tetrafluoroborate (1.000 g, 3.544 mmol, 1.00 equiv); {B|3a} but-3-en-2-one (0.298 g, 3.544 mmol, 1.20 equiv); {C} acetonitrile (5.00 mL); Yield {D|4a} = 64% (0.600 g, 2.268 mmol). The obtained crude product was purified via flash-chromatography on silica gel (cyclohexane/ethyl acetate 10:1, R_f = 0.44 (cyclohexane/ethyl acetate 4:1)).

^1H NMR (400 MHz, CDCl_3 , ppm) δ = 1.99 (s, 3H), 2.40–2.45 (m, 2H), 2.73–2.79 (m, 2H), 3.18–3.23 (m, 2H), 3.32–3.37 (m, 2H), 7.15–7.21 (m, 3H), 7.23–7.29 (m, 2H); ^{13}C NMR (100 MHz, CDCl_3 , ppm) δ = 29.8, 32.8, 37.8, 37.9, 41.5, 126.3, 127.1, 128.0 (2C), 128.3 (2C), 142.0, 208.1; EI (m/z , 70 eV, 50 °C): 264 (70) $[\text{M}]^+$, 207 (100), 181 (33), 131 (33), 105 (33); HRMS–EI (m/z): $[\text{M}]$ calcd for $\text{C}_{14}\text{H}_{16}\text{OS}_2$, 264.0643; found, 264.0642; IR (ATR, cm^{-1}): 1710, 1595, 1489, 1439, 1420, 1357, 1278, 1158, 1073, 1027, 899, 849, 791, 761, 705, 597, 534, 472, 430 cm^{-1} .

1

Name recognition &
Translation to structure

2

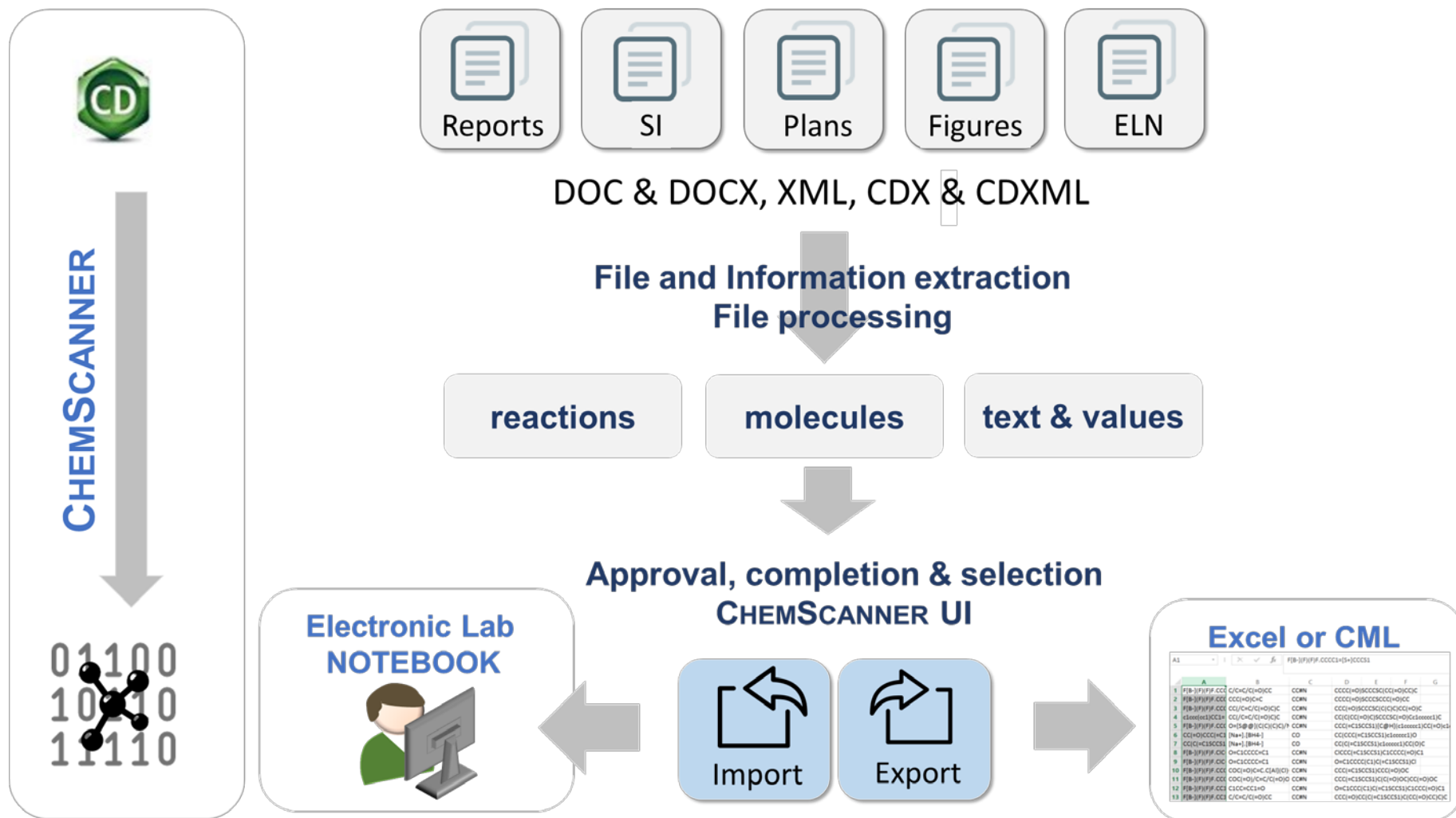
.cdx file analysis &
split to components

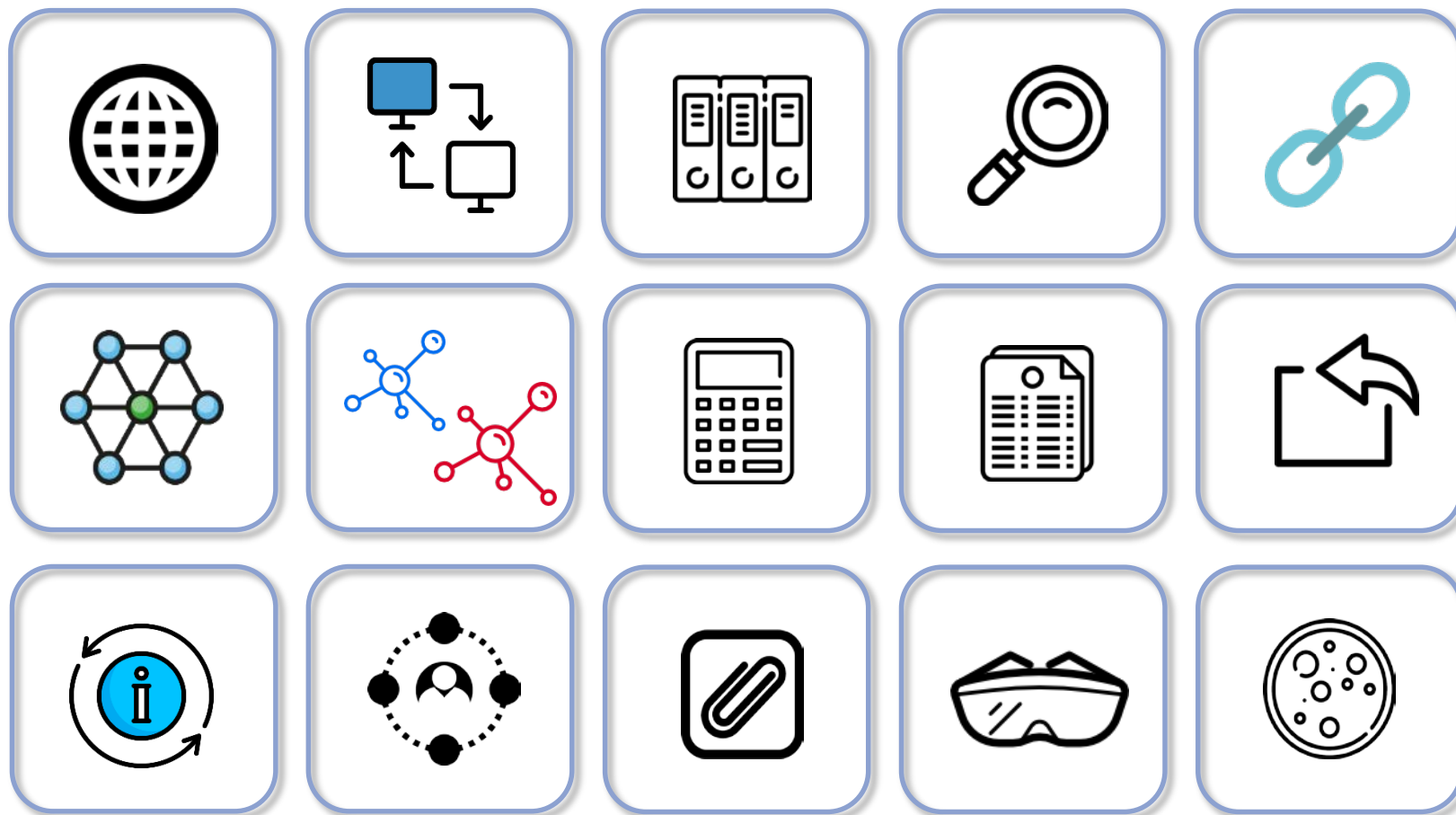
3

Additional information
store & extract

File types: .doc or .docx

ChemScanner – Current workflow





OPEN SOURCE

Source: Our findings &
S. Kanza, C. Willoughby, N. Gibbins, R. Whitby, J.G. Frey, J. Erjavac, K. Zupanic, M. Hren, K. Kovac, *J Cheminfo*. **2017**, 9:31.

ELN – Reactions

Secure | <https://eln.chemotion.net/#/collection/192/reaction/336>

Chemotion

All | IUPAC, InChI, SMILES, ...

Logged in as Nicole Jung

Collections

- All
- chemotion.net
- Collection 1
- BJOC data**
- Collection 2
- Collection 3
- Test Import
- My shared collections
- Shared with me
- Synchronized with me
- Inbox

219(0) 62(0) 1(0) 1(0) 1(0)

NJu-R243 According to General Procedure 2a

NJu-R242 According to General Procedure 2a

NJu-R241 According to General Procedure 2a

NJu-R240 According to General Procedure 2a

NJu-R239 According to General Procedure 2a

NJu-R238 According to General Procedure 2a

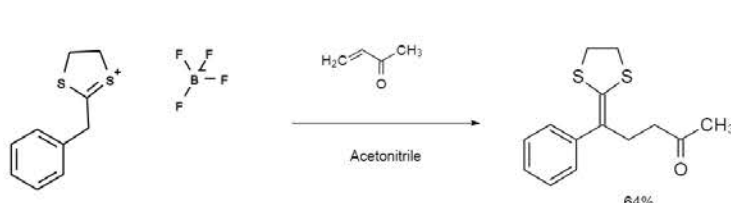
NJu-145 NJu-R24-A

NJu-R243 According to General Procedure 2a

NJu-R243 According to General Procedure 2a

From To

Reaction Scheme:



64%

Scheme Properties References Analyses Green Chemistry Zotero

Starting materials	Ref	T/R	Amount	Conc	Equiv
NJu-772					
2-benzyl-4,5-dihydro-1,3-dithiol-1-ylidene			1000 mg	0.00 ml	3.544 mmol
					708.9 mmol
					1.000 mmol

Reactants	Reagents	T/R	Amount	Conc	Yield
but-3-en-2-one			298.1 mg	0.00 ml	4.253 mmol
					850.6 mmol
					1.200 mmol

Products	T/R	Amount	Conc	Yield
NJu-773 NJu-R243-A				
5-(1,3-dithiolan-2-ylidene)-5-phenylpentan-3-one		599.8 mg	0.00 ml	2.268 mmol
				453.7 mmol
				64% mmol

Solvents

Name: According to General Procedure 2a Status: Select... Temperature: Temperature... °C

Role: Parts of GP According to

ELN – Reporting and Supporting Information

NJu-773 NJu-R243-A



$C_{14}H_{16}OS_2$

5-(1,3-dithiolan-2-ylidene)-5-phenylpentan-2-one

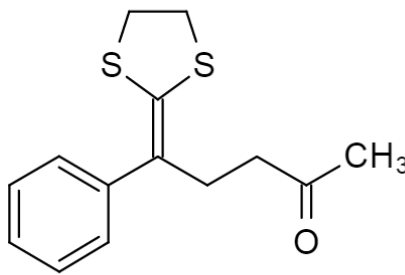
264.406240 g/mol

Exact mass: 264.064257 g/mol



2731035047





Properties

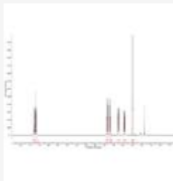
Analyses

Results

NMR

Edit mode

Add analysis



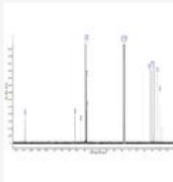
new

Type: 1H NMR

Status:

Content: 1H NMR (400 MHz, $CDCl_3$, ppm) δ = 1.99 (s, 3H), 2.40–2.45 (m, 2H), 2.73–2.79 (m, 2H), 3.18–3.23 (m, 2H), 3.32–3.37 (m, 2H), 7.15–7.21 (m, 3H), 7.23–7.29 (m, 2H).

☒ Add to Report 



new

Type: ^{13}C NMR

Status:

Content: ^{13}C NMR (100 MHz, $CDCl_3$, ppm) δ = 29.8, 32.8, 37.8, 37.9, 41.5, 126.3, 127.1, 128.0 (2C), 128.3 (2C), 134.3, 142.0, 208.1.

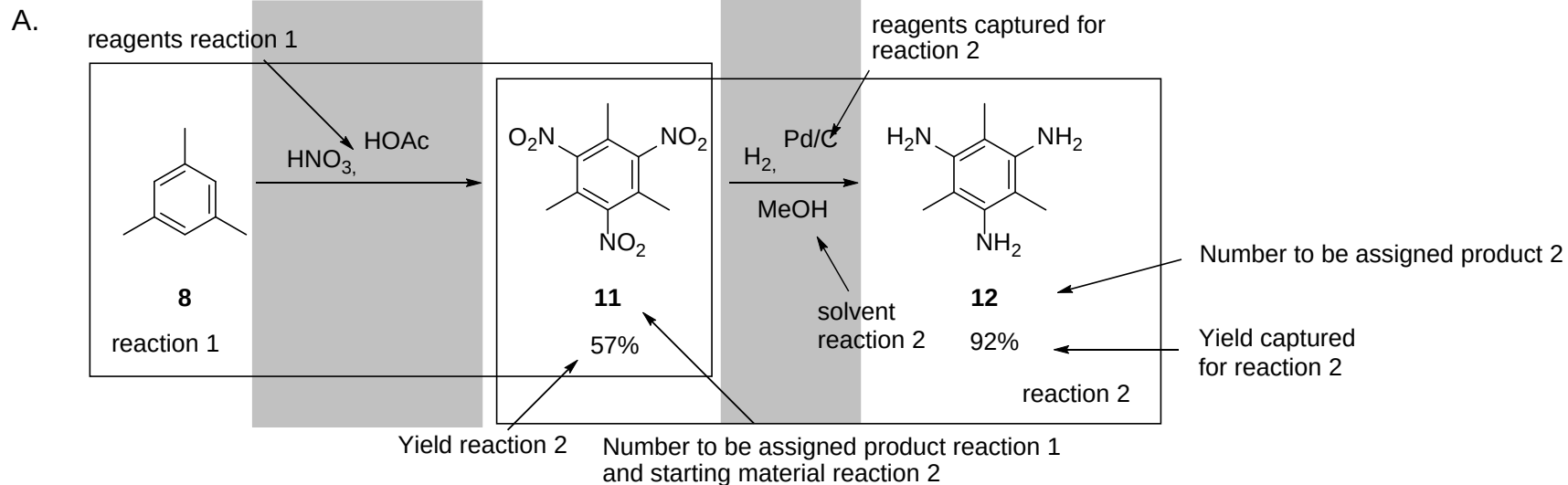
☒ Add to Report 

Features Stand Alone Software

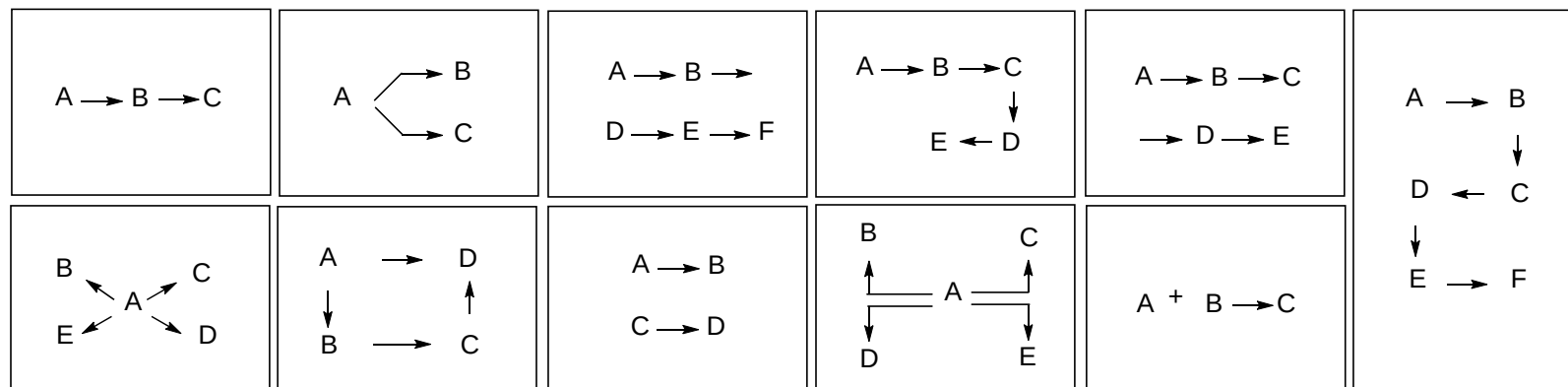
- Add superatoms (such as protective groups)
- Define abbreviations
- Add solvents or reagents
- Accept or decline
- Export to Excel or SDF

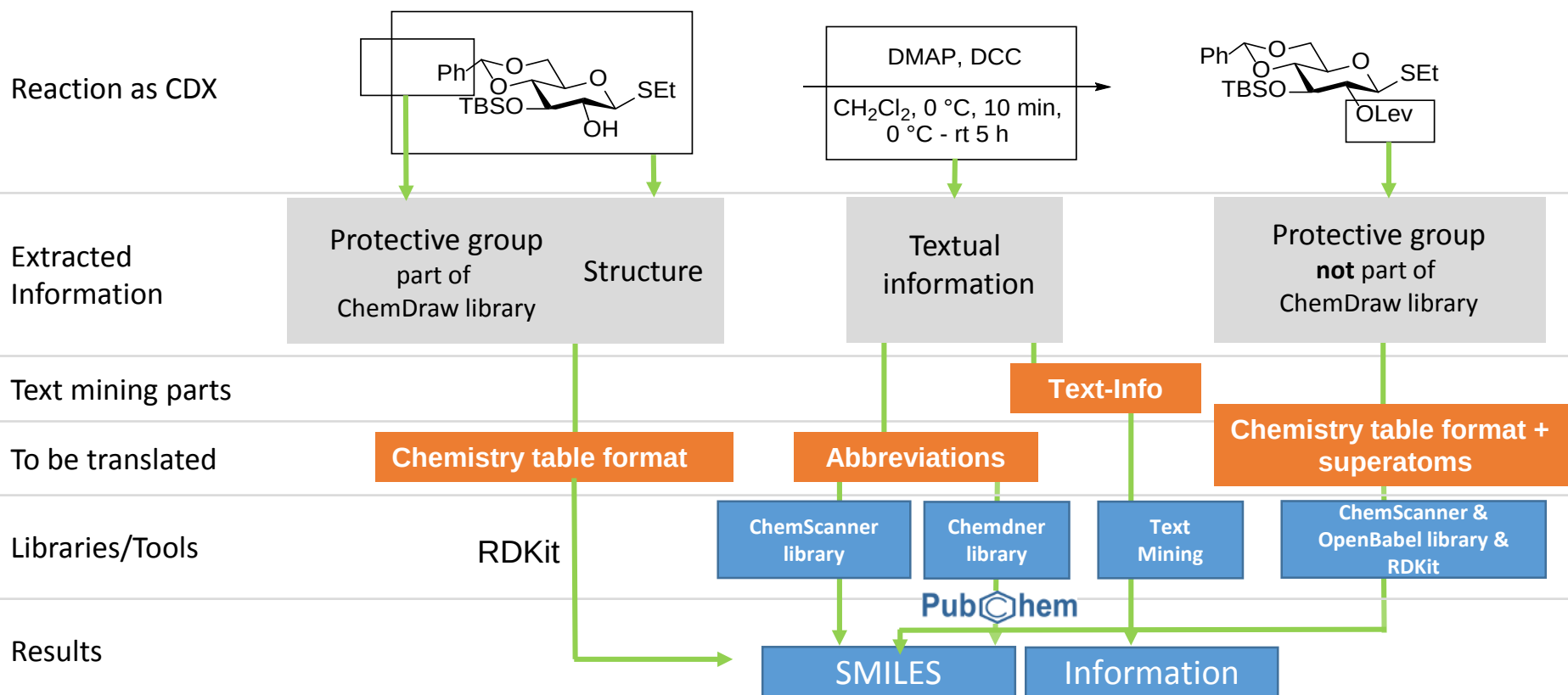
Features ELN embedded version

- Add non-predefined reagents
- Change order
- Correct mistakes in the structures
- Store and complete with other data



B.





ChemScanner – Import of data from *.doc or *.docx

ChemScanner – Extraction

Import Reactions from Docs

View

Abbreviations

Export

CML

Excel

SMILES Editing

Solvents

Select Solvents

Reagents

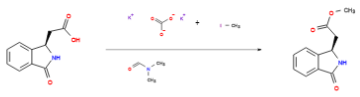
Select All options

Molecules Reactions

Drop Files, or Click to add File.

Export Example No1.cdx

Export Example No1.cdx



reaction

- description: . K2CO3, CH3I DMF
- status: **Success**
- yield: 98%

reactant 1

- text: 95% ee
- label: 1

product 1

- text: , 98%
- label: 2



reaction

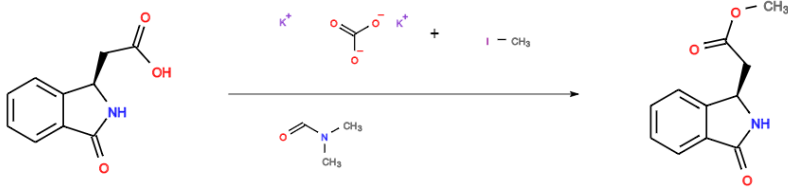
- description: THF, 0 °C LiBH4
- status: **Success**
- temperature: 0 °C
- yield: 88%

reactant 1

- text: , 98%
- label: 2

product 1

- text: , 88%
- label: 3



reaction

- description: . K2CO3, CH3I DMF
- status: **Success**
- yield: 98%

reactant 1

- text: 95% ee .
- label: 1

product 1

- text: , 98%
- label: 2

ChemScanner – Import to ELN

Chemotion

All | IUPAC, InChI, SMILES, ...

Logged in as Template Moderator

Collections

- All
- chemotion.net
- Test for Import/Extraction +
- My shared collections
- Shared with me
- Synchronized with me
- Inbox

4(0) 2(0)

From To

☐	TMo-R111	
☐	TMo-R110	

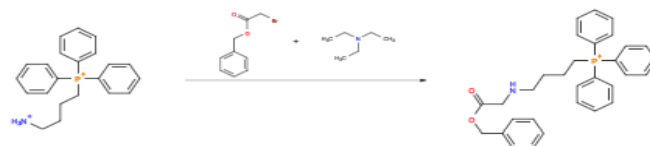
Show 15

KIT
Karlsruher Institut für Technologie



ChemScanner – Import of Data from CD-ELN

ChemScanner – Import of Data from CD-ELN



Hide Details

Reaction Conditions

- **Reaction Molarity:** 0,014 molar
- **Preparation:** Zu einer Lösung aus (4-ammoniumbutyl)triphenylphosphonium (0,5500 g, 1,640 mmol) und triethylamine (1,371 ml, 9,84 mmol) in Dichlormethane (80 ml) wird über ca. 12 Stunden langsam eine Lösung aus Benzylbromacetat (0,260 ml, 1,640 mmol) in Dichlormethane (40 ml) zugegeben. Die Lösung wird zur Aufarbeitung mit Wasser gewaschen und die wässrige Phase wird mit Dichlormethan extrahiert. Die vereinigten organischen Phasen werden über Natriumsulfat getrocknet, filtriert und vom Lösungsmittel befreit. Zur Reinigung wird eine Säulenchromatographie durchgeführt. DC: DCM/MeOH 10:1 Saule: DCM/MeOH 10:1 Produkt, aber nicht sauber, wird aber trotzdem weiter umgesetzt. (48-60)

Reaction Description

- **Name:** Benzylbromacetat
- **Limiting?:** false
- **Molecular Weight:** 229,07
- **Equivalents:** 1,000
- **Moles:** 1,640 mmol
- **Sample Mass:** 0,376 g
- **Volume:** 0,260 ml
- **Density:** 1,446 g/ml
- **Formula Mass:** 229,07
- **Reactant Mass:** 0,376 g

Solvents

- **Name:** Dichlormethane
- **Ratio:** 2,000
- **Volume:** 80 ml

reactant 1

- **Reaction Label:** I
- **Name:** (4-ammoniumbutyl)triphenylphosphonium
- **Molecular Formula:** C₂₂H₂₆NP²⁺
- **Limiting?:** true
- **Molecular Weight:** 495,23
- **Equivalents:** 1,000
- **Moles:** 1,640 mmol
- **Sample Mass:** 0,5500 g
- **Formula Mass:** 335,430
- **Reactant Mass:** 0,550 g

reagent 2

- **Reaction Label:** II
- **Name:** triethylamine
- **Molecular Formula:** C₆H₁₅N
- **Limiting?:** false
- **Molecular Weight:** 101,193
- **Equivalents:** 6,000
- **Moles:** 9,84 mmol
- **Sample Mass:** 0,996 g
- **Volume:** 1,371 ml
- **Density:** 0,726 g/ml
- **Formula Mass:** 101,193
- **Reactant Mass:** 0,996 g

product 1

- **Reaction Label:** III
- **Product ID:** P2
- **Name:** (4-((2-(benzyloxy)-2-oxoethyl)amino)butyl)triphenylphosphonium
- **Molecular Formula:** C₃₁H₃₃NO₂P⁺
- **Actual Mass:** 0,1620 g
- **Actual Moles:** 0,336 mmol
- **% Yield:** 20,47 %
- **Molecular Weight:** 482,583
- **Equivalents:** 1,000
- **Theoretical Moles:** 1,640 mmol
- **Theoretical Mass:** 0,791 g
- **Formula Mass:** 482,583

reaction

- **description:** [II]
- **status:** Success

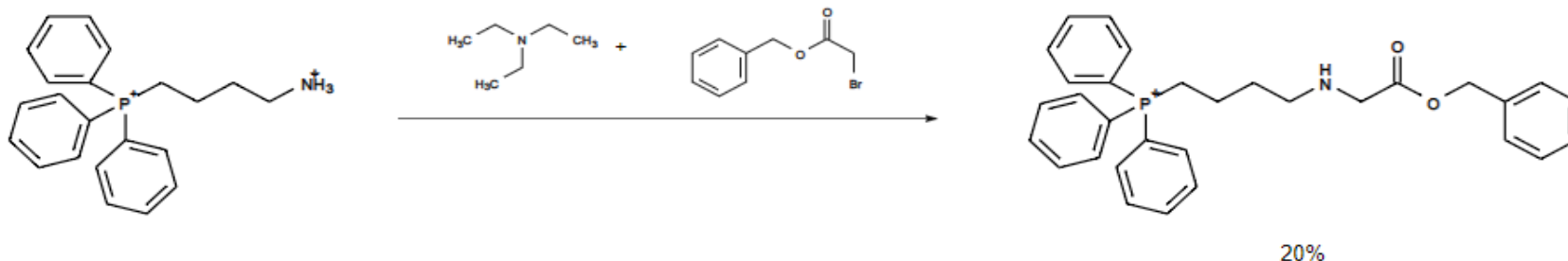
reactant 1

- **text:** [I]

product 1

- **text:** [III]

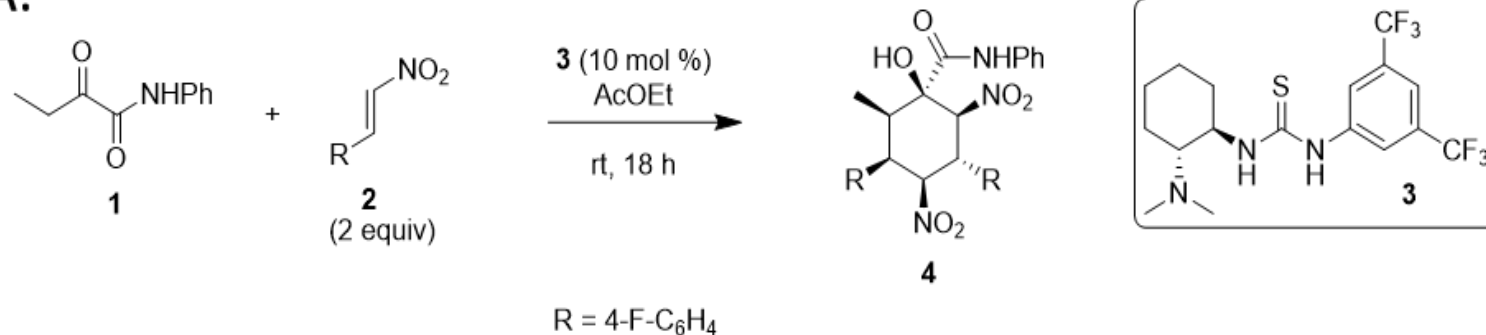
ChemScanner – Import of Data from CD-ELN



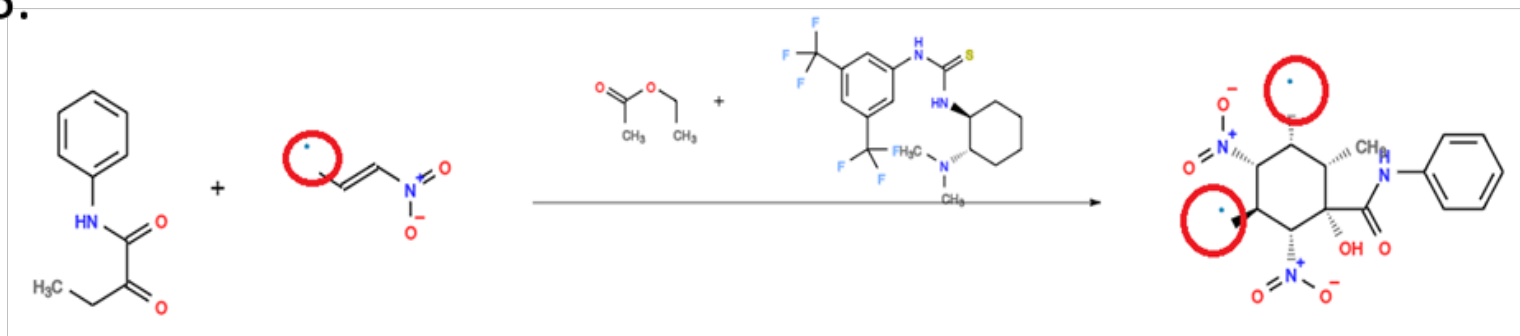
Scheme Properties References Analyses Green Chemistry

+	Starting materials	Ref	T/R	Amount		Conc		Equiv					
+	A TMO-258 (4-ammoniobutyl)triphenylphosphonium C22H26NP++		r	550.1	mg	0.00	ml	1.640	mmol	n.d.	mmol/l	1.000	
+	Reactants	Reagents											
+	B N,N-diethylethanamine		r	995.7	mg	1.37	ml	9.840	mmol	n.d.	mmol/l	6.000	
+	C benzyl 2-bromoacetate		t	0.000	mg	0.00	ml	0.000	mmol	n.d.	mmol/l	0.000	
+	Products		T/R	Amount		Conc		Yield					
+	P1 TMO-259 (4-((2-(benzyloxy)-2-oxoethyl)amino)butyl)triphenylphosphonium C31H33NO2P+		r	162.1	mg	0.00	ml	0.3360	mmol	n.d.	mmol/l	20%	

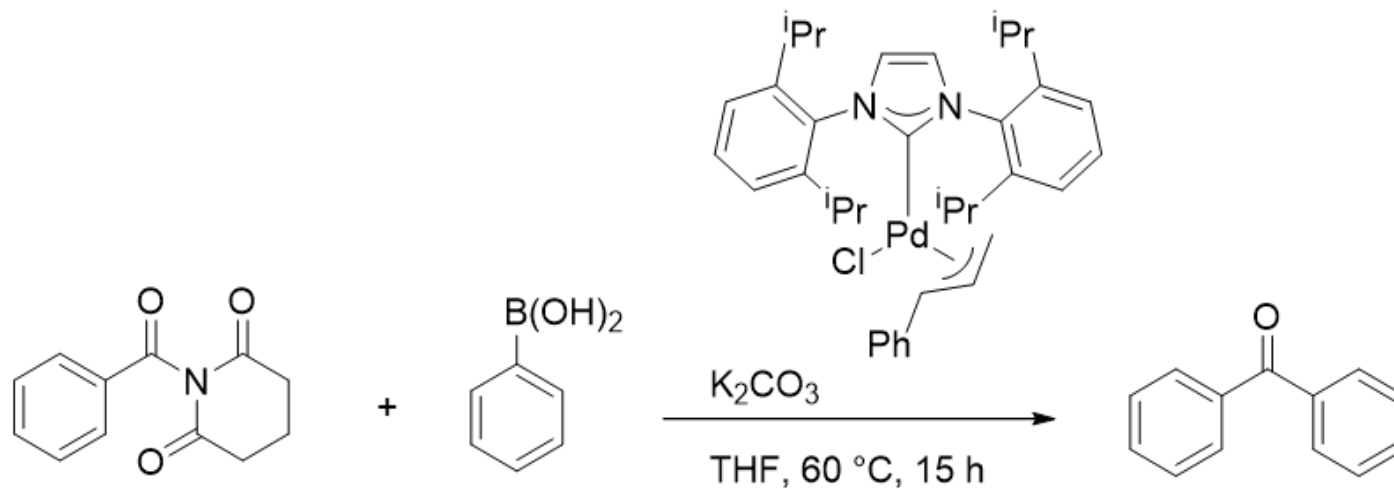
A.



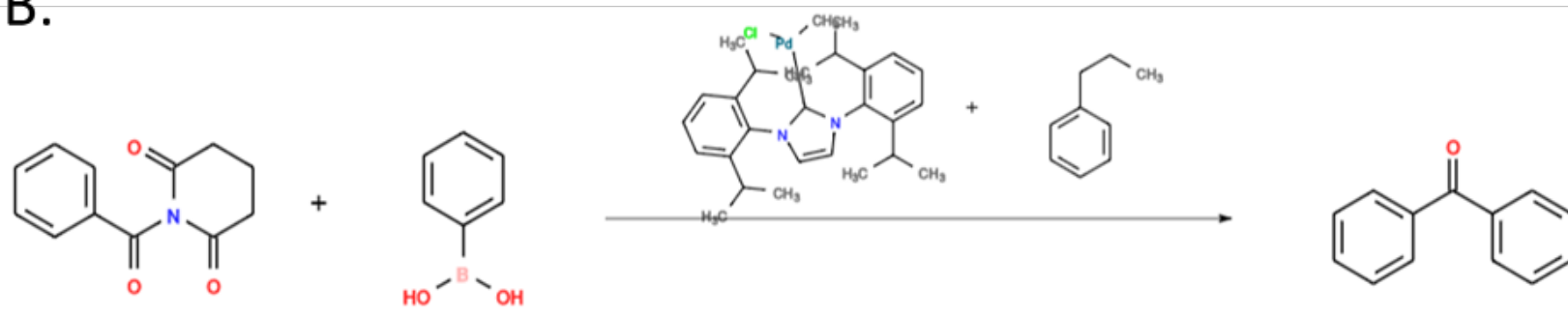
B.

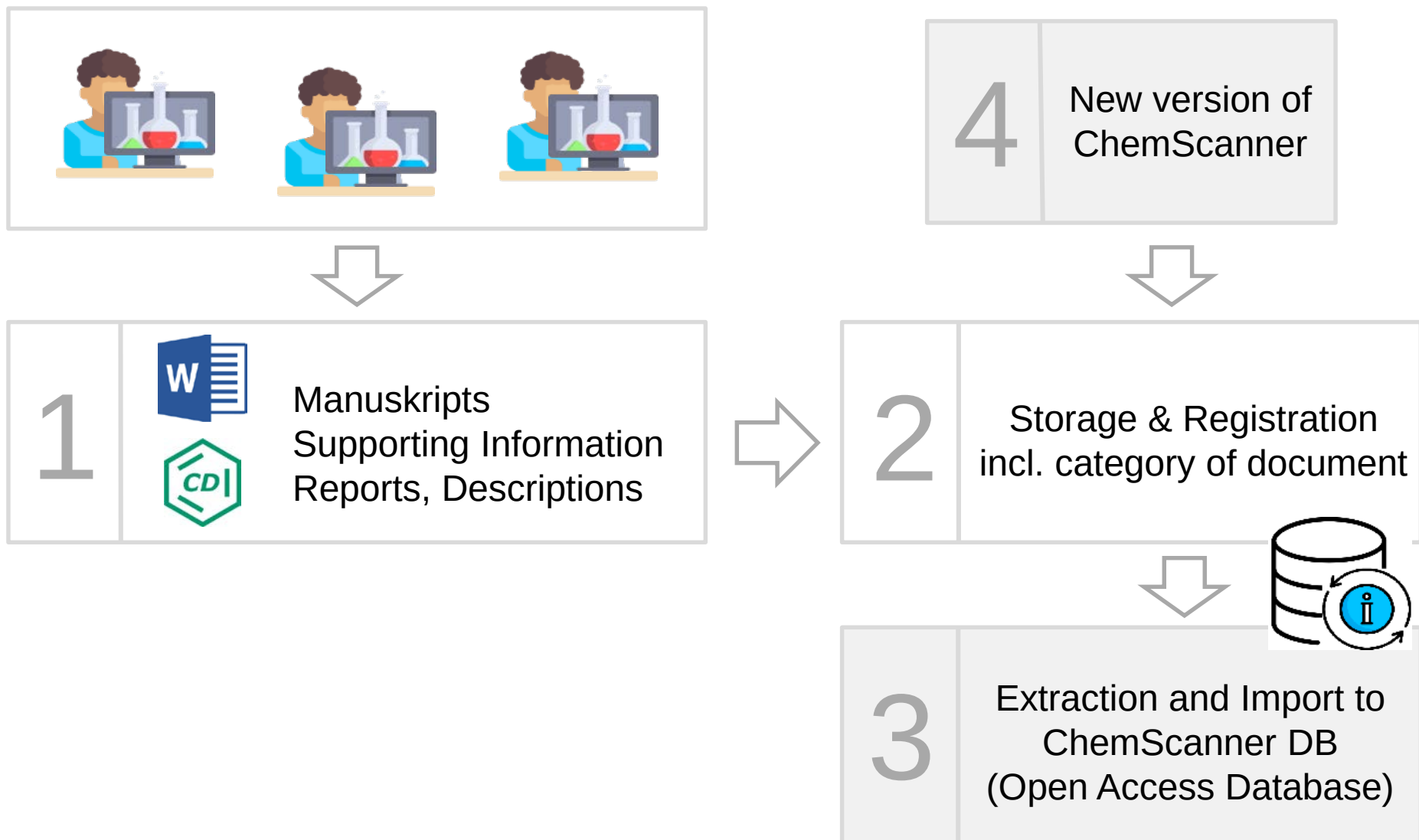


A.



B.

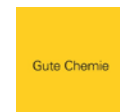




Thank you for your attention



KIT-library
KIT-SCC



Icons made by Freepik, Madebyoliver, Anatoli, Dinosoft, Vectors market, Zlatko Najdenovski, and Gregor Cresnar from www.flaticon.com

