CDD VAULT®
Complexity Simplified™

Capturing Mixtures:

Bringing Informatics to the World of Practical Chemistry

Moderated by: Samantha Jeschonek, PhD of **CDD**

Expert Panelists:

Leah McEwen, MS, MLS of **Cornell University**

Chris Jakober, PhD of **John Hopkins University**

Alex Clark, PhD of **CDD**

December 19, 2019

CDD VAULT®
Complexity Simplified

Smart Software Saves Time ™

- **Complete, intuitive solution for R&D data**
  - Deployed and trained in 90 min
  - Activity & Registration, ELN, Visualization, Inventory
- **Fully hosted, always available**
  - Nothing to install or support
  - All devices, OS's, browsers, locations
- **Collaboration-ready**
  - Configurable access for partners & CROs
- **Ecosystem compatible**
  - Uses common file formats & RESTful API
  - ChemAxon cartridge, full stereochem
- **Time-tested performance, security**
  - Deployed since 2004
- **Most affordable solution on the market**
  - Single annual fee
  - No services or upgrade costs

Today's Moderator and Panelists

Samantha Jeschonek, PhD
Research Scientist
Collaborative Drug Discovery

Leah McEwen, MS, MLS
Chemistry Librarian
Cornell University

Area of Focus:
IUPAC & Mixtures

Chris Jakober, PhD
Graduate Instructor
Johns Hopkins University,
Whiting School of Engineering

Area of Focus:
Safety & Mixtures

Alex Clark, PhD
Research Scientist
Collaborative Drug Discovery

Area of Focus:
AI & Mixtures

Have a question to ask our panel?
Open the **ZOOM Q&A** and type in your question at anytime!

We'll be sure to save time for them later!

Today's Moderator and Panelists

Samantha Jeschonek, PhD
Research Scientist
Collaborative Drug Discovery

Leah McEwen, MS, MLS
Chemistry Librarian
Cornell University

Area of Focus:
IUPAC & Mixtures

Chris Jakober, PhD
Graduate Instructor
Johns Hopkins University,
Whiting School of Engineering

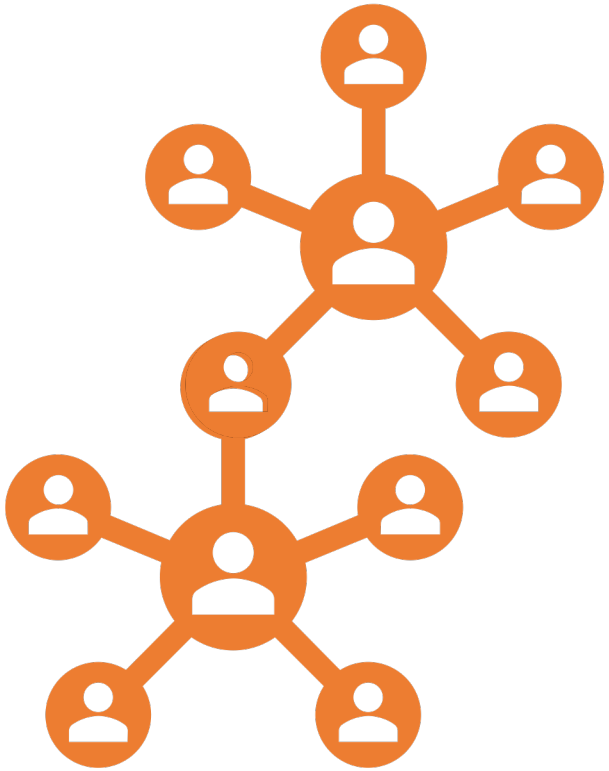
Area of Focus:
Safety & Mixtures

Alex Clark, PhD
Research Scientist
Collaborative Drug Discovery

Area of Focus:
AI & Mixtures

MInChI Project Team

- **Gerd Blanke**, StructurePendium Technologies GmbH, DEU
- **Alex Clark**, Collaborative Drug Discovery, CAN
- **John Duffus**, Edinburgh Centre for Toxicology, GBR
- **Richard Hartshorn**, University of Canterbury, NZL
- **Chris Jakober**, University of California, USA
- **Jon LaRue**, MilliporeSigma, USA
- **Leah McEwen**, Cornell University, USA, *Chair*
- **Andrey Yerin**, ACD/Labs, RUS



CDD VAULT® Complexity Simplified™

Real World Mixtures: Chemist View

Mixture representation today…

Phenol:Chloroform:Isoamyl Alcohol 25:24:1 Saturated with 10 mM Tris, pH 8.0, 1 mM EDTA

0.5 µg/mL B_2 and G_2 in acetonitrile

2 µg/mL B_1 and G_1 in acetonitrile

Description
n-Paraffins 18.9 % (w/w)
Isoparaffins 18.8 % (w/w)
Aromatics 23.3 % (w/w)
Naphthenes 20.5 % (w/w)
Olefins 18.5 % (w/w)

(each component in approx proportion indicated)

Butane, 15 % (w/w)
Heptane, 15 % (w/w)
Hexane, 15 % (w/w)
Nonane, 15 % (w/w)
Octane, 15 % (w/w)
Pentane, 15 % (w/w)
Propane, 10 % (w/w)

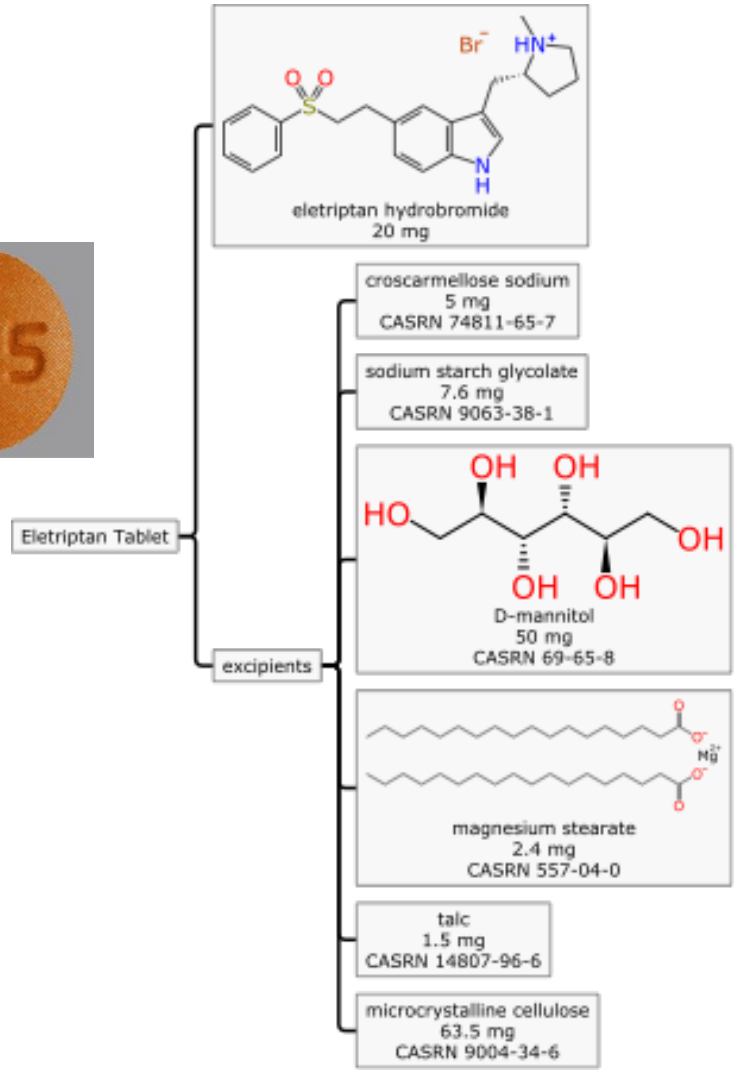
Ingredient	Wt. %
Phase A	
1. Lauryl PEG/PPG-18/18 Methicone	2
2. Aminopropyl Phenyl Trimethicone	2
3. Jojoba Oil	1.25
4. Isohexadecane	11.25
Phase B	
5. Glycerin	3
6. Phenoxyethanol and Methylisothiazolinone	0.5
7. Water	80

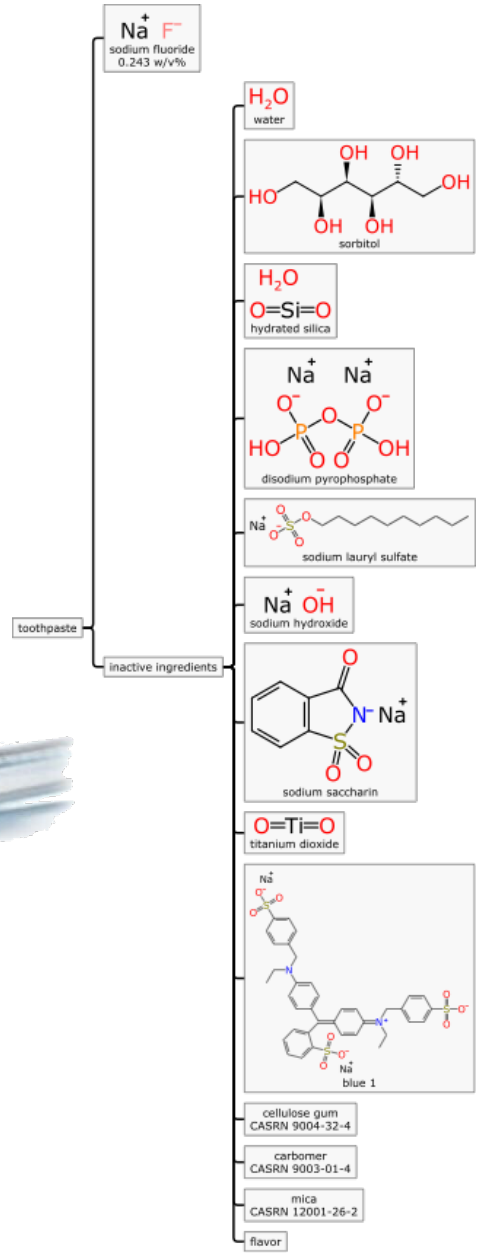
(equal weights of the hydrocarbons listed)

Heptadecane
Hexadecane
Pentadecane
Tetradecane

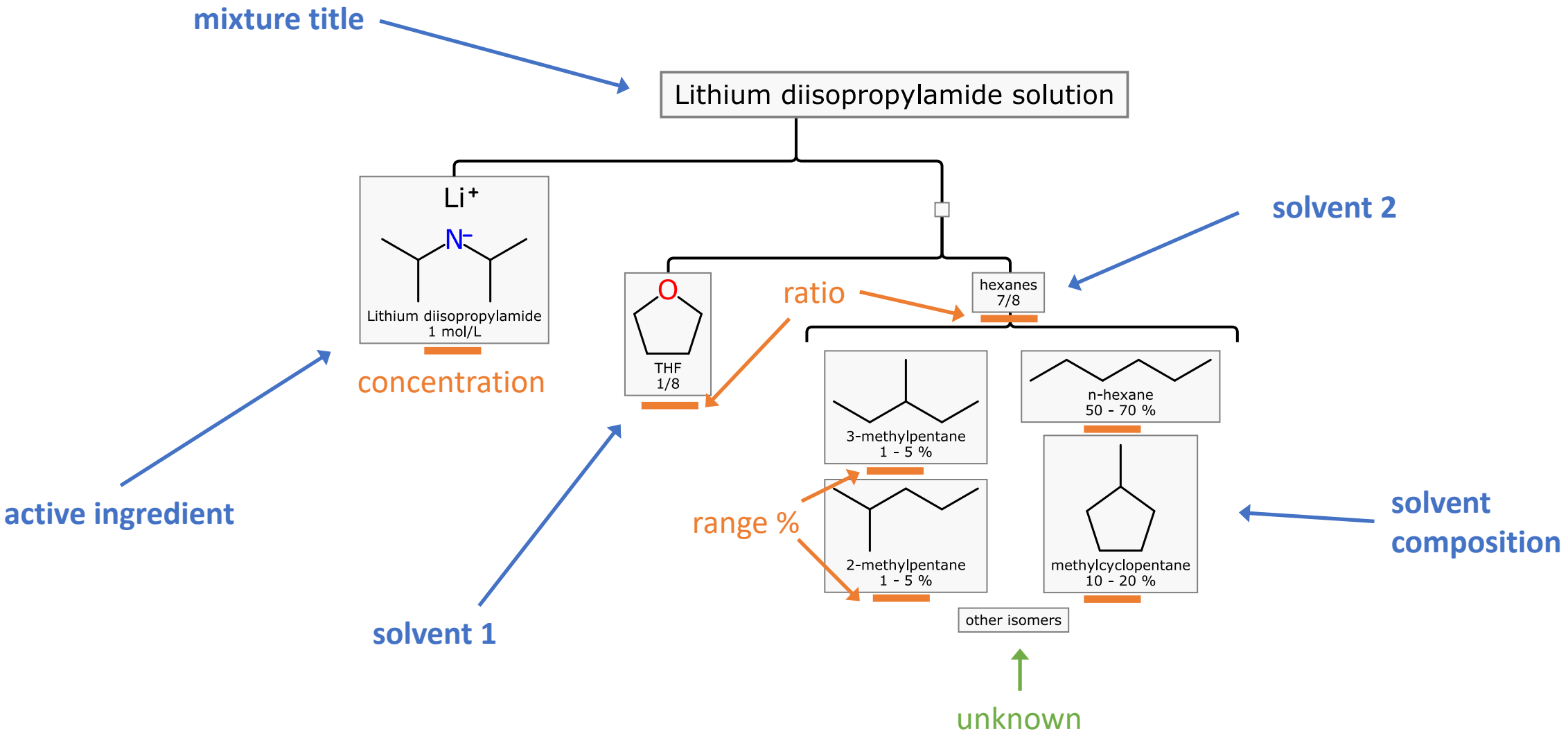
CDD VAULT®
Complexity Simplified™

Real World Mixtures: Consumer View





Mixtures Representation



1 M lithium diisopropyl amide in THF/hexanes (1:8)

CDD VAULT®
Complexity Simplified™

Multi-Component System Notation

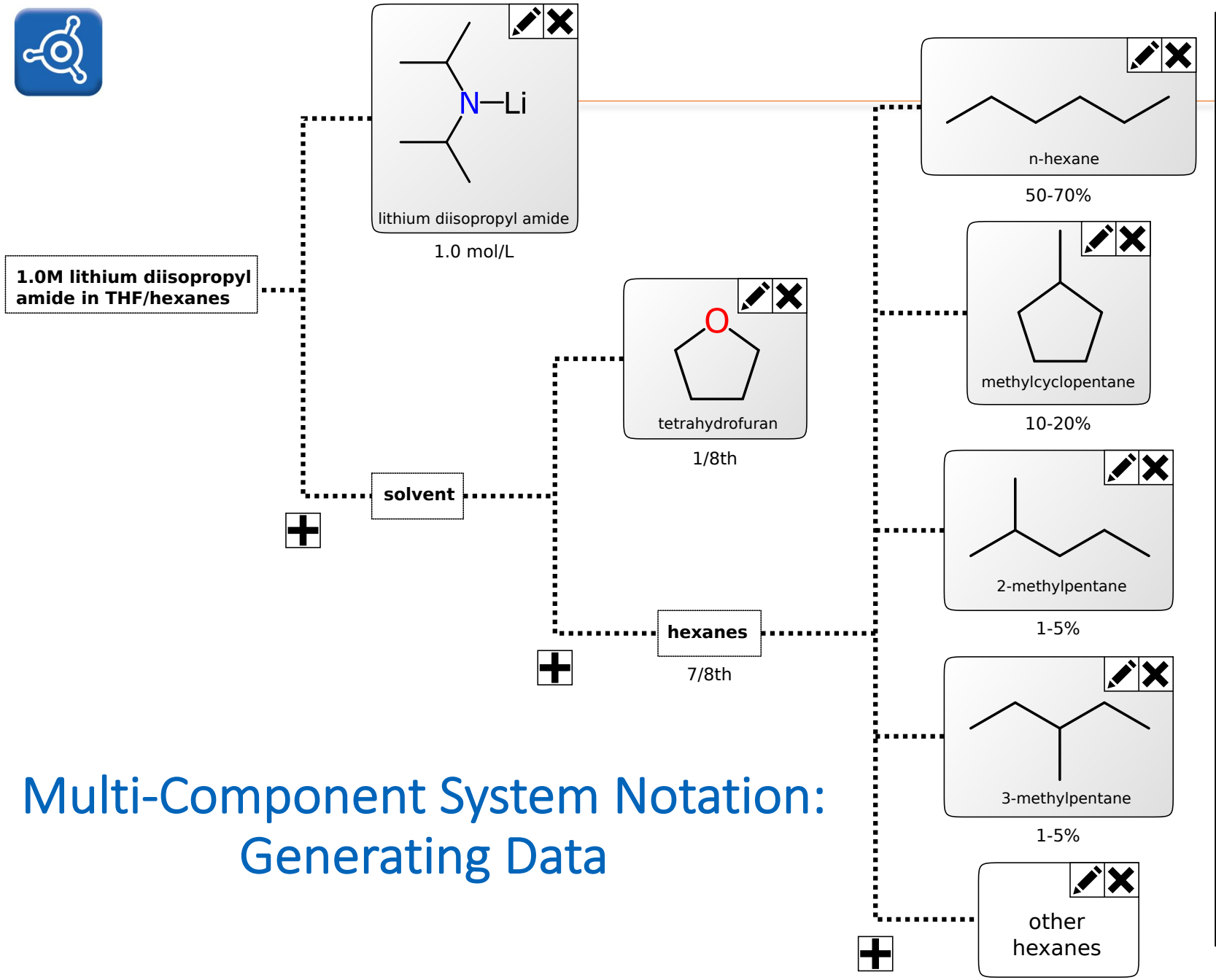
1.7M t-Butyllithium in Pentane:

MInChI=0.00.0S/
C4H9.Li/c1-4(2)3;/h1-3H3;/q-1;+1
&
C5H12/c1-3-5-4-2/h3-5H2,1-2H3&
/n{1&2}
/g{17mr-1&}

37% wt. Formaldehyde in Water with 10-15% Methanol:

MInChI=0.00.0S/
CH2O/c1-2/h1H2&
CH4O/c1-2/h2H,1H3&
H2O/h1H2
/n{1&2&3}
/g{37wf-2&10-15vf-2&}

- alphabetical order of components
- "**&**" separates components
- "**{}**" mixture groups (e.g., nested)
- "**/n**" indexes components (e.g., order)
- "**/g**" concentration (symbols detailed separately)



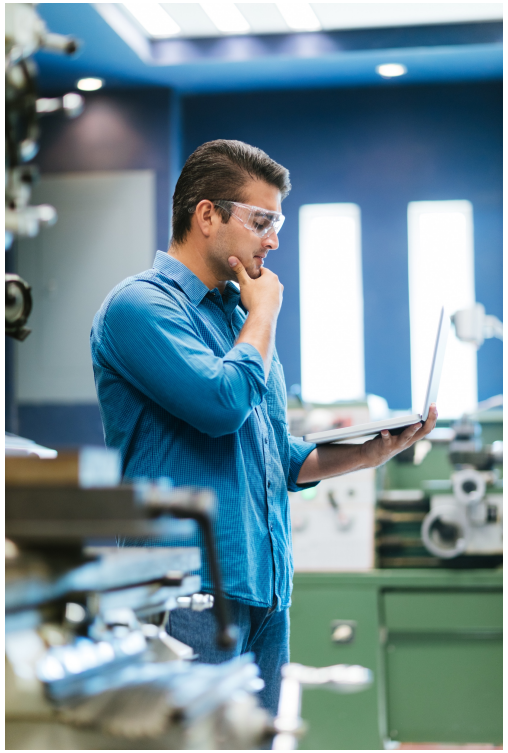
Multi-Component System Notation: Generating Data

1.0 M lithium diisopropyl amide in THF/hexanes :

MInChI=0.00.0S/C4H8O/c1-2-4-5-3-1/h1-4H2&C6H12/c1-6-4-2-3-5-6/h6H,2-5H2,1H3&C6H14/c1-3-5-6-4-2/h3-6H2,1-2H3&C6H14/c1-4-5-6(2)3/h6H,4-5H2,1-3H3&C6H14/c1-4-6(3)5-2/h6H,4-5H2,1-3H3&C6H14N.Li/c1-5(2)7-6(3)4;/h5-6H,1-4H3;/q-1;+1/n{6&{1&{2&3&4&5&nc}}/g{1mr&{1vp&{1-2vf-1&5-7vf-1&1-5vf-2&1-5vf-2&}7vp}}

CDD VAULT®
Complexity Simplified™

Poll the Audience!



How do you currently report mixtures in your industry?

a) Text based format, only listing the active ingredient details
b) Text based format that includes active ingredient and primary solvent details
c) Poorly...usually with someone's initials, date, and reference to a notebook page.
d) Structured format incorporating public standards (like InChI keys)
e) We don't since we never considered that inactive ingredients might be useful to capture.

CDD VAULT®
Complexity Simplified™

A Typical Research Inventory

Name	CAS NO		Phase	Bottles	Amount	Unit	
Butyllithium solution, 2.5M in hexanes	109-72-8		Liquid	1	100	mL
Butyllithium solution, 2.5M in hexanes	109-72-8		Liquid	2	50	mL
Butyllithium solution, 2.5M in hexanes	109-72-8		Liquid	4	25	mL
Butyllithium solution, n-, 2.5M in hexanes	109-72-8		Liquid	1	100	mL
Butyllithium solution,1.6 M in hexanes	109-72-8		Solid	1	25	mL
Butyllithium, (1.6 M in hexanes)	109-72-8		Liquid	1	100	mL
butyllithium, 0.14 mol/l in toluene	109-72-8		Liquid	1	90	mL
butyllithium, 0.16 mol/l in Hexane	109-72-8		Liquid	1	140	mL
butyllithium, 0.16 mol/l in hexane	109-72-8		Liquid	1	40	mL
butyllithium, 0.16 mol/l in Hexanes	109-72-8		Liquid	1	100	mL
butyllithium, 0.25 mol/l in Hexane	109-72-8		Liquid	1	50	mL
butyllithium, 0.25 mol/l in hexane	109-72-8		Liquid	1	80	mL
butyllithium, 0.25 mol/l in hexanes	109-72-8		Liquid	1	100	mL
butyllithium, 0.25 mol/l in Hexanes	109-72-8		Liquid	1	30	mL
butyllithium, 0.25 mol/l in Hexanes	109-72-8		Liquid	1	100	mL
butyllithium, 0.25 mol/l in hexanes	109-72-8		Liquid	1	100	mL
butyllithium, 0.25 mol/l in hexanes	109-72-8		Liquid	1	100	mL
butyllithium, 0.25 mol/l in hexanes	109-72-8		Liquid	1	100	mL
Butyllithium, 1.3M in cyclohexane/hexane (92/8), sec-	598-30-1		Liquid	1	100	mL
BUTYLLITHIUM, 1.6 M IN HEXANE	BUTYLLITHIUM, 1.6 M IN HEXANE-Mixture		Liquid	1	100	mL
BUTYLLITHIUM, 1.6 M IN HEXANE	BUTYLLITHIUM, 1.6 M IN HEXANE-Mixture		Liquid	1	400	mL
BUTYLLITHIUM, 1.6 M IN HEXANE	BUTYLLITHIUM, 1.6 M IN HEXANE-Mixture		Liquid	1	100	mL
BUTYLLITHIUM, 1.6 M IN HEXANE	BUTYLLITHIUM, 1.6 M IN HEXANE-Mixture		Liquid	2	250	mL
BUTYLLITHIUM, 1.6 M IN HEXANE	BUTYLLITHIUM, 1.6 M IN HEXANE-Mixture		Liquid	1	100	mL
BUTYLLITHIUM, 1.6 M IN HEXANE	BUTYLLITHIUM, 1.6 M IN HEXANE-Mixture		Liquid	1	50	mL
BUTYLLITHIUM, 1.6 M IN HEXANE	BUTYLLITHIUM, 1.6 M IN HEXANE-Mixture		Liquid	1	50	mL
BUTYLLITHIUM, 1.6 M IN HEXANE	BUTYLLITHIUM, 1.6 M IN HEXANE-Mixture		Liquid	1	100	mL
BUTYLLITHIUM, 1.6 M IN HEXANE	BUTYLLITHIUM, 1.6 M IN HEXANE-Mixture		Liquid	1	400	mL
BUTYLLITHIUM, 1.6 M IN HEXANE	BUTYLLITHIUM, 1.6 M IN HEXANE-Mixture		Liquid	1	100	mL
BUTYLLITHIUM, 1.6 M IN HEXANE	BUTYLLITHIUM, 1.6 M IN HEXANE-Mixture		Liquid	1	100	mL
BUTYLLITHIUM, 1.6 M IN HEXANE	BUTYLLITHIUM, 1.6 M IN HEXANE-Mixture		Liquid	1	50	mL
BUTYLLITHIUM, 1.6 M IN HEXANE	BUTYLLITHIUM, 1.6 M IN HEXANE-Mixture		Liquid	1	100	mL

Courtesy of the University of California

CDD VAULT®
Complexity Simplified™

Mixture Data

- Lots of mixtures, but mostly text
- Active ingredient often separated
  - purity
  - solvent/conc.
- Semi-structured databases also

n-Butyllithium solution

7 Product Results | Match Criteria: Property, Product Name

Synonym: ***n*-BuLi, Butyl lithium, Butyllithium solution, Lithium-1-butanide**

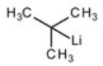
Linear Formula: **$CH_3(CH_2)_3Li$** | Molecular Weight: **64.06** | CAS Number: **109-72-8**

230707	2.5 M in hexanes	Sigma-Aldrich	SDS Pricing
302120	2.0 M in cyclohexane	Sigma-Aldrich	SDS Pricing
20159	2.7 M in heptane	Sigma-Aldrich	SDS Pricing
230715	11.0 M in hexanes	Sigma-Aldrich	SDS Pricing
186171	1.6 M in hexanes	Sigma-Aldrich	SDS Pricing

Show All 7 Results

tert-Butyllithium

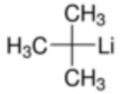
1 Product Result | Match Criteria: Product Name



8.14147	(approx.15% solution in n-pentane) for synthesis	Sigma-Aldrich	SDS Pricing

tert-Butyllithium solution

1 Product Result | Match Criteria: Keyword



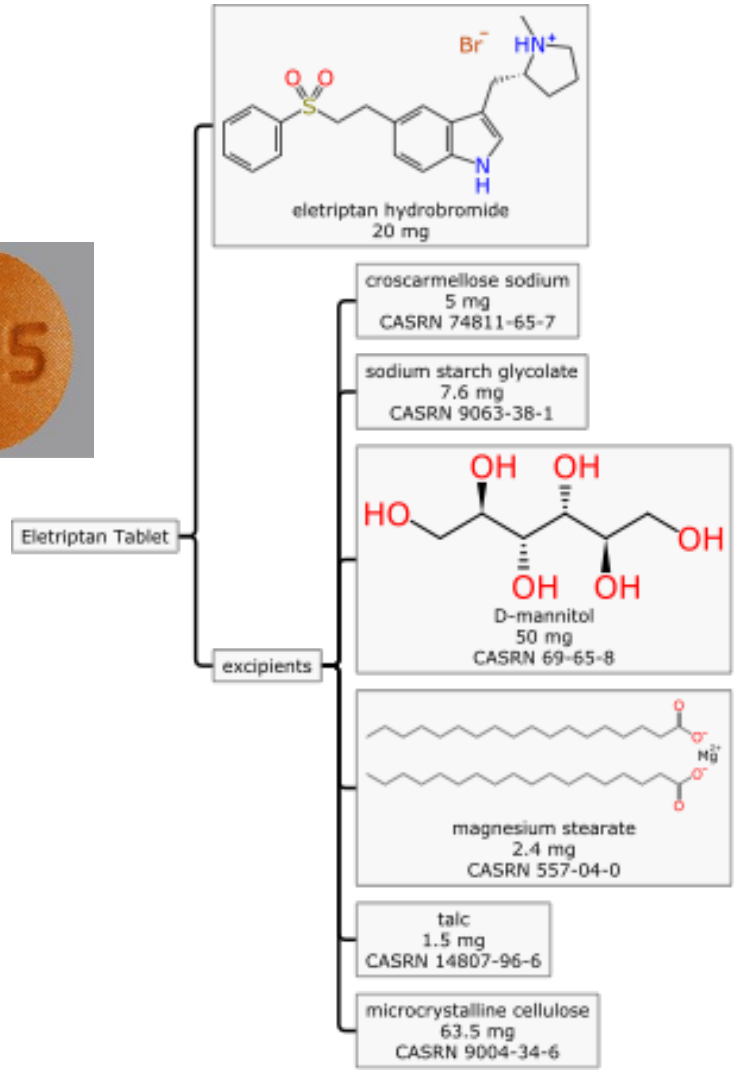
Synonym: **Lithium-2-methyl-2-propanide, t-BuLi**

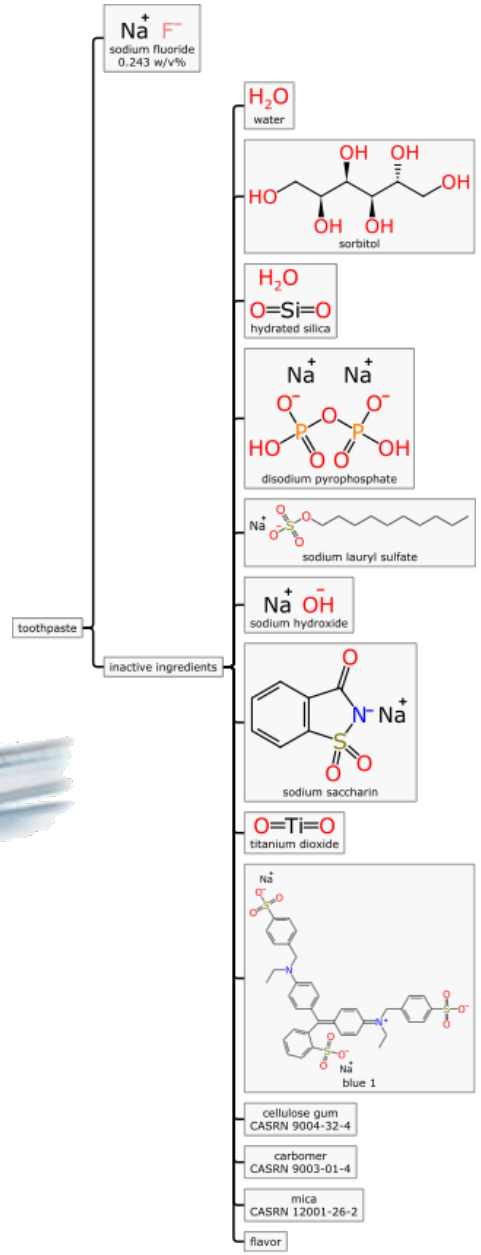
Linear Formula: **$(CH_3)_3CLi$** | Molecular Weight: **64.06** | CAS Number: **594-19-4**

186198	1.6 M in pentane	Sigma-Aldrich	SDS Pricing

CDD VAULT®
Complexity Simplified™

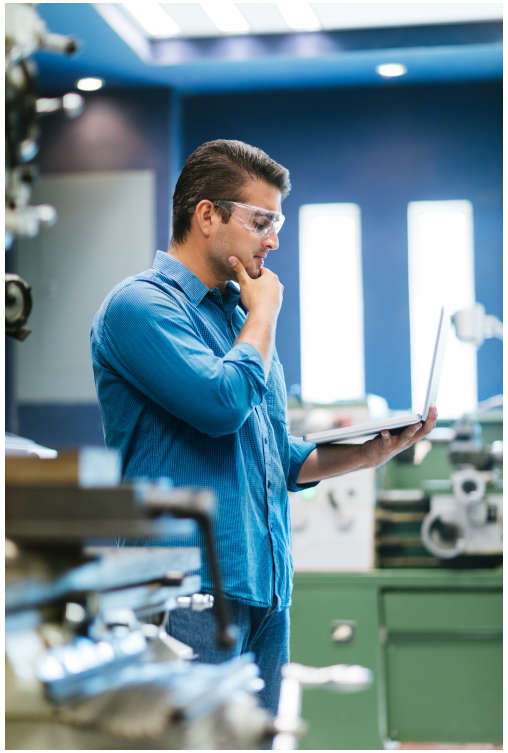
Real World Mixtures: Consumer View





CDD VAULT®
Complexity Simplified™

Poll the Audience!



Where would mixtures standards (like MInChI)

have the greatest impact on you or your company?

a) Scientist level - being able to unambiguously identify details about a mixture for assays
b) Informatics level - being able to improve models by knowing more about the mixture composition
c) Inventory level - ensuring that duplicate solutions are not purchased
d) Transport and safety level - improving risk management
e) Regulatory level - increasing efficiency in reporting

Have a question to ask our panel?
Open the **ZOOM Q&A** and type in your question at anytime!

We'll be sure to save time for them later!

Today's Moderator and Panelists

Samantha Jeschonek, PhD
Research Scientist
Collaborative Drug Discovery

Leah McEwen, MS, MLS
Chemistry Librarian
Cornell University

Area of Focus:
IUPAC & Mixtures

Chris Jakober, PhD
Graduate Instructor
Johns Hopkins University,
Whiting School of Engineering

Area of Focus:
Safety & Mixtures

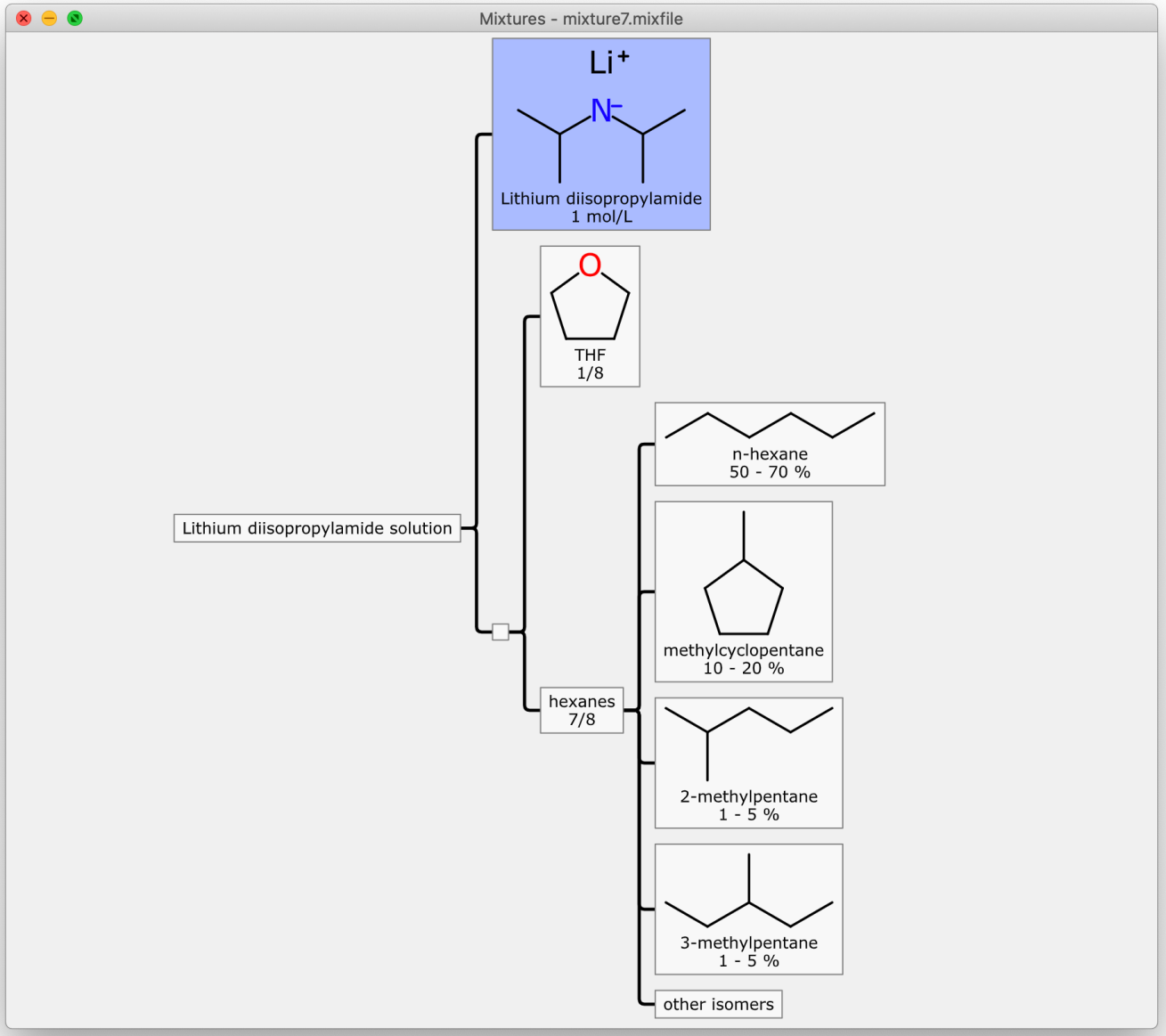
Alex Clark, PhD
Research Scientist
Collaborative Drug Discovery

Area of Focus:
AI & Mixtures

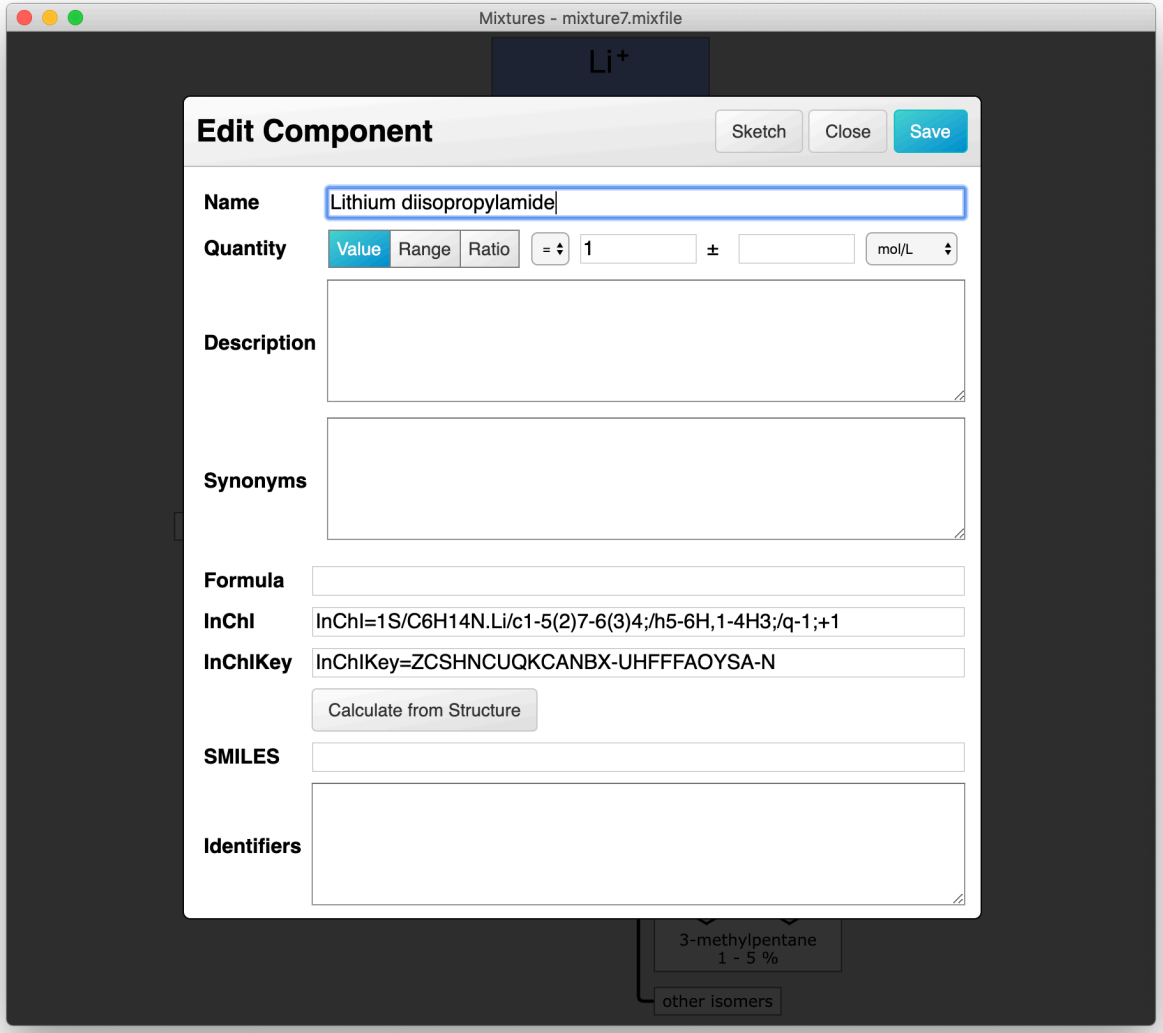
CDD VAULT®
Complexity Simplified™

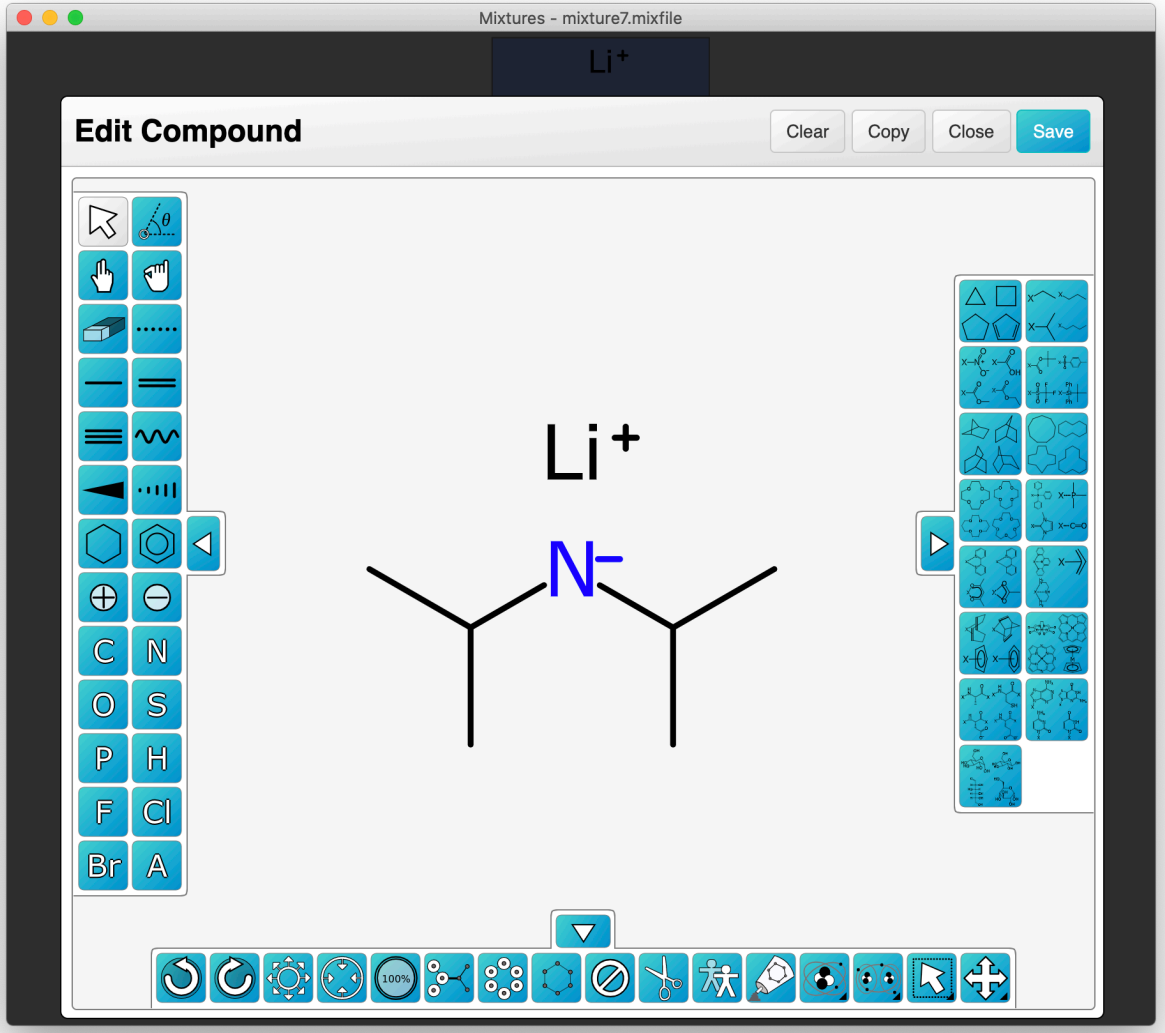
Editing Mixture Trees

github.com/cdd/mixtures



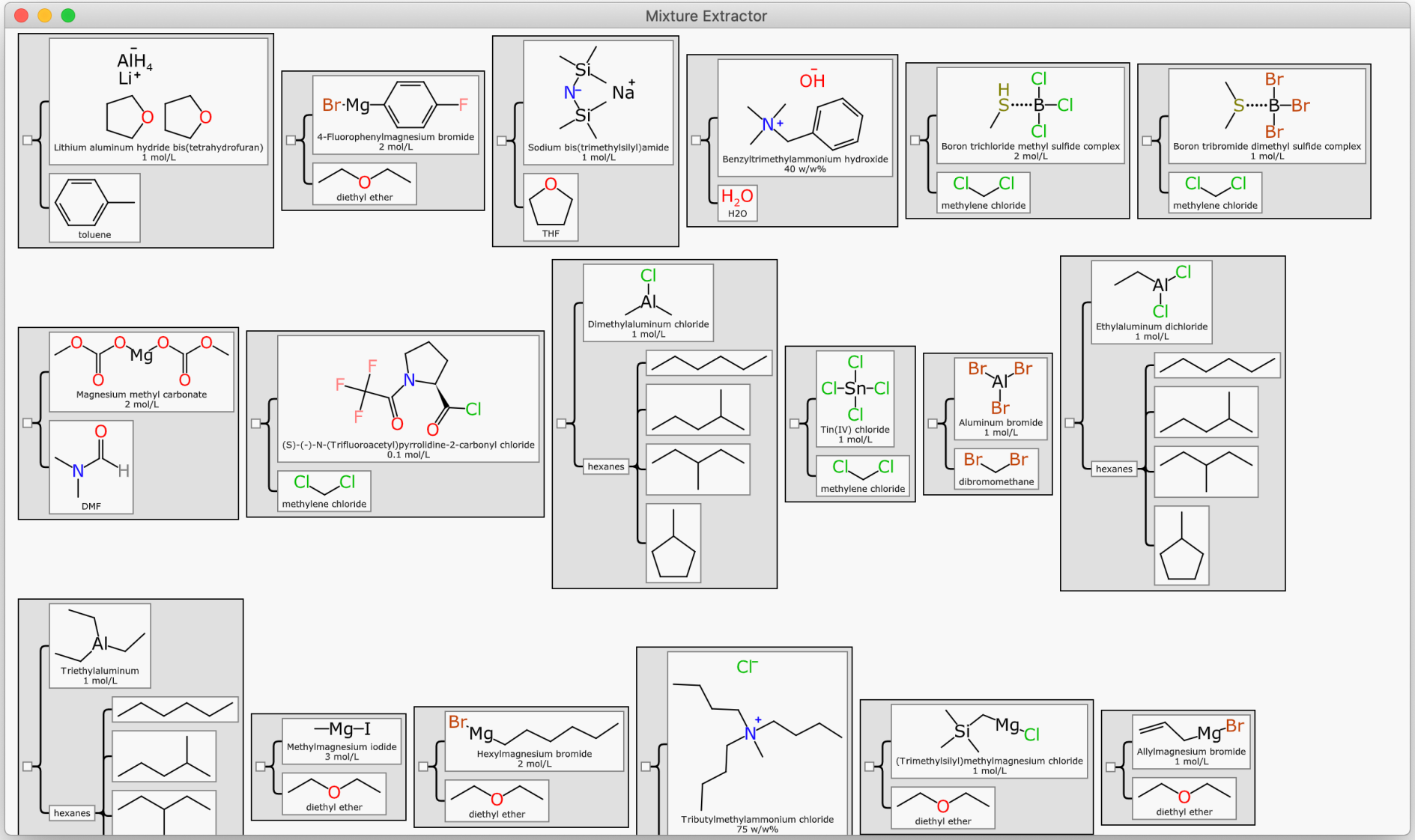
Editing Components





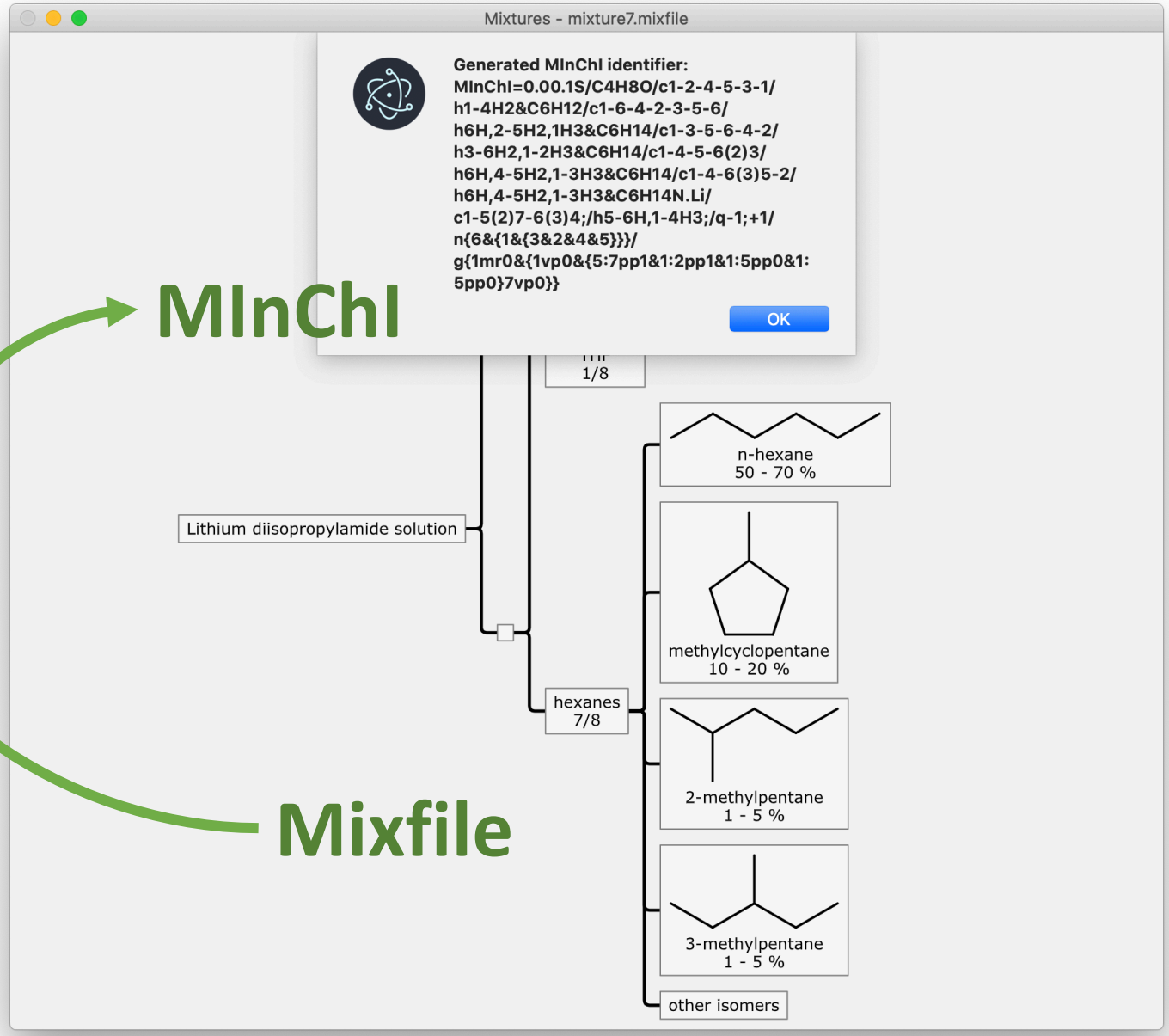
CDD VAULT®
Complexity Simplified™

Extracted Mixtures



CDD VAULT®
Complexity Simplified™

Calculating MInChI strings



DOI: 10.1186/s13321-019-0357-4

Today's Moderator and Panelists

Samantha Jeschonek, PhD
Research Scientist
Collaborative Drug Discovery

Leah McEwen, MS, MLS
Chemistry Librarian
Cornell University

Area of Focus:
IUPAC & Mixtures

Chris Jakober, PhD
Graduate Instructor
Johns Hopkins University,
Whiting School of Engineering

Area of Focus:
Safety & Mixtures

Alex Clark, PhD
Research Scientist
Collaborative Drug Discovery

Area of Focus:
AI & Mixtures

Practical Demonstration of CDD Vault

Hosted by: Robert Thorn, PhD of **CDD**

Thursday, January 09, 2020

- 12:00 PM EST
- 11:00 AM CST
- 9:00 AM PST

Questions?

CDD VAULT®

Complexity Simplified

*Subscribe to the **CDD Blog** to get all up-to-date events including links to webinar content, interviews and all other CDD activities…*

https://www.collaborativedrug.com/blog