



ELSEVIER

Helping Medicinal Chemists Discover New Opportunities during Lead ID and Optimisation

Leveraging ChemAxon's technology for future product
developments

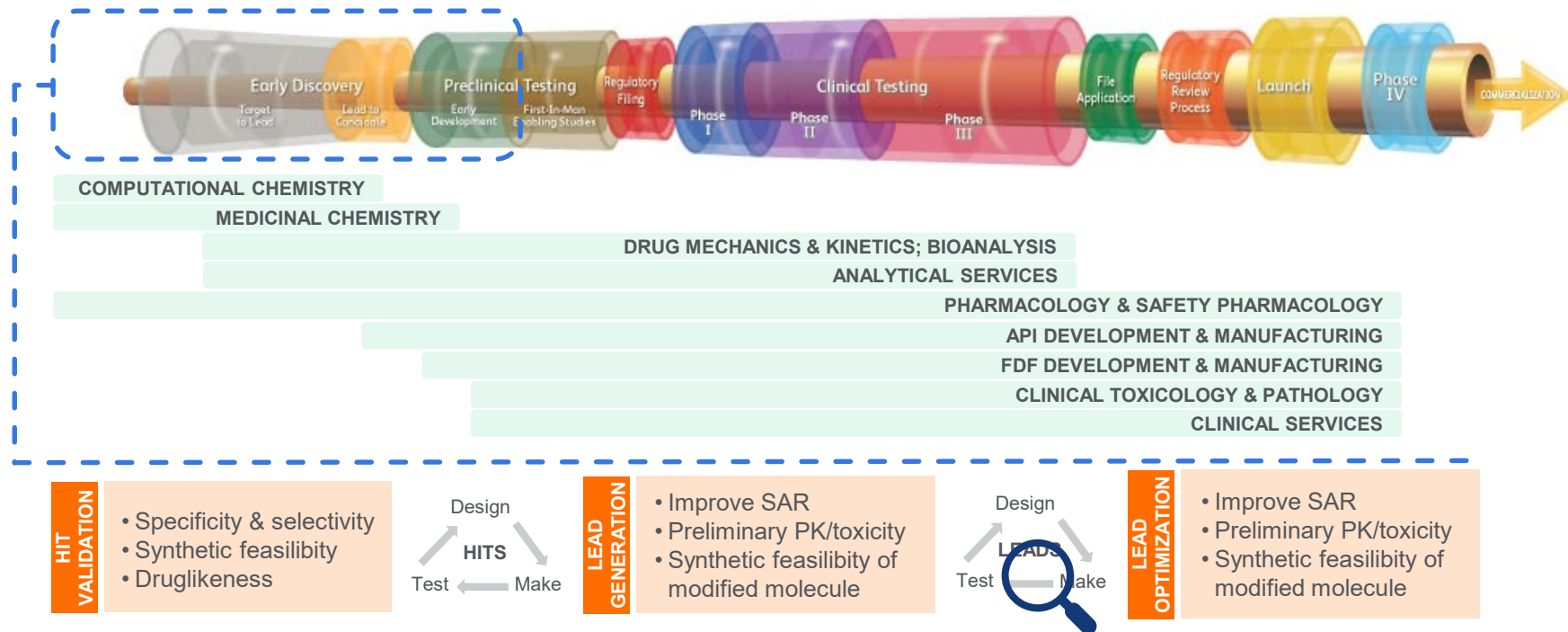
ChemAxon User Group Meeting

May 2019

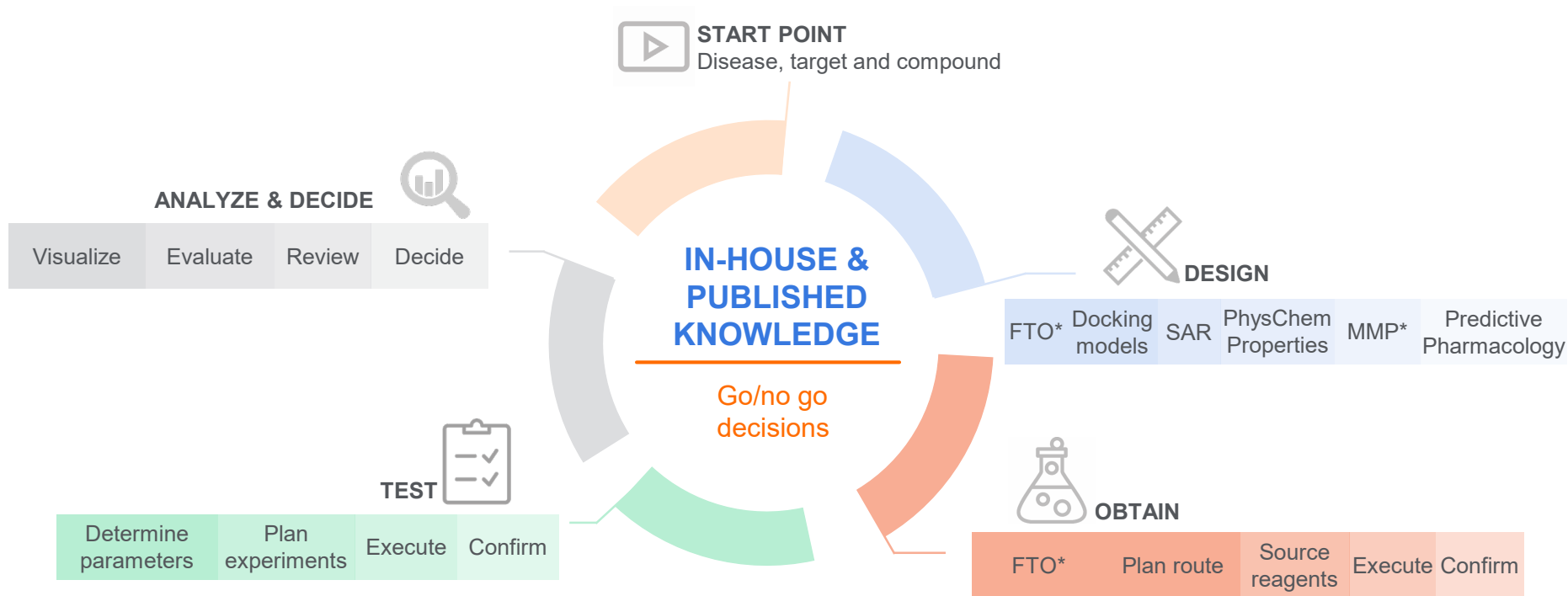
Dr Rosalind Sankey
Senior Product Manager – Reaxys Medicinal Chemistry
r.sankey@elsevier.com



The requirements for effective drug discovery are complex and command a multi-disciplinary approach



Drug design cycle leads to optimized candidates, but necessitates both in-house and published knowledge and data



Reaxys aims to deliver actionable answers during the drug design cycle

Reaxys

- An integrated solution that augments understanding and catalyzes action in drug development
- One source of advanced approaches and best practices to reduce attrition
- Designed to help researchers identify and progress drug candidates as quickly and as safely as possible

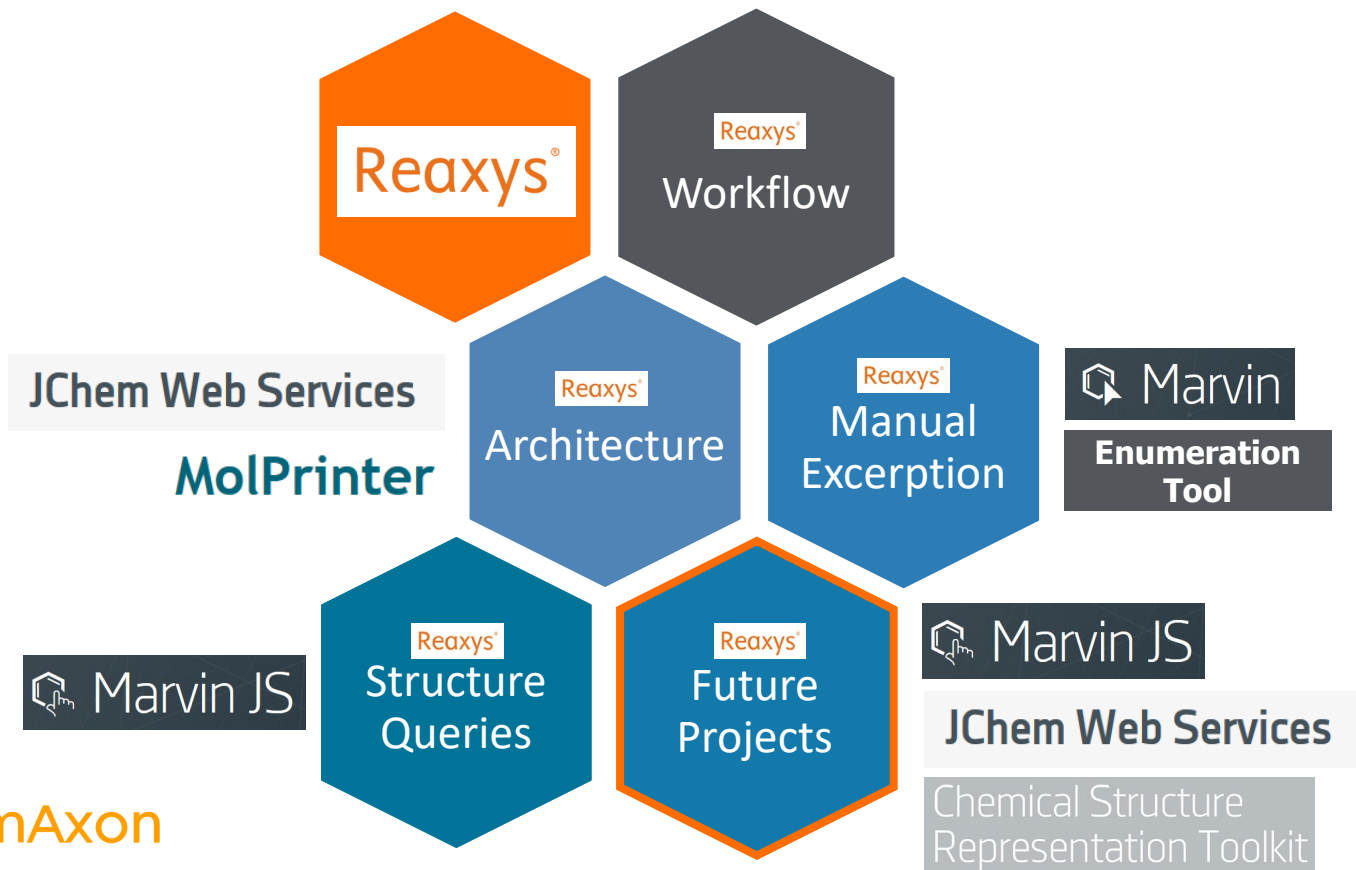


REAXYS & RMC

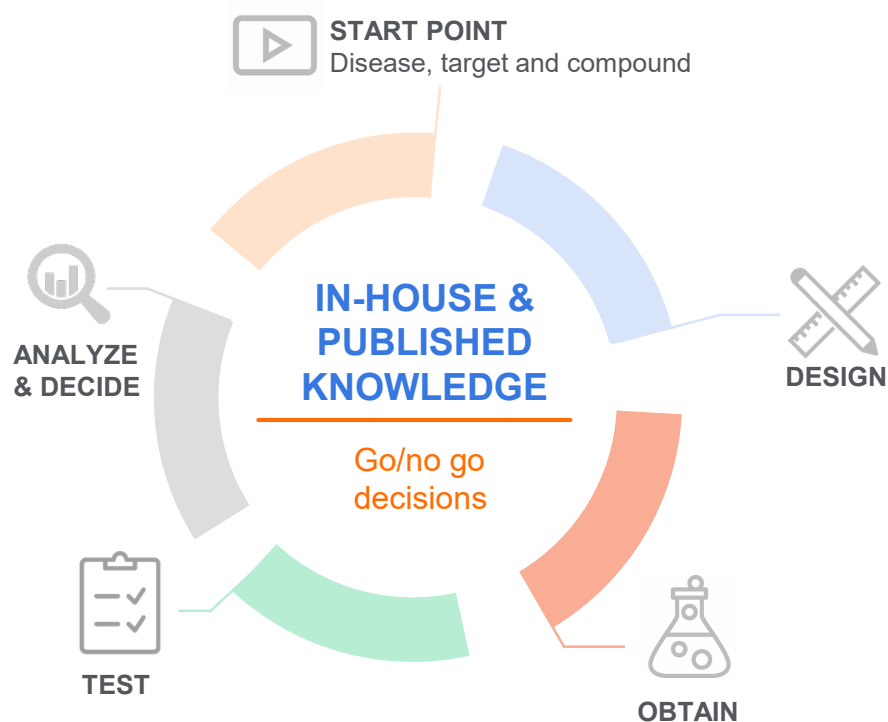
		REAXYS & RMC
DESIGN	Docking models	Red square
	SAR	Green square
	Properties	Green square
	MMP	Black square
OBTAIN	Predictive pharmacology	Black square
	Plan route	Green square
	Source reagents	Green square
	Execute	NA
TEST	Confirm	Green square
	Determine parameters	Yellow square
	Plan experiments	Yellow square
	Execute	NA
ANALYZE	Confirm	Green square
	Visualize	Yellow square
	Evaluate	Green square
	Review	Yellow square
	Decide	NA

- Green square: Excellent support
- Yellow square: Good/fair support
- Red square: Limited support
- Black square: Currently under development

Reaxys' end-to-end processing workflow embeds and employs ChemAxon's technology in various stages and processes



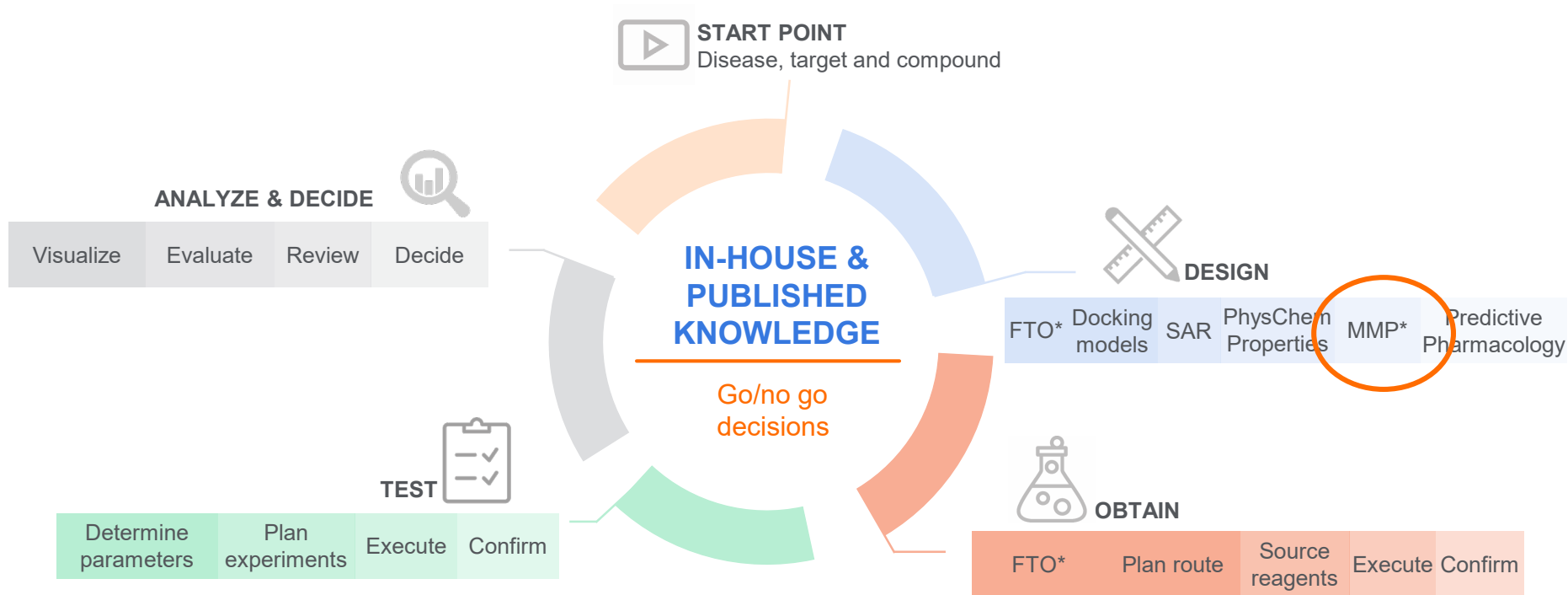
Development of new capabilities within Reaxys aim to enhance the ability to make informed and actionable decisions



		REAXYS & RMC
DESIGN	Docking models	Red
	SAR	Green
	Properties	Green
	MMP	Black
	Predictive pharmacology	Black
OBTAIN	Plan route	Green
	Source reagents	Green
	Execute	NA
	Confirm	Green
TEST	Determine parameters	Yellow
	Plan experiments	Yellow
	Execute	NA
	Confirm	Green
ANALYZE	Visualize	Yellow
	Evaluate	Green
	Review	Yellow
	Decide	NA

■ Excellent support
■ Good/fair support
■ Limited support
■ Currently under development

Using data from Reaxys we want to strengthening our support for the drug design cycle



Typical problems in lead optimization

Optimizing PK and ADME properties



“A part of my current lead structure is thought to be responsible for poor cell permeability. My plan is to **replace this substructure** with something that is:

- chemically feasible and
- likely keeping the activity against the primary target.”

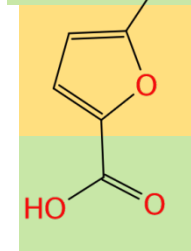
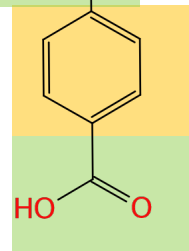
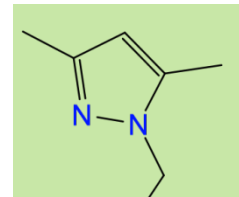
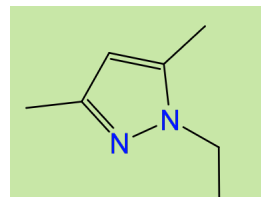
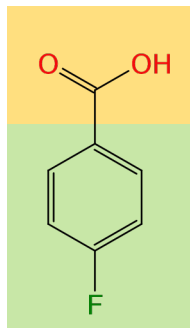
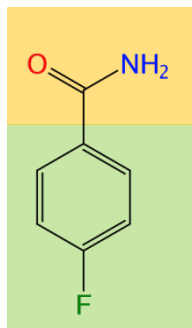
Researchers must identify, optimize and make lead compounds with **less PK and ADME issues** whilst **maintaining high affinity** for the identified target.

Critical to understand is **how to best modify a compound to achieve desired activity, PK or PD properties**

“What should we make next?”

Matched Molecular Pair (MMP) analysis

Pairs of compounds with a small sub-structural exchange



Used to understand the impact of sub-structural replacements on a given parameter of interest – particularly popular with bioactivity data

Matched Molecular Pair (MMP) analysis – a collaboration with the Swiss Institute of Bioinformatics using RMC data

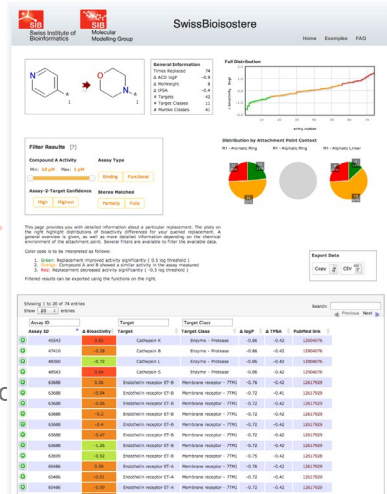
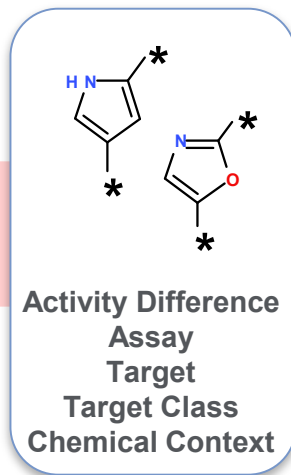
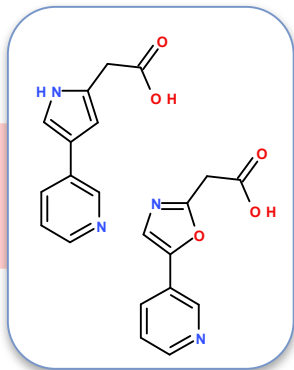
Data Source

Pair Identification

Replacement Extraction

Database Storage

Outcome



Reaxys
Medicinal
Chemistry

Back end: **JChem**

Matched Molecular Pair
The computational engine

Front end: **JChem Web Services**



Swiss Institute of
Bioinformatics

With, M.; Zoete, V.*; Michelin, O.*; Sauer, W. H. B*. SwissBioSistere: a database of molecular replacements for ligand design. *Nucleic Acids Res.* **2013**, *41*, D1137–43.

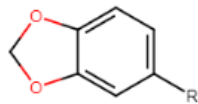
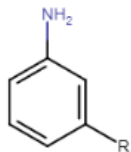
RMC data derived Matched Molecular Pair web application – user feedback



Medicinal Chemist in Pharma

- Used on a current drug design project
 - “Gave me new ideas and helped creativity”
 - New ideas were handed to my computational chemistry colleagues for docking and initiated further discussion
- Expected impact on research decisions:
 - Extend chemical space due to larger knowledge base
 - Trigger decisions more quickly when identifying and evaluating good/bad replacements

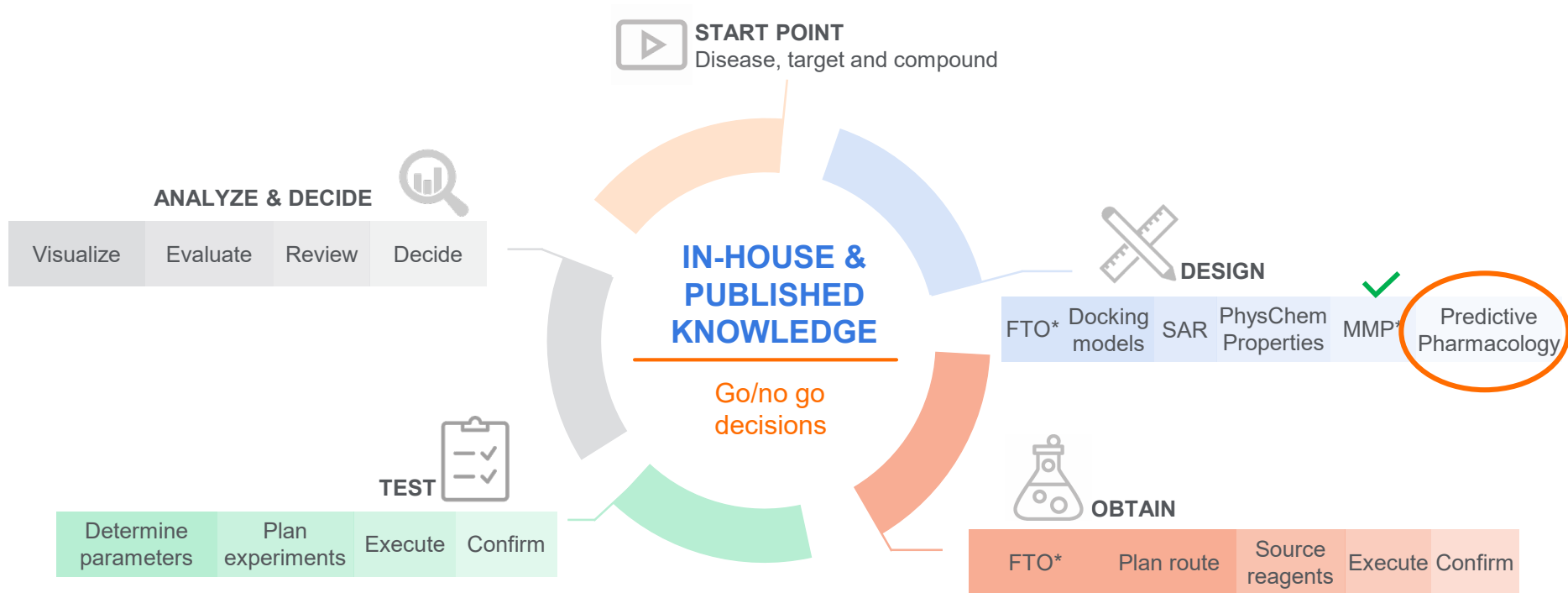
Illustrative example:



Δ ACD logP	0.7
Δ MolWeight	29
Δ tPSA	-7.6

1197662	-1.54	Cyclooxygenase-2	Enzyme	0.47	-7.56
Reaxys Compound_ID:	7725904	Reaxys Compound_ID:	7729355		
Activity [pIC50]:	5.23		6.77		
Structure:		→			

Using data from Reaxys we want to strengthening our support for the drug design cycle



Typical problems in lead optimization

Safety pharmacology issues



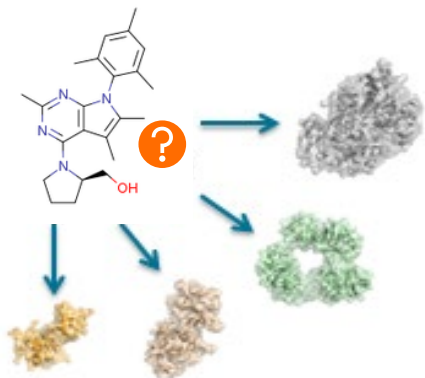
“I know my lead compound has good activity on my primary target but what other targets could it be active on? I need to understand if my compound could be responsible for any **unwanted effects** in a **safety assay**.”

Researchers must identify, optimize and make lead compounds with **less toxicity or safety pharmacology issues** whilst **maintaining high affinity** for the identified target.

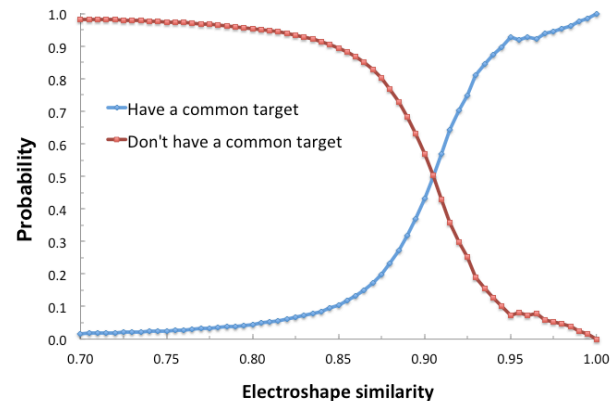
Critical to understand if a compound has potential to be active on **secondary targets** so that any **unwanted interactions/effects can be assessed and addressed**

“What can we expect from our compound?”

Target Prediction by molecular similarity



Assumption: if two molecules are very similar, they are more likely to be active on the same target



Target prediction (reverse screening):

Molecule with unknown target

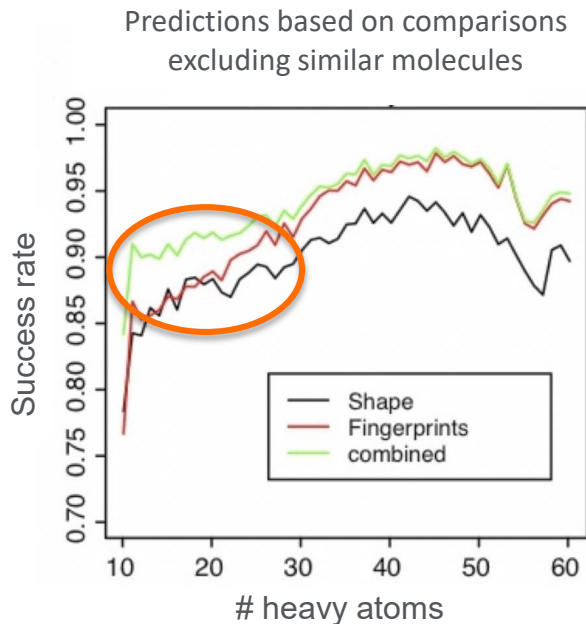
Similarity

RMC
Library of molecules with known targets

List of possible targets

Target Prediction by molecular similarity

The dual scoring function, based on both 2D and 3D molecular similarity, gives high performing predictions



RMC data*

Unique small molecules	521,445
Datapoints	745,106
Total protein targets included (human, rat & mouse)	2,590

* Only compounds having activity < 10 nM

Back end:

JChem

Front end:

Marvin

JChem Web Services



Swiss Institute of
Bioinformatics

Gfeller, D.; Michielin, O.; Zoete, V. Shaping the Interaction Landscape of Bioactive Molecules. *Bioinformatics*. **2013**, *29*, 3073–3079.

Gfeller, D.; Grosdidier, A.; Wirth, M.; Daina, A.; Michielin, O.; Zoete, V. SwissTargetPrediction: a Web Server for Target Prediction of Bioactive Small Molecules. *Nucleic Acids Res.* **2014**, *42(Web Server issue)*, W32-8.

Gfeller D., Zoete V. Protein homology reveals new targets for bioactive small molecules. *Bioinformatics*. **2015**, *31*, 2721-7.

RMC data derived Target Prediction web application – user feedback



Medicinal Chemist in Pharma

- Used on a current drug design project
 - Looked for potential **off-target** pharmacology and assess safety

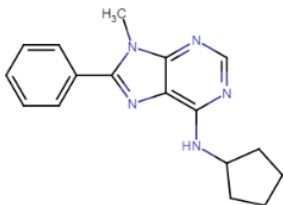
Used to **prioritise**
 “I took the predic

Target prediction

- Expected impact on ...
 - Would be used
 - Doing **drug safety** later on
 - Doing the test

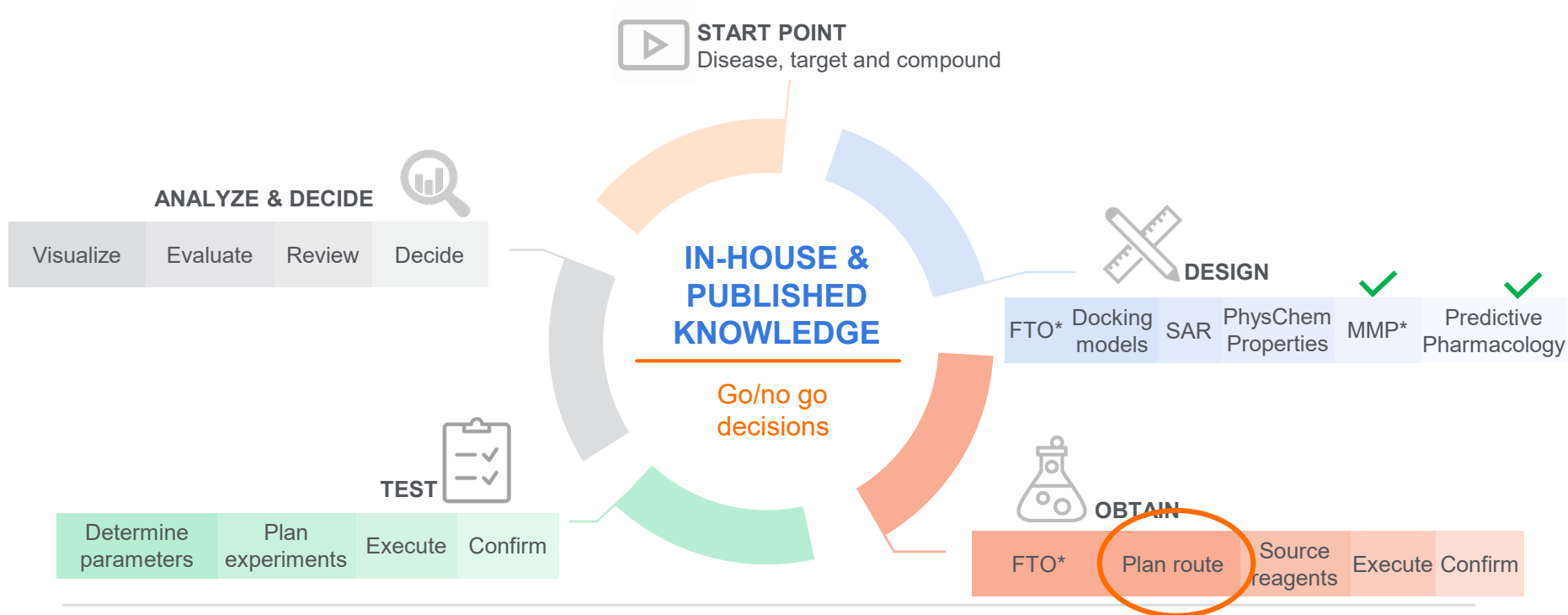
Target	Common name	Uniprot ID	Target Class	Probability*	Known actives (3D / 2D)
Adenosine receptor A2a	ADORA2A	P29274	Membrane receptor	<div style="width: 100%; height: 10px; background-color: green;"></div>	584 / 465
Adenosine receptor A2b	ADORA2B	P29275	Membrane receptor	<div style="width: 100%; height: 10px; background-color: green;"></div>	170 / 232
Adenosine receptor A1	ADORA1	P30542	Membrane receptor	<div style="width: 100%; height: 10px; background-color: green;"></div>	488 / 591
Adenosine receptor A3	N/A	P33765	Membrane receptor	<div style="width: 100%; height: 10px; background-color: green;"></div>	690 / 639
Neuropeptide Y receptor type 5	NPY5R	Q15761	Membrane receptor	<div style="width: 100%; height: 10px; background-color: green;"></div>	299 / 149
Cannabinoid receptor 1	CNR1	P21554	Membrane receptor	<div style="width: 80%; height: 10px; background-color: green;"></div>	1418 / 305
5-hydroxytryptamine receptor 2A	HTR2A	P28223	Membrane receptor	<div style="width: 60%; height: 10px; background-color: green;"></div>	108 / 42
5-hydroxytryptamine receptor 2C	HTR2C	P28335	Membrane receptor	<div style="width: 60%; height: 10px; background-color: green;"></div>	54 / 34
Cannabinoid receptor 2	CNR2	P34972	Membrane receptor	<div style="width: 50%; height: 10px; background-color: green;"></div>	1633 / 133
Corticotropin-releasing factor receptor 1	CRHR1	P34998	Membrane receptor	<div style="width: 30%; height: 10px; background-color: green;"></div>	216 / 123
Sodium-dependent serotonin transporter	SLC6A4	P31645	Transporter	<div style="width: 20%; height: 10px; background-color: green;"></div>	12 / 33
Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit delta isoform	PIK3CD	O00329	Enzyme	<div style="width: 10%; height: 10px; background-color: green;"></div>	100 / 135
Cathepsin L1	CTSL	P07711	Cysteine Protease	<div style="width: 5%; height: 10px; background-color: green;"></div>	159 / 65
Epidermal growth factor receptor	EGFR	P00533	Tyr Kinase	<div style="width: 5%; height: 10px; background-color: green;"></div>	674 / 53
Cyclin-dependent kinase 2	CDK2	P24941	Ser_Thr Kinase	<div style="width: 5%; height: 10px; background-color: green;"></div>	160 / 518

Illustrative example:



Taken from *Lan Chem.* 151 (2011) 100-105
 adenosine rece

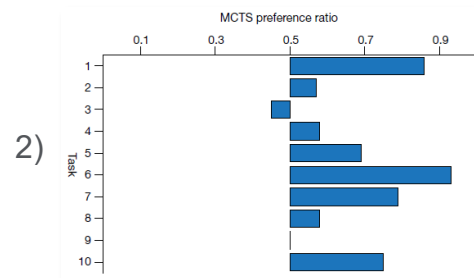
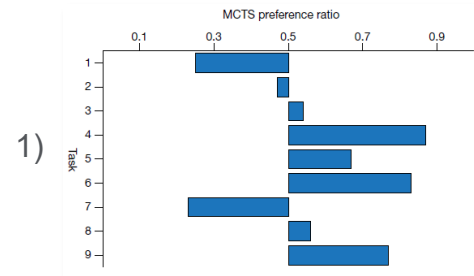
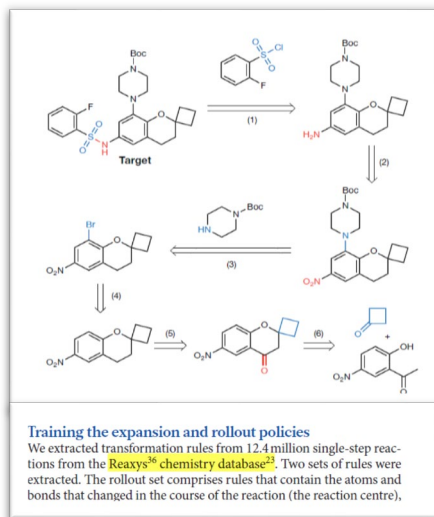
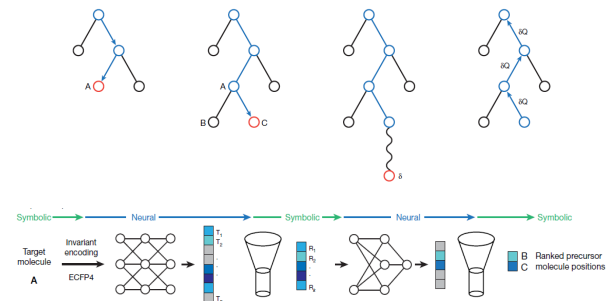
Using data from Reaxys we want to strengthening our support for the drug design cycle



*Freedom to operate
*Matched Molecular Pairs

Predictive Retrosynthesis: rewiring chemistry and redesigning synthetic routes

The challenge: Merck quotes that 55% of the time, a benchmarked catalytic reaction fails to deliver the desired product*. Therefore a radical and innovative step change in synthesis is needed.



*Science 02, Jan 2015

JChem

Marvin

- 1) Performance vs literature routes
- 2) Performance vs other predictive models

Predictive Retrosynthesis: rewiring chemistry and redesigning synthetic routes

The solution: Reaxys will offer a new approach to retrosynthesis in 2019 using deep neural networks and symbolic AI in collaboration with Prof. Mark Waller. The new approach will extend syntheses of small organic molecules into predictive modeling of previously unpublished synthetic pathways

Retrosynthesis benefits...



Augment chemist knowledge

- Via scrutinization of millions of synthetic pathways and presenting a highly customized route to the researcher



Reduce time to discovery / development / synthesis

- Design successful synthesis pathway for novel molecules where previous attempts may have failed



Reduce cost

- Design improved lead series, reduce risk of late stage failure
- Design more efficient synthesis routes (improved yield, reduced production costs)



Support IP strategy

- Identify alternative synthesis routes to a patented route for a commercial product
- Identify novel synthesis routes for a new molecule

..leading to improved business outcomes



Improve speed to market

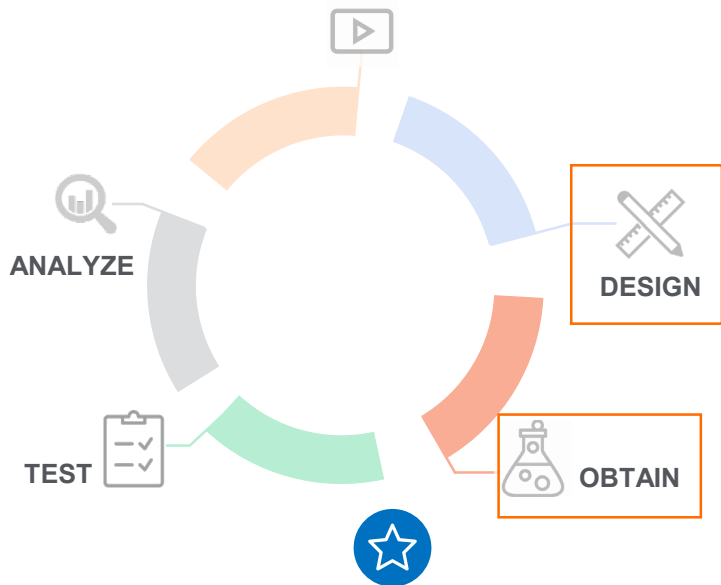


Embed / Leverage next-gen technology to out-compete



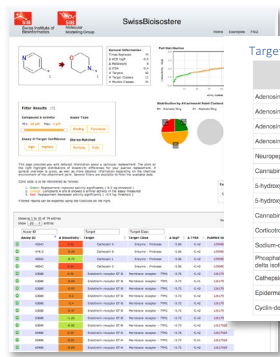
Increase productivity

Summary and outlook



Curated, clean, linked and normalized Reaxys/RMC data underpins all three applications

Matched Molecular Pair analysis



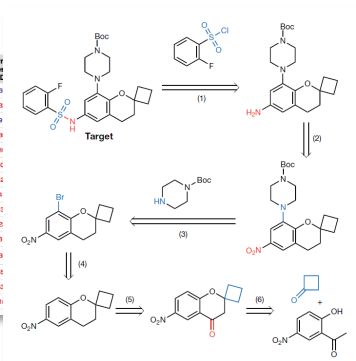
AI in pharma – will enable researchers to focus on innovation

Target Prediction

Target	Common name	Uniprot ID	Target Class	Probability*	Known actives (DZ/ZE)
Adenosine receptor A2a	ADORA2A	P08074	Membrane receptor		584 / 46
Adenosine receptor A2b	ADORA2B	P08075	Membrane receptor		173 / 23
Adenosine receptor A1	ADORA1	P05042	Membrane receptor		468 / 58
Adenosine receptor A3	NIA	P33765	Membrane receptor		693 / 63
Neurokinin Y receptor type 5	NPY5R	Q15761	Membrane receptor		269 / 14
Caeruloplasmin receptor 1	CNR1	P21584	Membrane receptor		1418 / 30
5-hydroxytryptamine receptor 2A	HTR2A	P08223	Membrane receptor		106 / 40
5-hydroxytryptamine receptor 2C	HTR2C	P08235	Membrane receptor		54 / 34
Caeruloplasmin receptor 2	CNR2	P34072	Membrane receptor		1600 / 11
Coricotropin-releasing factor receptor 1	CRFR1	P34988	Membrane receptor		216 / 12
Sodium-dependent serotonin transporter	SLC6A4	P31845	Transporter		19 / 23
Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit delta isoform	PIK3CD	Q03529	Enzyme		500 / 13
Cathepsin L1	CTSL	P07711	Cysteine Protease		159 / 65
Epidermal growth factor receptor	EGFR	P05533	Tyrosinase		674 / 51
Cyclin-dependent kinase 2	CDK2	P30461	Ser_Thr_Kinase		149 / 50

ChemAxon's technology and tools underpin these developments

Predictive Retrosynthesis



Acknowledgments

Antoine Daina

Christophe Bovigny

Vincent Zoete

Molecular Modelling Group – SIB



Swiss Institute of
Bioinformatics



Pieder Caduff

Olivier Barberan

Ivan Kristic

Jürgen Swienty-Busch

Elena Herzog

Abhinav Kumar

Mark Waller

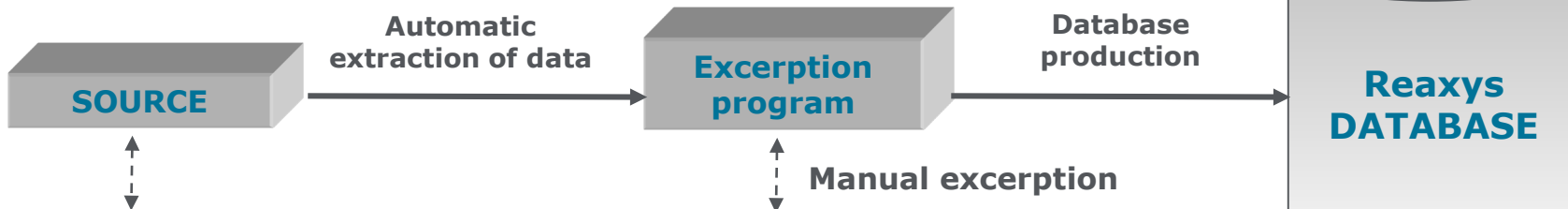
Marwin Segler



For more information please contact Dr Rosalind Sankey
r.sankey@elsevier.com

Additional Slides

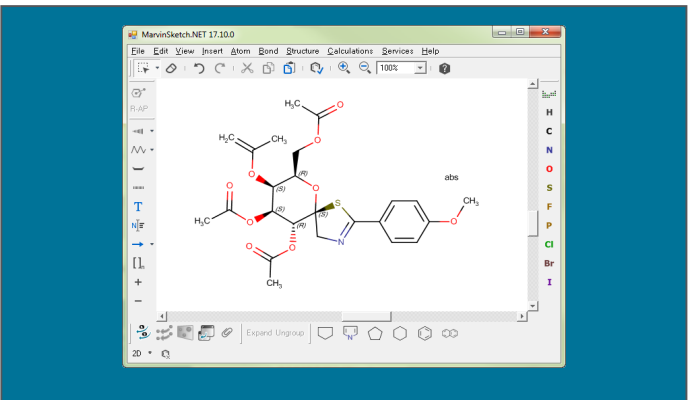
Reaxys Workflow (simplified)



Paper or electronic file



Structures, Markush, chemical names, NMR, IR, MP, USE, ...
Reactions, temp., solvent, ...
Targets, MedChem data



ieI (intuitive Excerption Interface)

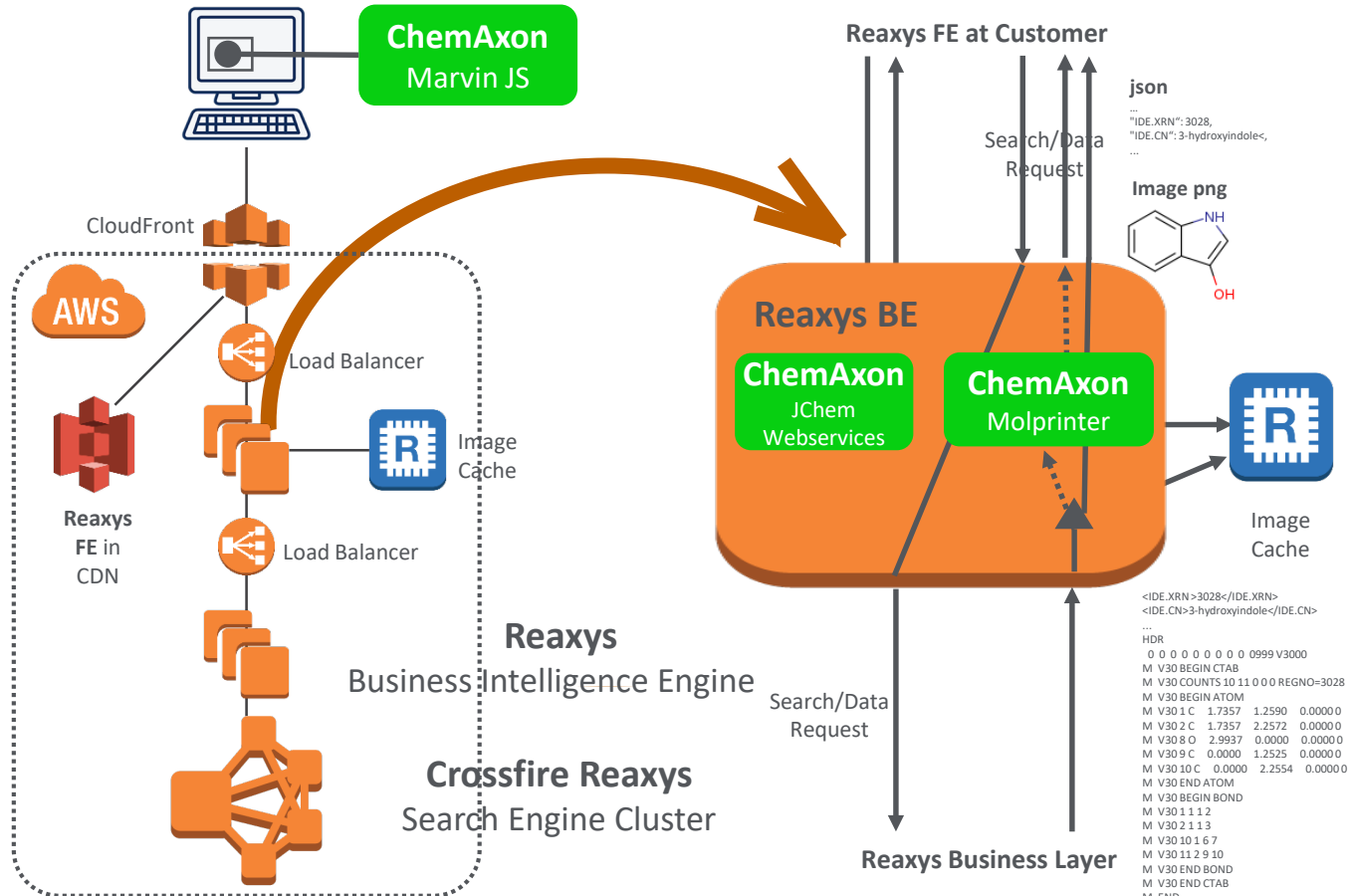


ChemAxon Marvin Sketch in Reaxys excerption

The screenshot displays the ChemAxon Reaxys software interface, version 3.71.21.0, with the following components:

- Top Bar:** File Edit View Exception Checks Extras Window Help. A menu bar with tabs for ASSAY, TARGET, BIOM, MEAS, PH, META..., IDE, RX, NMR, MS, COL, MP, IR, UV.
- Target Pane:** inorg. chim. acta, 2018, 473, 121-132, tetranuclear cubane cu4o4 complexes as prospective anticancer agents: design, synthesis, structural ...
- Compounds List:** A list of 11 items, with IDE #11: naloxone selected. Other items include 2-amino-2-(hydroxymethyl), 2-hydroxy-3-methoxybenza, H4L, 1, 2, copper acetate monohydrat, triethylamine, piperazine, and calf thymus DNA.
- Compound Identification (IDE) #11:** A detailed view for the selected compound, naloxone. It includes:
 - Working name:** (empty field)
 - Label:** (empty field)
 - Structure:** A 3D ball-and-stick model of naloxone, a morphine derivative with a hydroxyl group and an allyl group. The label "abs" is next to the structure.
 - Molecular formula:** C19H21NO4
 - Main name:** naloxone
 - Synonym(s):** (-)-17-allyl-4,5 α -epoxy-3,14-dihydroxy-6-morphinan-6-one and (5 α)-4,5-epoxy-3,14-dihydroxy-17-(2-propenyl)morphinan-6-one
 - Modification:** A set of dropdown menus.
 - Macroscopic type:** A dropdown menu.
 - Comments:** A text area for notes.
- Left Panel:** A sidebar showing a search result for "Inorganica Chimica Acta" with a thumbnail and a brief description of tetranuclear cubane Cu₄O₄ complexes.

Reaxys Architecture Diagram (simplified)



Key Reaxys Use Case: Reaction Searching and Synthesis Planning

Using MarvinJS to enter even complex structure/reaction queries

The image displays two overlapping screenshots of the Reaxys software interface. The background screenshot shows the main search page with the following elements:

- Navigation menu: Quick search, Query builder, Results, Synthesis planner, History.
- Search area: "Search substances, reactions, documents and bioactivity data" with a sub-note "in Reaxys, Reaxys Medicinal Chemistry, PubChem, eMolecules, LabNetwork and SigmaAldrich".
- Search input: "Search Reaxys **New**" with a "Find >" button.
- Example text: "Documents, e.g. publications about quasicrystals AND" with a "Draw" button.

The foreground screenshot shows the MarvinJS structure editor with the following elements:

- Header: "Reaxys" and navigation links: Quick search, Query builder, Results, Synthesis planner, History.
- Editor title: "Structure editor ChemAxon's MarvinJS".
- Editor content: A large empty canvas with the "Marvin JS by ChemAxon" logo in the center.
- Right sidebar: "Search this structure as:" with options: As drawn (selected), As substructure, Similar, Tautomers, Stereo, Additional ring closures, Related Markush, Salts, Mixtures, Isotopes, Charges, Radicals, and More options.
- Bottom bar: "Clear", "Cancel", "Transfer to query >", and "Feedback" buttons.