Analyzing Building Blocks Diversity for DNA Encoded Library Design

Cresset User Group Meeting

Nik Stiefl & Finton Sirockin, Novartis

2016.06.16

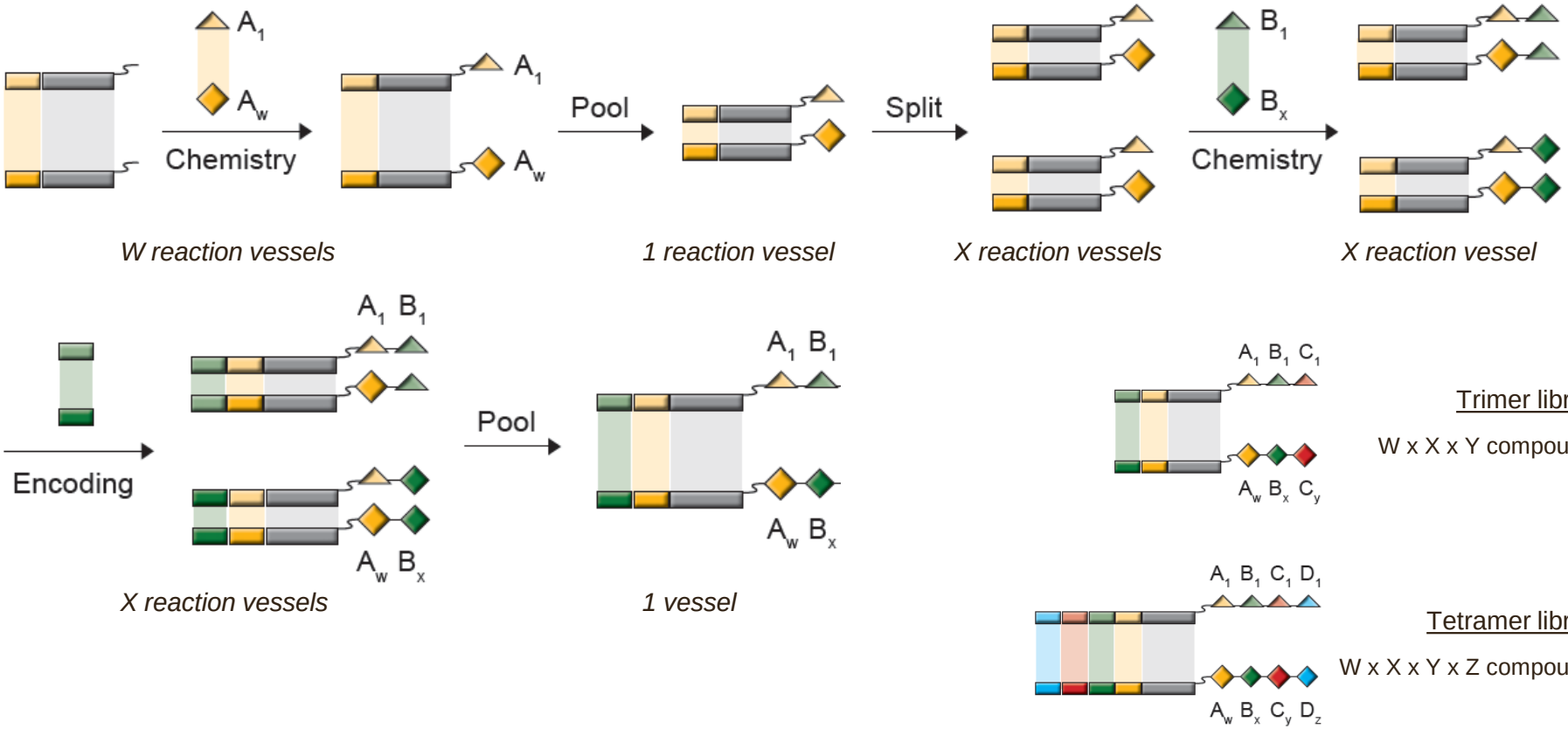
NOVARTIS

Outline

- DNA Encoded Libraries (DEL)
- Building blocks selection process
- Capping group selection
  - 2D fingerprints
  - Cresset fields
- Caveats
- Conclusion

DEL technology uses DNA oligonucleotides to record the combinatorial synthesis of organic molecules...

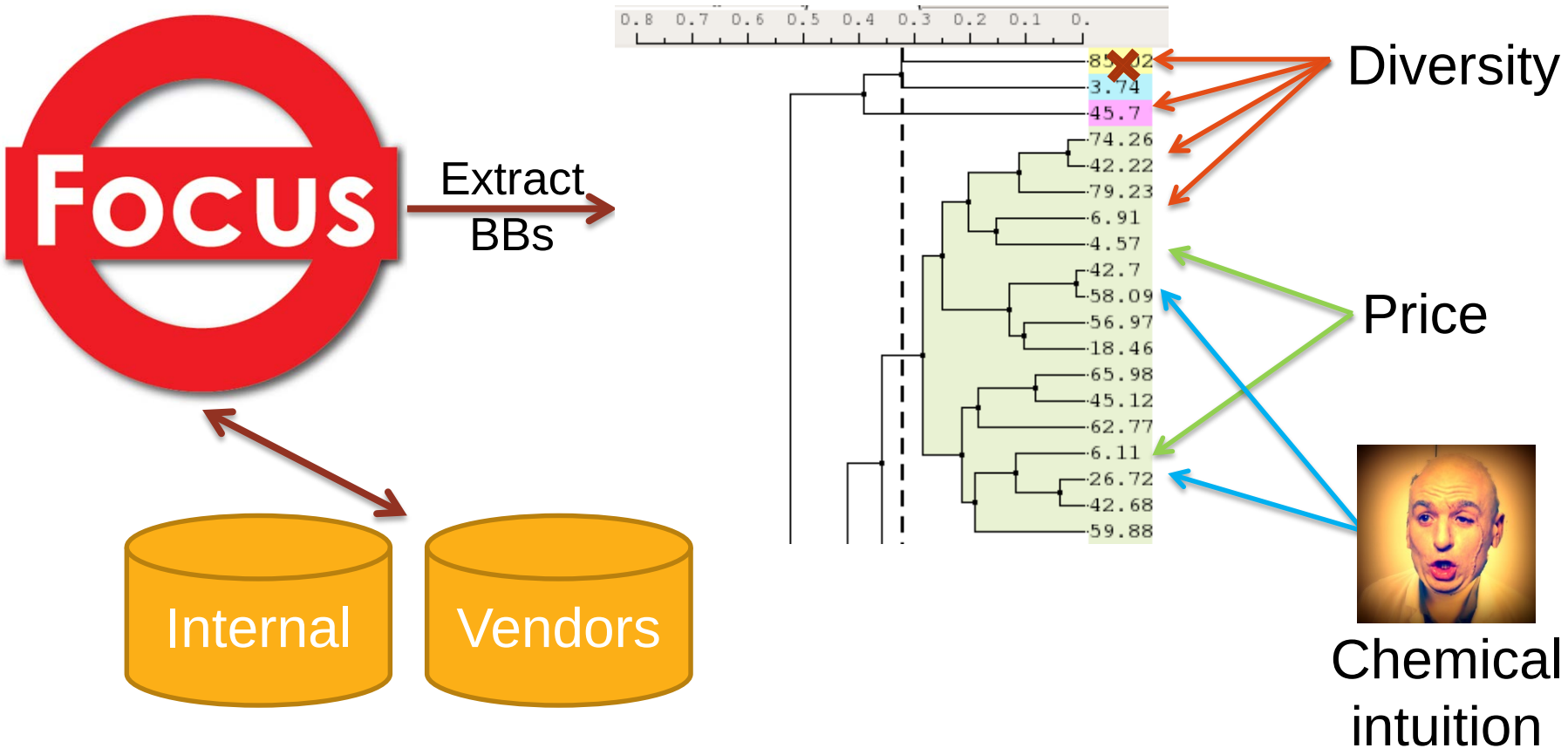
- Dimer library, W x X compounds
  - Pos 1: W building blocks, A_1 to A_W
  - Pos 2: X building blocks, B_1 to B_X



NOVARTIS

Building Blocks selection process

General process



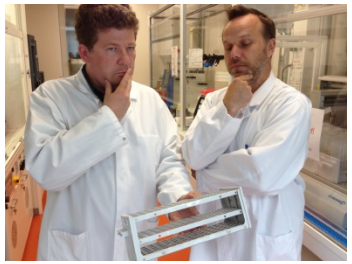
Stiefl, N.; Gedeck, P.; Chin, D.; Hunt, P.; Lindvall, M.; Spiegel, K.; Springer, C.; Biller, S.; Buenemann, C.; Kanazawa, T.; Kato, M.; Lewis, R.; Martin, E.; Polyakov, V.; Tommasi, R.; van Drie, J.; Vash, B.; Whitehead, L.; Xu, Y.; Abagyan, R.; Raush, E.; Totrov, M. FOCUS — Development of a Global Communication and Modeling Platform for Applied and Computational Medicinal Chemists. *J. Chem. Inf. Model.* **2015**, *55* (4), 896–908.

NOVARTIS

Building Blocks selection process

What makes DNA Encoded Libraries different ?

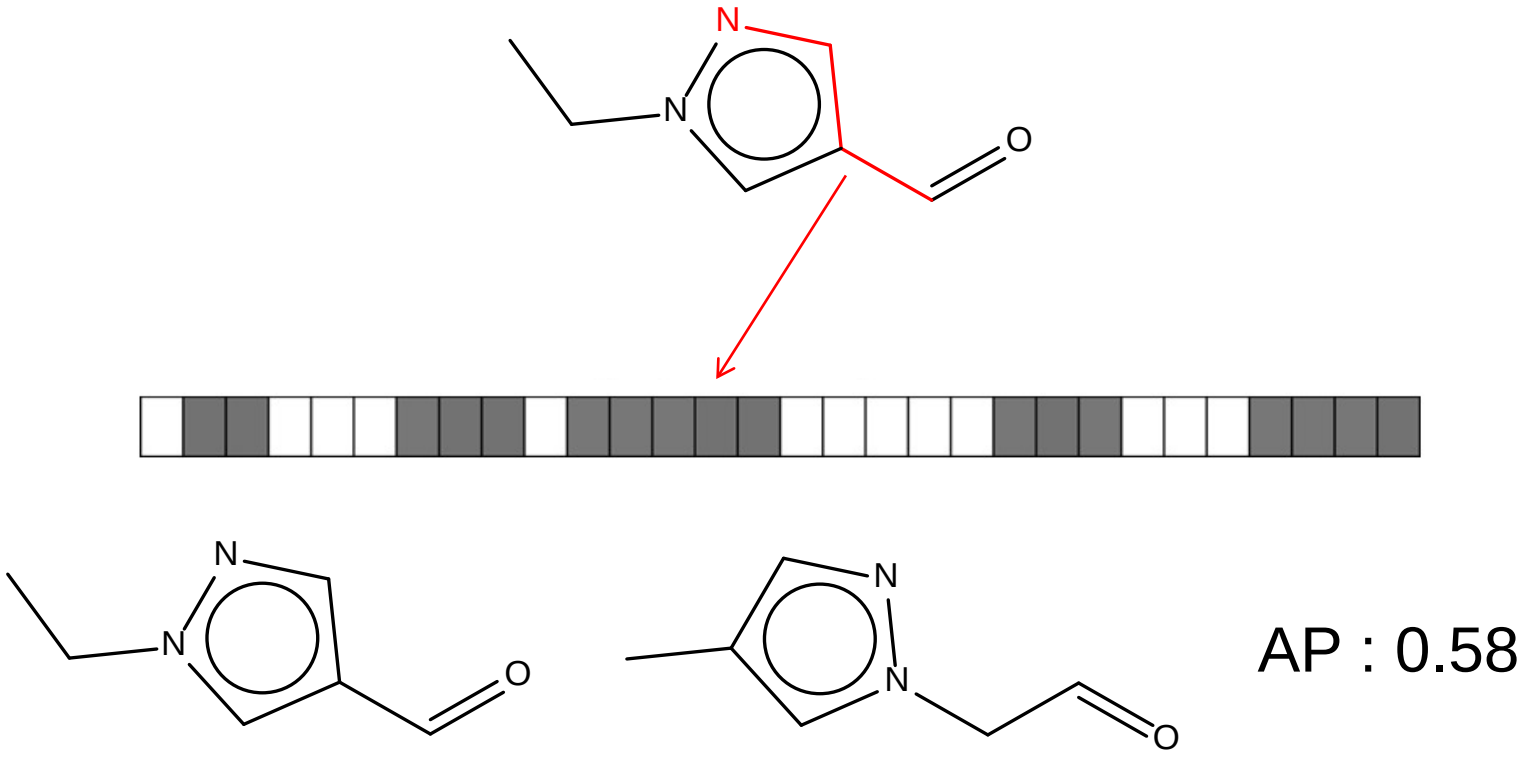
	Project libraries	DNA Encoded Libraries
# of building blocks	10-50	100s – 1000s
Design principle	Target knowledge	Diversity & Density
Priority on	Scaffold	Combinator'ics'
Property space	Target driven and limited	Simultaneous SAR and physico-chemical exploration



2D descriptions

Example: Atom Pairs[1]

- Atom properties (atomic number, degree, number pi electrons) and inter-atom distances encode molecules



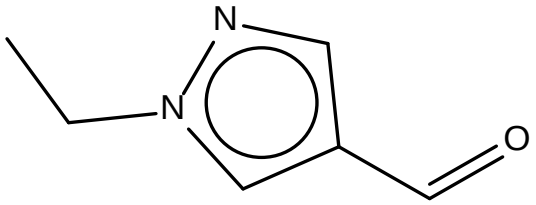
AP : 0.58

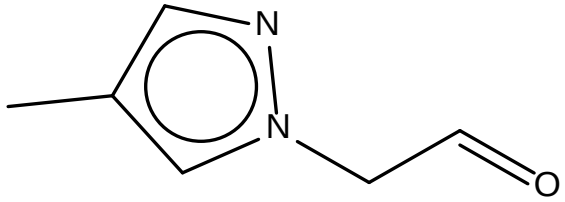
[1] Sheridan, R. P.; Miller, M. D.; Underwood, D. J.; Kearsley, S. K. *J. Chem. Inf. Comput. Sci.* **1996**, 36, 128 – 136

2D descriptions

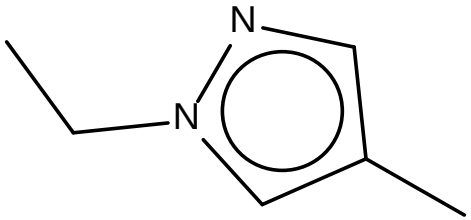
Influence of reacting group on descriptors

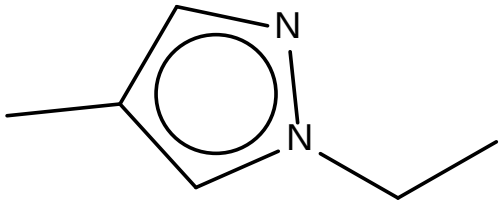
- Modifying atom invariants and Rooted Fingerprints [1]



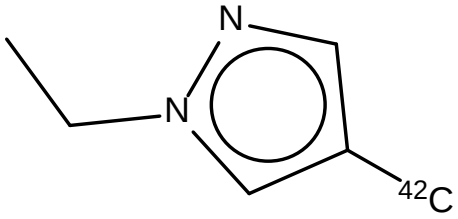


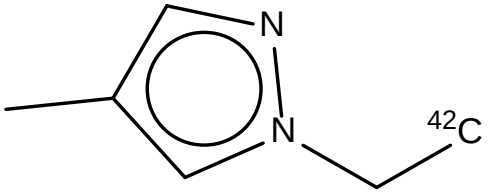
AP : 0.58





AP : 1.0





APMI : 0.43
APRooted : 0.29

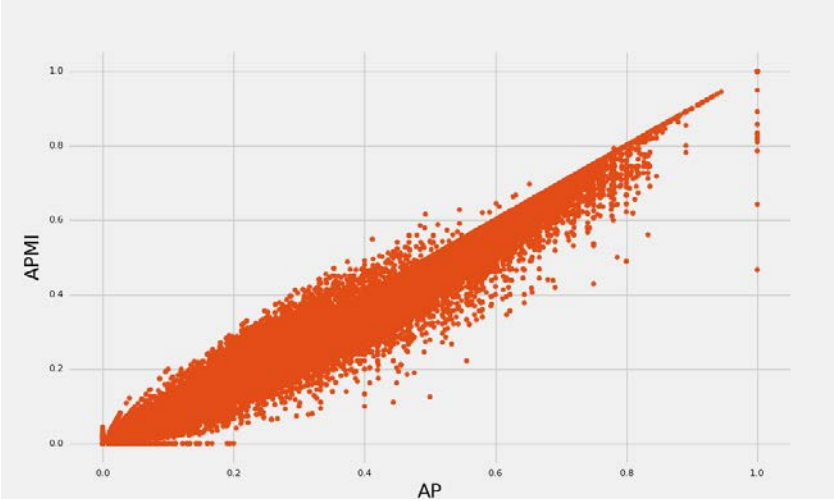
APMI : Atom Pairs Modified Invariant

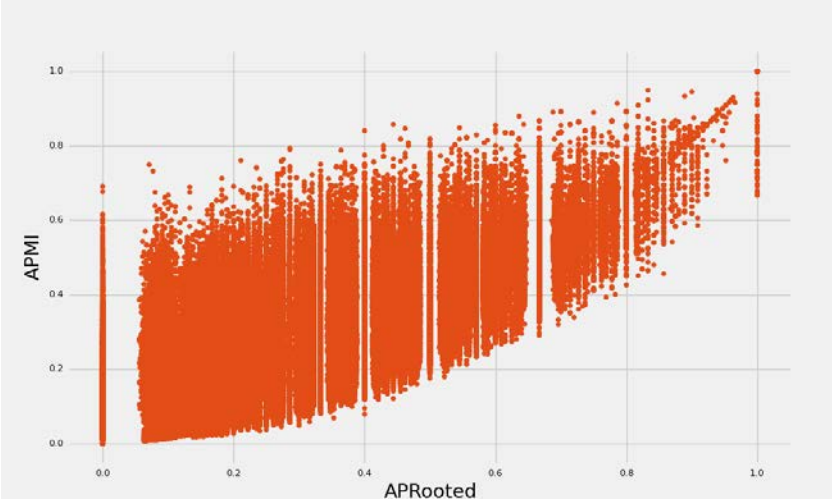
[1] Vulpetti, A.; Hommel, U.; Landrum, G.; Lewis, R.; Dalvit, C. *J. Am. Chem. Soc.*, **2009**, 131, 12949-12959

NOVARTIS

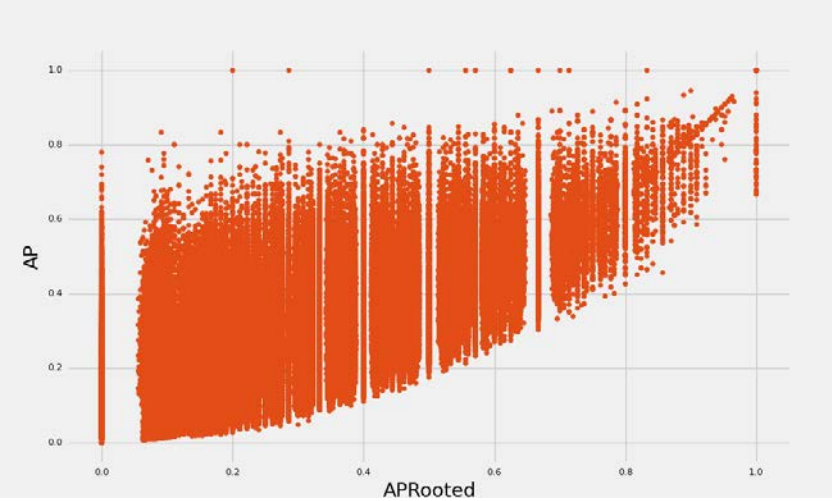
Capping groups specific fingerprints

Clockwise : AP vs. APMI ; APRooted vs. APMI ; APRooted vs. AP.





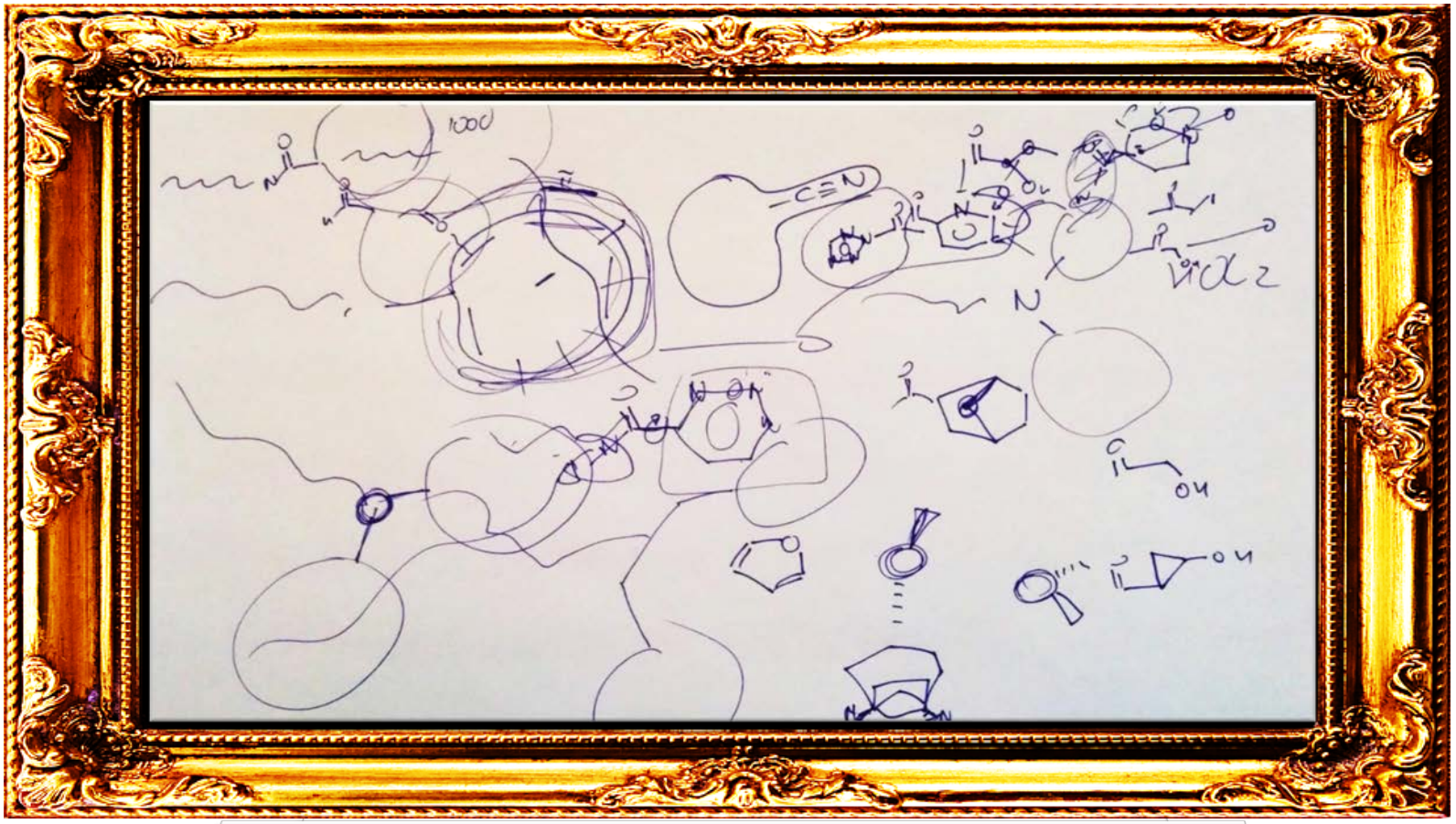
Similarity plots of pairs of 1832 aldehydes



NOVARTIS

From 2D to 3D

Brainstorming in Novartis

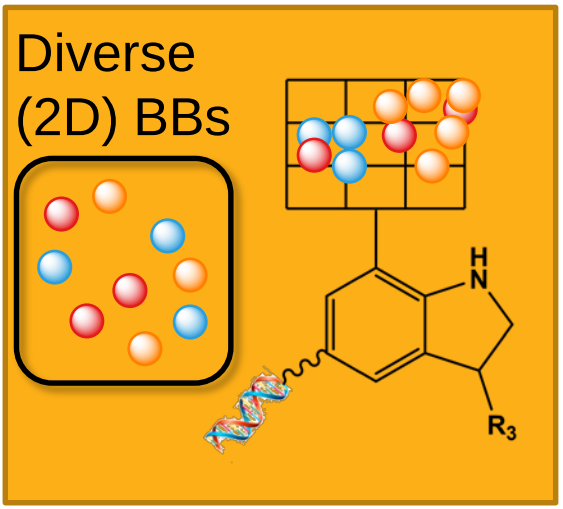


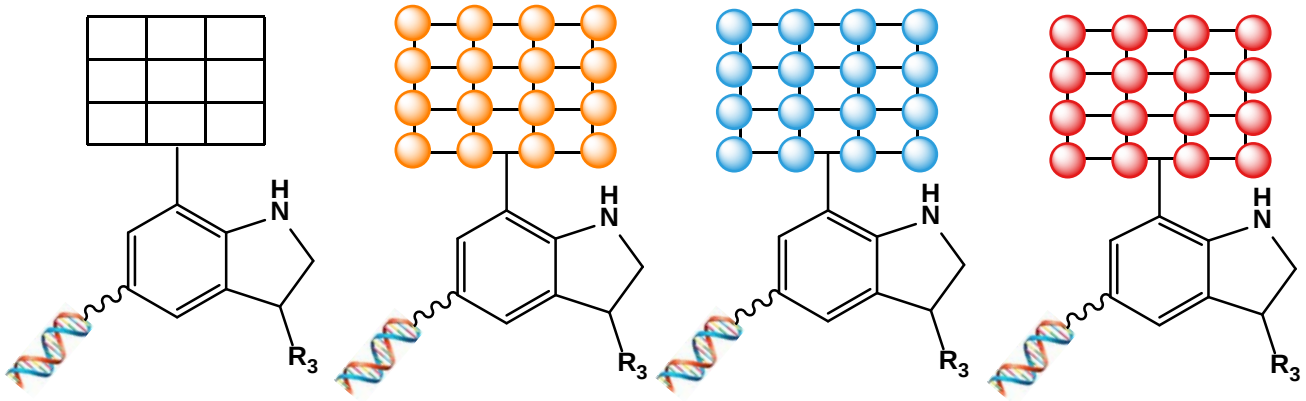
A chemist and two modellers : Le Brainstorming

NOVARTIS

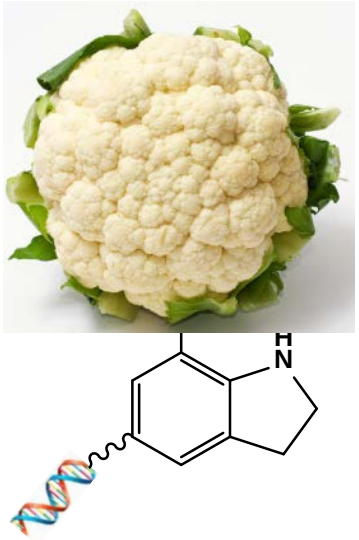
The Cauliflower™ (Frédéric Berst - Novartis)

Distributing pharmacophores (Cresset fields) in space





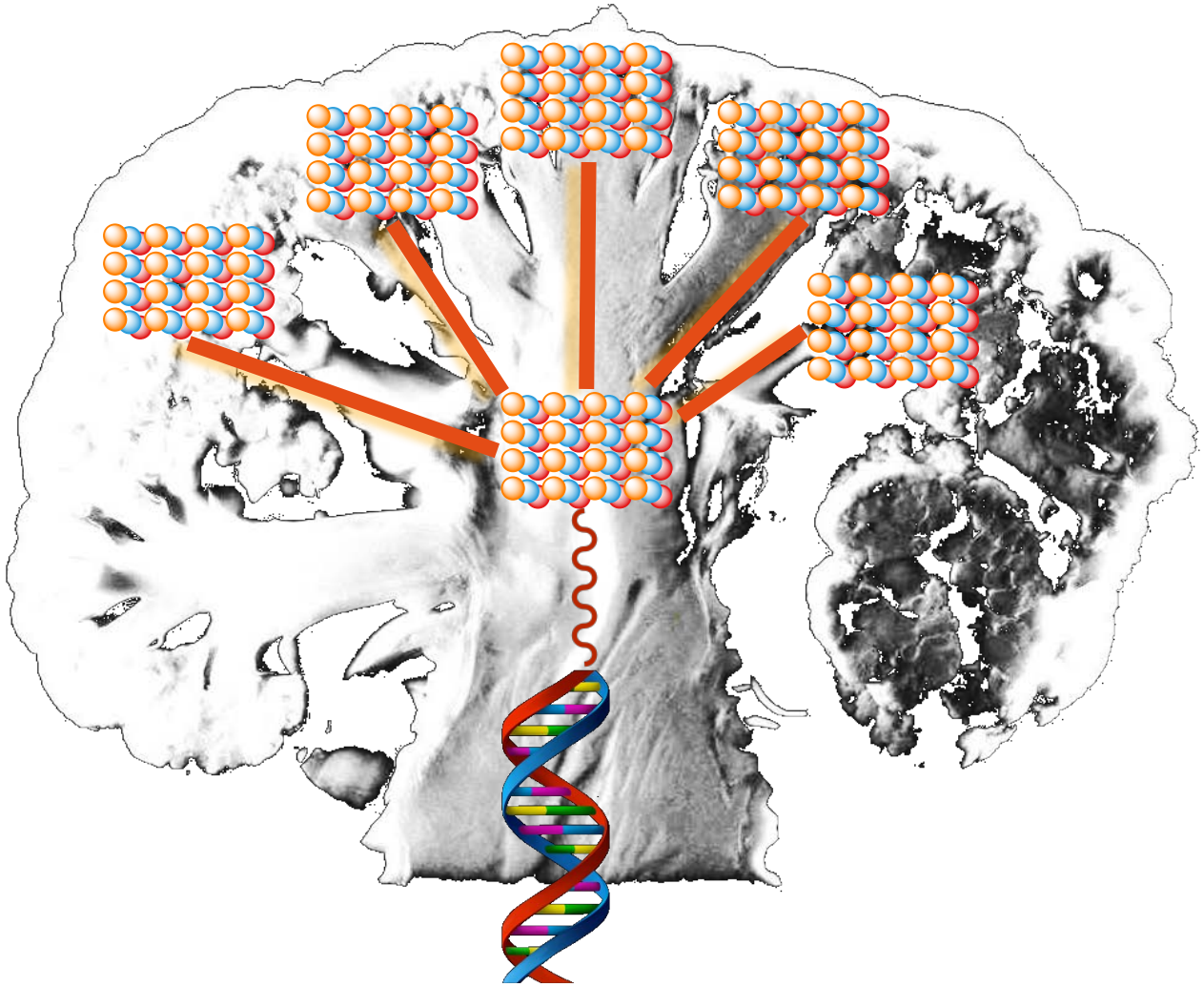
Distributing different PH4 (fields) types as evenly as possible in space



Hydrophobic
Polar +
Polar –
Space around derivatisation point

NOVARTIS

Le «Caulitree»



NOVARTIS

Building Blocks clustering with Cresset technology

Workflow developed by Paolo Tosco (Cresset)

BBs preparation

- Cleaning
- Reacting group labelling

BBs conformations

- 3D coordinates generation and protonation
- Conformations generation

BBs Cresset processing

- Alignment on bond vector
- Cresset similarity matrix calculation
- Clustering

BBs selection

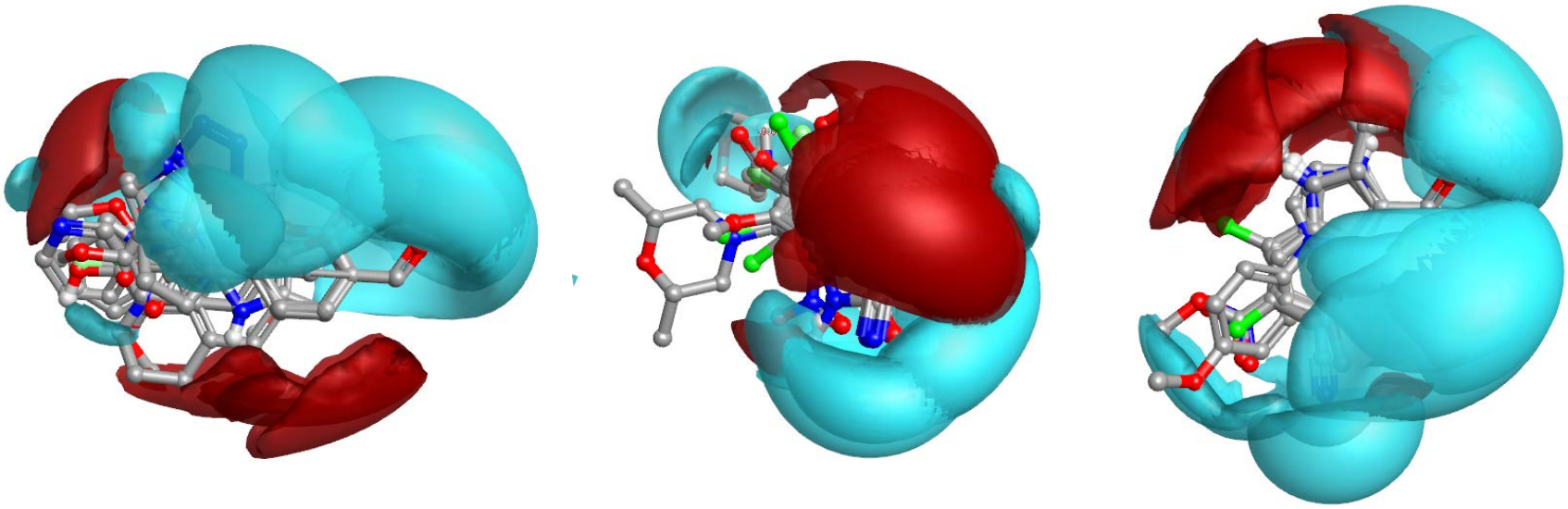
- Cluster visualisation
- BBs selection by clusters using multiple parameters (price, reactivity profiling, density of in-house availablility, etc)

Non Cresset

CRESSET

Pharmacophore distribution clustering

*Help chemists choosing BBs to maximise pharmacophore coverage **and** BB attractiveness*

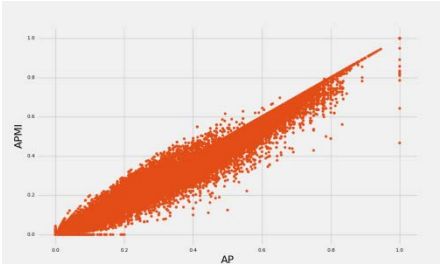


Example clusters from the processing of 1832 aldehydes (~ 12 hours on a cluster)

Handling the reactive handle

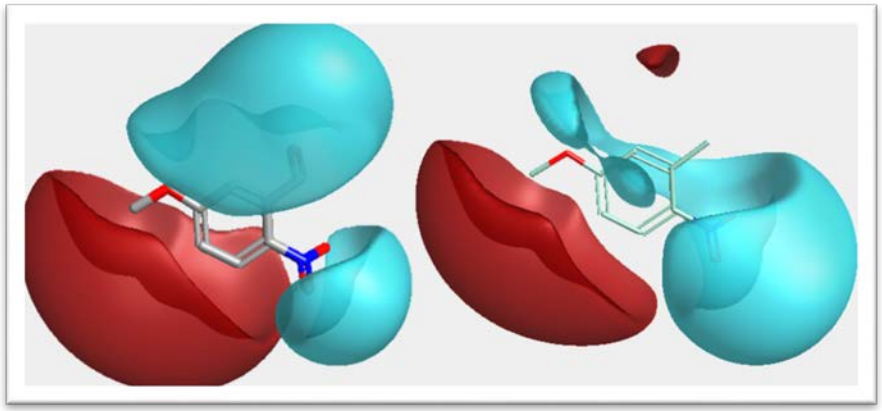
Functional group vs. isotope labelling

- Minor influence on 2D description depending on BB size and functionality



AP vs. APMI

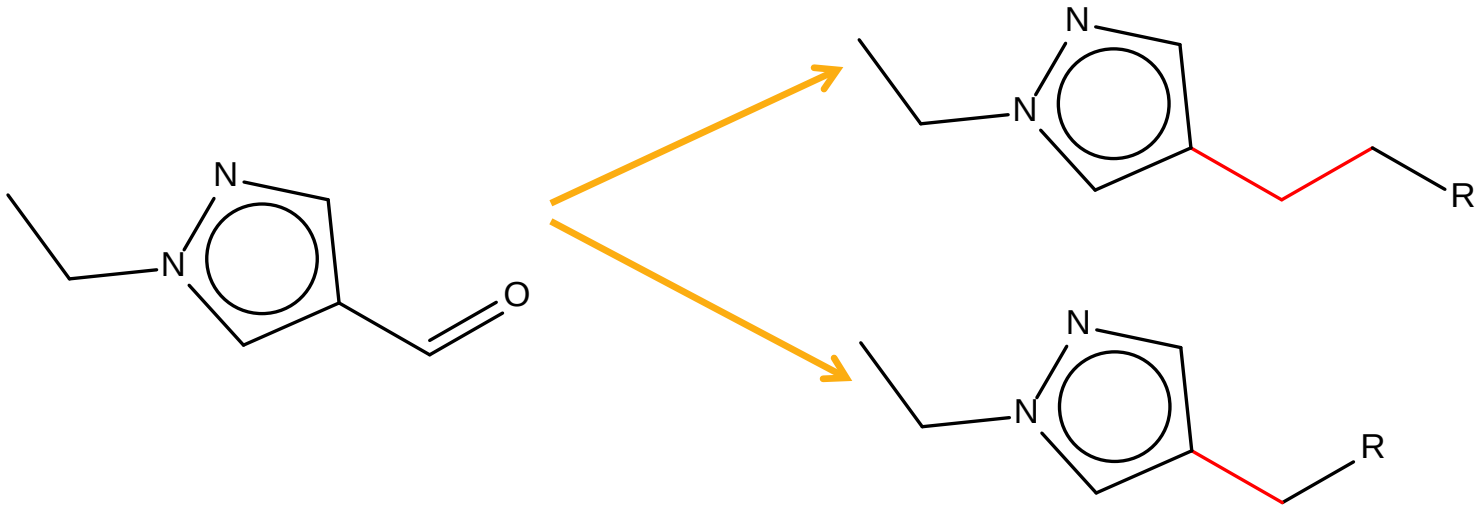
- Major influence on Cresset fields (→ transform with respect to resulting product)

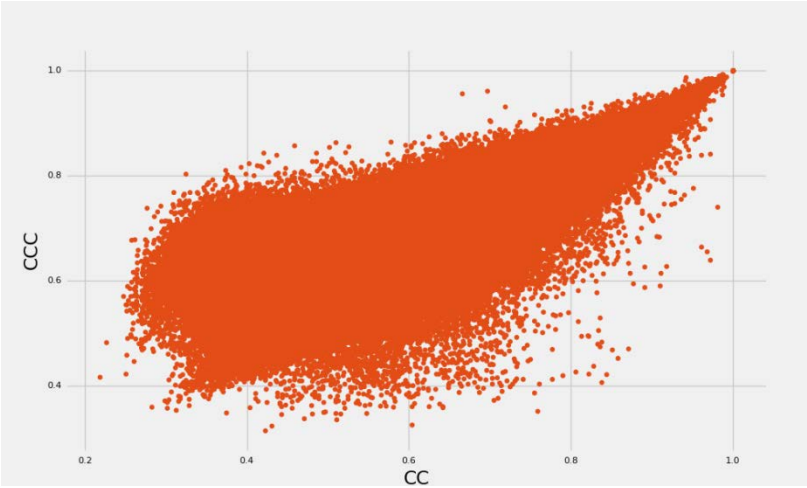


NOVARTIS

Handling the reactive handle

Replace vs. add (“CC” vs. “CCC”)

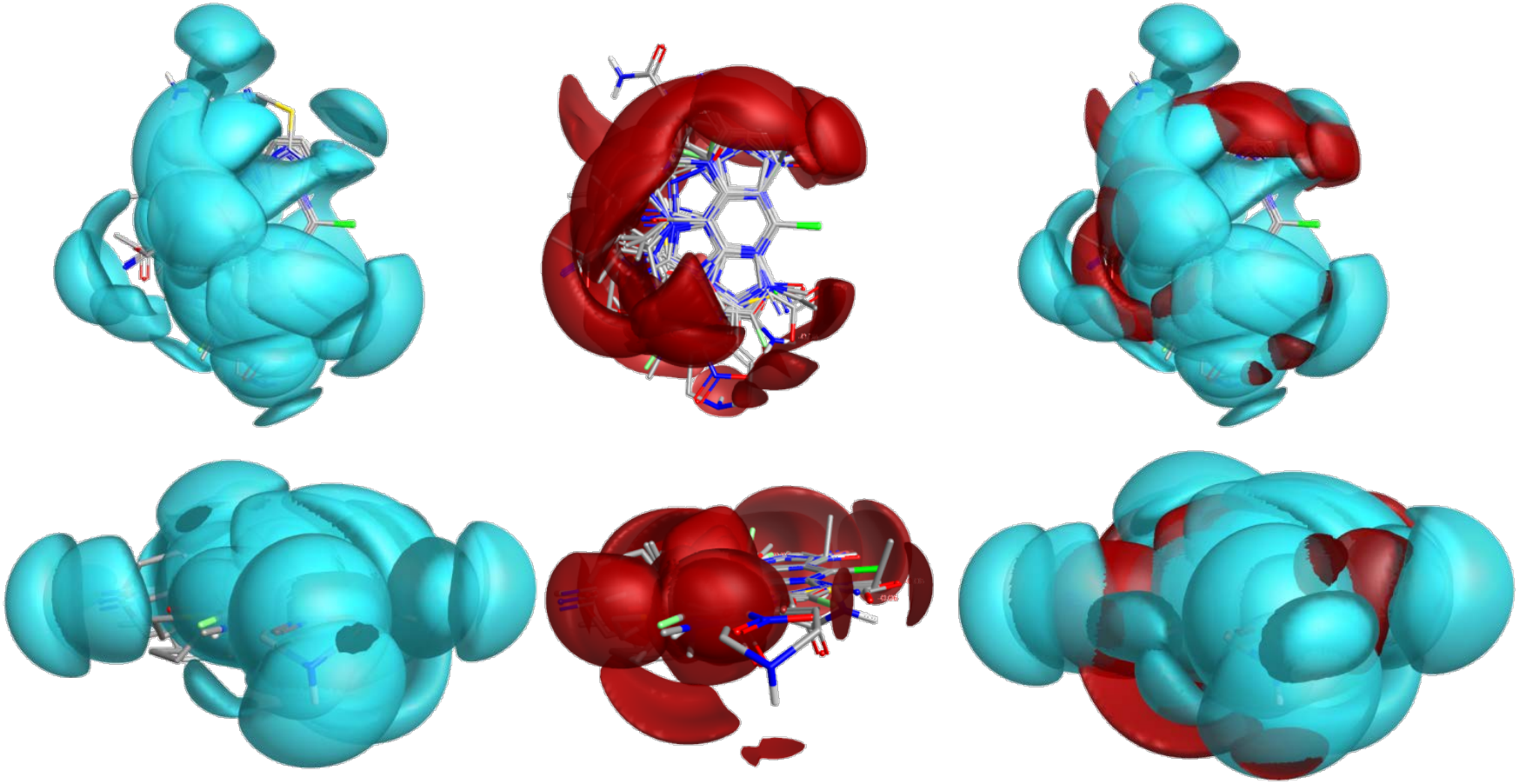




NOVARTIS

Pharmacophore spaces

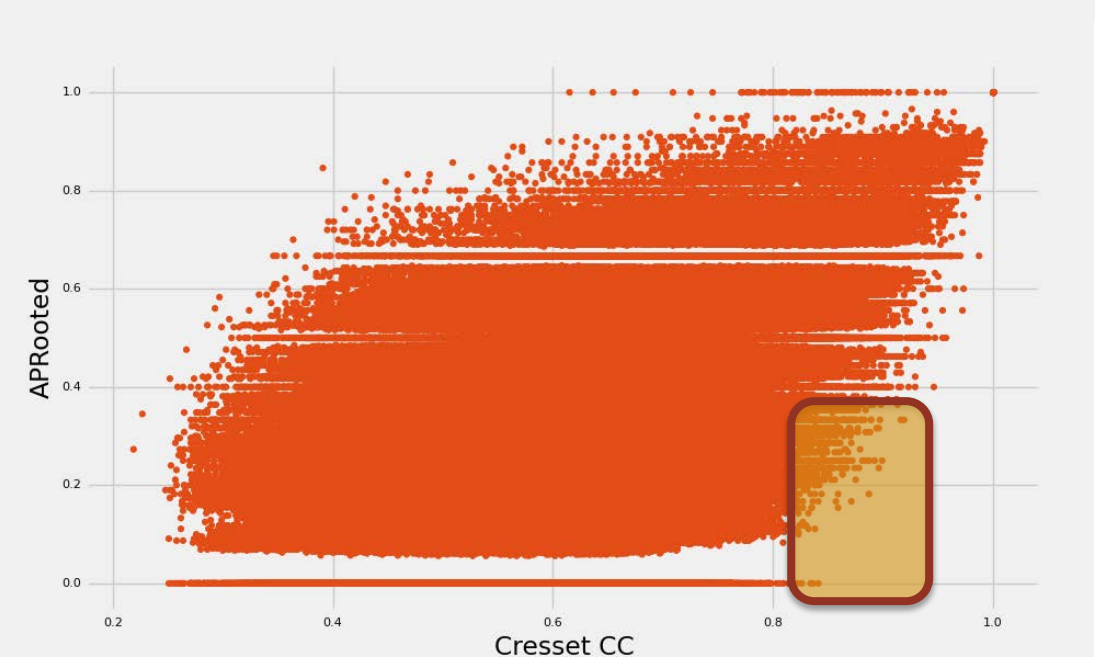
Identifying missing spots

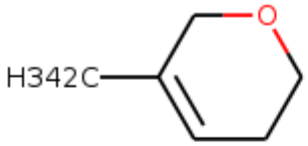


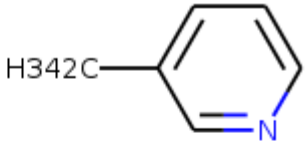
1300 NAS reagents aligned on the derivatisation vector. Top and side views
Disc-like electropositive density with obvious hole on top and bottom

2D vs 3D - does it make a difference?

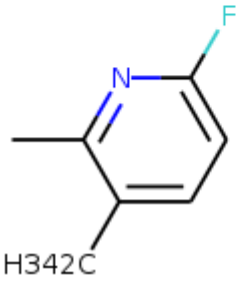
Cresset vs. APRooted

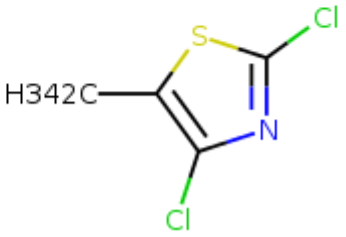






Cresset: 0.92 APRooted: 0.33



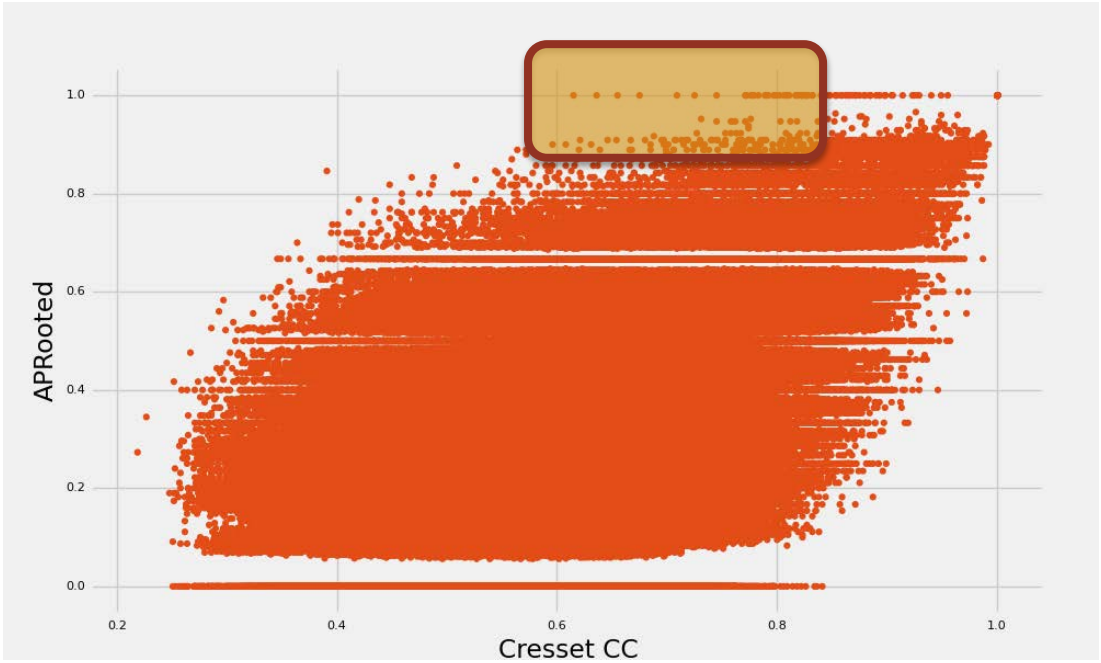


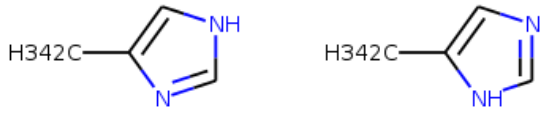
Cresset: 0.90 APRooted: 0.40

NOVARTIS

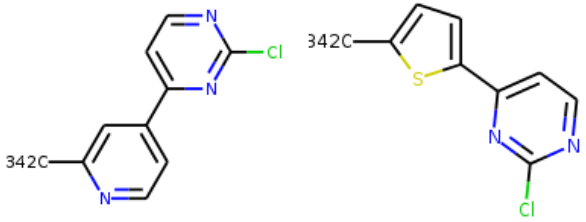
2D vs 3D - does it make a difference?

Cresset vs. APRooted

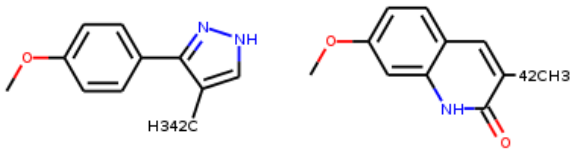




Cresset: 0.66 APRooted: 1.00



Cresset: 0.57 APRooted: 0.88

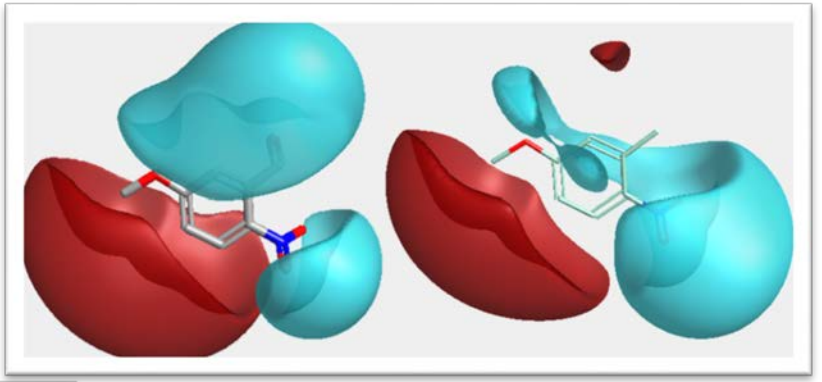


Cresset: 0.39 APRooted: 0.85

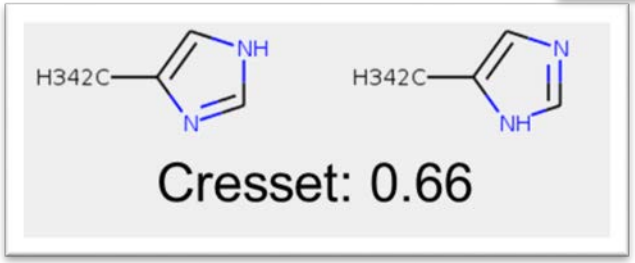
NOVARTIS

Caveats

- Reactive group treatment



- Tautomers



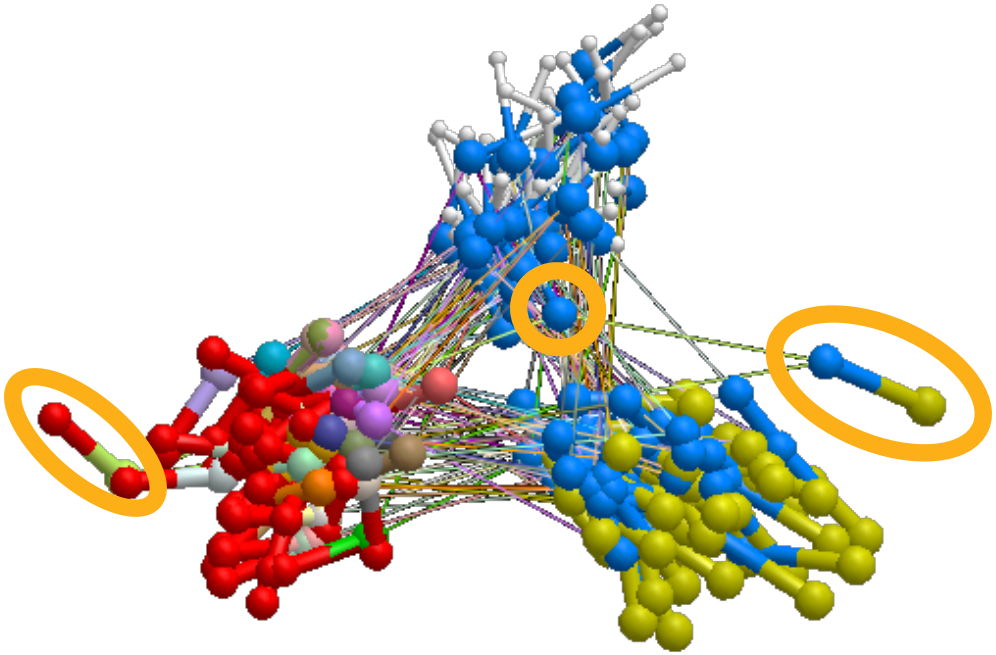
- Capping group only (for the moment)

NOVARTIS

Outlook

Core scaffolds – exit vector distribution analysis

Identification of unexplored exit vector space (branches in the caulitree)



To be combined with spatial distribution of cresset fields

Conclusions

- Library design tool to distribute pharmacophoric constraints evenly around linking vector
- Used to aid chemists visually select BBs from pharmacophoric clusters
- Modified 2D descriptors should be used in conjunction
- Applied in-house to large BB datasets for prioritization

Acknowledgments

F. Berst

J. Ottl

C. Freslon

R. Martinez

P. Fille

A. Marzinzik

X. Pelle

Y. Ruff

S. Liverneau

S. Pickett

P. Lehr

M.-H. Bellance

V. Acker

P. Drueckes

K. Twesten

S. Haddad

C. Dumelin

M. Huebscher

J. Giovannoni

D. Ouzineb

N. Schneider

C. Gaul

P. Ertl

R. Lewis

H. Hemmerle

J. Duca

G. Landrum (now @ Knime)

P. Tosco

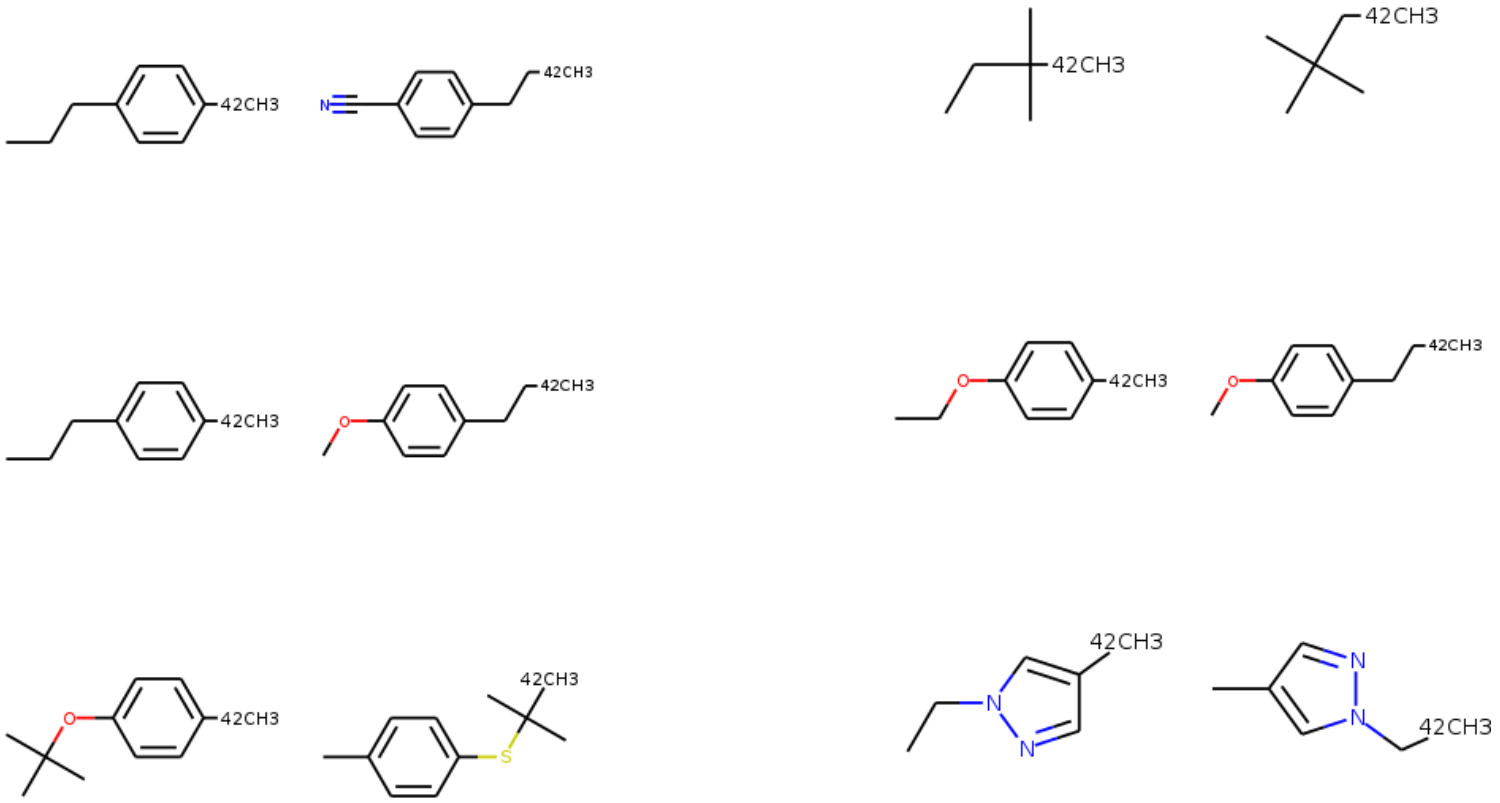
T. Cheeseright

M. Mackey

R. Chauhan

- BACKUPS

Large discrepancies between AP and APRooted

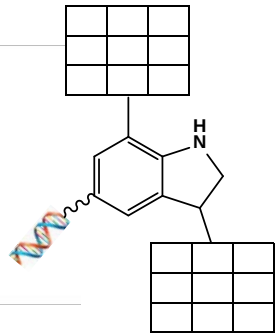


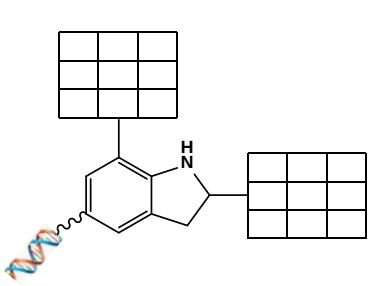
NOVARTIS

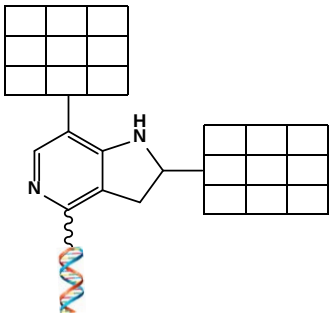
The Cauliflower™ (Frédéric Berst)

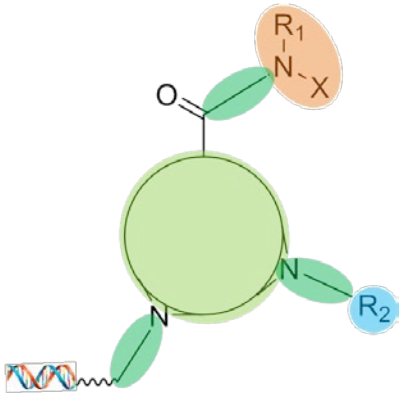
It's more like a trunk (scaffold) with cauliflowers fruits (capping) : a caulitree ?

- A DELibrary is **not limited to a single central scaffold**



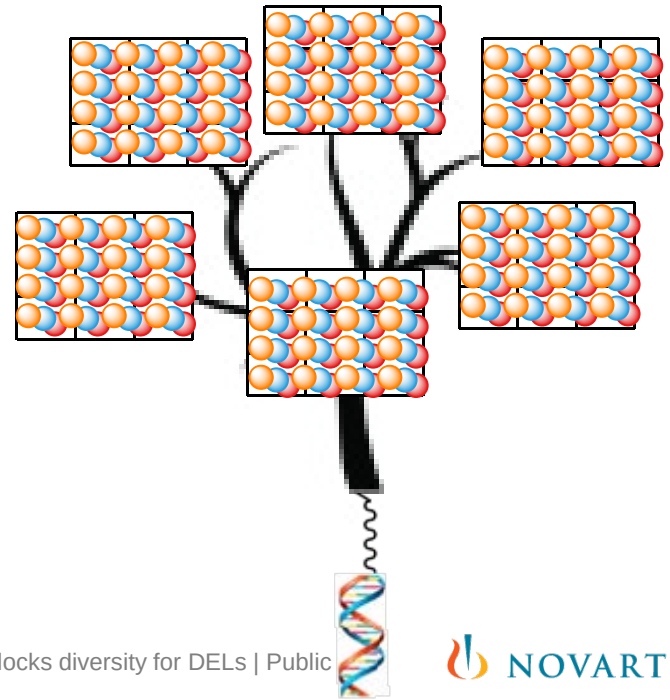






Elements of design

- Scaffold fragment set
  - Derivatization vectors
  - Pharmacophoric elements
- Derivatisation fragment set
  - Pharmacophoric elements



NOVARTIS