PERFORMANCE OF SUB-PHARMACOPHORE MODELS AS SEEDS IN DRUG DISCOVERY

Workshop in Cheminformatics
Sep 2014
EPFL Lausanne

Modest Korff, Joel Freyss, Thomas Sander

ACTELION

TOPICS

- Task
- Flexophore descriptor
- Sub-Flexophores
- Improving pharmacophore models
- Results and conclusions

TASK

Improve pharmacophore models

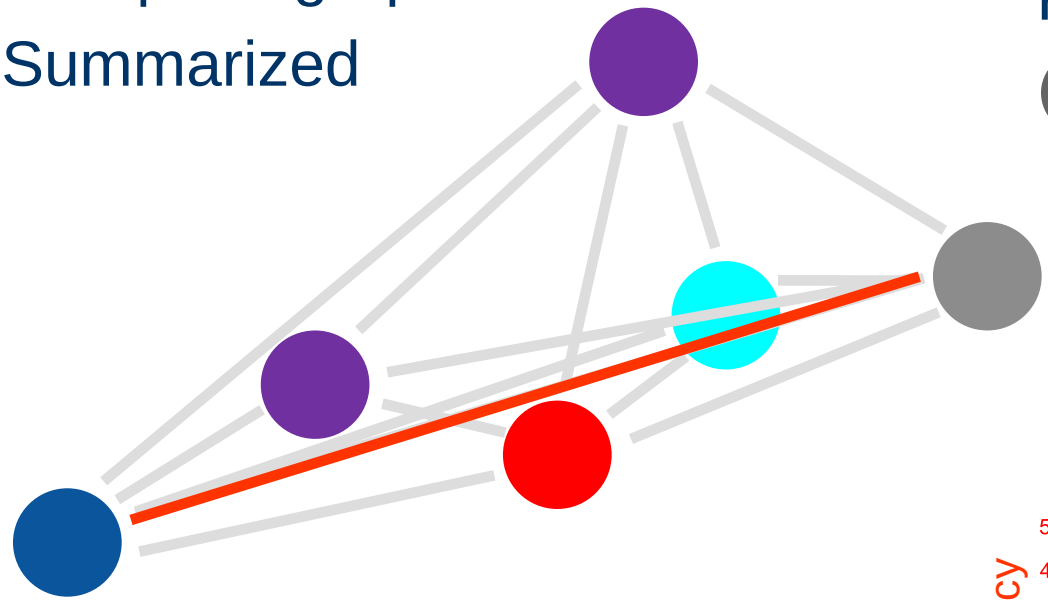
for

successful virtual screening

THE FLEXOPHORE DESCRIPTOR

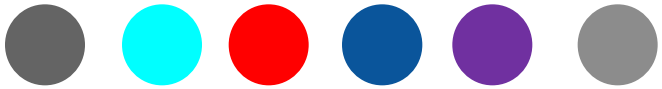
Complete graph

Summarized



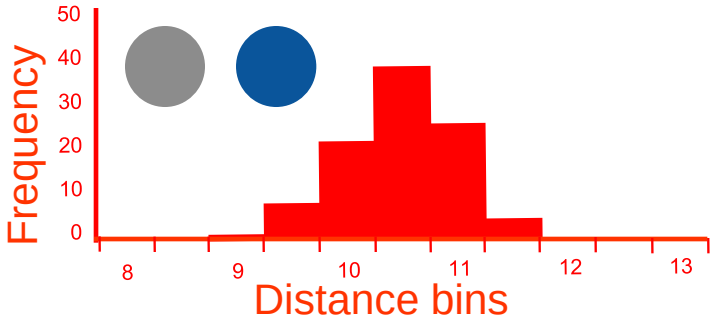
Nodes:

Pharmacophore points

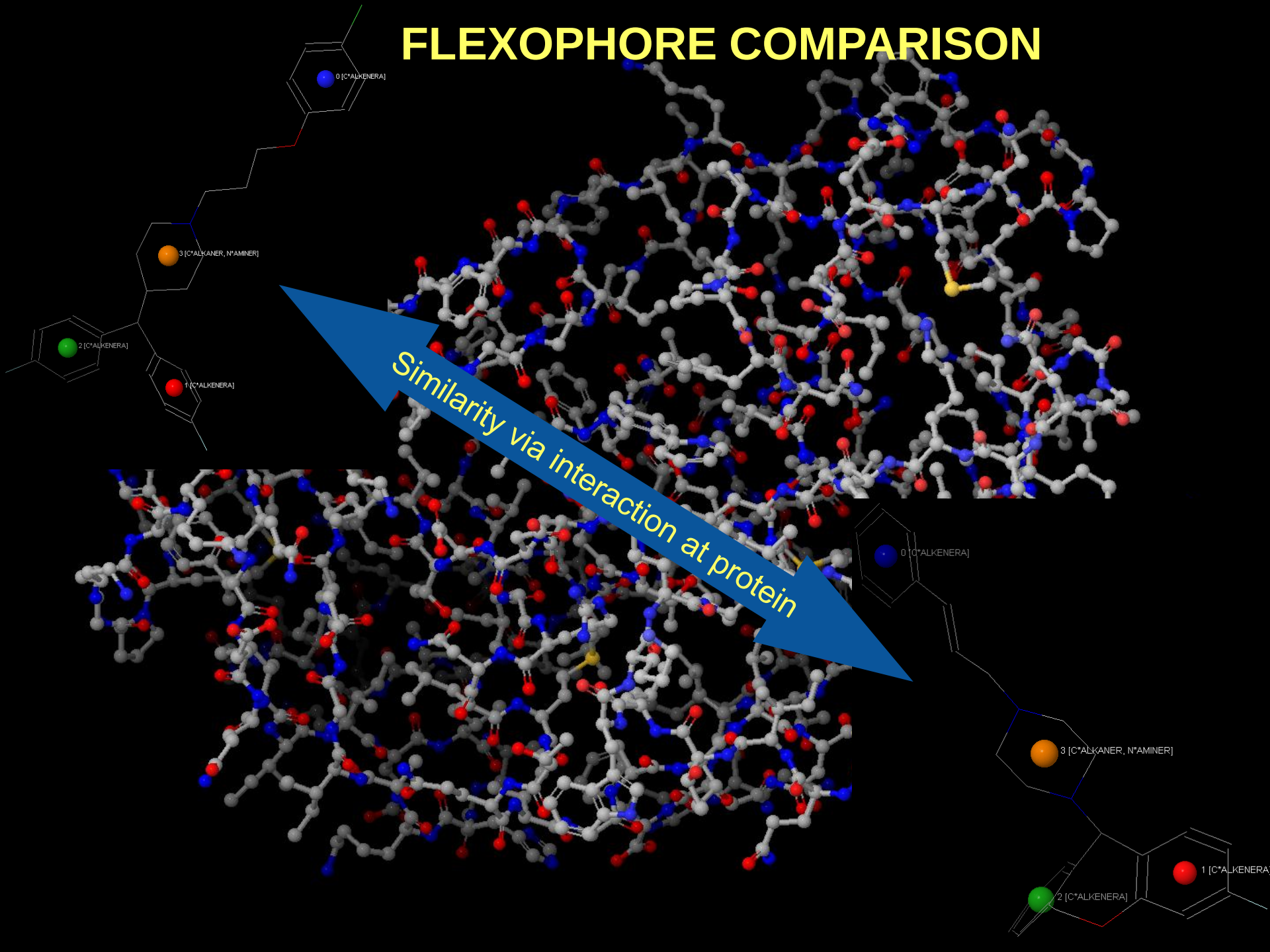


Edges:

Distance histograms



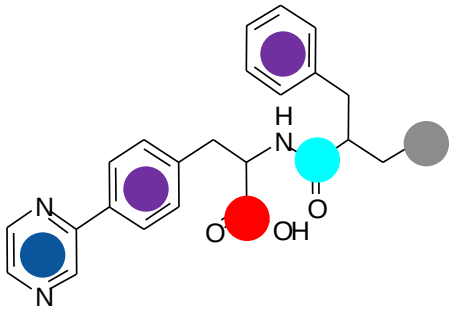
FLEXOPHORE COMPARISON



SIMILARITY METRIC FOR BIOISOSTERS

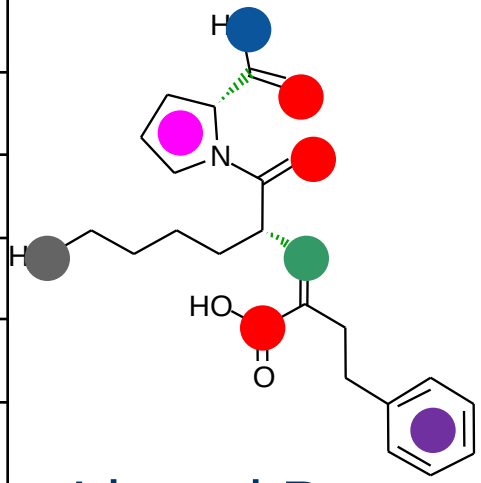
a

Protein-ligand interaction statistics



Ligand A

PP Types	1.	2.	3.	4.	5.	6.
1.C Alkane (prim.)	1.00	0.72	0.83	0.36	0.37	0.30
2. C Alkane (sec.)	0.72	1.00	0.51	0.38	0.40	0.38
3. C Carbonyl	0.83	0.51	1.00	0.41	0.45	0.45
4. O Carbonyl	0.36	0.38	0.41	1.00	0.98	0.88
5. O Ether	0.37	0.40	0.45	0.98	1.00	0.85
6. O Alcohol	0.30	0.38	0.45	0.88	0.85	1.00

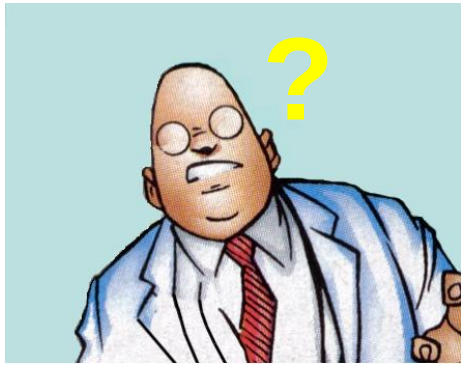


Ligand B

$Sim_{A,B}$ = f(Pharmacophore point pairs bioisosterism)

ACTELION

HOW TO IMPROVE A PHARMACOPHORE MODEL?

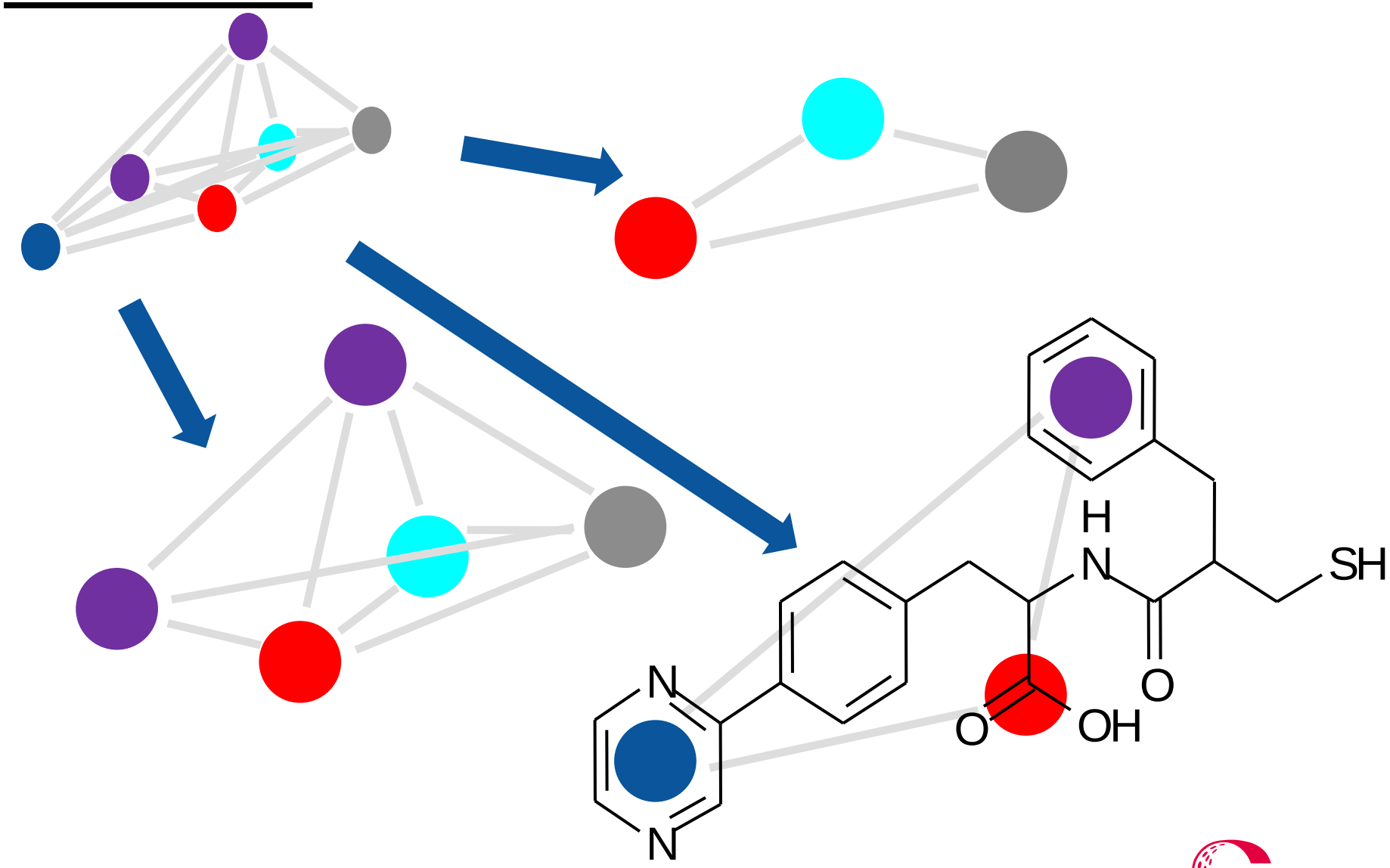


Two changeable components

- Pharmacophore points
  - Substructure based
  - Easy to manipulate on/off
- Distance histograms
  - Edges of complete graph ((n * n) - n) / 2
  - I.e. 7 PPNodes → 21 distance histograms
  - Histograms, not ranges
  - Meaningful manipulation difficult

ACTELION

SUB PHARMACOPHORES



COMBINATORIAL EXPLOSION

How many subpharmacophores?

Possible combinations $\binom{n}{k}$

 # PPoints in query molecule

# PPoints in Sub Pharmacophore	4	6	8	10	12	14
3	4	20	56	120	220	364
5		6	56	252	792	2002
7			8	120	792	3432
9				10	220	2002
11					12	364

 # Possible combinations

ACTELION

HOW TO MEASURE IMPROVEMENT?

Need for validation

DUD dataset

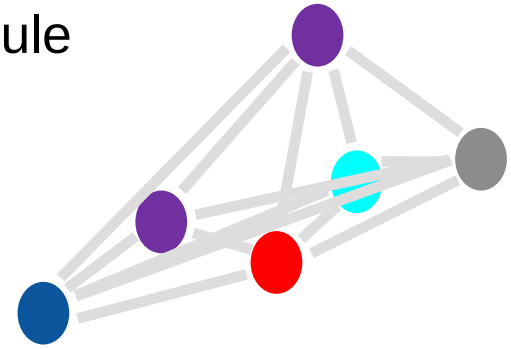
- 39 targets
- 241 data sets
- 10-50 seeds (actives)
- Filled up to 1000 with decoys (assumed inactives)
  - Physico chemical properties like seeds

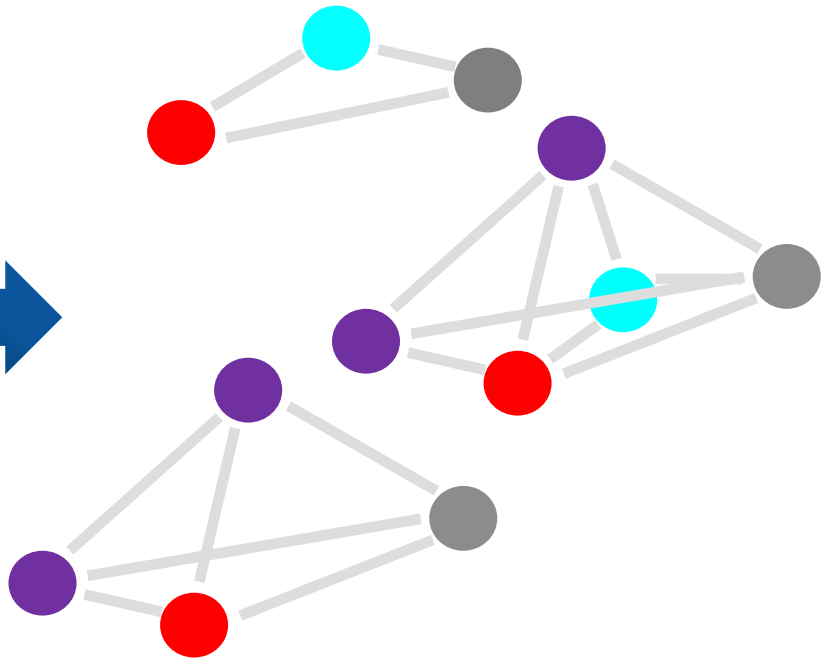
Huang, N.; Shoichet, B. K.; Irwin, J. J., Benchmarking Sets for Molecular Docking. *J. Med. Chem.* ***2006, 49, (23), 6789-6801.S***

ACTELION

TEST SETUP

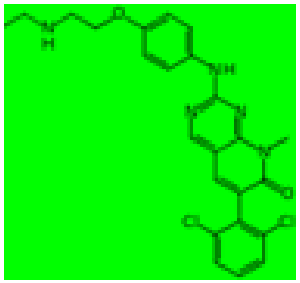
Probe molecule

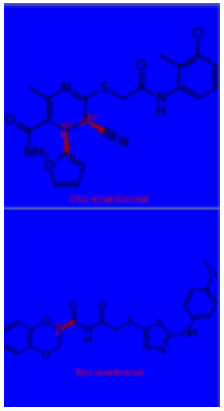




- Seeds (actives)
- Decoys (inactives)

1000





Split

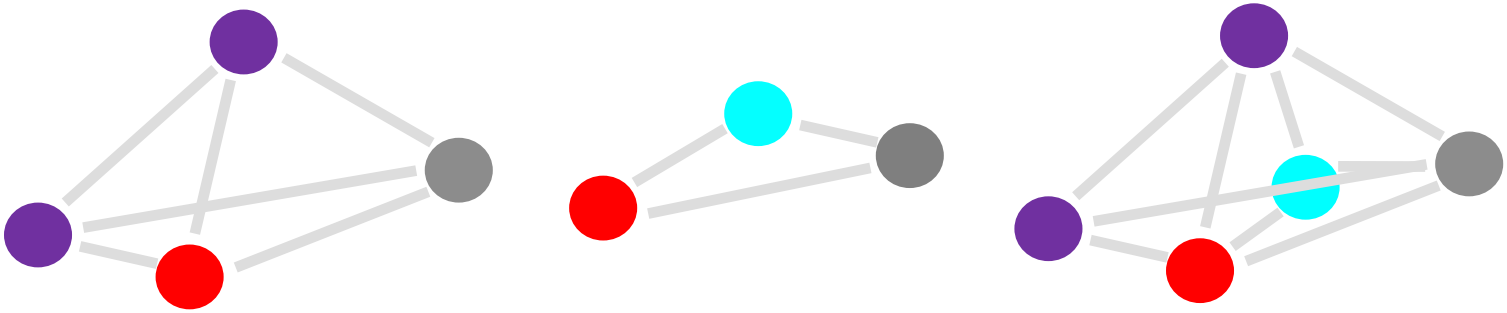
Train (500)

Test (500)

ACTELION

RUNNING A TEST

Probe SubFlexophores



Compare SubFlexophores with train Flexophores

Compare best performing SubFlexophores with test Flexophores

Calculate enrichment

ACTELION

RESULTS ALL TESTS

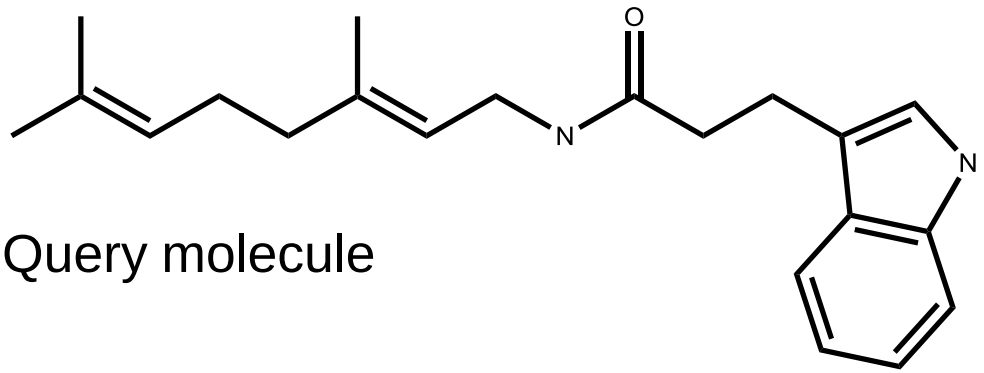
- Data sets 241
- Probe SubFlexophores: 20-200 per probe molecule
- Compare to train
- Top SubFlexopores 1-10 per probe molecule
- Compare to test

Summary

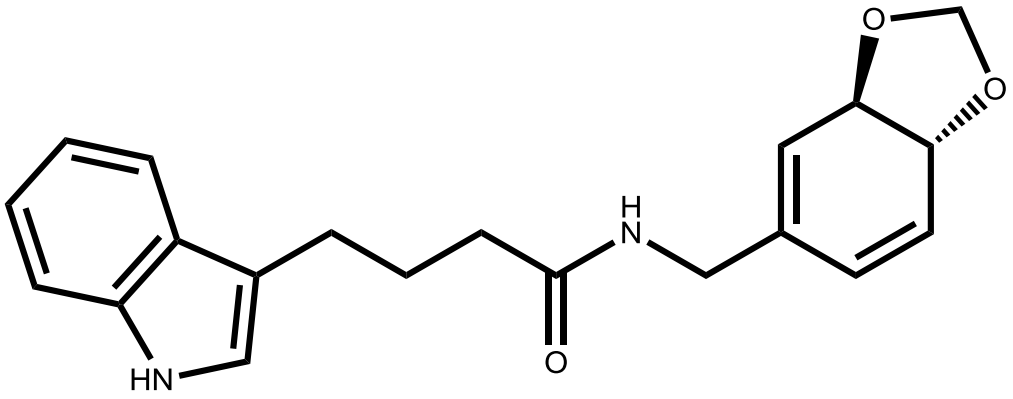
Enrichment SubFlexophore test	**127**
Enrichment full Flexophore test	**86**
Best SubFlexophore better than full Flexophore	**91**

ACTELION

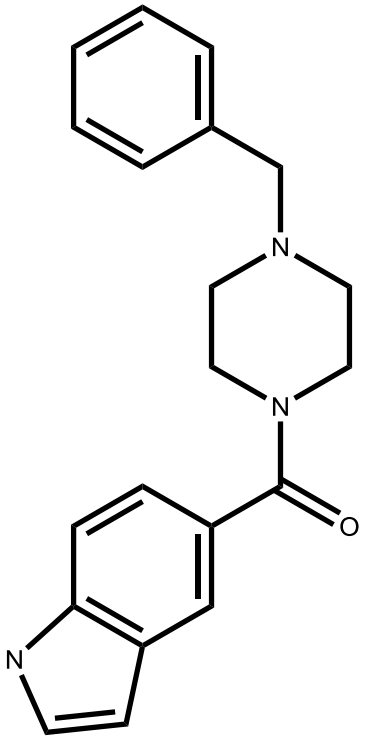
SUCCESSFUL SCAFFOLD HOPPING



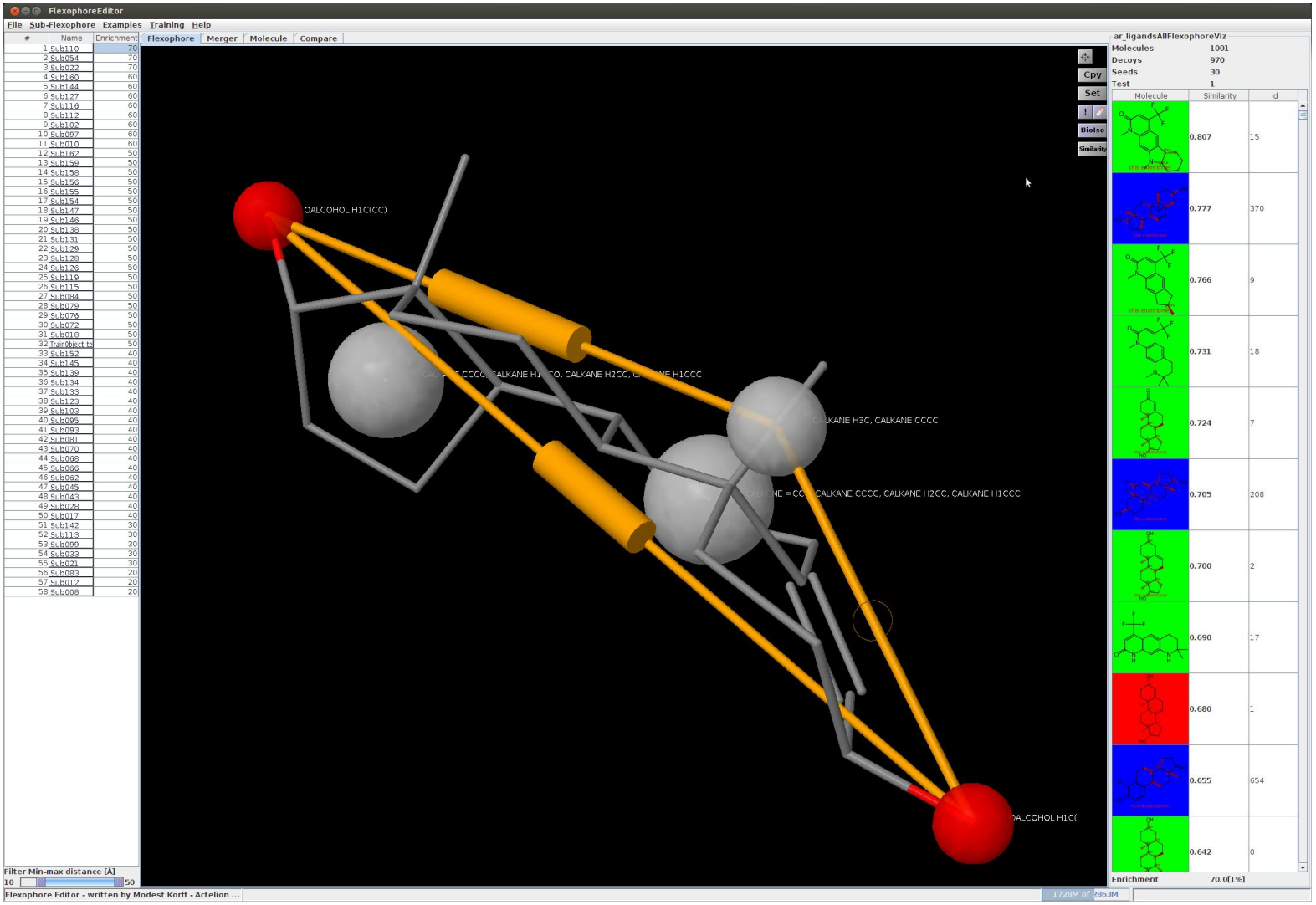
Query molecule



Found seeds



FLEXOPHORE EDITOR: PLAYGROUND FOR SCIENTISTS

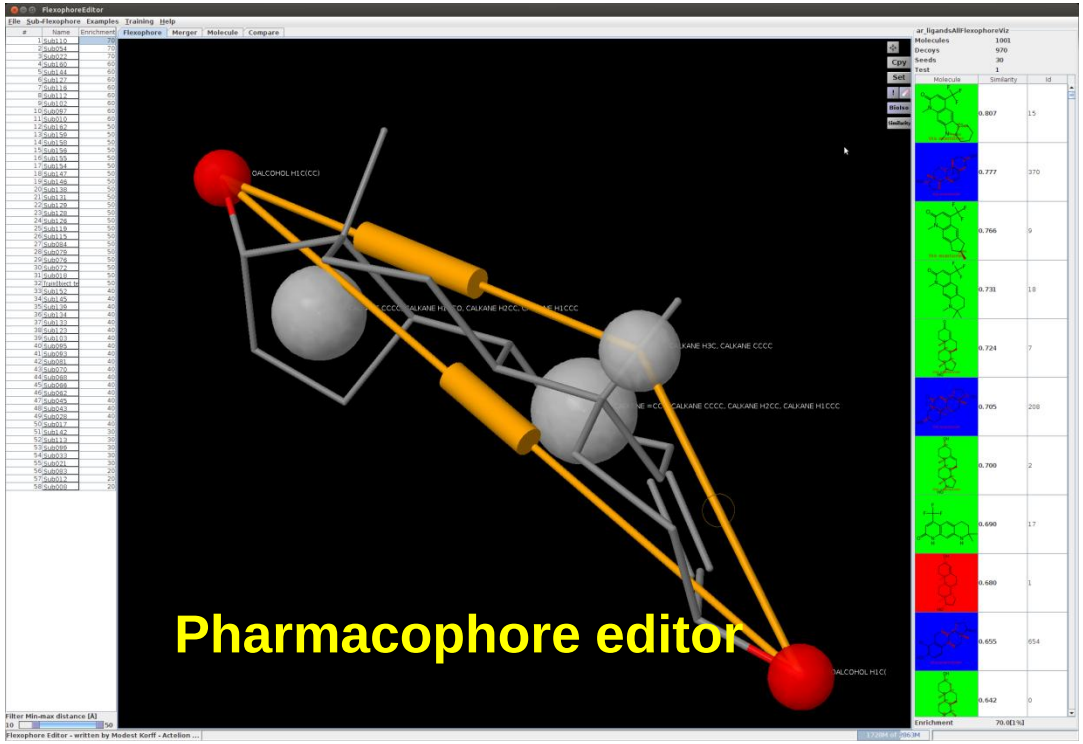


IMPLEMENTATION DETAILS

- In-house development
- Java programming language
- Jmol for pharmacophore visualization
- Multi thread similarity calculations

Training data

SubFlexophores



ACTELION

CONCLUSIONS

- Impossible to outsmart combinatorial explosion of possible pharmacophore models
- Medicinal chemist may construct test and validation data
- Combination of brute force computer power plus medicinal chemist know how resulted in improved pharmacophore models

ACTELION

THANK YOU!