Enumerating, Mapping and Scoring Chemical Space

Jean-Louis Reymond

FBLD 2010, Philadelphia

11 October 2010

10^{60} (?)

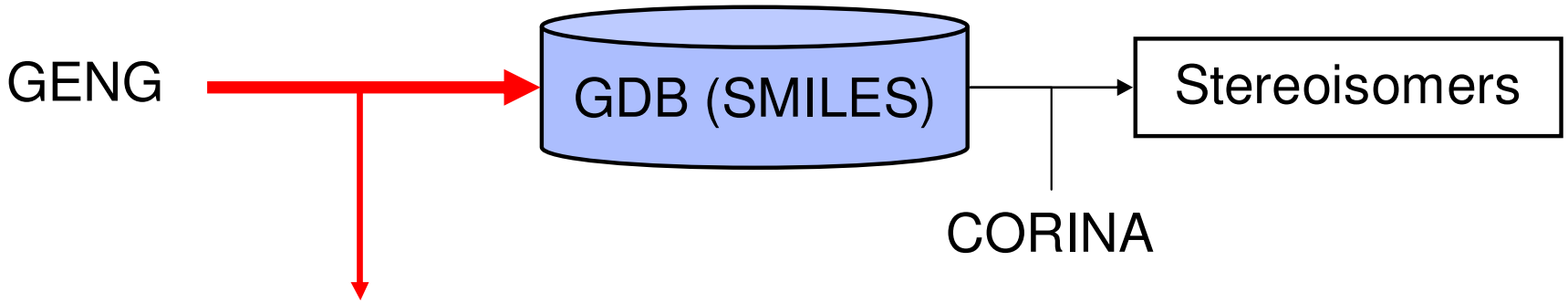
1. GDB, CST

2. MQN

3. Scoring

J.-L. Reymond et al.. Chemical space as a source for new drugs. *Med. Chem. Commun.* **2010**, *1*, 30-38

GDB Assembly



I. Select hydrocarbon graphs (**ring strain**, topology)

II. C-C to C=C, C#C following valency rules, no allenes, no DB in bridgehead or 3- and 4-rings, no TB in <9-rings, **no DB torsion**

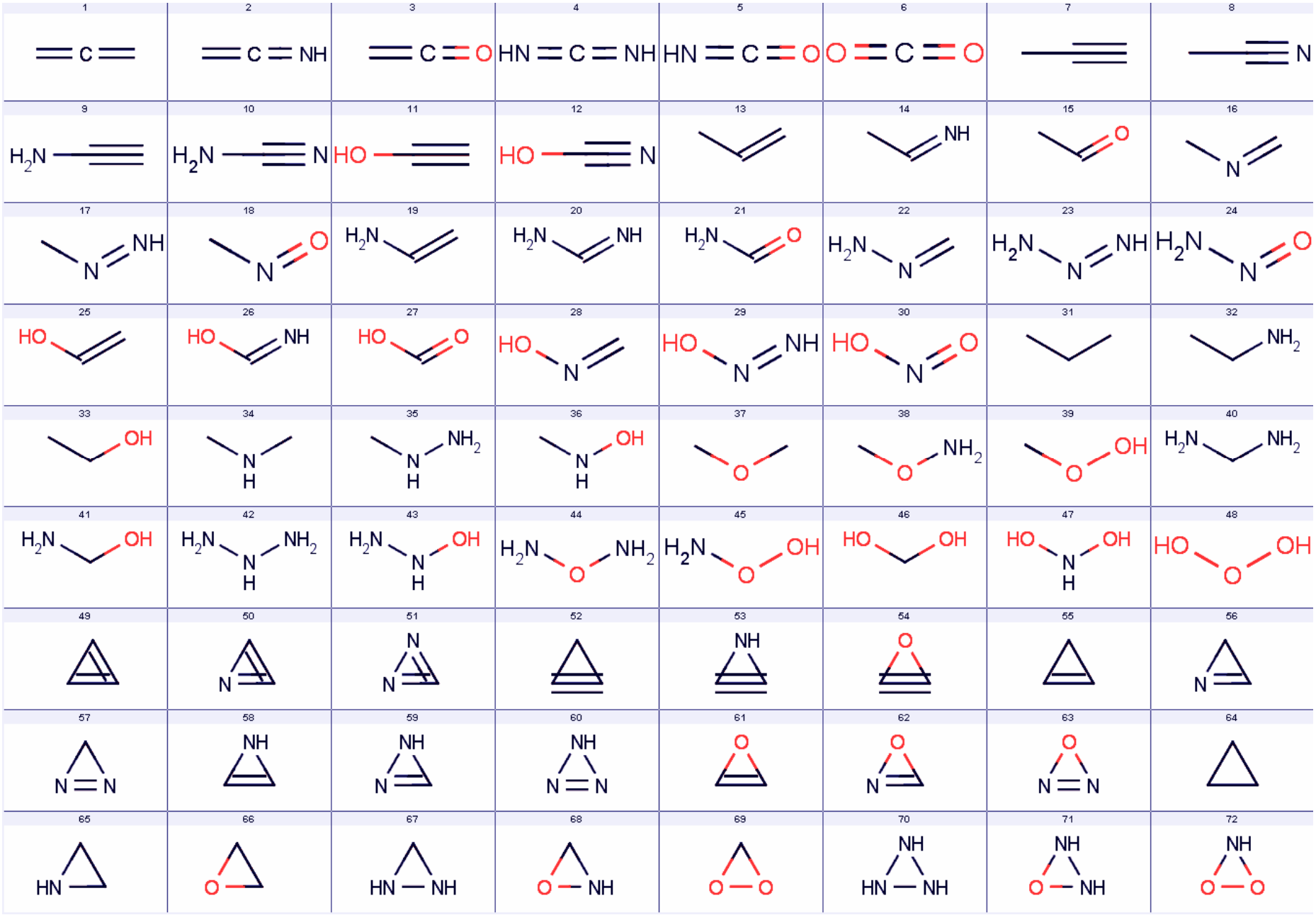
III. C to N or O following valency rules

IV. Filter bad functional groups, select tautomers

V. Post-processing: Halogens, S, etc.

T. Fink et al. *Angew. Chem. Int. Ed.* **2005**, *44*, 1504-1508, *J. Chem. Inf. Model.* **2007**, *47*, 342-353 (GDB-11)
L. C. Blum, J.-L. Reymond, *J. Am. Chem. Soc.* **2009**, *131*, 8732-3 (GDB-13)

DMU (valency rules only)

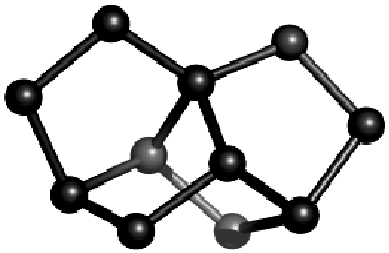


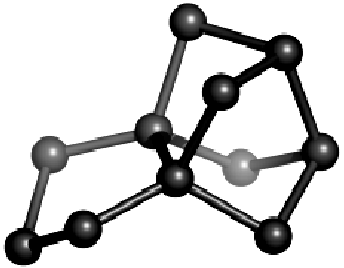
GDB-11 (CNOF)

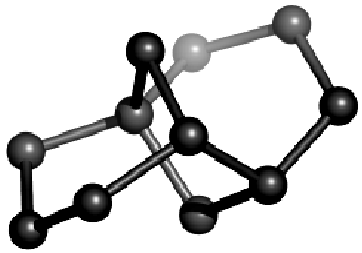
Table 2. Overview of the structure generation process.

Nodes	Graphs[a]	Generated[b]	Accepted[c]	Unique Tautomers (GDB)[d]	All Tautomers	Stereoisomers[e]
1	1	4	4	4	4	4
2	1	10	9	9	9	9
3	2	52	20	20	21	20
4	4	332	80	80	88	87
5	8	2'294	357	352	397	469
6	20	18'066	1'906	1'850	2'135	2'911
7	57	154'542	10'953	10'568	12'438	19'904
8	194	1'445'073	69'563	66'706	79'899	153'601
9	705	14'213'741	464'402	444'313	540'002	1'258'963
10	2'822	146'004'340	3'259'036	3'114'041	3'827'907	10'898'065
11	11'912	1'558'491'448	23'875'101	22'796'628	28'240'425	98'645'474
Total	**15'726**	**1'720'329'902**	**27'681'431**	**26'434'571**	**32'703'325**	**110'979'507**

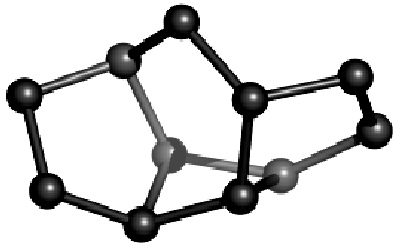
99.8 % are unknown

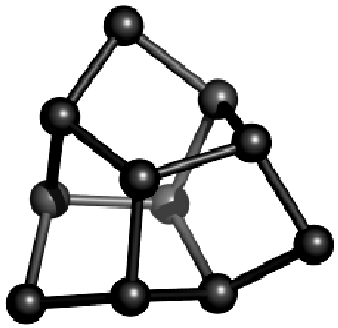






	Number of 4-membered rings			
Number of 3-membered rings	0	1	2	**Total**
0	124 [3]	189 [60]	103 [67]	**416 [130]**
1	225 [50]	238 [177]	20 [19]	**483 [246]**
2	201 [88]	55 [48]	-	**256 [136]**
3	53 [26]	-	-	**53 [26]**
Total	**603 [167]**	**482 [285]**	**123 [86]**	**1'208 [538]**

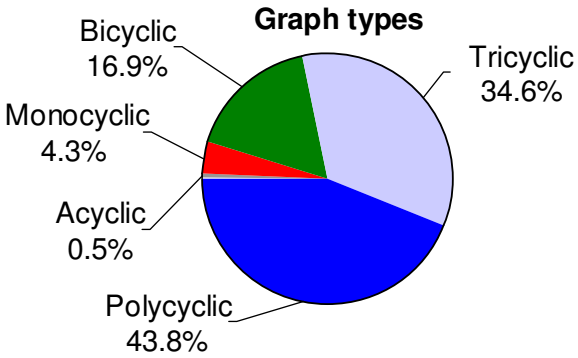


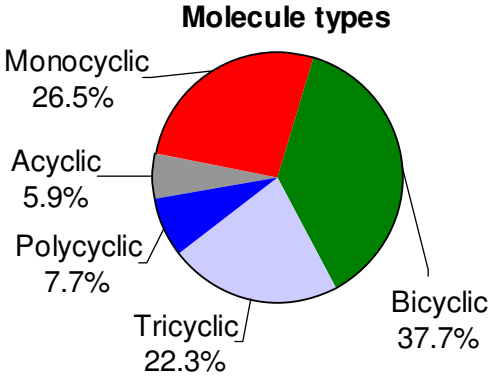


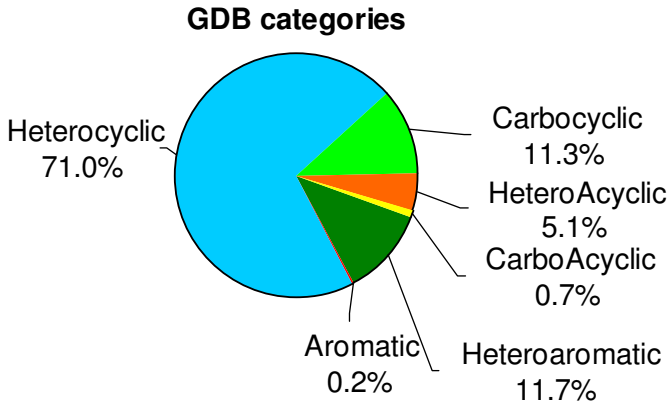
GDB-13 (CNOSCl, max. heteroatom ratio)

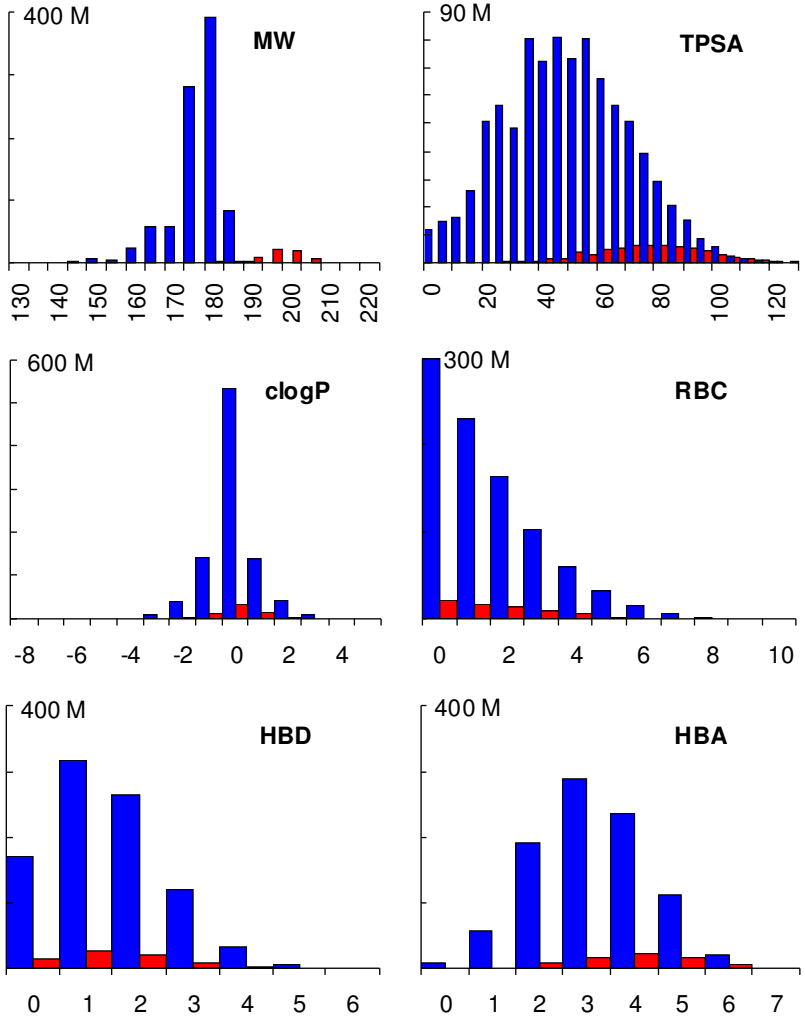
nodes[a]	graphs[b]	GDB[c]	Cl/S[d]
1	1	1	0
2	1	3	0
3	2	12	0
4	4	43	0
5	8	155	3
6	20	934	19
7	57	5 726	315
8	194	37 151	2 438
9	706	255 542	17 056
10	2 831	1 784 626	130 465
11	12 011	12 961 686	938 704
12	53 789	99 821 343	7 240 108
13	250 268	795 244 451	59 027 533
Total	**319 892**	**910 111 673**	**67 356 641**

L. C. Blum, J.-L. Reymond, *J. Am. Chem. Soc.* **2009**, *131*, 8732-3


Graph types
Bicyclic
16.9%
Tricyclic
34.6%
Monocyclic
4.3%
Acyclic
0.5%
Polycyclic
43.8%


Molecule types
Monocyclic
26.5%
Acyclic
5.9%
Polycyclic
7.7%
Tricyclic
22.3%
Bicyclic
37.7%


GDB categories
Heterocyclic
71.0%
Carbocyclic
11.3%
HeteroAcyclic
5.1%
CarboAcyclic
0.7%
Aromatic
0.2%
Heteroaromatic
11.7%


400 M
MW
130 140 150 160 170 180 190 200 210 220
90 M
TPSA
0 20 40 60 80 100 120
600 M
clogP
-8 -6 -4 -2 0 2 4
300 M
RBC
0 2 4 6 8 10
400 M
HBD
0 1 2 3 4 5 6
400 M
HBA
0 1 2 3 4 5 6 7

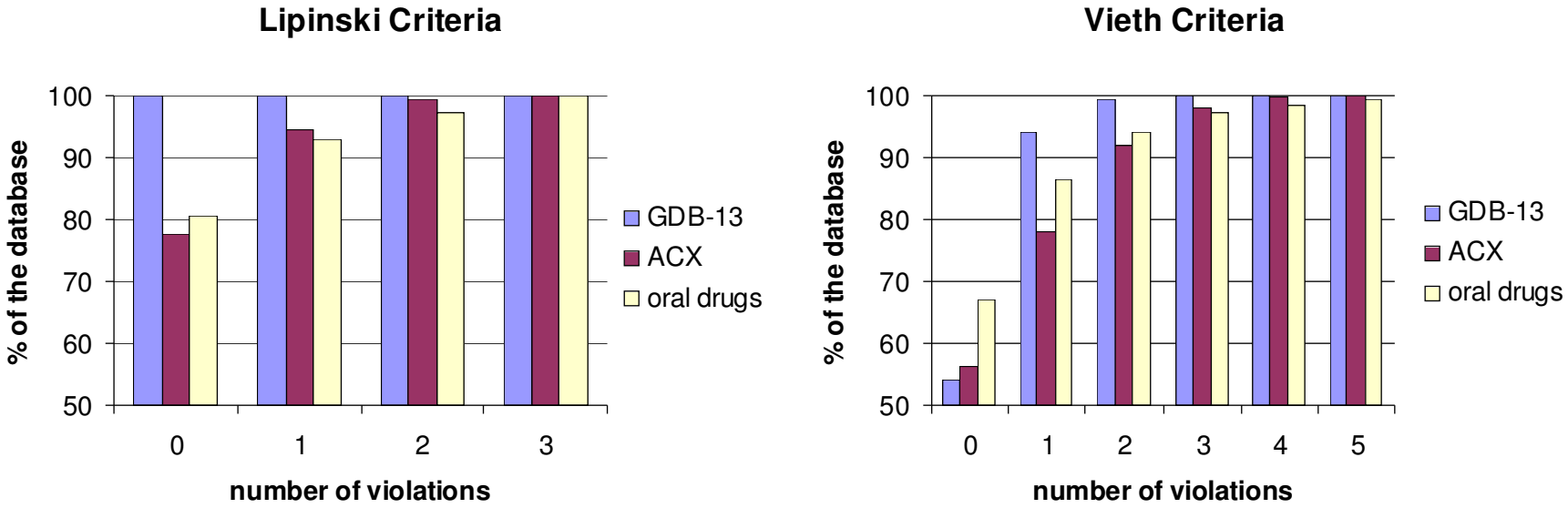
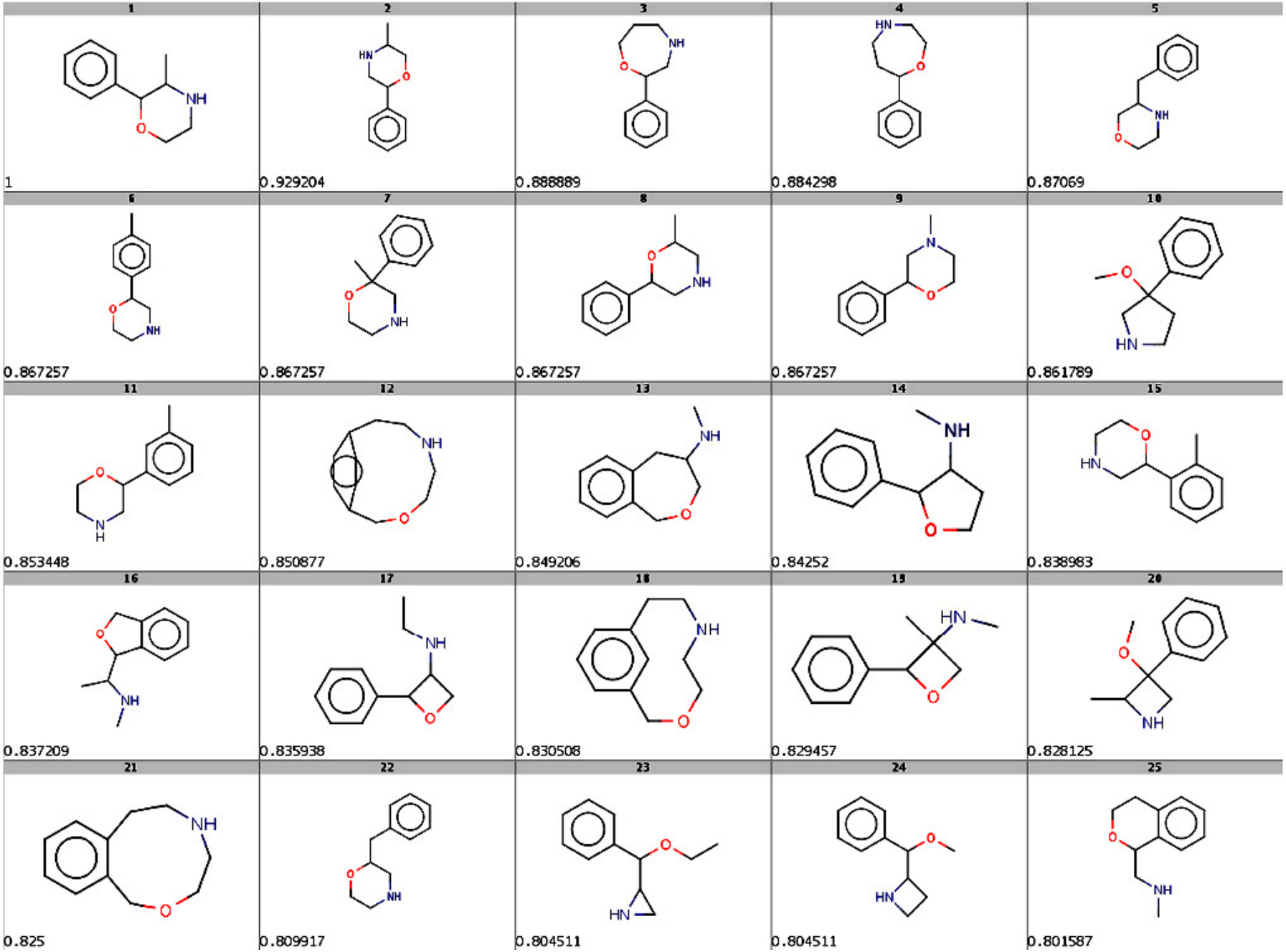


Figure S4 A and B. Bar-plot of the percentage of compounds from GDB-13 (blue bars), 216 188 commercial compounds from the ACX directory (purple bars) and 1193 oral drugs listed in ref. 10b (yellow bars) that pass the Lipinski or Vieth criteria as a function of the number of violations allowed. The Lipinski's rule of 5 states that 90% of drug candidates (entering phase II) have no more than one violation in the four criteria.

Drugs and isomers in GDB-13

Name	Formula	Same Formula	T_{SF} AVG	T_{SF} > 0.7
Aspirin	$C_9H_8O_4$	804 153	0.23	178
Benzocaine	$C_9H_{11}NO_2$	1 846 579	0.24	74
L-Tyrosine	$C_9H_{11}NO_3$	9 276 529	0.46	24 952
Levetiracetam	$C_8H_{14}N_2O_2$	2 154 955	0.28	35
Memantine	$C_{12}H_{21}N$	2 872 586	0.31	10 912
Menadione	$C_{11}H_8O_2$	233 715	0.44	112 186
Metaraminol	$C_9H_{13}NO_2$	2 920 516	0.26	30
Mexiletine	$C_{11}H_{17}NO$	18 371 393	0.25	119
Propofol	$C_{12}H_{18}O$	5 263 227	0.25	240
Rasagiline	$C_{12}H_{13}N$	1 323 525	0.13	411
Rimantadine	$C_{12}H_{21}N$	2 872 586	0.26	173

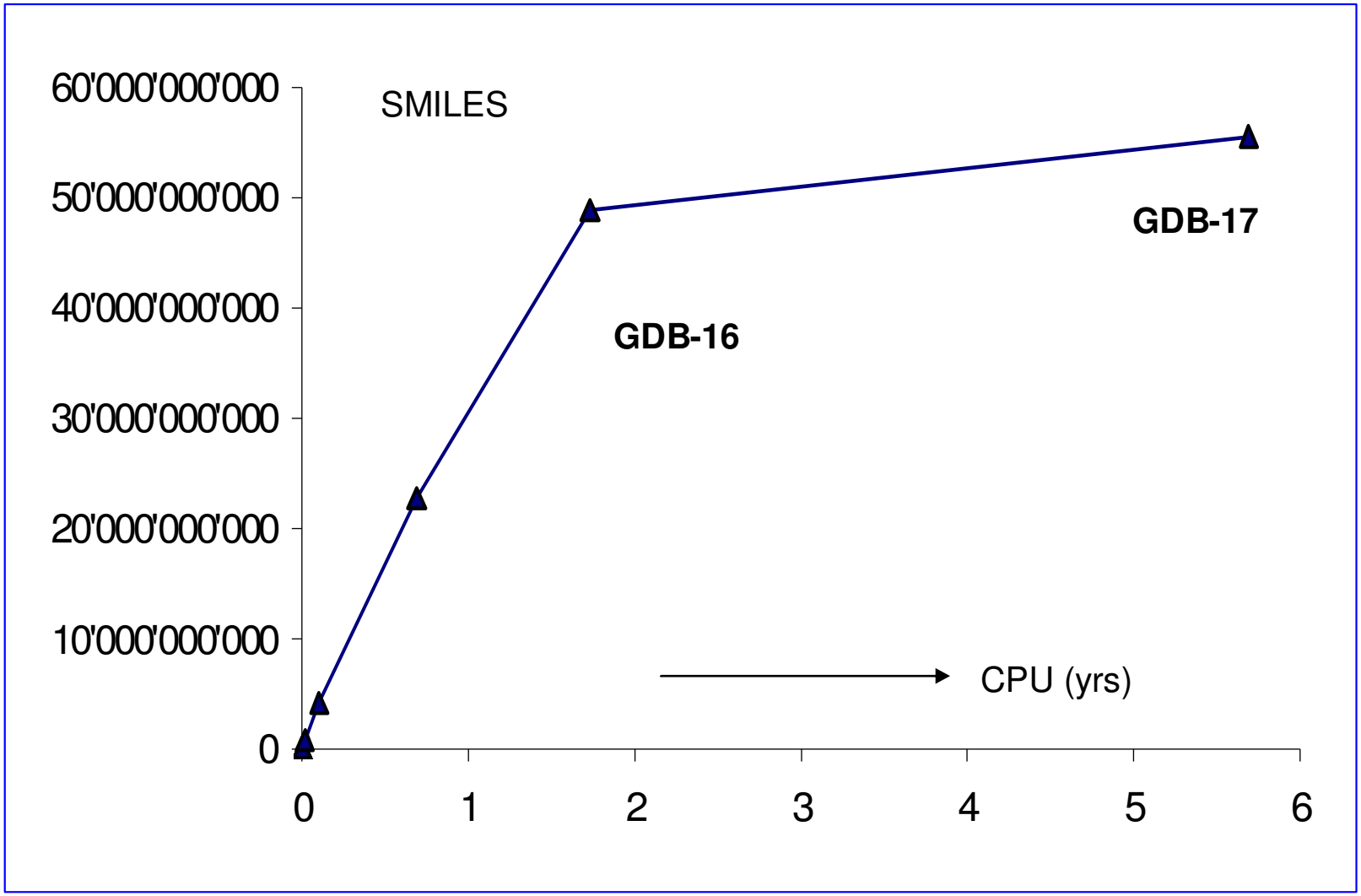


Phenmetrazine isomers

GDB-17 (CNOSHal + gradual filters)

HAC	SMILES	Database	Total	Filters
1	3			FG+strain+torsion
2	6			
3	19			
4	23			
5	268			
6	1'320			
7	7'488			
8	46'608			
9	309'243			
10	2'175'788			
11	16'189'774	**GDB11**	18'730'540	
12	96'915'932		115'646'472	no atom in 2 small rings
13	794'711'725	**GDB13**	910'358'197	
14	4'141'604'653		5'051'962'850	no bridgehead in three rings
15	*22'721'659'277*		*27'773'622'127*	
16	*48'849'650'795*		*76'623'272'922*	one small ring
17	*55'516'921'640*	**GDB17**	*132'140'194'562*	no small rings, no non-arom. C=C

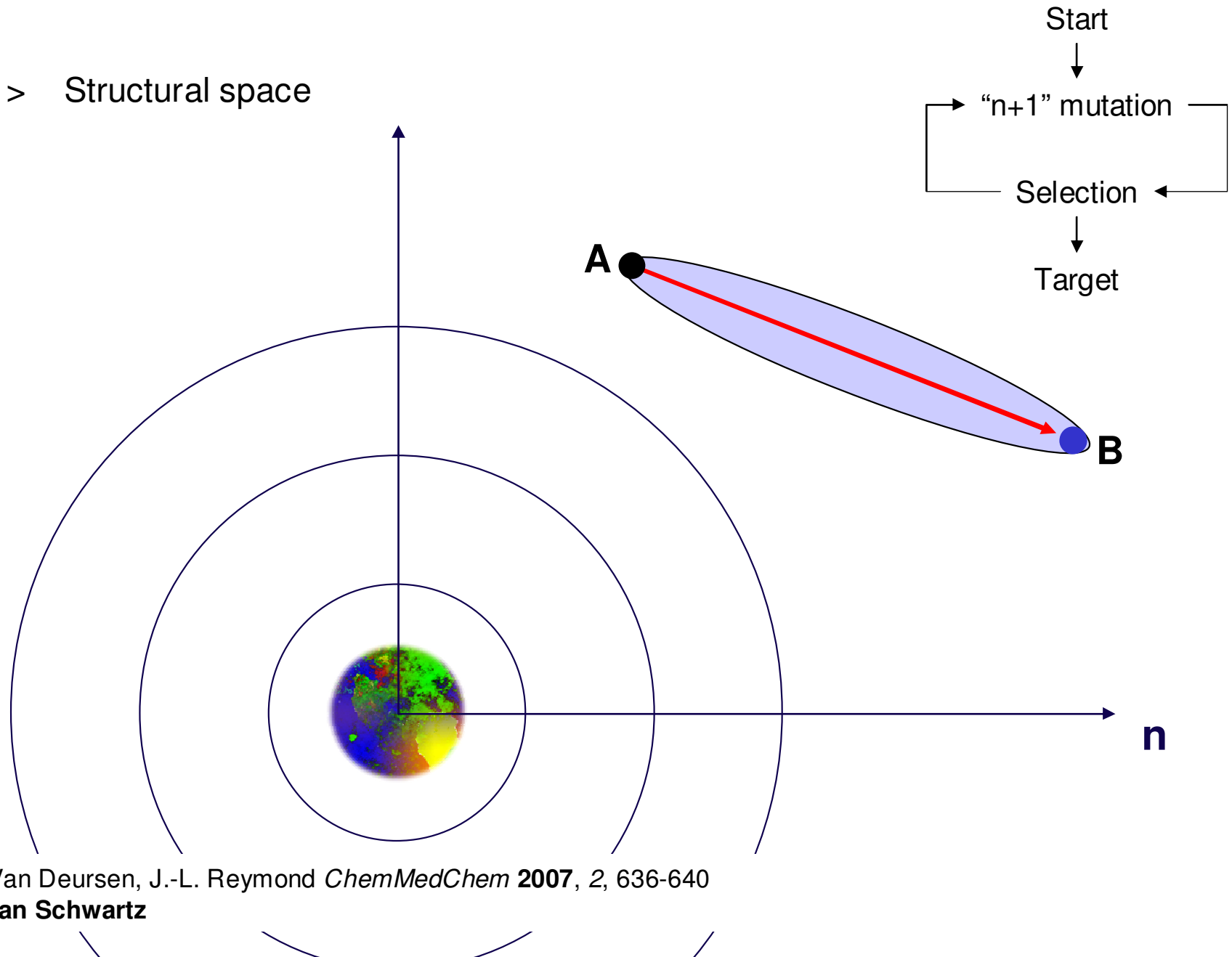
Processing times



Lars Ruddigkeit, Ruud van Deursen, Lorenz Blum, **ongoing work**

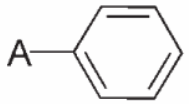
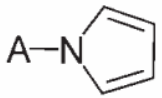
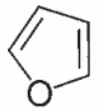
Chemical Space Travel

> Structural space



R. Van Deursen, J.-L. Reymond *ChemMedChem* **2007**, *2*, 636-640

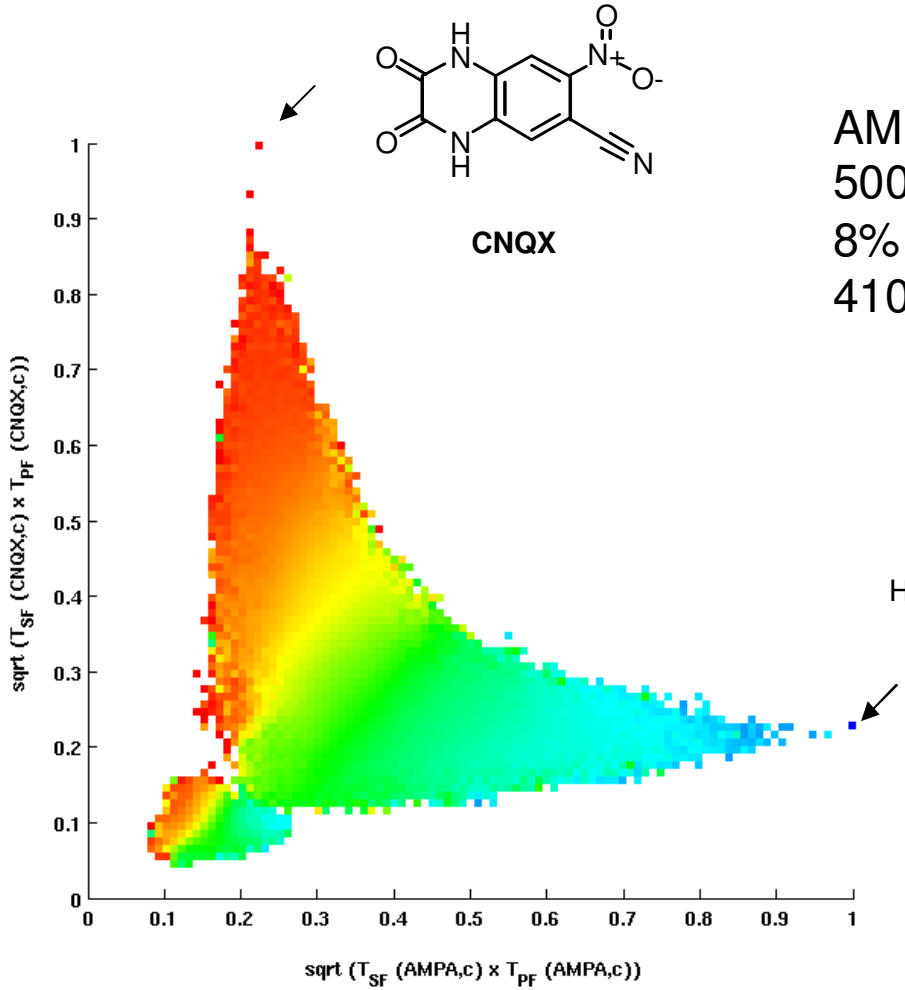
Julian Schwartz

Nearest neighbour mutations[a]	
Atom type exchange[b,c]	Replaces any atom by another atom type
Atom inversion[c]	Inverts two neighbouring atoms
Atom removal[c]	Primary: A–X → A Secondary: A–X-A → A–A Tertiary: XA_3 → A–A–A (max. 6 combinations if 3 different A's) $A_2CH–CHA_2$ or $A_2C=CA_2$ → CA_4 Quaternary: XA_4 → A–A–A–A or $A(A)_3$ (max. 16 combinations if 4 different A's)
Atom addition[b,c]	On terminal atoms: A → A–X In any bond: A–A → A–X–A In chains: A–A–A → XA_3; A–A–A–A → XA_4 Quaternary centres: CA_4 → $A_2CH–CHA_2$ and $A_2C=CA_2$ (max. 6 combinations if 4 different A's)
Bond saturation[c]	Breaks a cyclic σ- or any π-bond
Bond unsaturation	Makes a cyclic σ- or π-bond
Bond rearrangement[c]	Breaks a σ- or π-bond and inserts it anywhere else in the molecule
Non-nearest neighbour mutations	
	A-CH_3 → 
Aromatic ring addition[c,d]	A-NH_2 → 
	H_2O → 

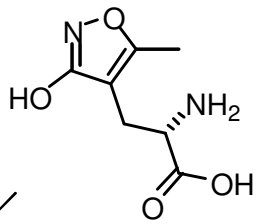
Cross-Trajectories

From: \ To:	Cubane	Aspirine	VX	Adenosine	Sucrose	Penicillin G	Strychnine	Colchicine	Tetracycline	Vitamin K
Cubane	-	10	18	23 (1)	19	18 (1)	18 (1)	22 (1)	24 (1)	26 (1)
Aspirine	10*	-	14	21	15	16	24	22	22	33
VX	13	17 (1)	-	31 (1)	18	15 (1)	21 (1)	20 (2)	24* (1)	25* (1)
Adenosine	17*	27	18*	-	14	15	24	23	27*	29
Sucrose	18*	22 (1)	22*	29 (1)	-	25	26 (1)	31 (1)	25 (1)	25 (1)
Penicillin G	19*	13*	14*	23	19*	-	20	19*	21*	29
Strychnine	21*	17*	20	26	22	16*	-	30*	17*	22*
Colchicine	27	22*	21	26	18	22	23	-	22*	21*
Tetracycline	28*	20	25*	49	19	19*	16	28	-	17
Vitamin K	30*	24*	30*	34*	28*	27*	19*	30*	22*	-

AMPA ↔ CNQX



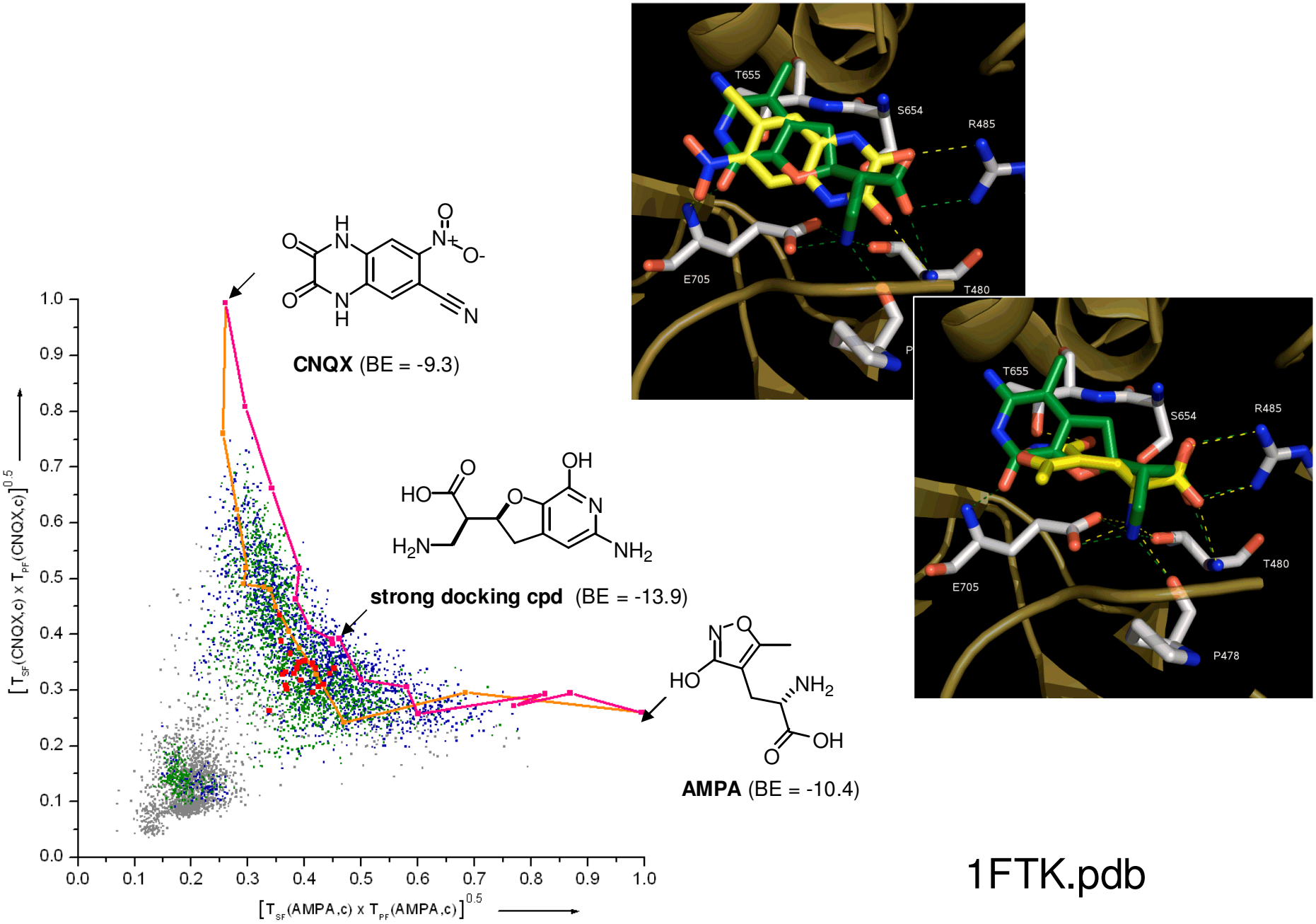
CNQX



AMPA

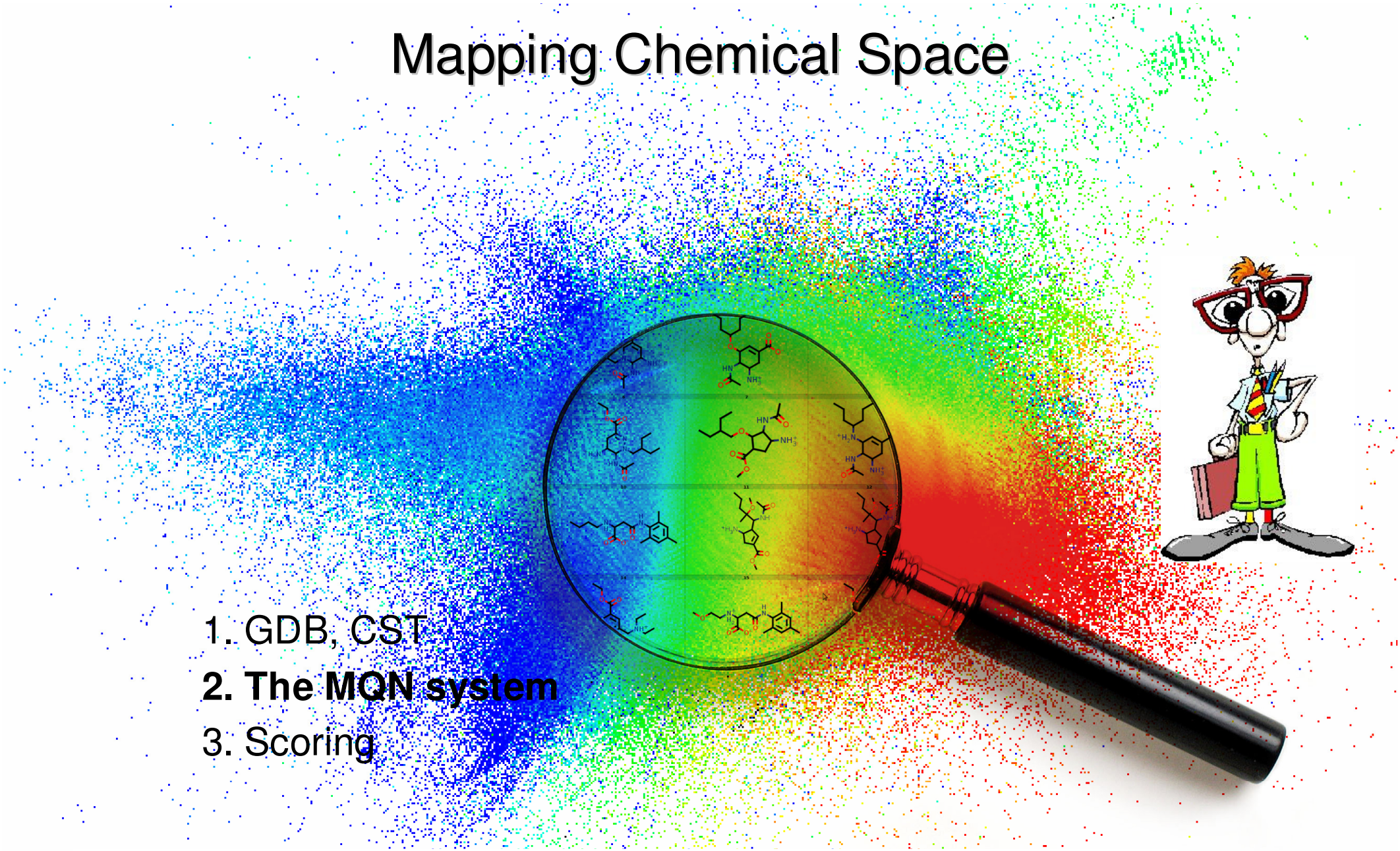
AMPA ↔ CNQX
500 runs, 559,658 cpds
8% from 6.9 million generated
410 cpds (0.1%) also in ZINC

0 **0.50** **1.0**


T655
S654
R485
E705
T480
P478
CNQX (BE = -9.3)
strong docking cpd (BE = -13.9)
AMPA (BE = -10.4)
[T_SF(CNQX,c) x T_PF(CNQX,c)]^0.5
[T_SF(AMPA,c) x T_PF(AMPA,c)]^0.5
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

1FTK.pdb

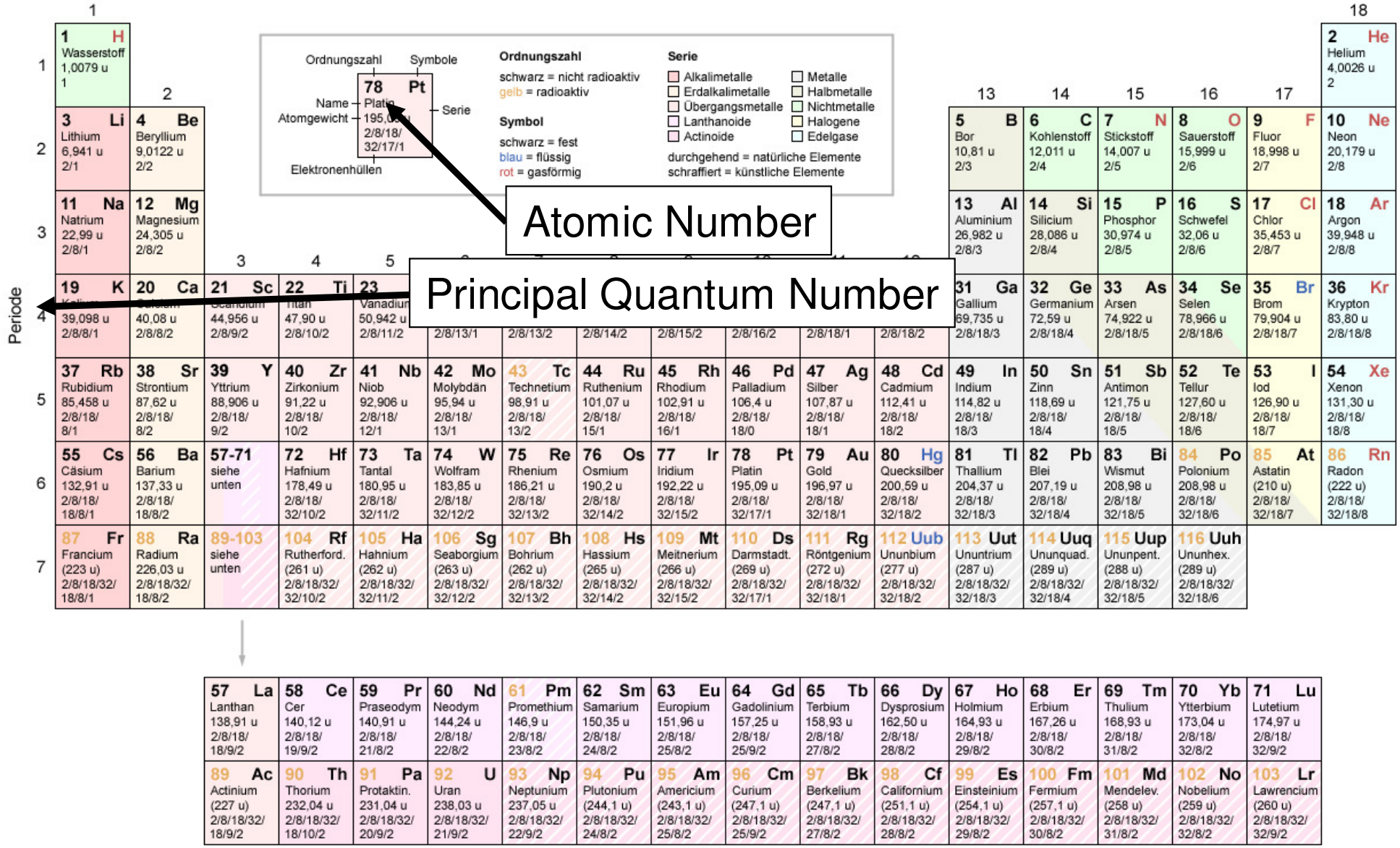
Mapping Chemical Space



1. GDB, CST
2. **The MQN system**
3. Scoring

K. T. Nguyen, L. C. Blum, R. van Deursen, J.-L. Reymond *ChemMedChem* **2009,** *4*, 1803-5
R. van Deursen, L. C. Blum, J.-L. Reymond, *J. Chem. Inf. Model.* **2010**, in press

The Periodic System of the Elements



Molecular Quantum Numbers

- integer value descriptors of atoms, bonds, polarity, topology
- immediate meaning, the values can be determined "by hand"
- 42 MQNs define a 42-dimensional "Chemical Space"
- PCA to vizualize most of the diversity in 2D or 3D (Map)

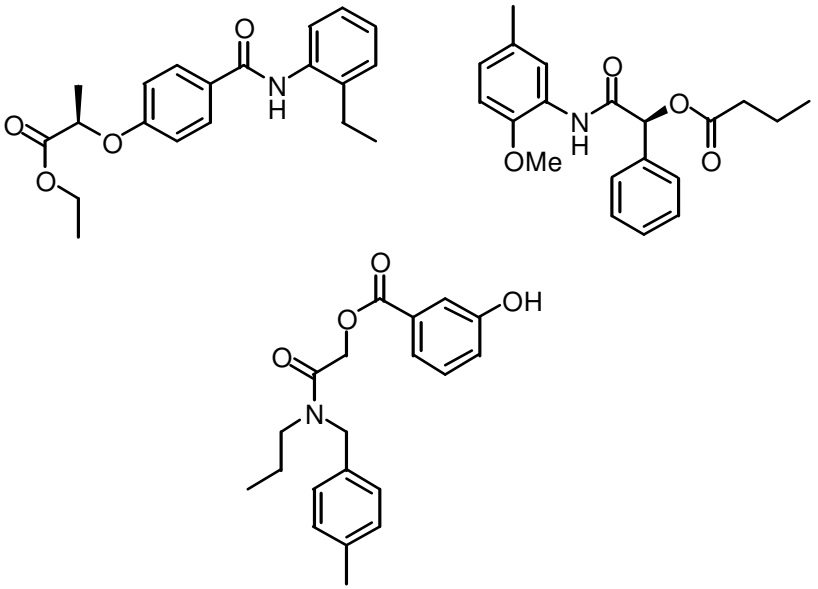
Category 1: Atoms
1. c (carbon)
2. f (fluorine)
3. cl (chlorine)
4. br (bromine)
5. i (iodine)
6. s (sulfur)
7. p (phosphorous)
8. an (acyclic nitrogen)
9. cn (cyclic nitrogen)
10. ao (acyclic oxygen)
11. co (cyclic oxygen)
12. thac (all non-H)

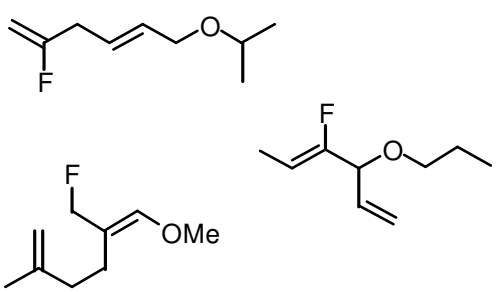
Category 2: Bonds
13. asb (acyclic single bonds)
14. adb (acyclic double bonds)
15. atb (acyclic triple bonds)
16. csb (cyclic single bonds)
17. cdb (cyclic double bonds)
18. ctb (cyclic triple bonds)
19. rbc (rotatable bonds)

Category 3: Polarity
20. hbam (H-bond acceptor sites)
21. hba (H-bond acceptor atoms)
22. hbdm (H-bond donor sites)
23. hbd (H-bond donor atoms)
24. negc (negative charges)
25. posc (positive charges)

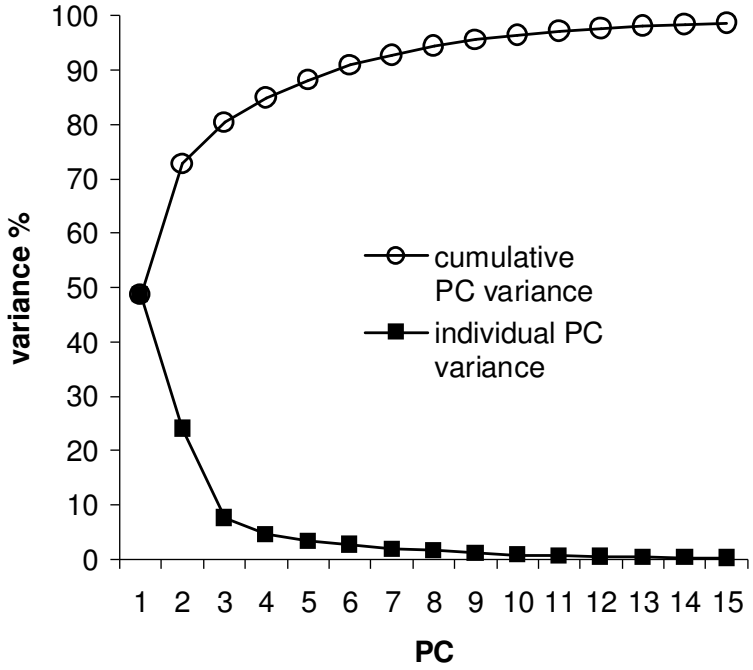
Category 4: Topology
26. asv (acyclic single valent nodes)
27. adv (acyclic divalent nodes)
28. atv (acyclic trivalent nodes)
29. aqv (acyclic tetravalent nodes)
30. cdv (cyclic divalent nodes)
31. ctv (cyclic trivalent nodes)
32. cqv (cyclic tetravalent nodes)
33. r3 (3-membered rings)
34. r4 (4-membered rings)
35. r5 (5-membered rings)
36. r6 (6-membered rings)
37. r7 (7-membered rings)
38. r8 (8-membered rings)
39. r9 (9-membered rings)
40. rg10 (≥10-membered rings)
41. afrc (nodes in ≥ 2 rings)
42. bfrc (edges in ≥ 2 rings)

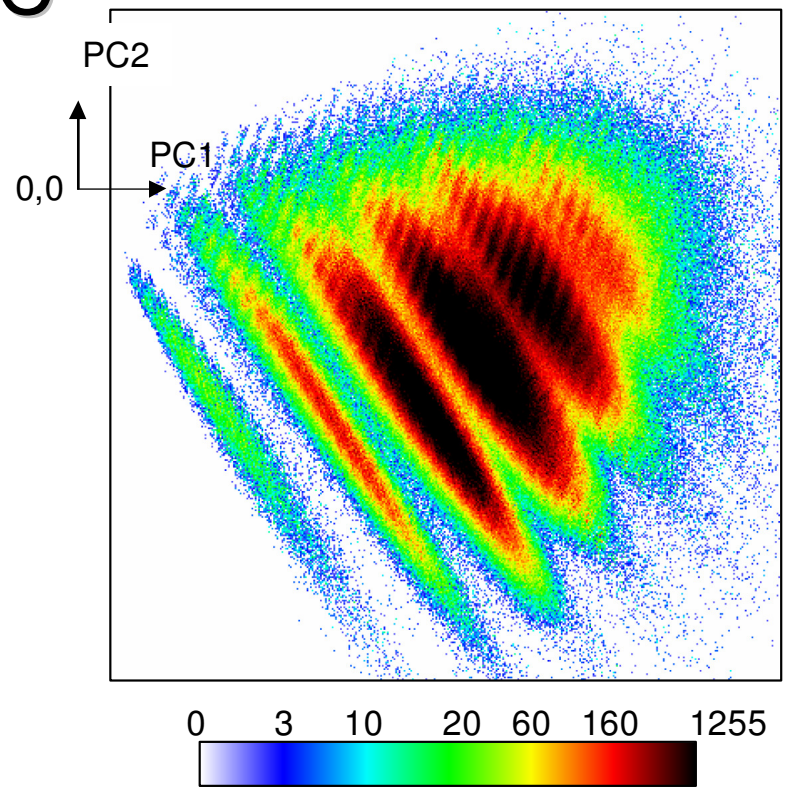
	ZINC	GDB-11
no. of cpds	8 436 272	26 434 567
no. of MQN-bins	3 654 836	2 859 938
no. of single occupied MQN-bins	1 832 566	660 851
no. of cpds in most occupied MQN-bin	300	1 982
no. of shared MQN-bins	13 769	13 769
no. of cpds in shared MQN-bins	30 779	254 604

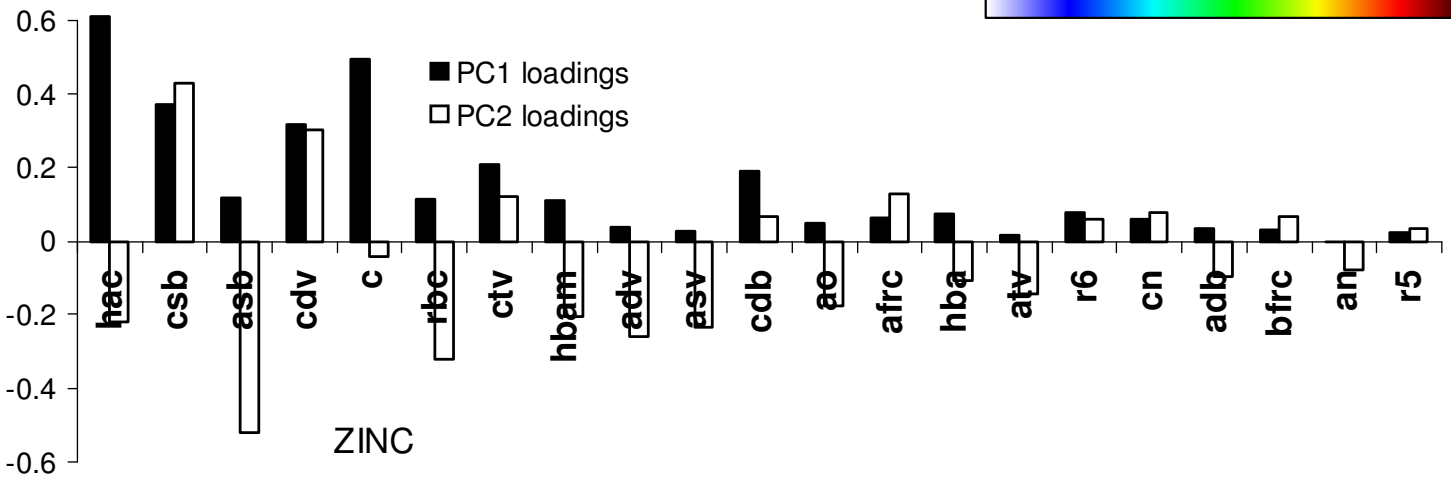


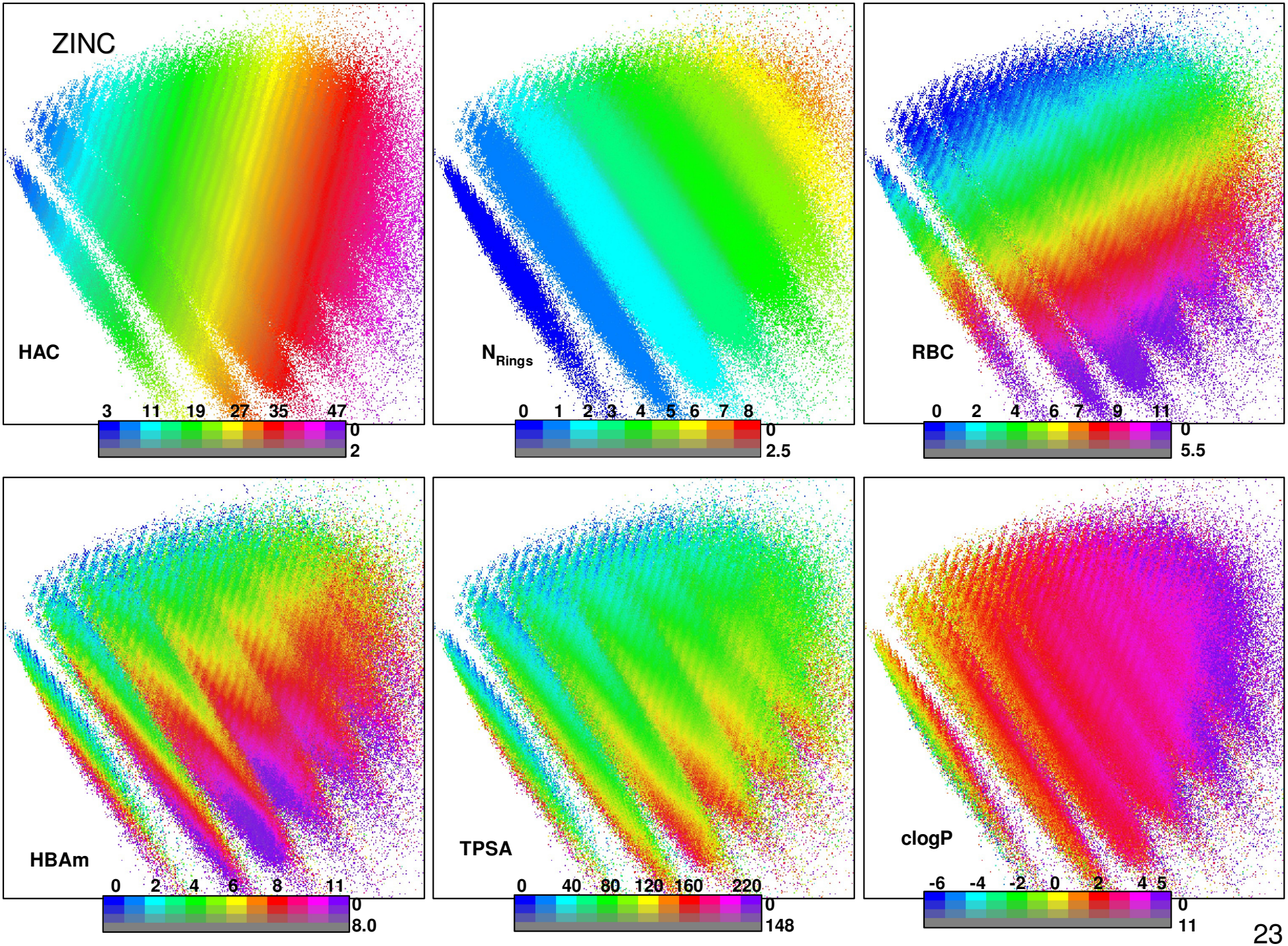


ZINC


variance %
100
90
80
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
PC
cumulative PC variance
individual PC variance

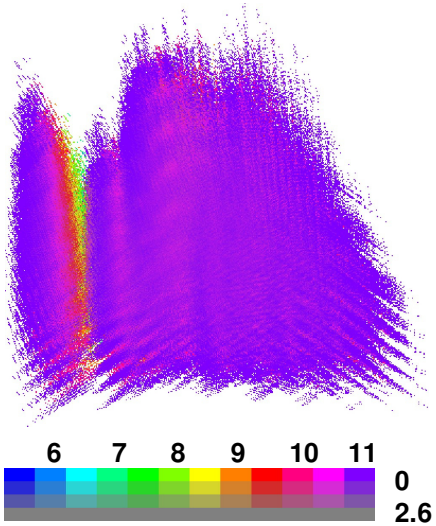

PC2
PC1
0,0
0 3 10 20 60 160 1255


0.6
0.4
0.2
0
-0.2
-0.4
-0.6
PC1 loadings
PC2 loadings
hac
csb
asb
cdv
c
rbc
ctv
hbam
adv
asv
cdb
ao
afrc
hba
atv
r6
cn
adb
bfrc
an
r5
ZINC

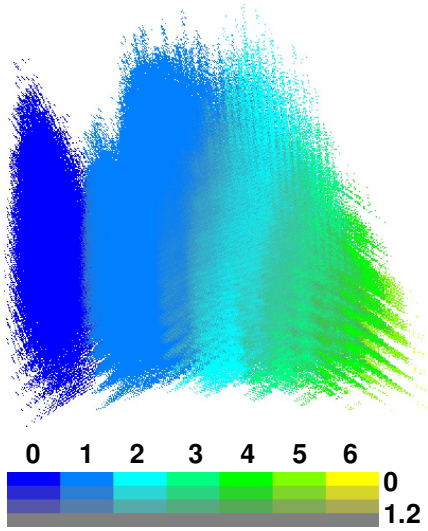

ZINC
HAC
3 11 19 27 35 47
0
2
N Rings
0 1 2 3 4 5 6 7 8
0
2.5
RBC
0 2 4 6 7 9 11
0
5.5
HBAm
0 2 4 6 8 11
0
8.0
TPSA
0 40 80 120 160 220
0
148
clogP
-6 -4 -2 0 2 4 5
0
11

GDB-11

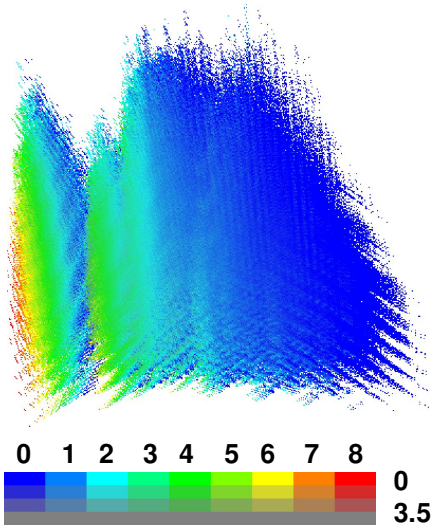
A. HAC



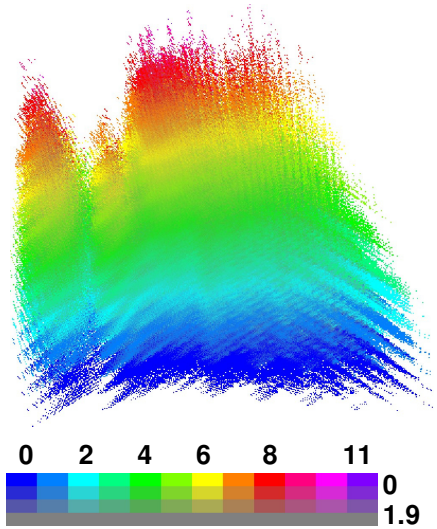
B. N_{Rings}



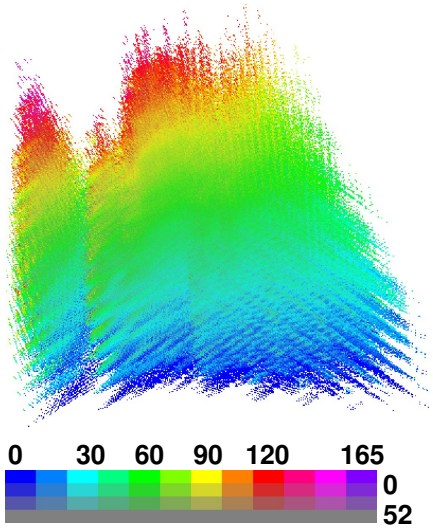
C. RBC



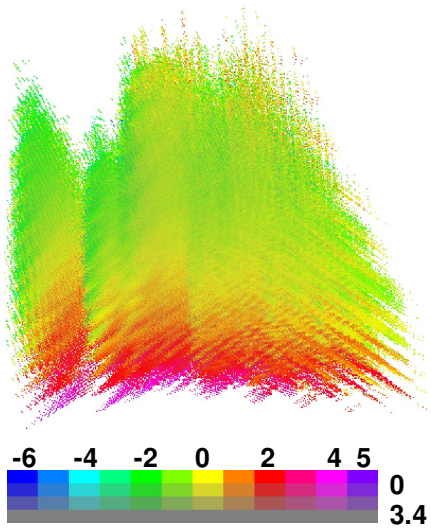
D. HBAm



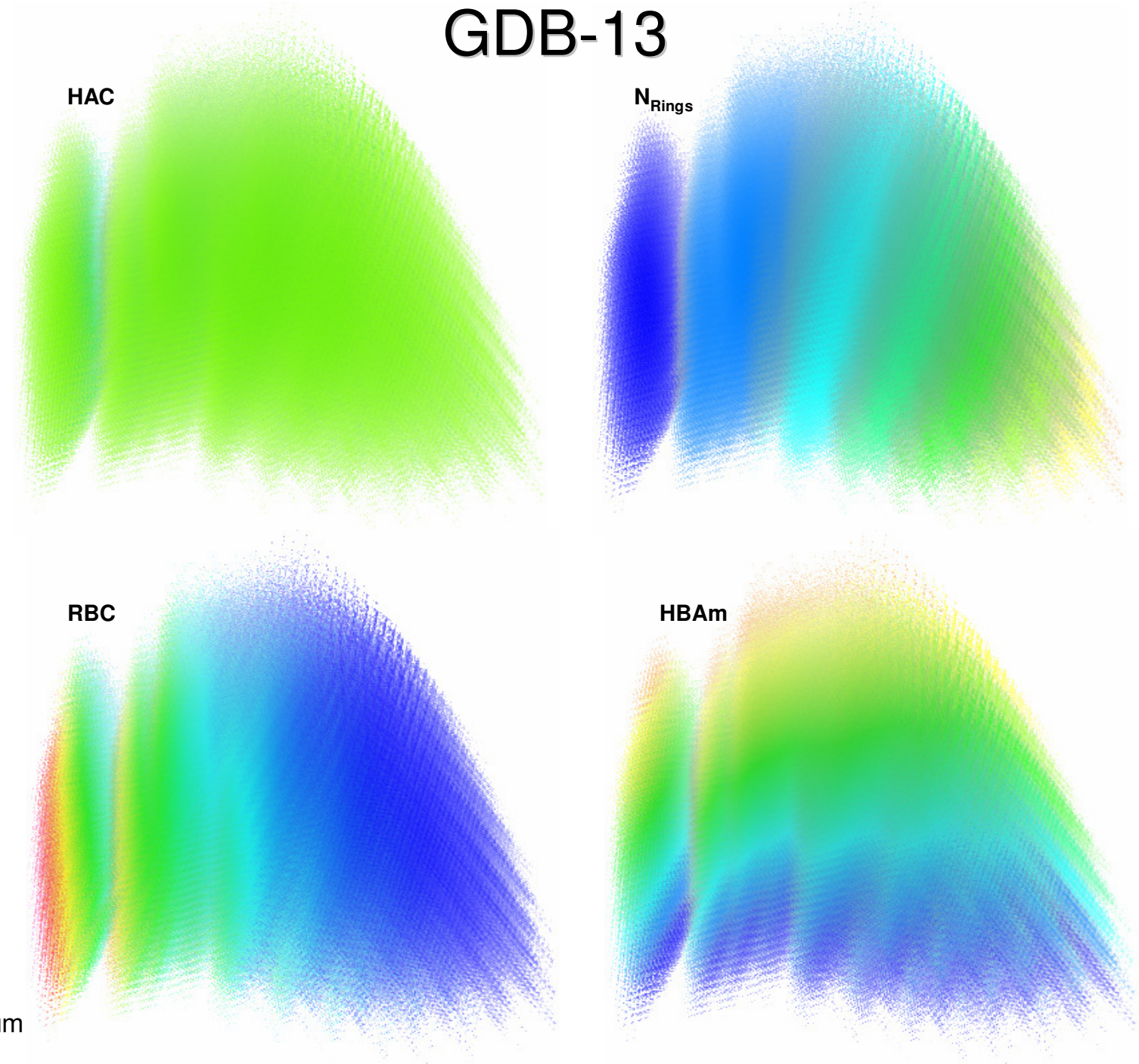
E. TPSA



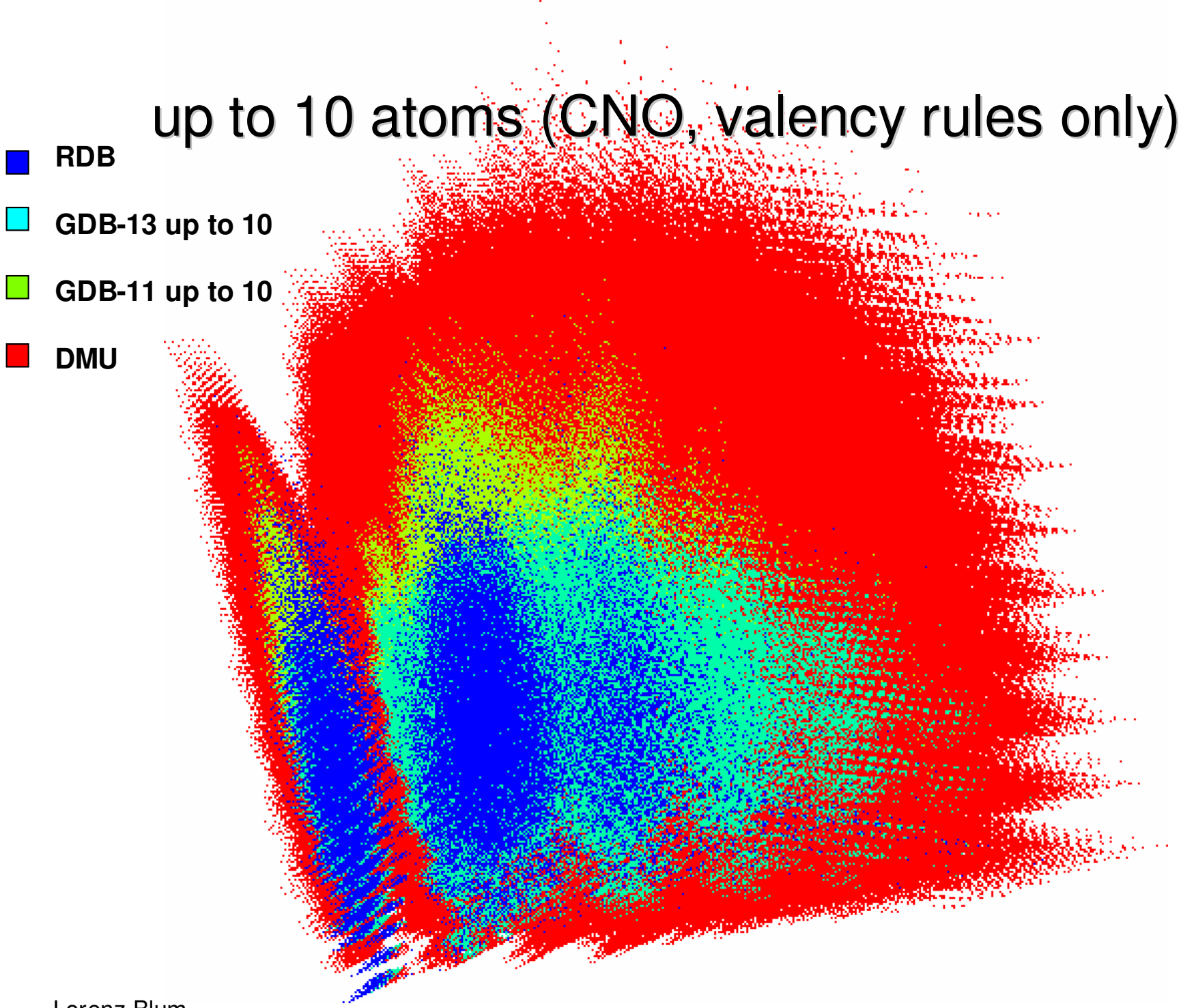
F. clogP



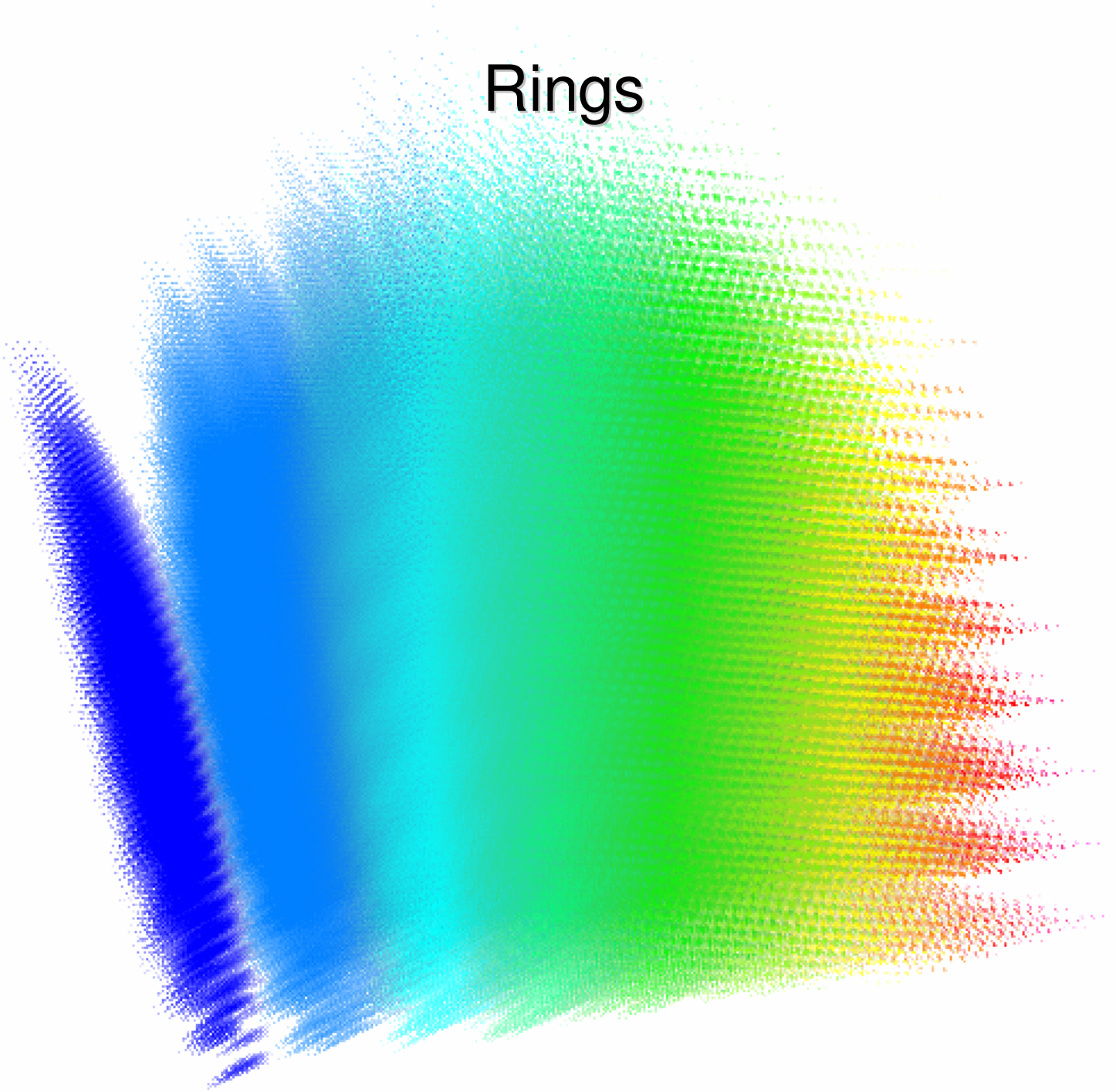
GDB-13



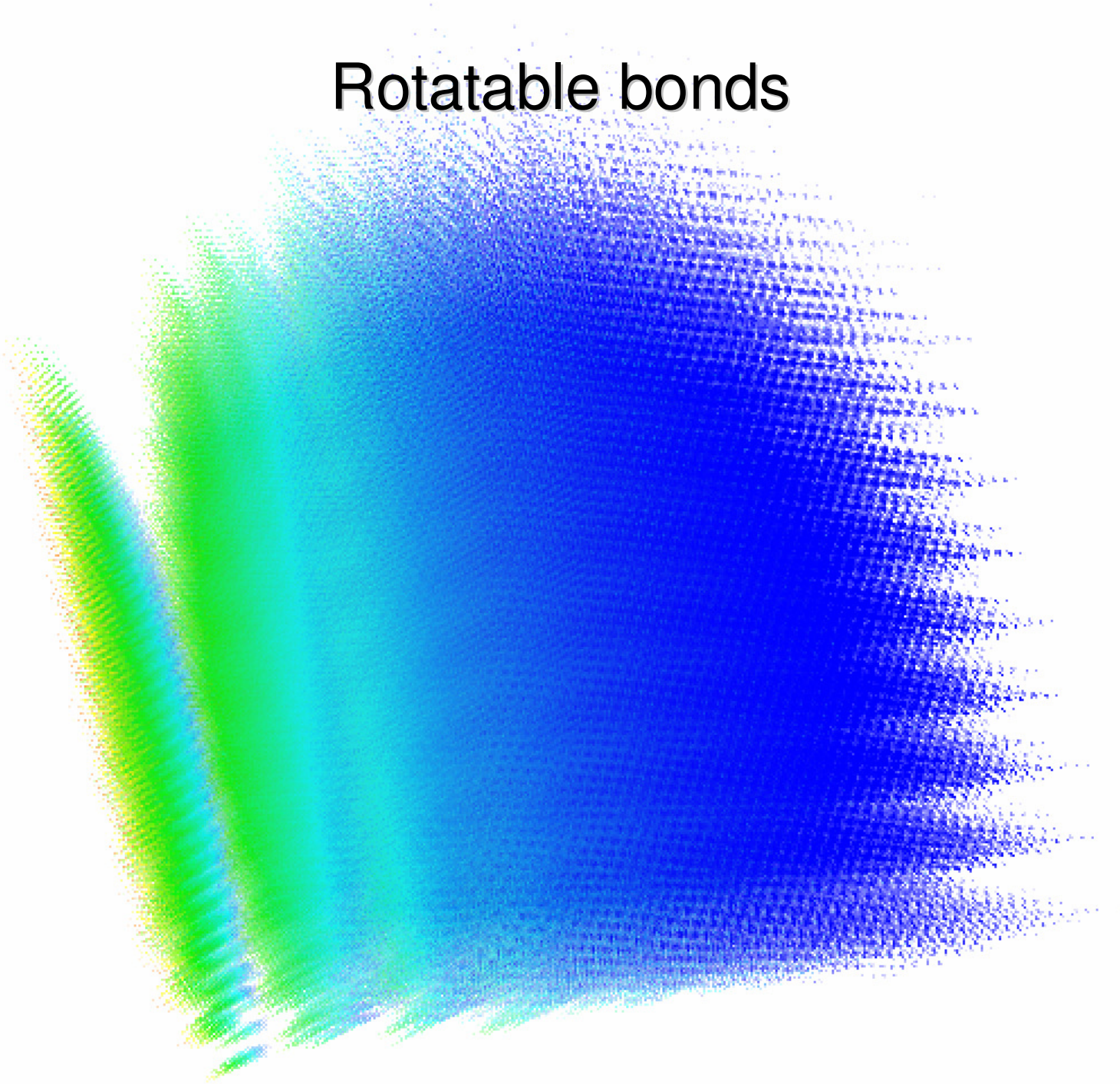
up to 10 atoms (CNO, valency rules only)



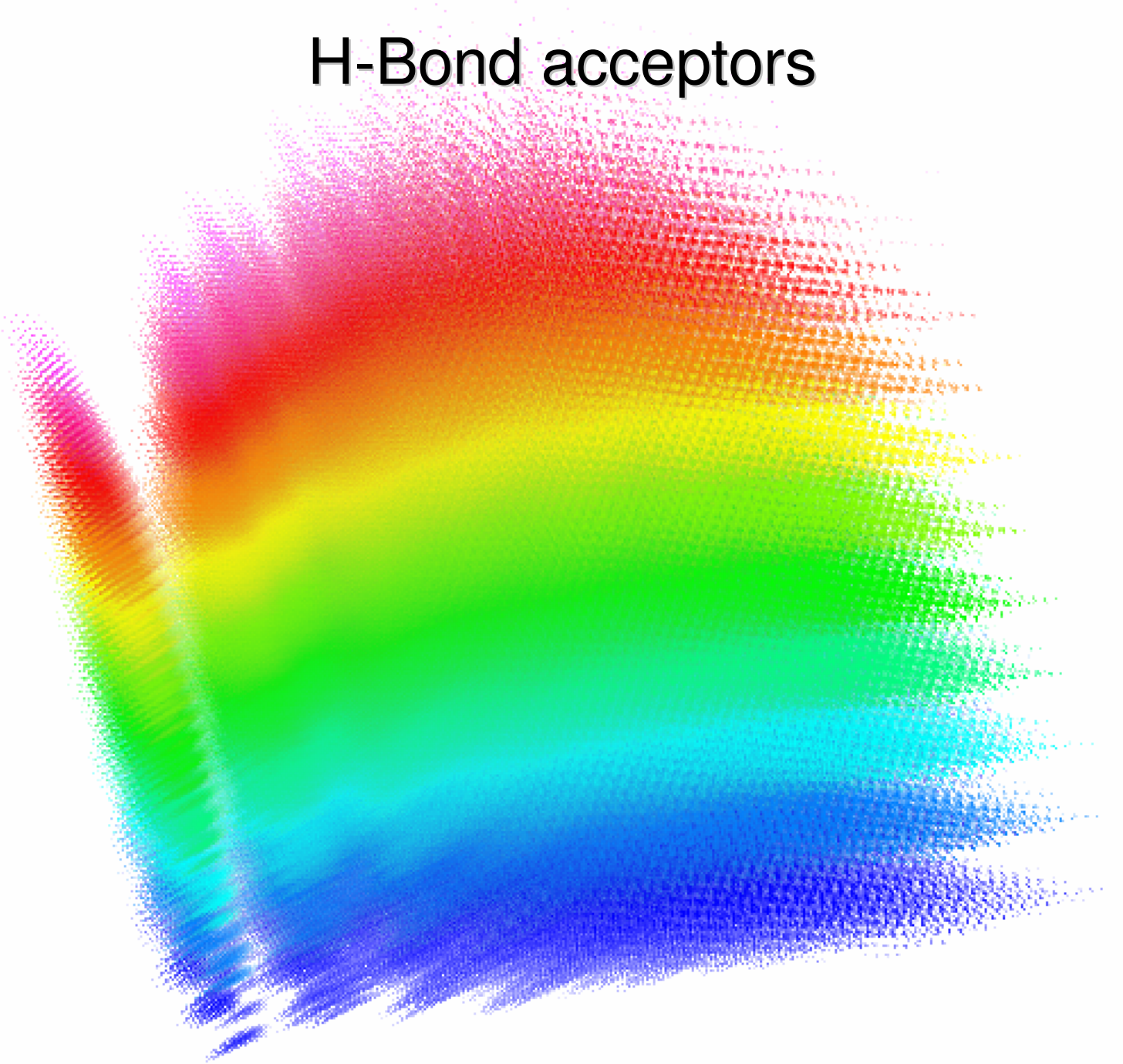
Rings



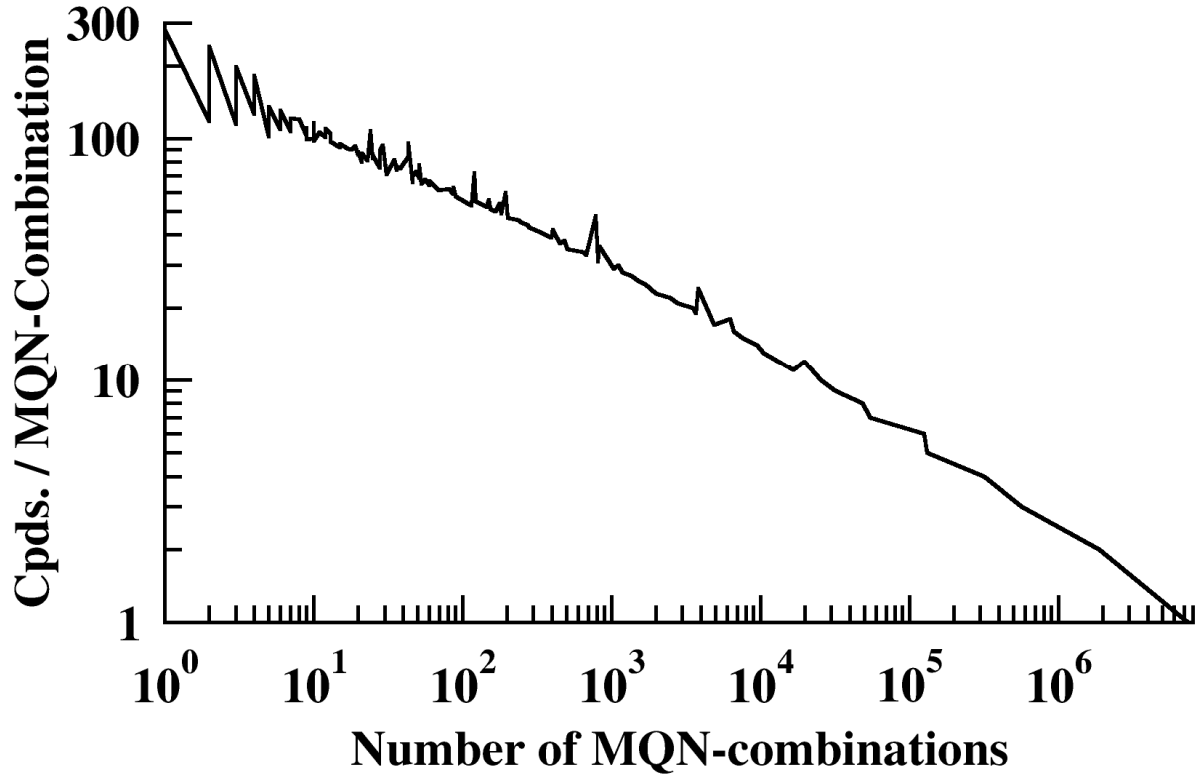
Rotatable bonds



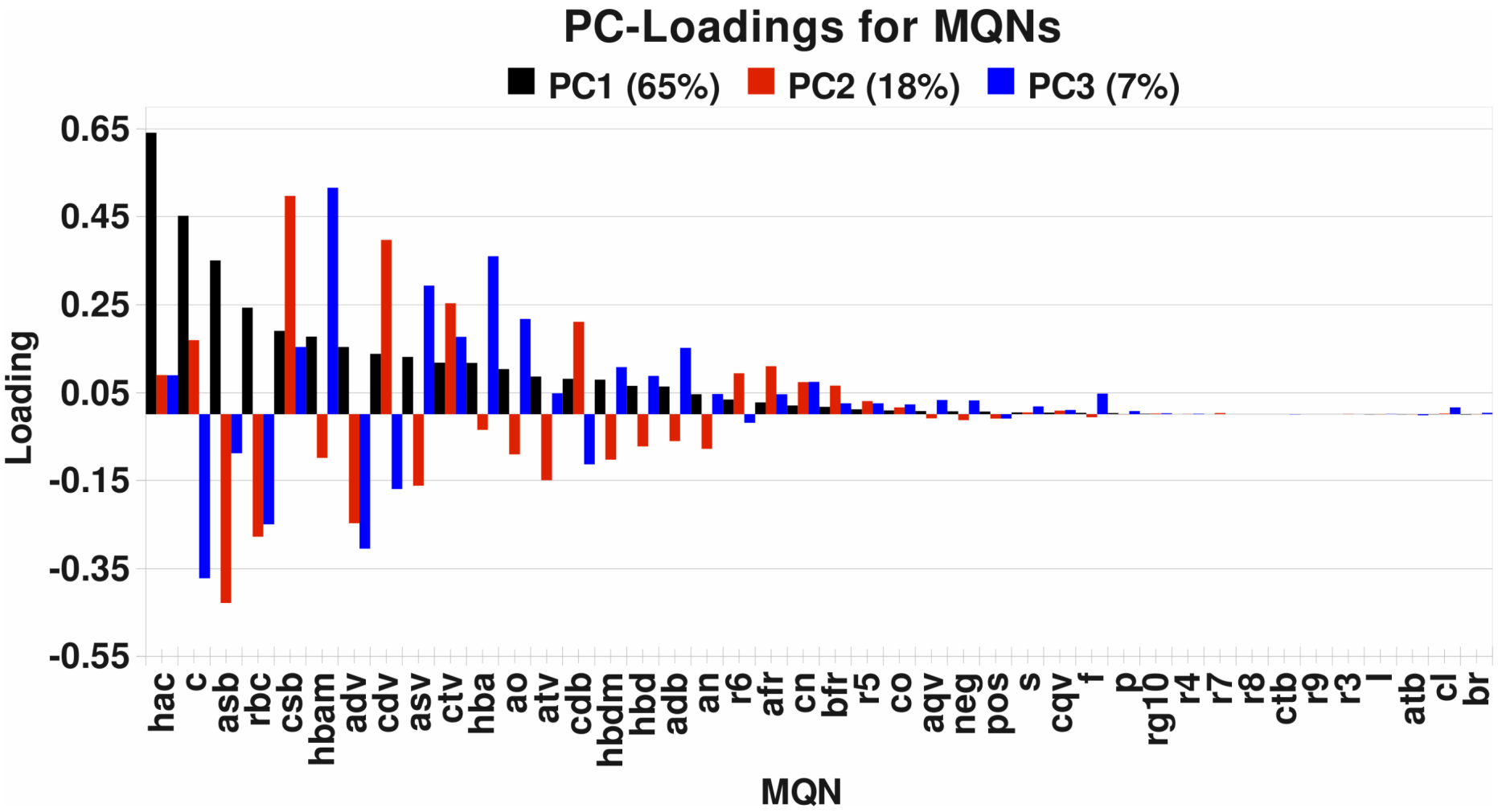
H-Bond acceptors



Analysis of PubChem (19.2 million SMILES)

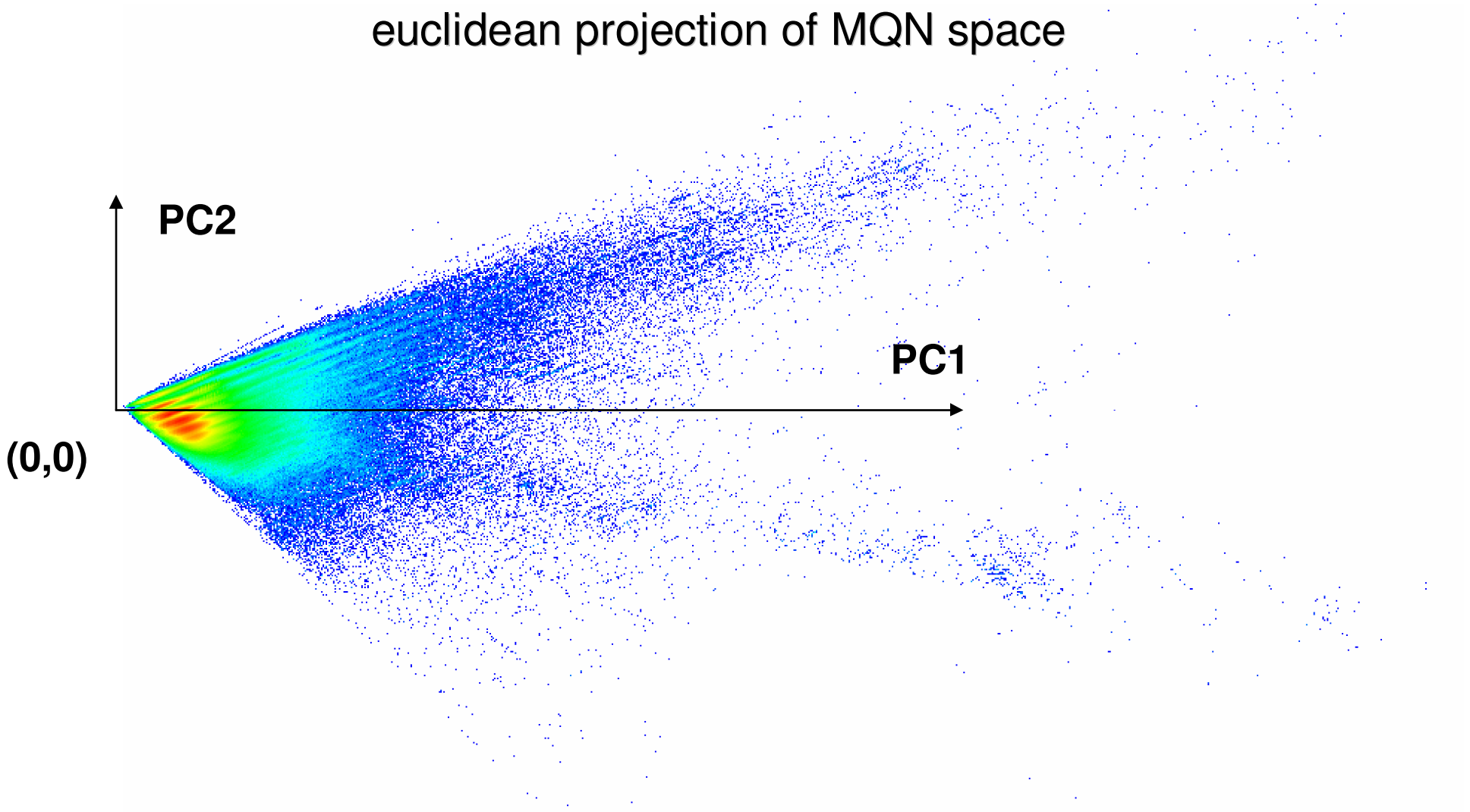


R. van Deursen, L. C. Blum, J.-L. Reymond, *J. Chem. Inf. Model.* **2010**, in press


PC-Loadings for MQNs
PC1 (65%)
PC2 (18%)
PC3 (7%)
Loading
0.65
0.45
0.25
0.05
-0.15
-0.35
-0.55
hac
c
asb
rbc
csb
hbam
adv
cdv
asv
ctv
hba
ao
atv
cdb
hbdm
hbd
adb
an
r6
afr
cn
bfr
r5
co
aqv
neg
pos
s
cqv
f
p
rg10
r4
r7
r8
ctb
r9
r3
l
atb
cl
br
MQN

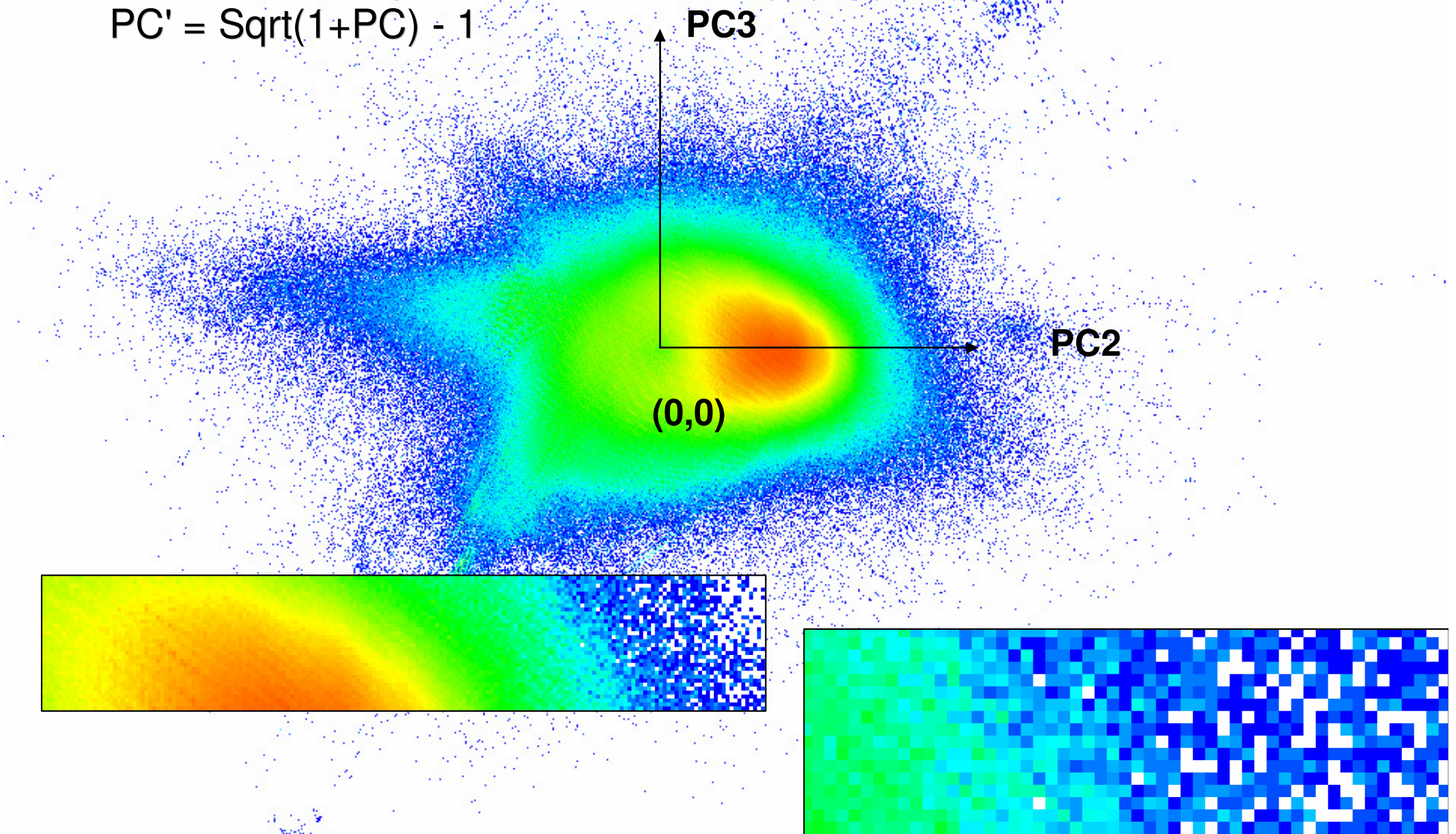
Frequency map

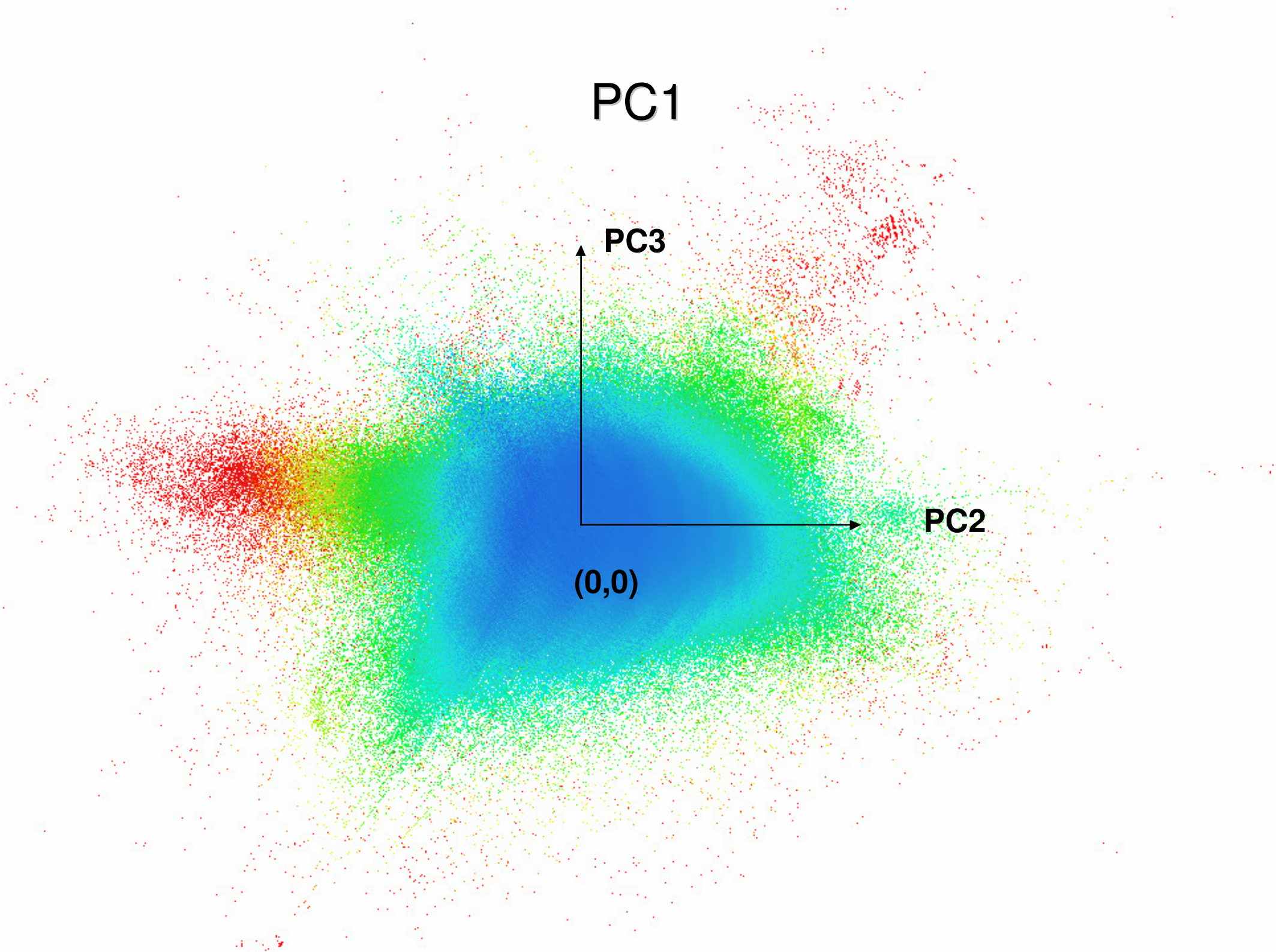
euclidean projection of MQN space



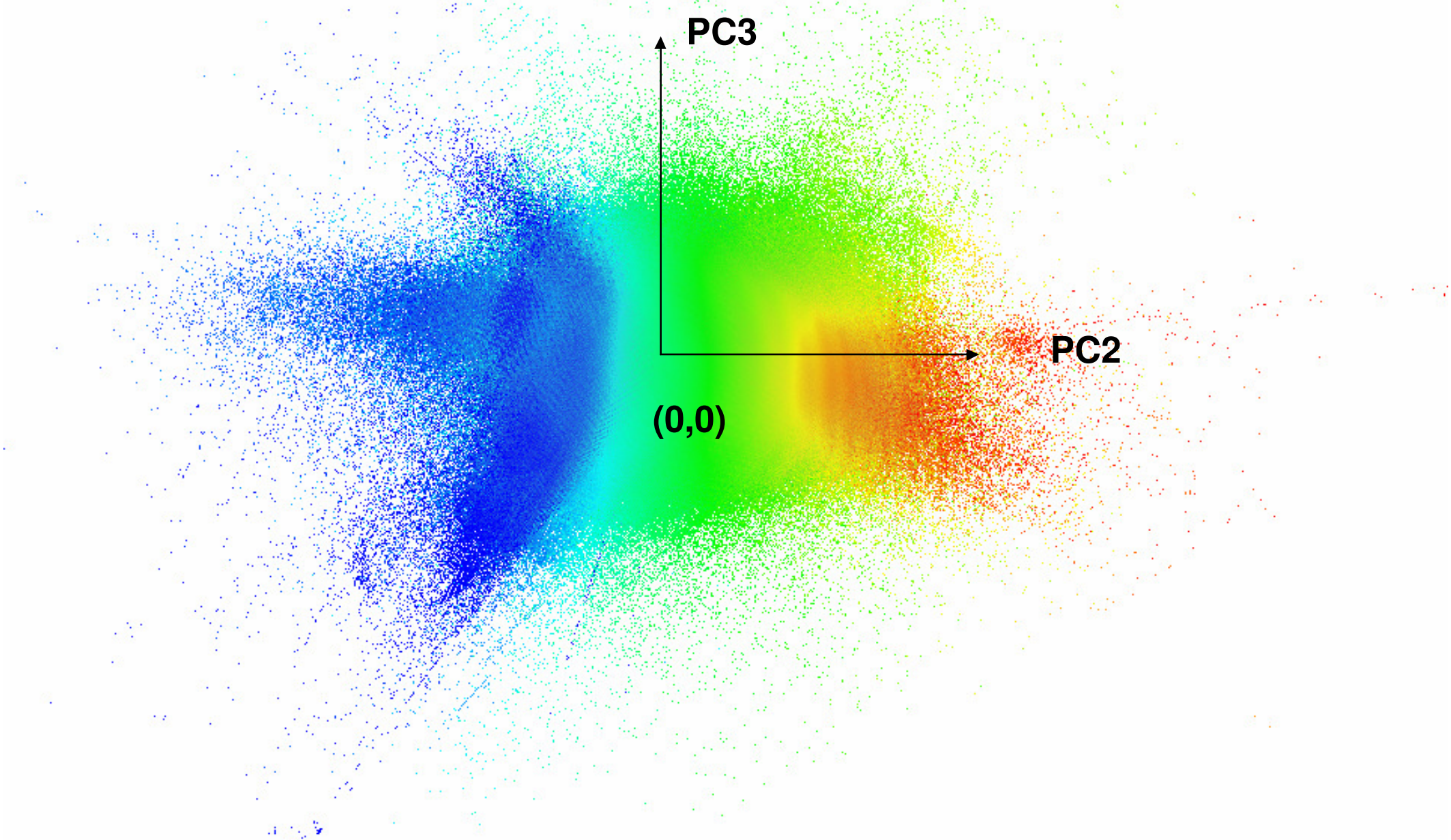
Frequency map

PC' = Sqrt(1+PC) - 1

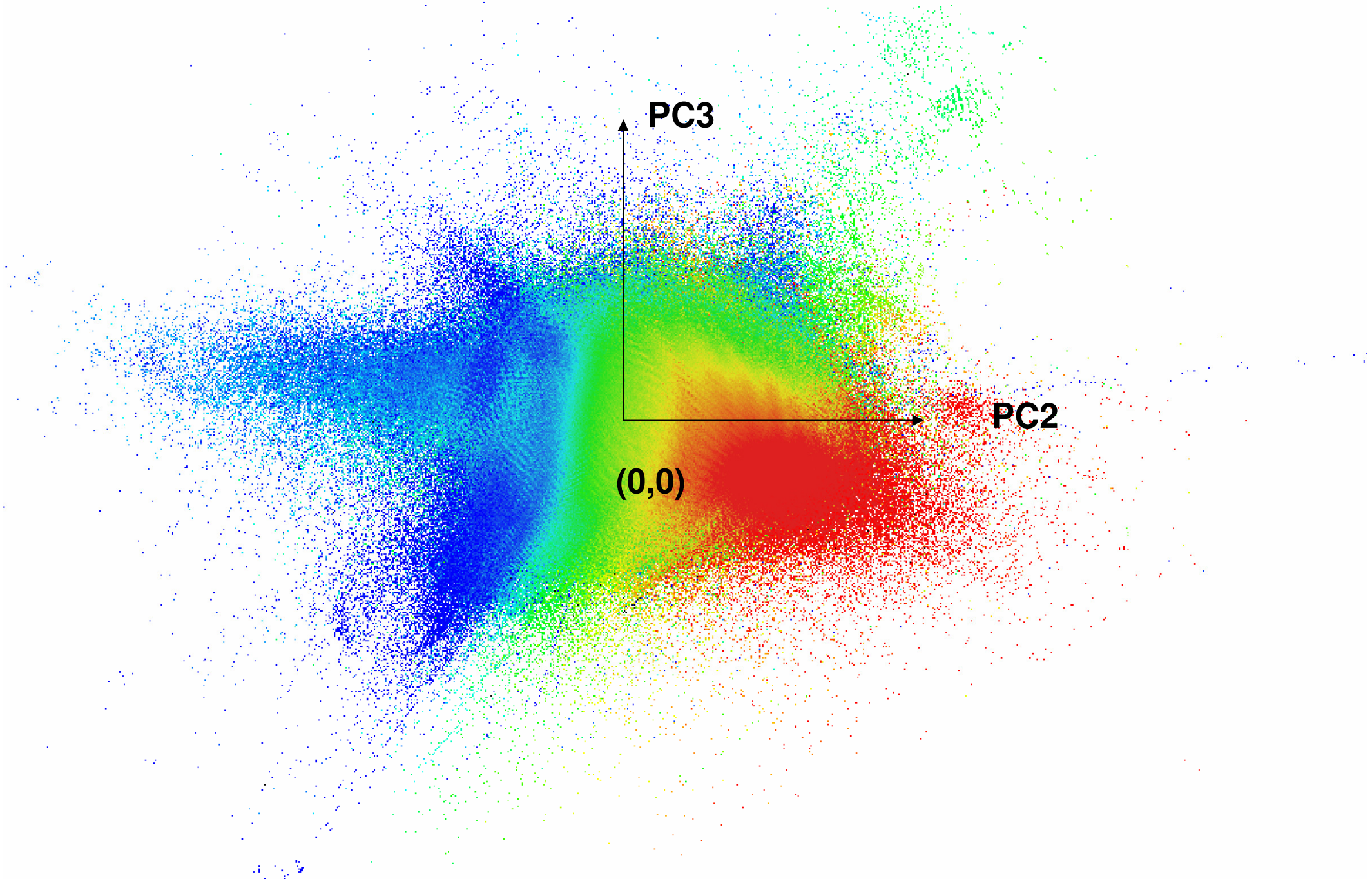



PC1
PC3
PC2
(0,0)

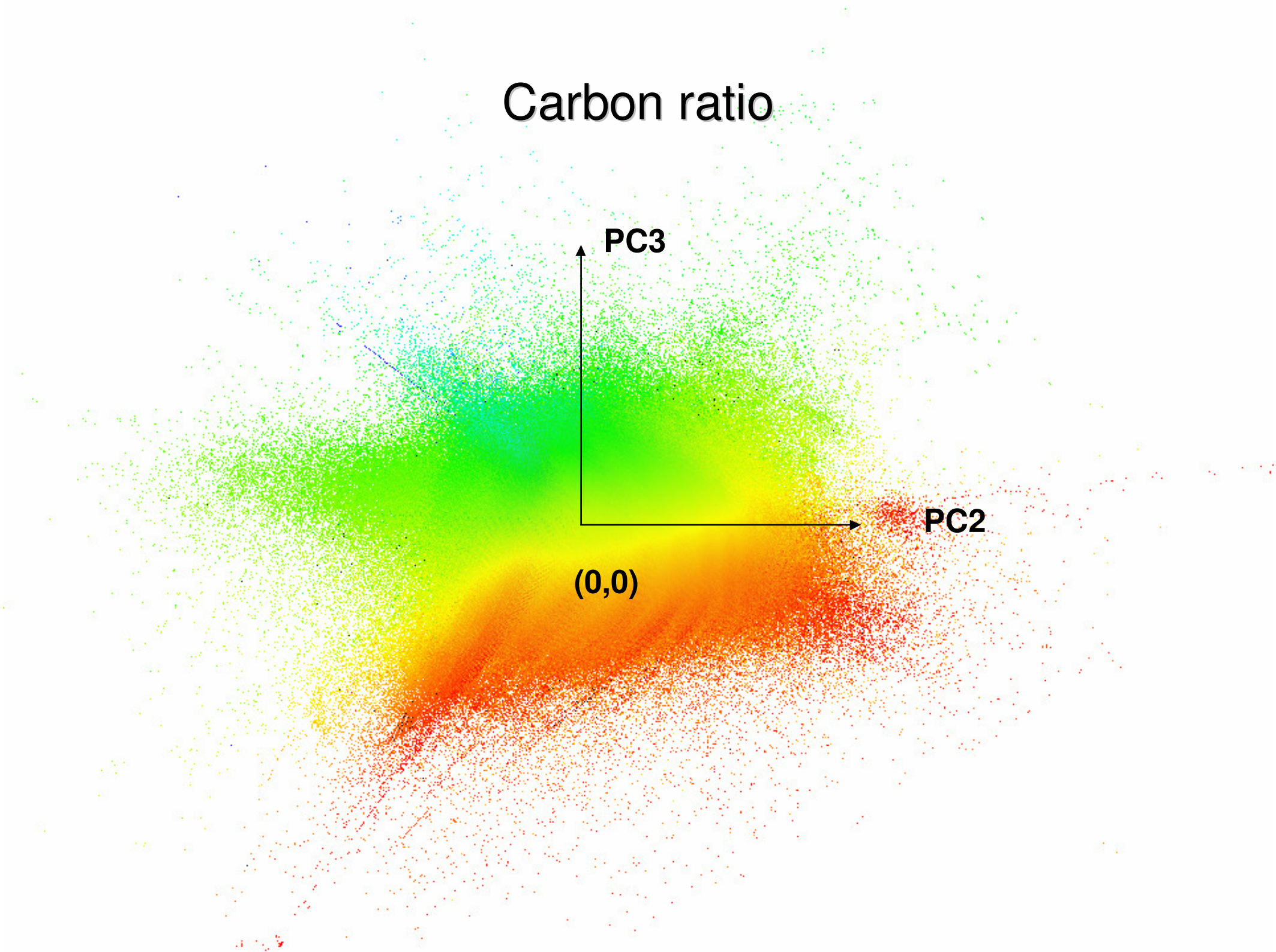
Ring atom ratio



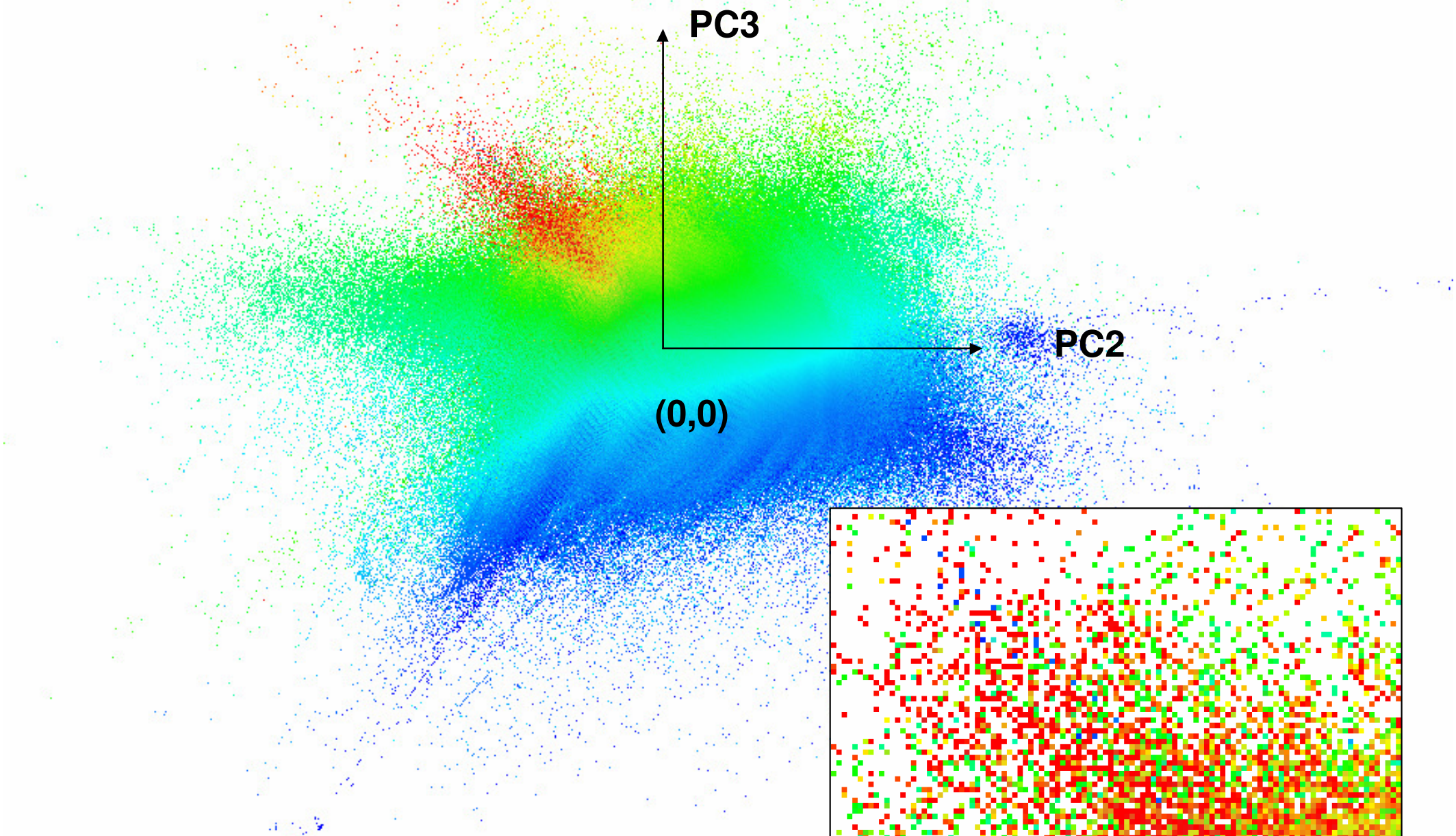
Cyclic double bond ratio

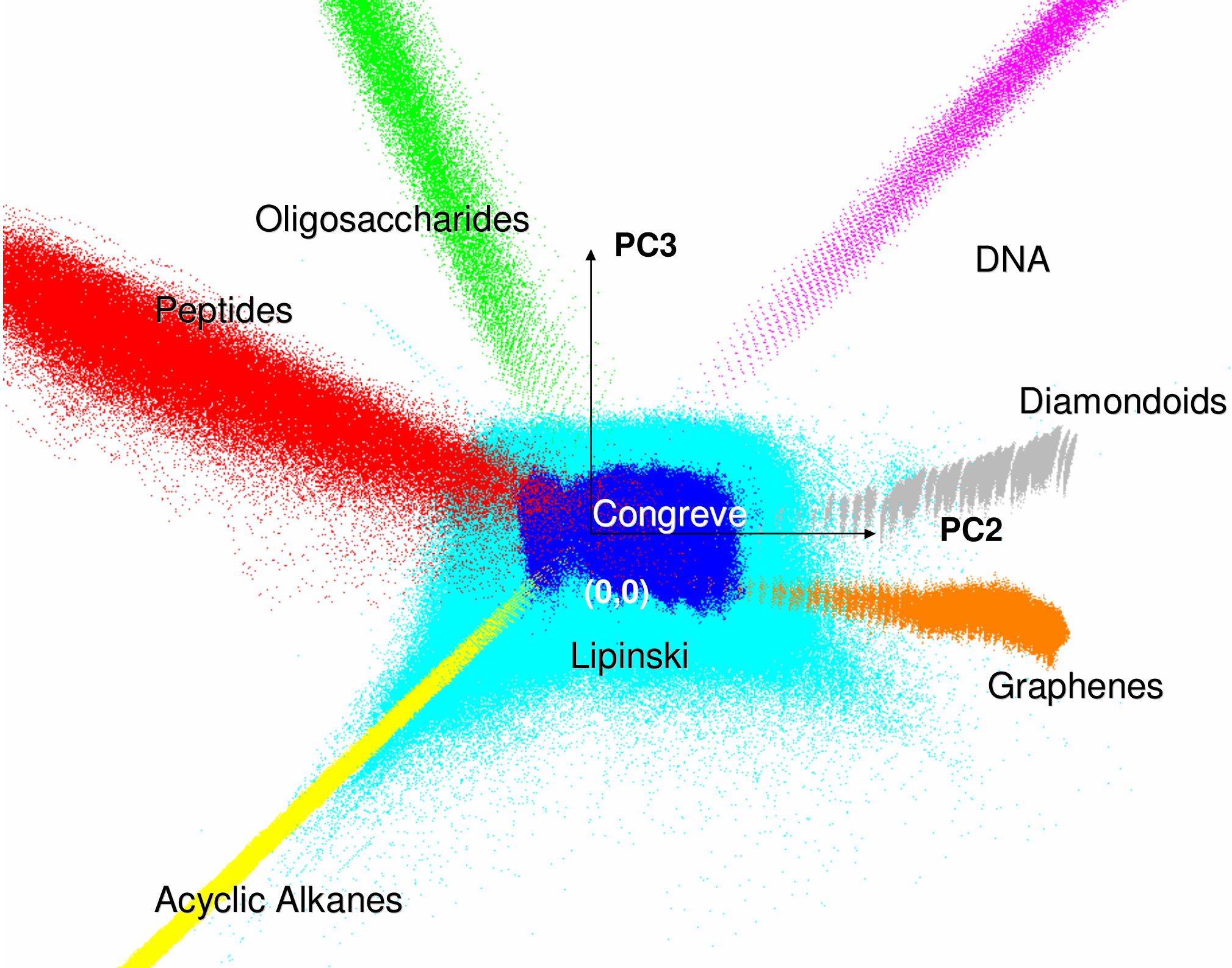


Carbon ratio

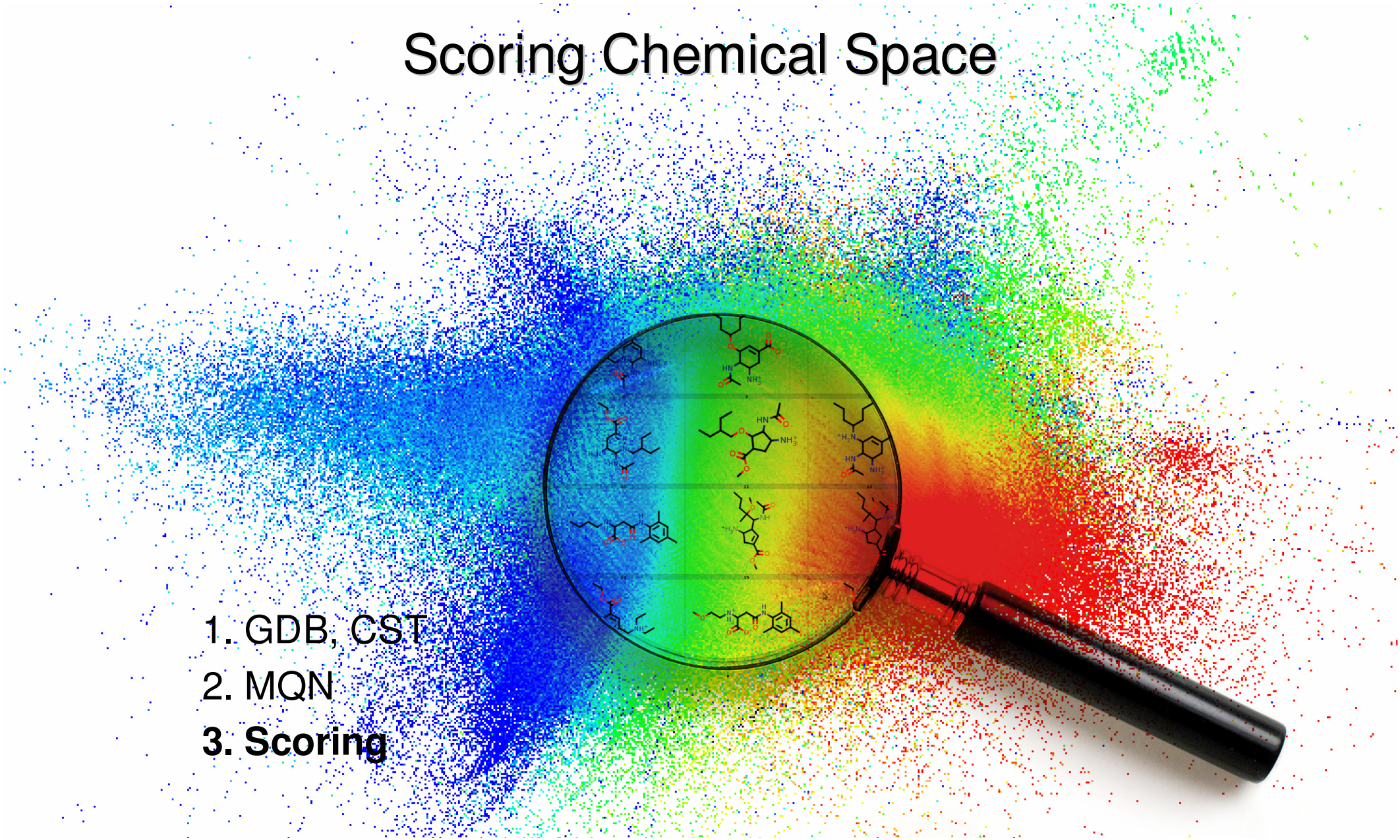


H-Bond acceptor ratio




Oligosaccharides
PC3
DNA
Peptides
Diamondoids
Congreve
PC2
(0,0)
Lipinski
Graphenes
Acyclic Alkanes

Scoring Chemical Space



1. GDB, CST
2. MQN
3. **Scoring**

R. van Deursen, L. C. Blum, J.-L. Reymond, *J. Chem. Inf. Model.* **2010**, in press

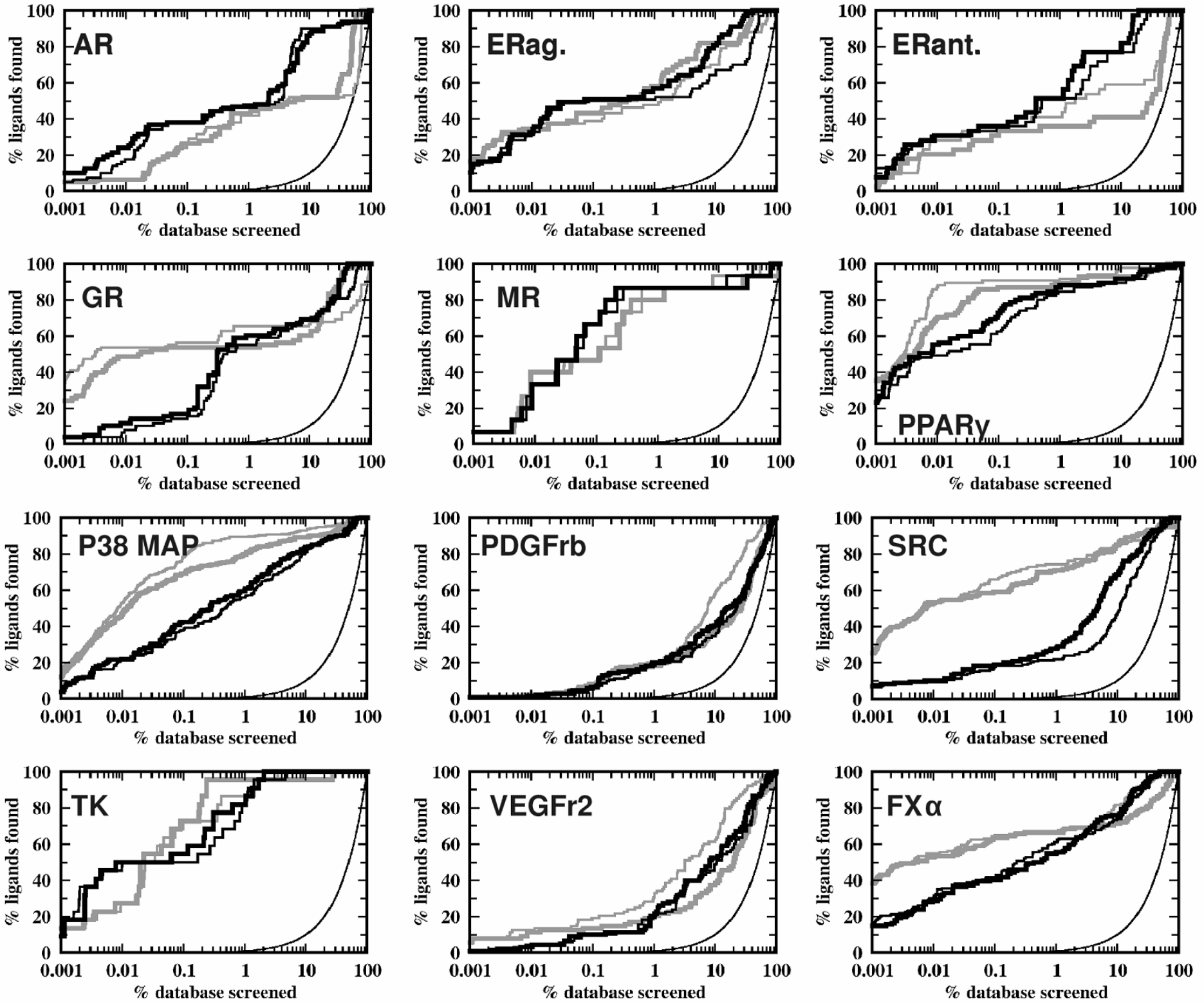
Lorenz Blum (GDB subsets)

E. Luethi et al., *J. Med. Chem.* **2010**, *53*, 7236, and N. Garcia-Delgado et al., *ACS Med. Chem. Lett.* **2010**, online

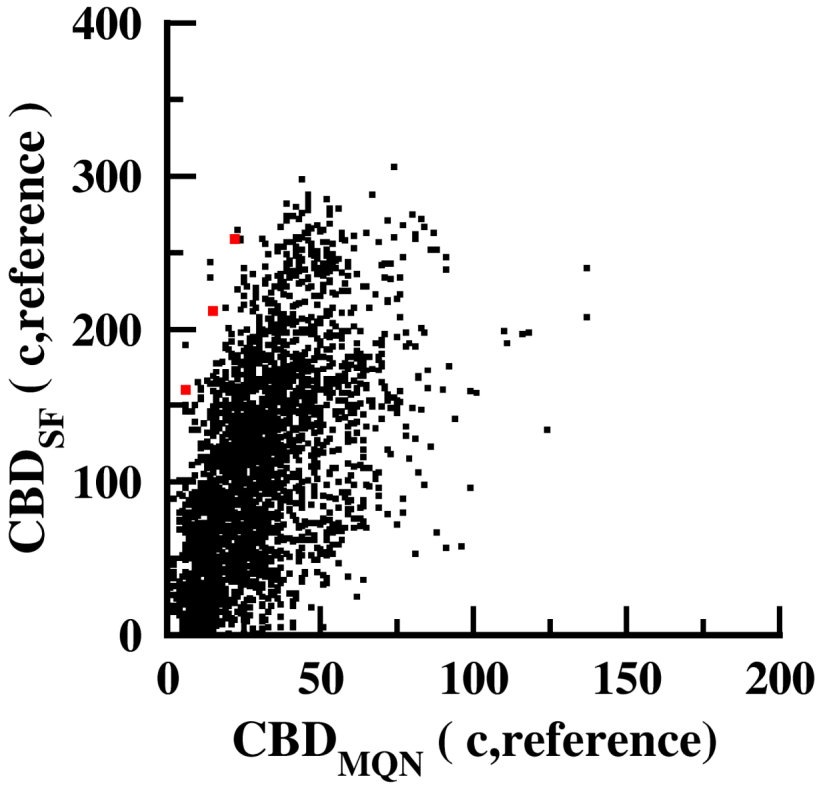
Enriching the DUD actives from Pubchem

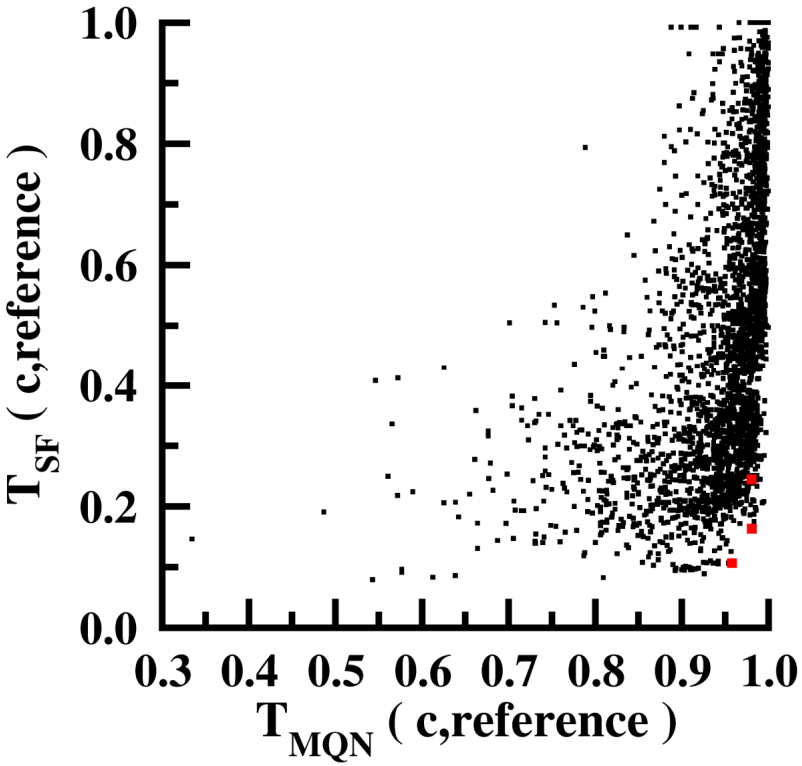
	nr. of actives[b)]	$EF_{0.1}$ CBD_{MQN}	T_{MQN}	CBD_{SF}	T_{SF}	EF_1 CBD_{MQN}	T_{MQN}	CBD_{SF}	T_{SF}
Nuclear Hormone Receptors									
AR	79	379.5	379.5	265.6	265.6	46.8	48.1	43.0	41.8
ERagonist	67	507.1	507.1	432.5	387.8	56.7	50.7	58.2	47.8
ERantagonist	39	358.7	358.7	333.1	333.1	51.3	51.3	35.9	41.0
GR	78	166.6	140.9	538.1	563.7	60.2	55.1	53.8	65.4
MR	15	666.2	666.2	466.3	466.3	86.7	86.7	80.0	86.7
PPARg	85	728.9	623.1	870.0	905.3	87.0	84.7	89.4	91.7
PR	27	592.2	555.2	592.2	629.2	59.2	59.2	59.2	70.4
RXRa	20	849.4	599.6	849.4	849.4	95.0	85.0	85.0	100.0
Kinases									
CDK2	72	111.0	83.3	138.8	138.8	20.8	18.1	15.3	16.7
EGFR	475	90.5	67.3	126.2	132.5	25.5	20.8	20.2	27.2
FGFr1	120	191.5	183.2	624.6	724.5	29.2	22.5	74.2	85.0
HSP90	37	378.1	378.1	648.2	702.2	54.0	37.8	70.3	70.3
P38 MAP	454	424.8	380.8	691.2	783.6	59.5	55.9	79.5	89.4
PDGFrb	170	64.7	52.9	82.3	82.3	20.0	18.2	18.2	20.0
SRC	159	188.5	182.3	590.8	659.9	27.7	22.0	71.1	74.2
TK	22	545.1	499.7	726.8	726.8	81.8	81.8	95.4	86.3
VEGFr2	88	102.2	102.2	136.3	193.0	20.5	19.3	20.5	30.7

Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking.
J. Med. Chem. **2006,** *49*, 6789-6801.

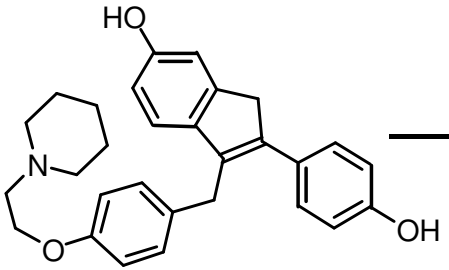

AR
ERag.
ERant.
GR
MR
PPARγ
P38 MAP
PDGFrb
SRC
TK
VEGFr2
FXα
% ligands found
% database screened
0
20
40
60
80
100
0.001
0.01
0.1
1
10
100

Lead Hop(p)ing

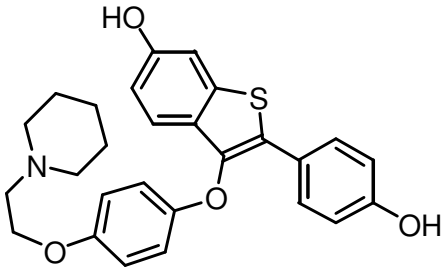




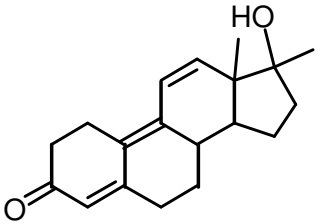
Lead Hop(p)ing



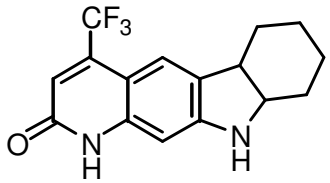
Lead ER antagonist



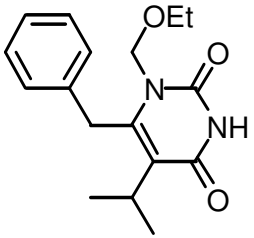
CBD_{MQN} = 6
(CBD_{SF}= 160)



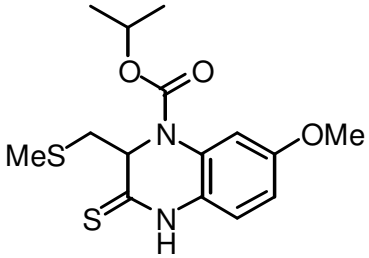
Lead AR inhibitor



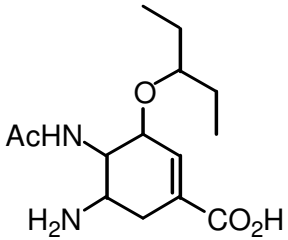
T_{MQN} = 0.958
(T_{SF} = 0.107)



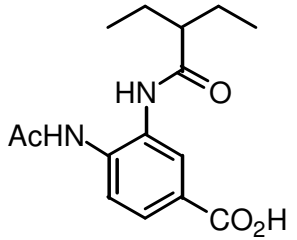
Lead HIV RT inhibitor



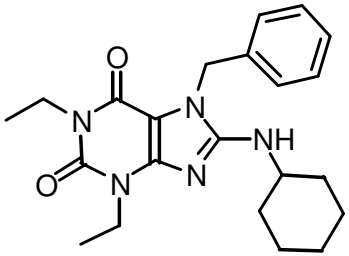
CBD_{MQN} = 15
(CBD_{SF}= 212)



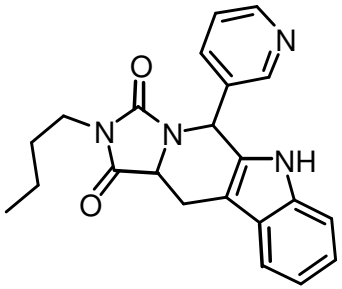
Lead NA inhibitor



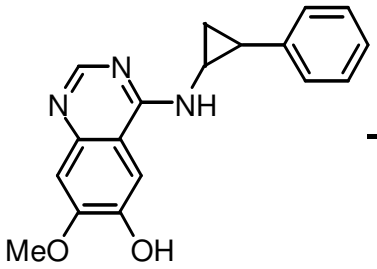
T_{MQN} = 0.981
(T_{SF} = 0.163)



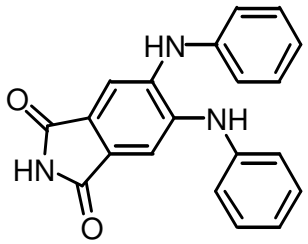
Lead pde5 inhibitor



CBD_{MQN} = 22
(CBD_{SF} = 259)



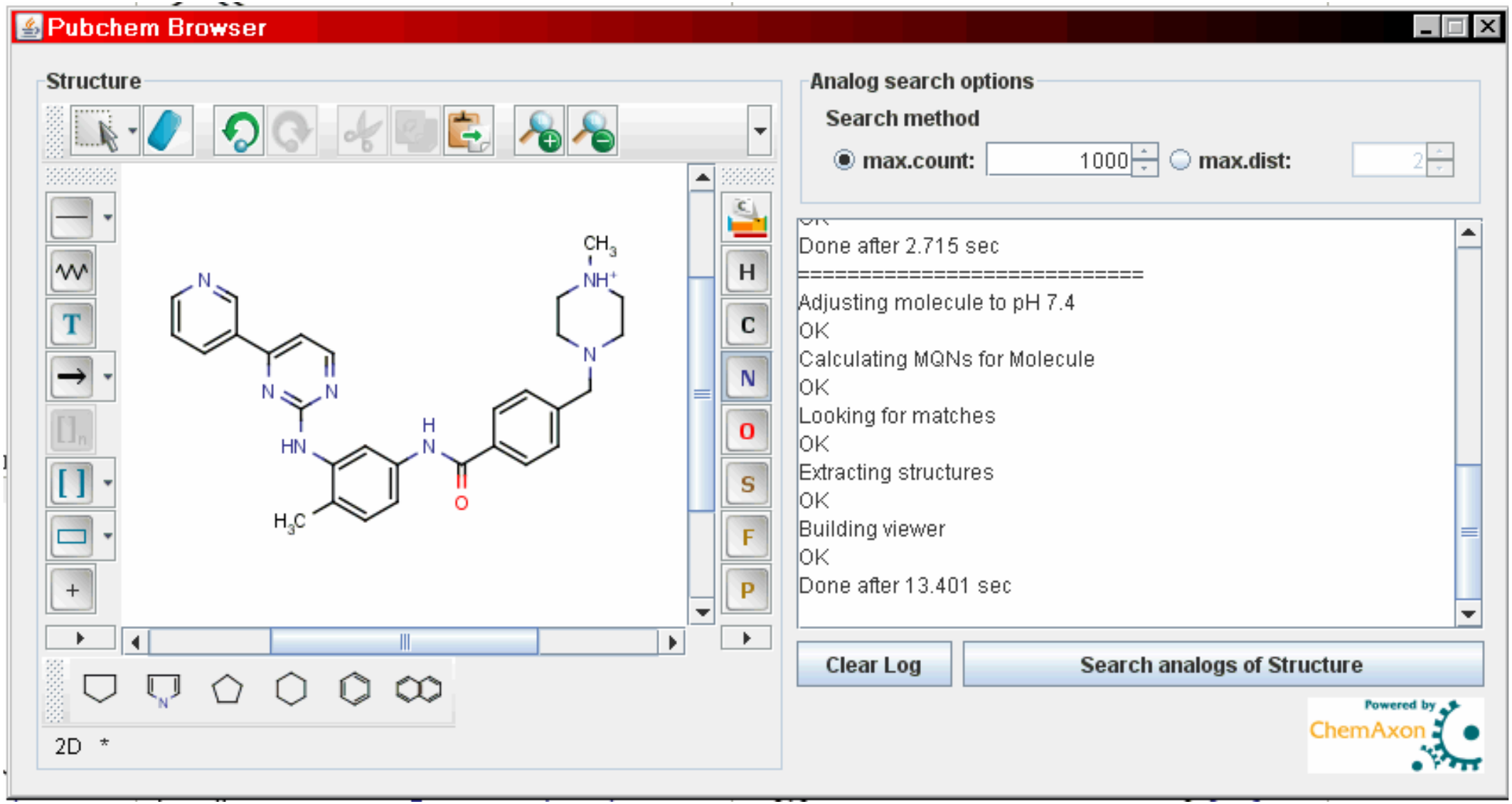
Lead EGFR inhibitor

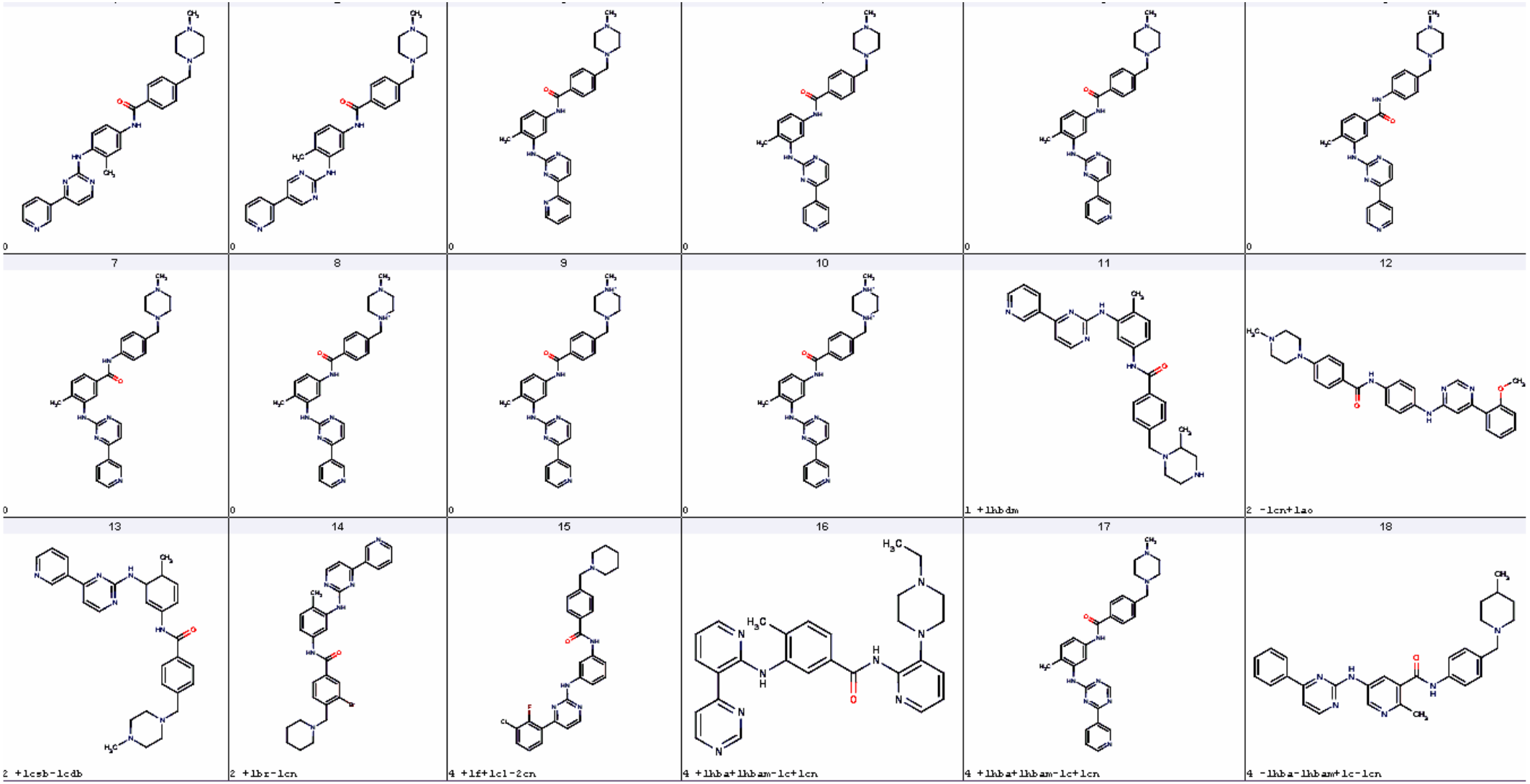


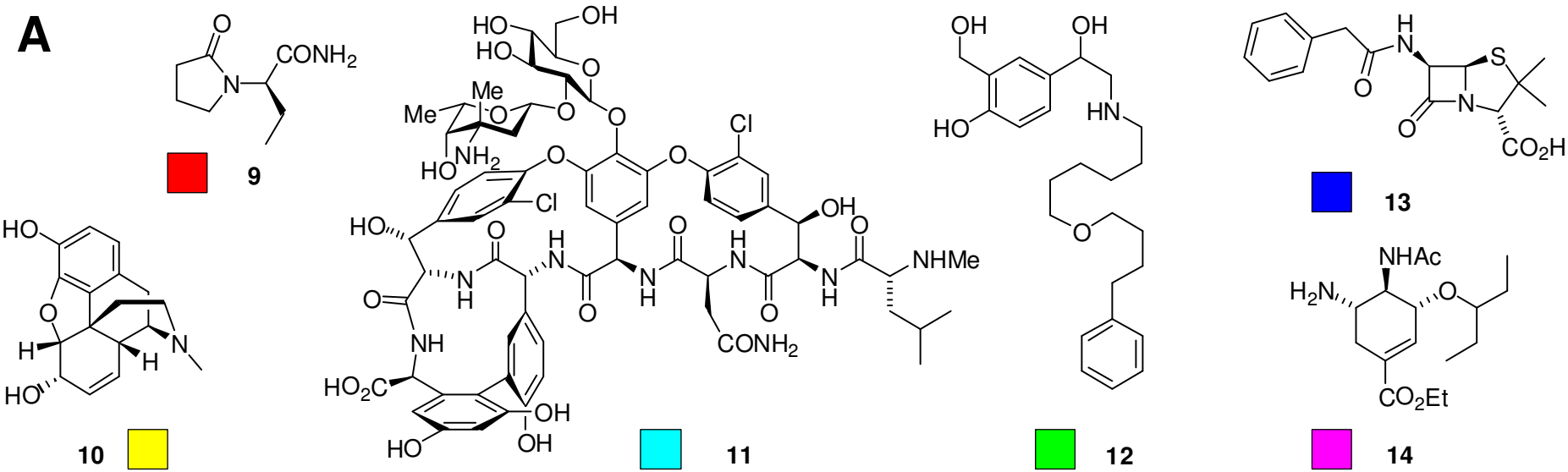
T_{MQN} = 0.980
(T_{SF} = 0.244)

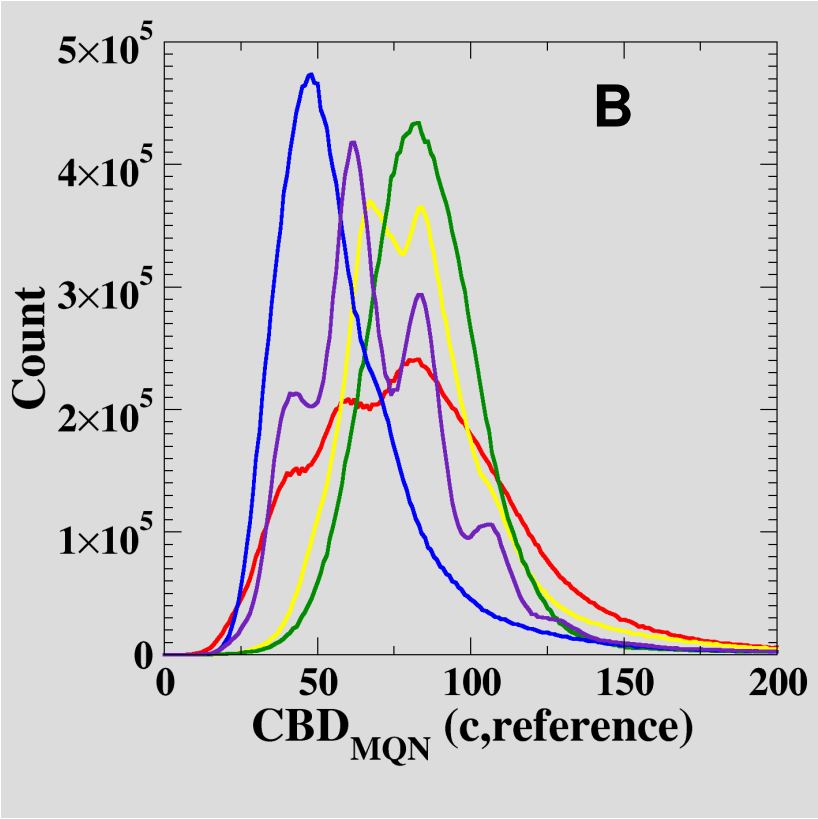
The PubChem-browser

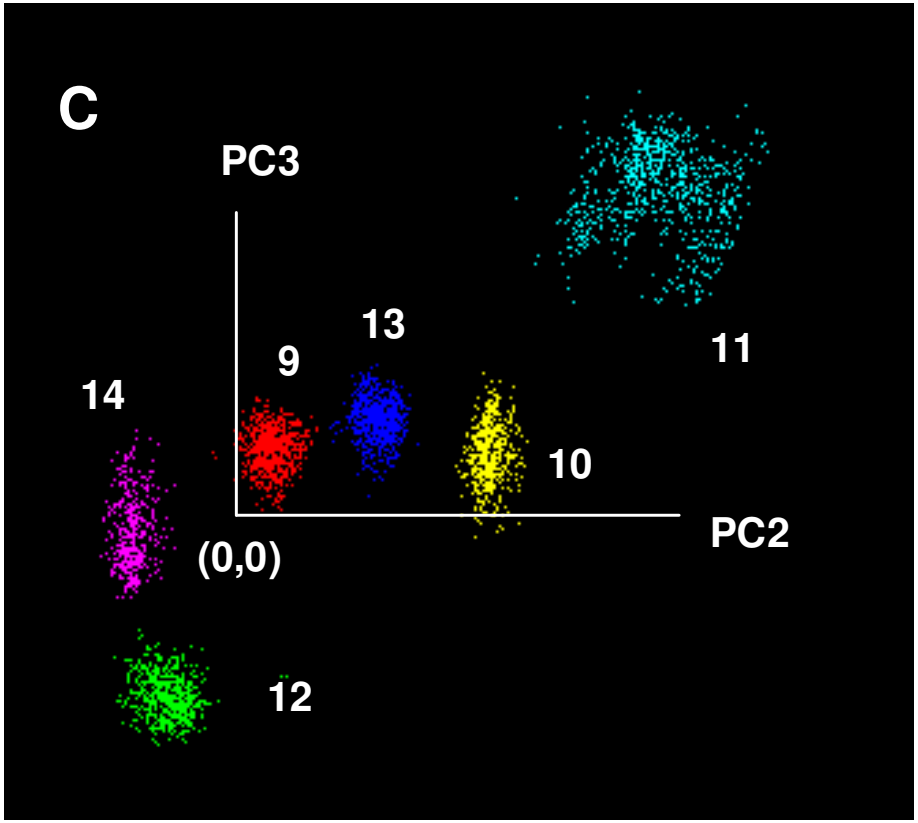
- Draw structure
- retrieve MQN_{CBD} neighbours in database




0
0
0
0
0
0
7
8
9
10
11
12
0
0
0
0
1 +lkbdm
2 -lcn+lao
13
14
15
16
17
18
2 +lcsb-lcdb
2 +lbr-lcn
4 +lf+lcl-2cn
4 +lkba+lkbam-lc+lcn
4 +lkba+lkbam-lc+lcn
4 -lkba-lkbam+lc-lcn


A
9
10
11
12
13
14


B
Count
5×10^5
4×10^5
3×10^5
2×10^5
1×10^5
0
0
50
100
150
200
CBD_MQN (c,reference)

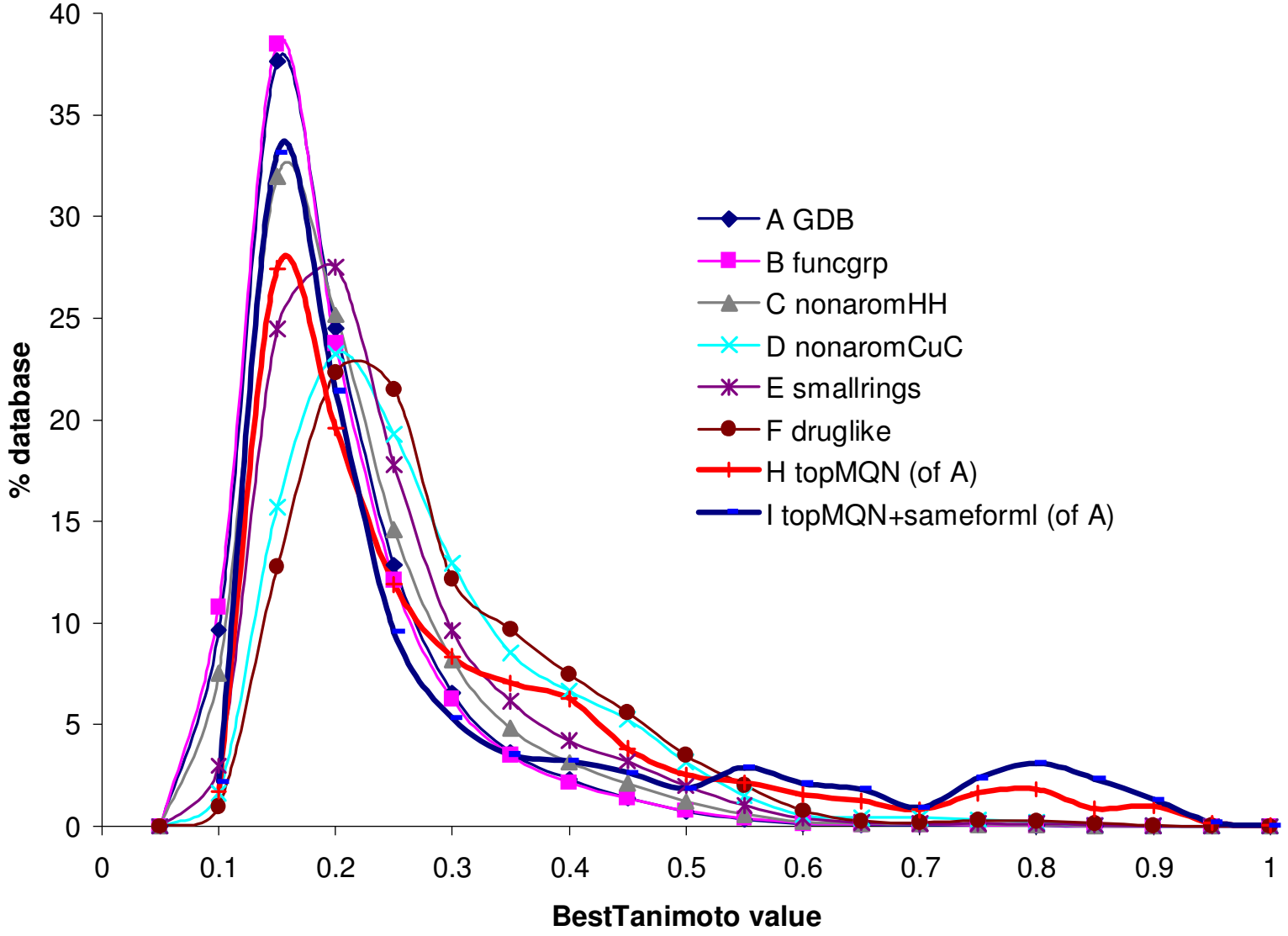

C
PC3
PC2
(0,0)
9
10
11
12
13
14

Subsets of GDB-13

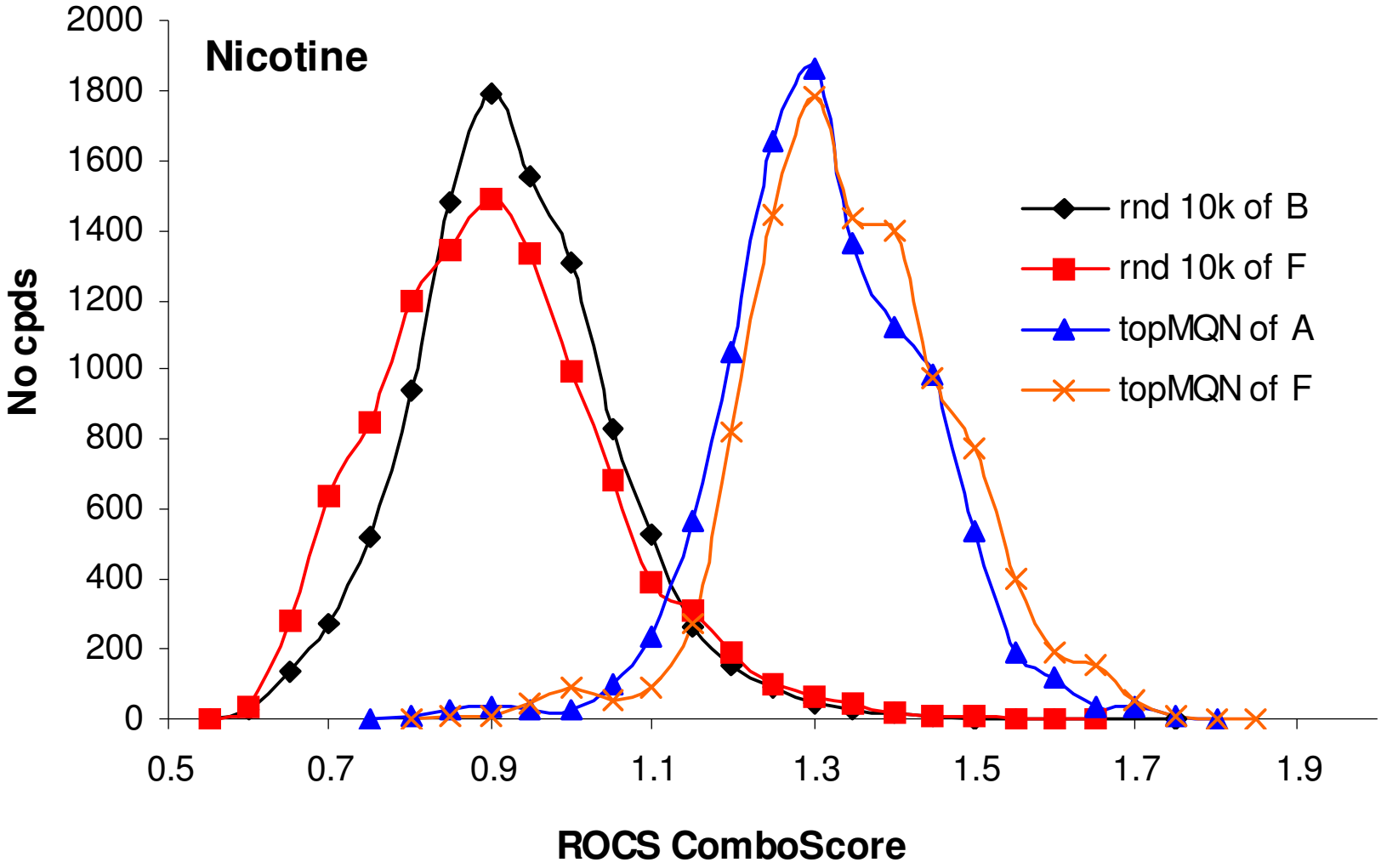
- A: full GDB-13
- B: A without esters, carbonates, sulfates, aldehydes, epoxides, aziridines
- C: B without non-aromatic het-het bonds
- D: C without non-aromatic C=C
- E: D without small rings
- F: fragment-like portion of E
- MQN sets: 10,000 MQN neighbours of query molecule in any of the subsets A-F

- Scoring:
  - structural similarity to query (Tanimoto of SF fingerprint)
  - shape similarity to query (ROCS)

MQN-sets yield high T_{SF} analogs



MQN-sets yield high ROCS scores



7236 *J. Med. Chem.* **2010,** *53,* 7236–7250
DOI: 10.1021/jm100959g

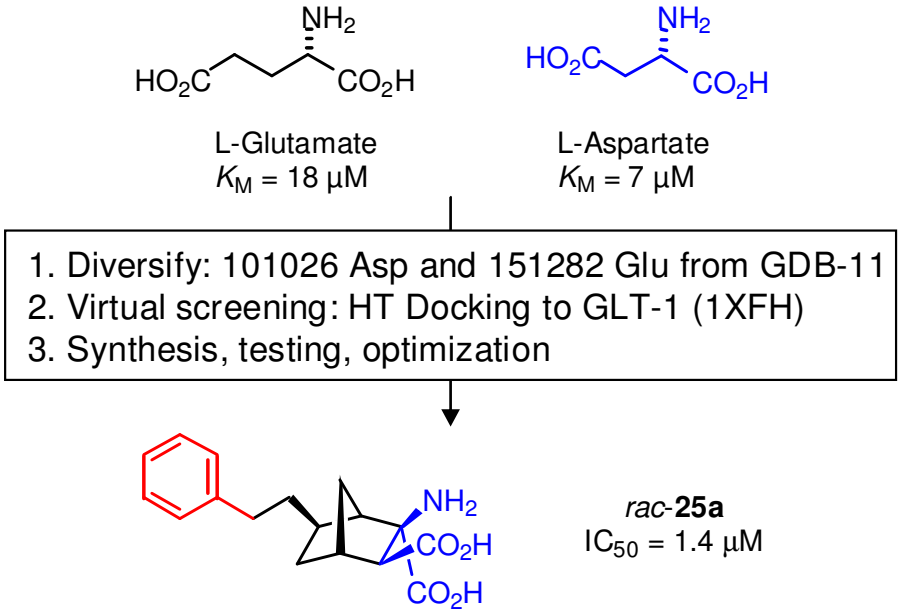
Journal of Medicinal Chemistry Article

Identification of Selective Norbornane-Type Aspartate Analogue Inhibitors of the Glutamate Transporter 1 (GLT-1) from the Chemical Universe Generated Database (GDB)

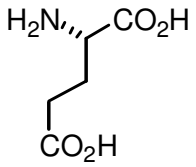
Erika Luethi,[†] Kong T. Nguyen,[†] Marc Bürzle,[‡] Lorenz C. Blum,[†] Yoshiro Suzuki,[‡] Matthias Hediger,[‡,§] and Jean-Louis Reymond*[,†,§]

[†]*Department of Chemistry and Biochemistry, University of Berne, Freiestrasse 3, 3012 Berne, Switzerland,*
[‡]*Institute of Biochemistry and Molecular Medicine, University of Berne, Bühlstrasse 28, 3012 Berne, Switzerland, and*
[§]*Swiss National Center of Competence in Research, NCCR-TransCure, University of Berne, Switzerland*

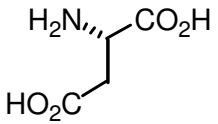
Received July 28, 2010



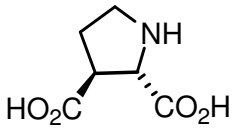
Known GLT-1 Ligands



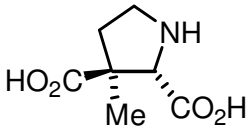
1 (L-Glutamate)
BE = -9.1 kcal/mol



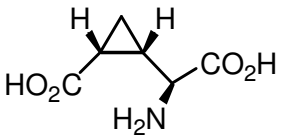
2 (L-Aspartate)
BE = -8.2 kcal/mol
Inh.% of control = 84@100μM



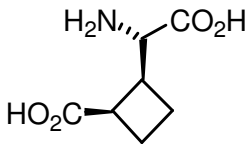
3 (trans-2,3-PDC)
BE = -9.5 kcal/mol
K_i=12 μM



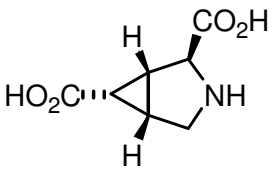
4 (3-Me-L-2,3-PDC)
BE = -9.4 kcal/mol
Inh.% of control = 88@100μM



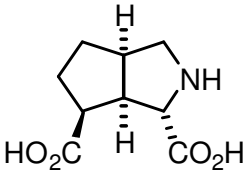
5 (L-CCG-III)
BE = -9.2 kcal/mol
IC_{50}=2 μM



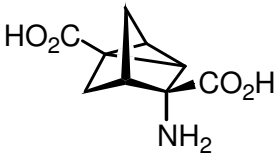
6 (L-CBG-IV)
BE = -10.7 kcal/mol
K_i=7 μM



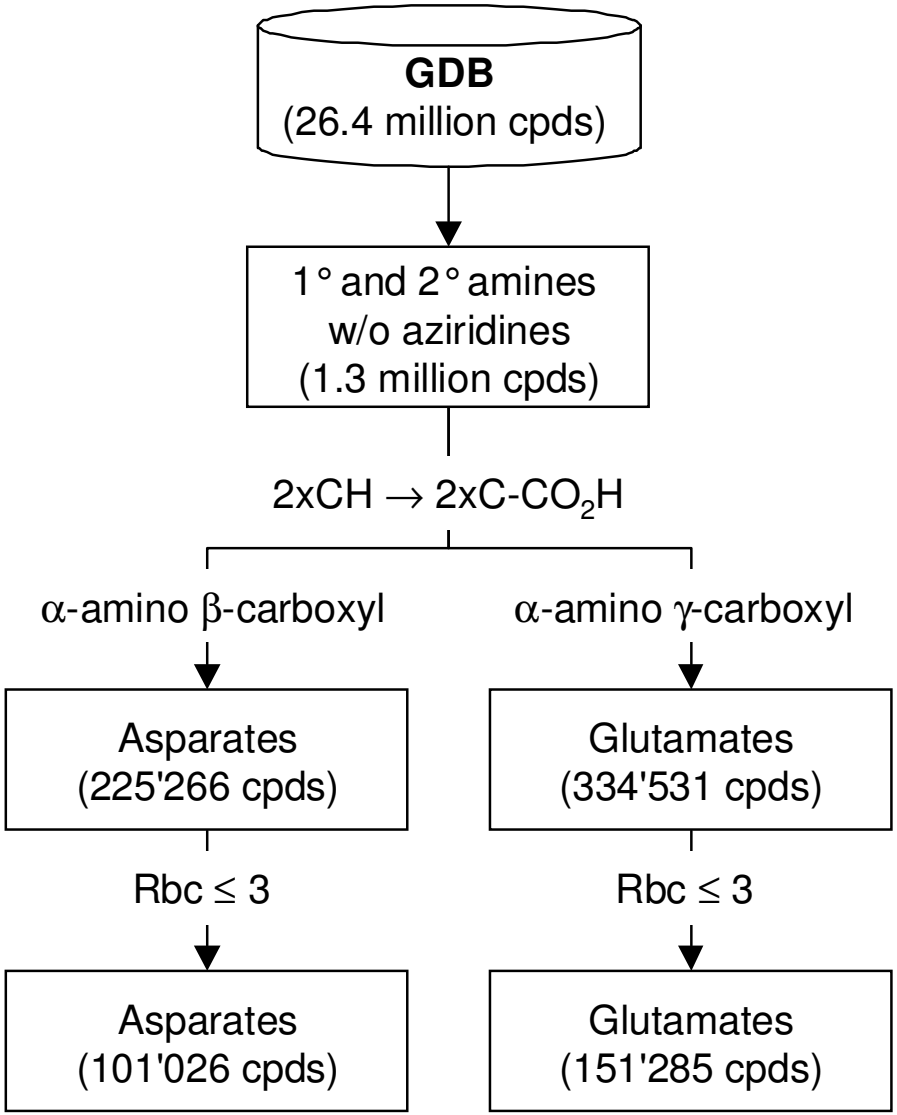
7 (L-3,4-MPDC)
BE = -9.5 kcal/mol
K_i=1 μM

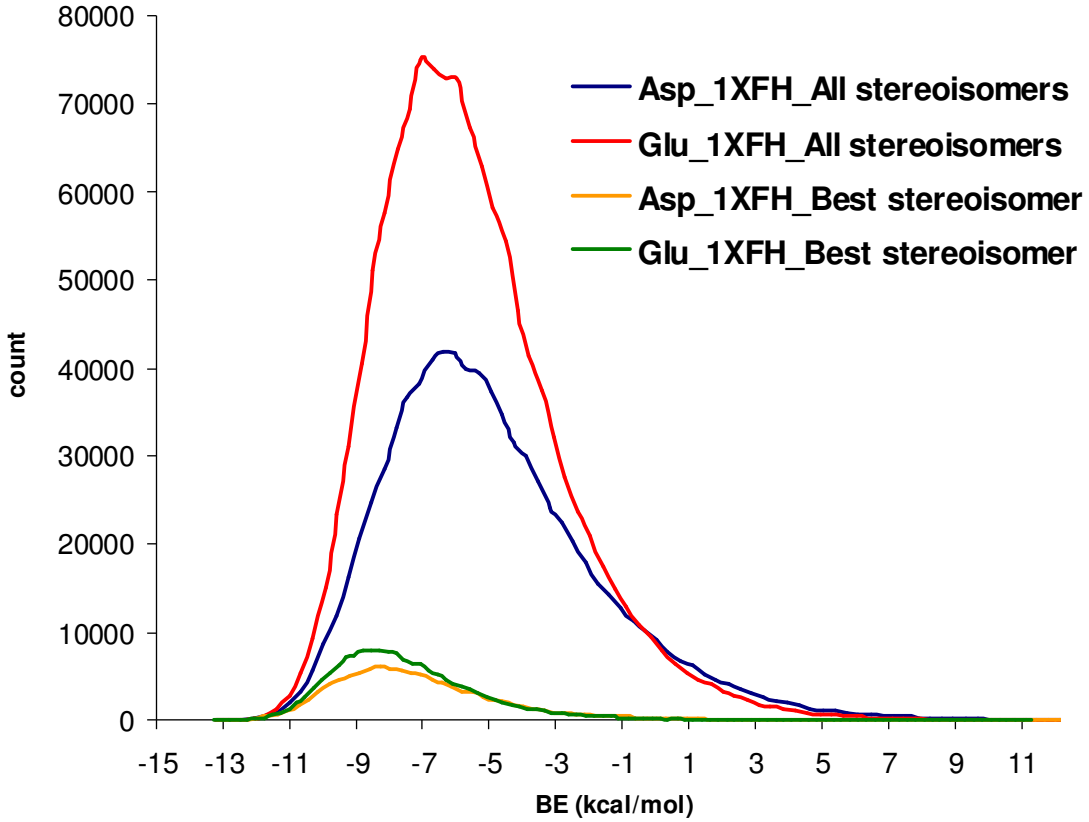


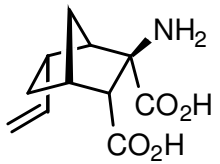
8 (Azabicyclo-ODD)
BE = -10.0 kcal/mol
K_i=52 μM



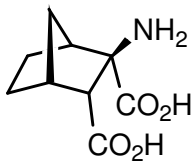
9 (WAY-855)
BE = -10.7 kcal/mol
IC_{50}=1μM


GDB
(26.4 million cpds)
1° and 2° amines
w/o aziridines
(1.3 million cpds)
2xCH → 2xC-CO2H
α-amino β-carboxyl
α-amino γ-carboxyl
Asparates
(225'266 cpds)
Glutamates
(334'531 cpds)
Rbc ≤ 3
Rbc ≤ 3
Asparates
(101'026 cpds)
Glutamates
(151'285 cpds)

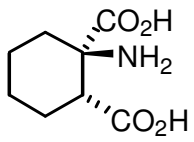




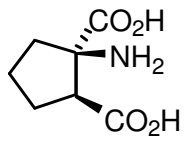
8 (rank 1)
-13.34 kcal/mol



9d (rank 19)
-12.52 kcal/mol

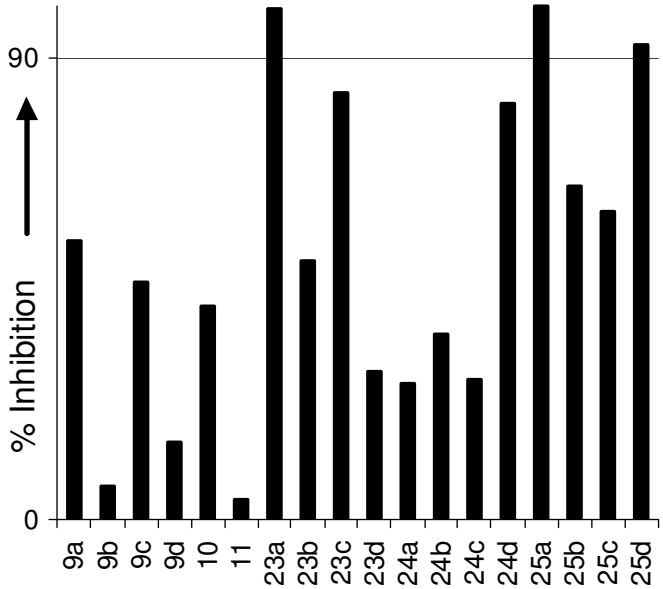


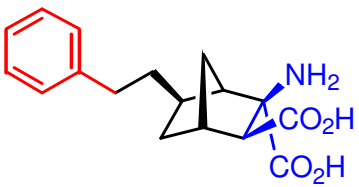
10 (rank 4888)
-10.46 kcal/mol



11 (rank 5176)
-10.42 kcal/mol

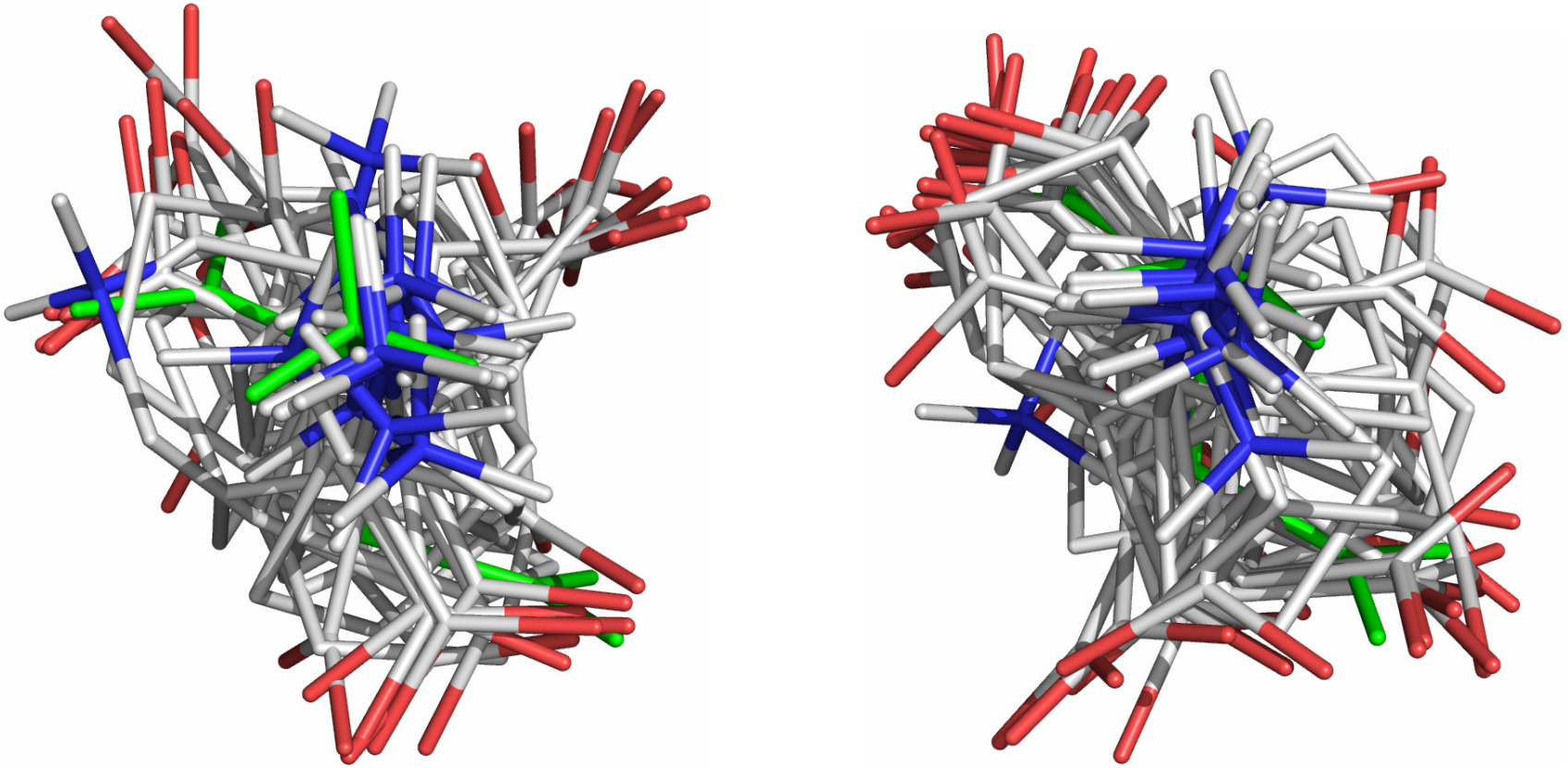
Glutamate uptake assay in Oocytes



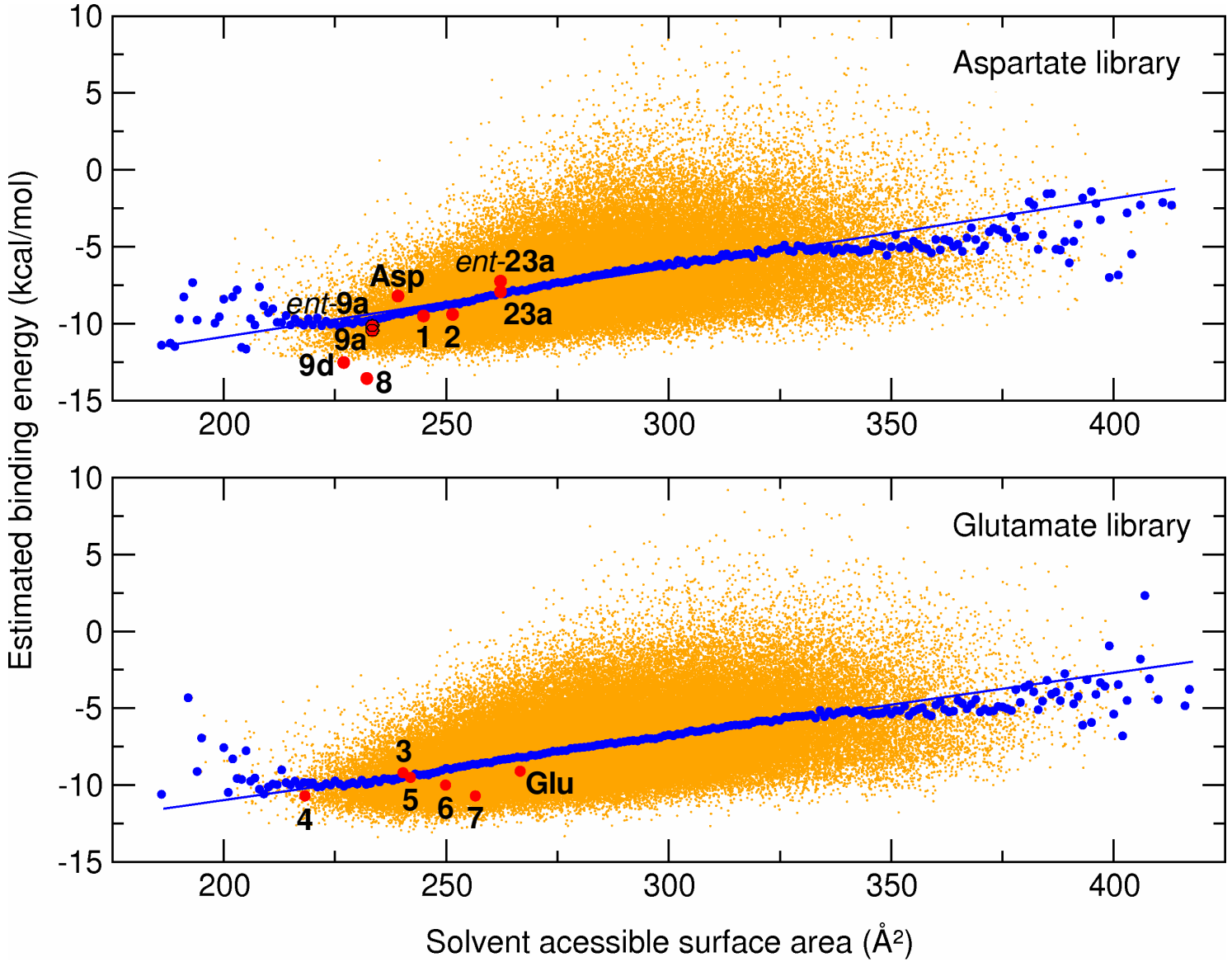


	IC_{50} (GLT-1) [μM]	IC_{50} (EAAC1) [μM]
rac-**23a** (vinyl)	130 ± 70	no inhibition
rac-**25a** (phenethyl)	1.4 ± 0.7	no inhibition
rac-**25d** (phenethyl)	19 ± 5	no inhibition
rac-**28a** (propyl)	25 ± 3	no inhibition
rac-**28b** (butyl)	14 ± 8	n. d.
rac-**28c** (*o*-HOPh)	21 ± 11	no inhibition
rac-**28d** (*p*-ClPh)	17 ± 11	no inhibition
7 (WAY-855)[22]	1.3	53
17 (L-TBOA)[41]	3.8	7.0
TFB-TBOA[43]	0.017	0.3

Docking found consistent binding modes



Docking selected compact ligands



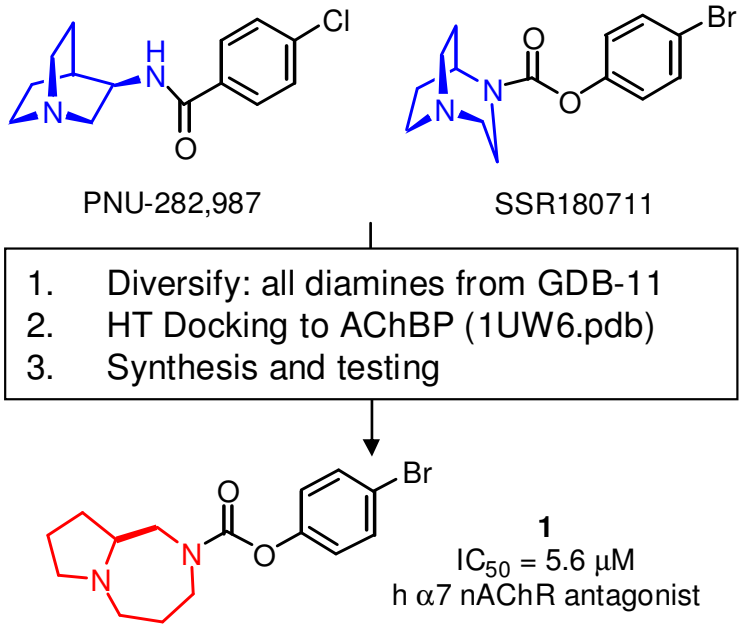
ACS Medicinal Chemistry Letters

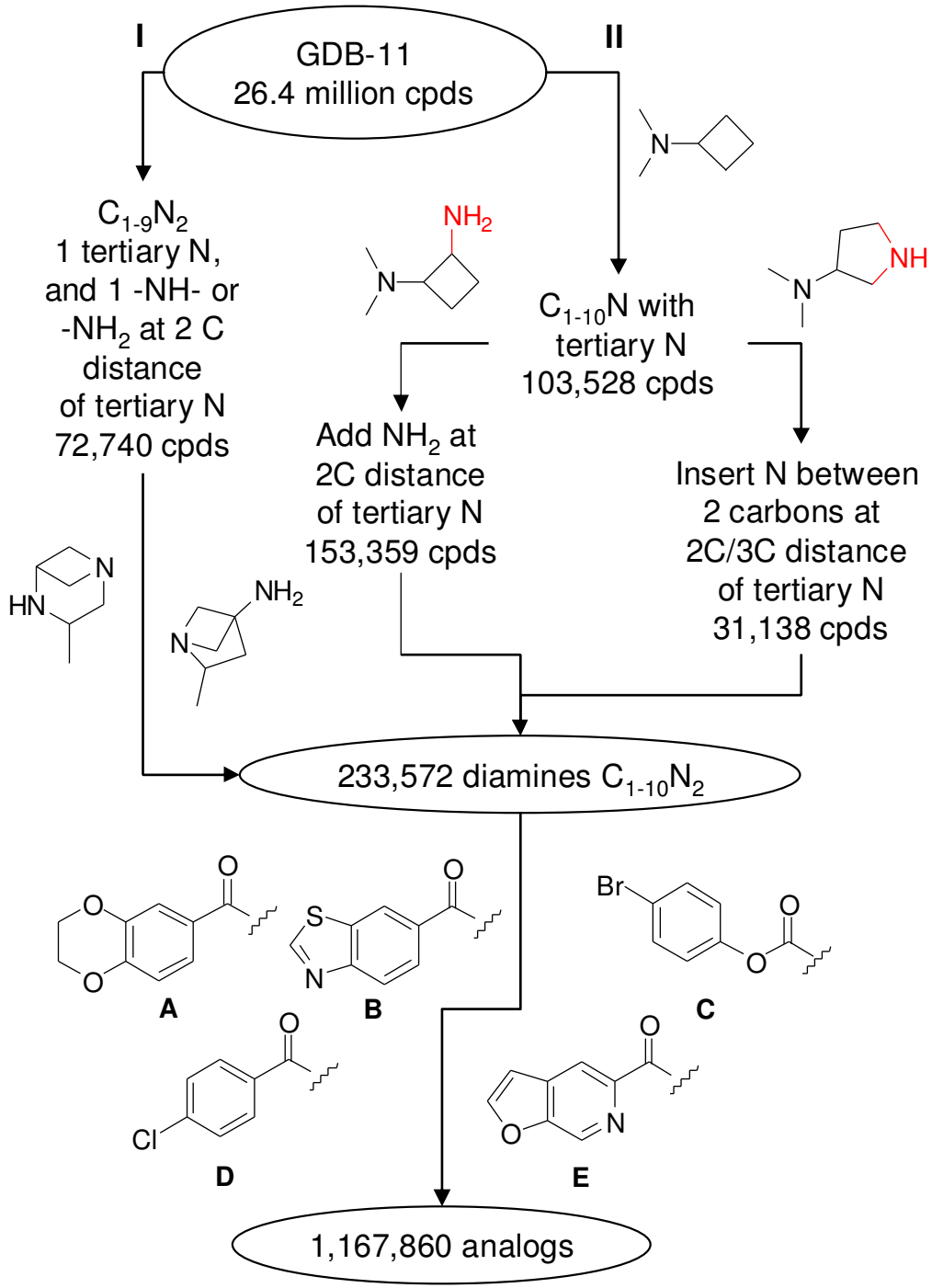
pubs.acs.org/acsmedchemlett

Exploring α7-Nicotinic Receptor Ligand Diversity by Scaffold Enumeration from the Chemical Universe Database GDB

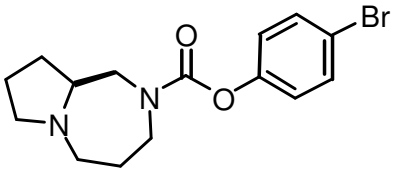
Noemi Garcia-Delgado,[†] Sonia Bertrand,[‡] Kong T. Nguyen,[†] Ruud van Deursen,[†] Daniel Bertrand,[‡] and Jean-Louis Reymond*,[†]

[†]Department of Chemistry and Biochemistry, University of Berne, Freiestrasse 3, 3012 Berne, Switzerland, and
[‡]Department of Neuroscience, Medical Faculty, 1, rue Michel Servet CH-1211 Geneva 4, Switzerland

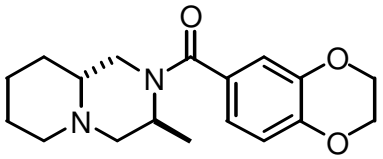



I
GDB-11
26.4 million cpds
II
$C_{1-9}N_2$
1 tertiary N,
and 1 -NH- or
-NH_2 at 2 C
distance
of tertiary N
72,740 cpds
C_{1-10}N with
tertiary N
103,528 cpds
Add NH_2 at
2C distance
of tertiary N
153,359 cpds
Insert N between
2 carbons at
2C/3C distance
of tertiary N
31,138 cpds
233,572 diamines $C_{1-10}N_2$
A
B
C
D
E
1,167,860 analogs

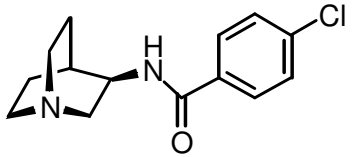
Compound	IC_{50} or EC_{50}	activity type
1	5.6 ± 1.7 μM	competitive antagonist to ACh
2	6.1 ± 1.5 μM	non competitive antagonist
3	7.0 ± 1.1 μM	mixed antagonist
4	7.2 ± 1.2 μM	mixed antagonist
6	4.4 μM[b)]	partial agonist[b)]



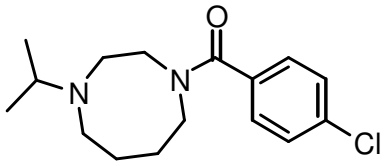
1
IC_{50} = 5.6 μM



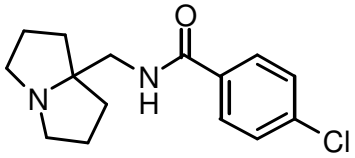
2



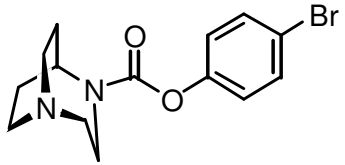
5 (PNU-282,987)



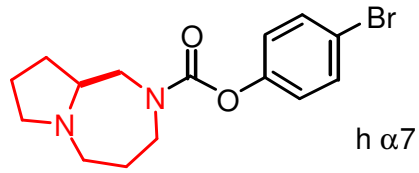
3



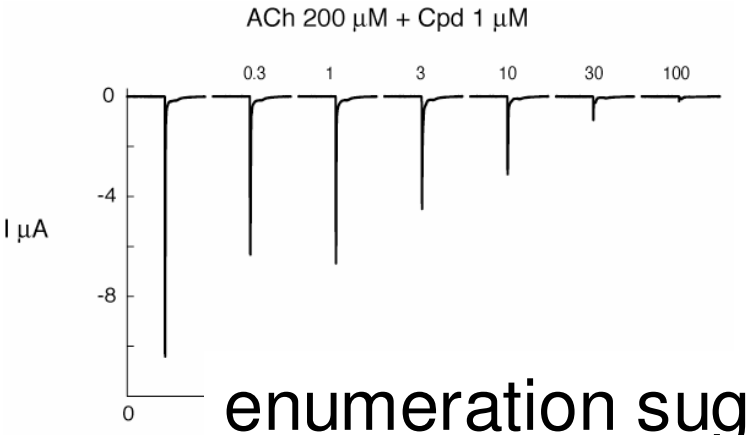
4

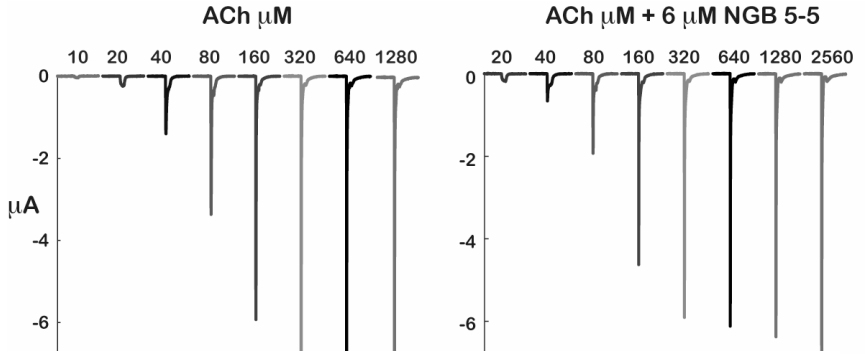


6 (SSR180711)
EC_{50} = 4.4 μM

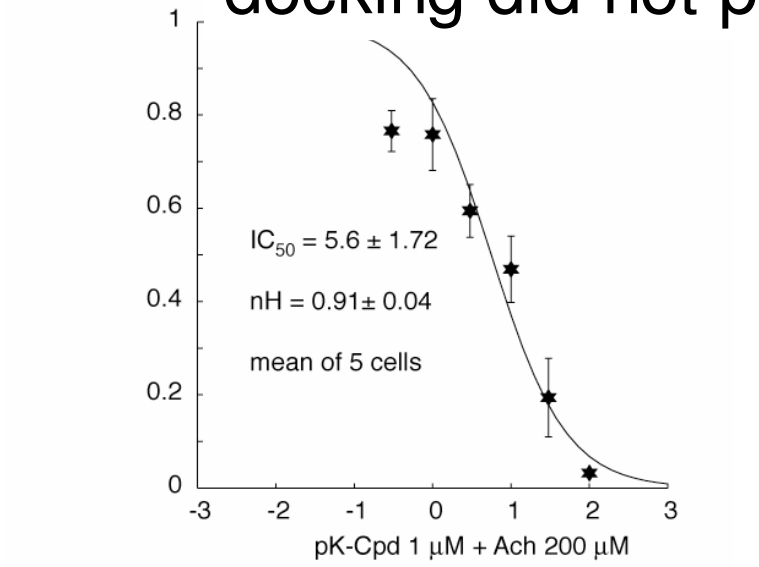


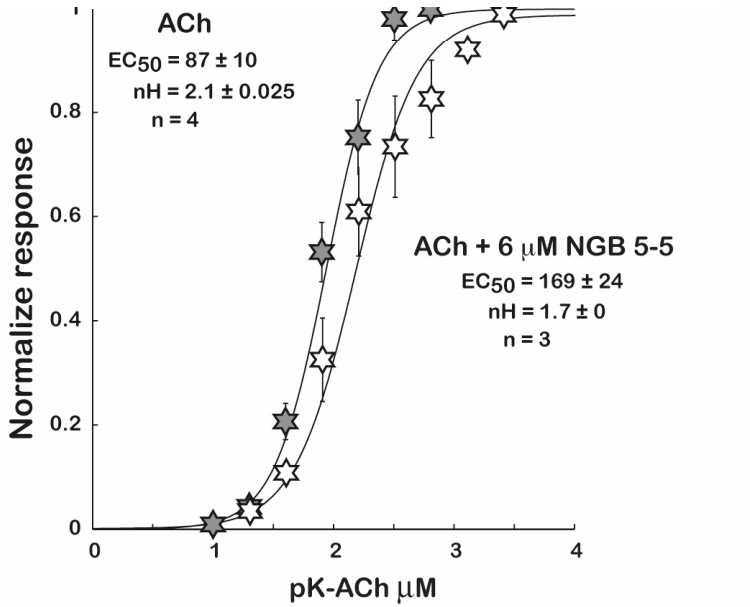
1
IC_{50} = 5.6 µM
h α7 nAChR antagonist





enumeration suggested new structures
docking did not predict binding mode and activity






Salahuddin Syed
Erika Lüthi
Noemi Garcia-Delgado
Justus Bürgi
Lise Brethous
Tobias Fink
Kong Nguyen
Ruud van Deursen
Lorenz Blum
Lars Ruddigkeit
Julian Schwartz
www.gdb.unibe.ch
Matthias Hediger, Yoshiro Suzuki (UniBE)
Daniel and Sonia Bertrand (UniGE)
University of Berne
Swiss National Science Foundation
ChemAxon
NCCR
TransCure
Excellence in Membrane Transport Research