

Supplementary material for 'LoFT: Similarity-driven focused library design'

J. Robert Fischer, Uta Lessel, and Matthias Rarey

Center for Bioinformatics Hamburg (ZBH) and Boehringer Ingelheim Pharma, Biberach

August 3, 2009

Contents

1	Parameters and data sets	4
2	Case studies	5
2.1	H3 query 1 (Sch49384)	5
2.1.1	Random	5
2.1.2	FTreesFS	6
2.1.3	LoFT	7
2.1.4	Leadlike	8
2.1.5	Diverse	9
2.1.6	Diverse Leadlike	10
2.2	H3 query 2	11
2.2.1	Random	11
2.2.2	FTreesFS	12
2.2.3	LoFT	13
2.2.4	Leadlike	14
2.2.5	Diverse	15
2.2.6	Diverse Leadlike	16
2.3	H3 query 3	17
2.3.1	Random	17
2.3.2	FTreesFS	18
2.3.3	LoFT	19
2.3.4	Leadlike	20
2.3.5	Diverse	21
2.3.6	Diverse Leadlike	22
2.4	H3 query 4	23
2.4.1	Random	23
2.4.2	FTreesFS	24
2.4.3	LoFT	25
2.4.4	Leadlike	26
2.4.5	Diverse	27
2.4.6	Diverse Leadlike	28
2.5	CDK2 query ZINC00582575	29
2.5.1	Random	29
2.5.2	FTreesFS	30
2.5.3	LoFT	31
2.5.4	Leadlike	32
2.5.5	Diverse	33
2.5.6	Diverse Leadlike	34
2.6	CDK2 query ZINC03591113	35
2.6.1	Random	35
2.6.2	FTreesFS	36
2.6.3	LoFT	37
2.6.4	Leadlike	38
2.6.5	Diverse	39
2.6.6	Diverse Leadlike	40
2.6.7	Diverse Leadlike Scoring	41
2.7	CDK2 query ZINC03814433	42
2.7.1	Random	42
2.7.2	FTreesFS	43
2.7.3	LoFT	44
2.7.4	Leadlike	45
2.7.5	Diverse	46
2.7.6	Diverse Leadlike	47
2.8	5HT2a query 1	48

2.8.1	Random	48
2.8.2	FTreesFS	49
2.8.3	LoFT	50
2.8.4	Leadlike	51
2.8.5	Diverse	52
2.8.6	Diverse Leadlike	53
2.9	5HT2a query Ketanserin	54
2.9.1	Random	54
2.9.2	FTreesFS	55
2.9.3	LoFT	56
2.9.4	Leadlike	57
2.9.5	Diverse	58
2.9.6	Diverse Leadlike	59

1 Parameters and data sets

The following parameters were used to generate the 5x5 and 4x4x4 sublibraries:

Settings:

```
Scoring:                arithmetic
Steps to go:            200000
Nof max unsuccessful steps: 200000
Nof solutions per core: 1
Random link selection:  every step
Seed mode:              take user defined seed
Seed:                   1
FTree similarity size filter
(min_nof_a(query) < > max_nof_a(query)): 0.2000 1.800000
Select start reagents by: random choice
Max nof reagents for each
    core link to consider: 1000000
Algorithm:              simulated annealing
Temperature:            1.000000000
Expon. cooling:          0.99992000
Alterate start value:   every step
```

For FTrees-FS runs, we used the following script:

```
MOL
  SELINIT !*
  CONVERT $(queries) 1 $(nof_queries) $(queries) n
FTREES
  READ % $(queries) 1 $(nof_queries)
FFSPACE
  READ $(fspace)
FFSEARCH
  FOREACH $(query) FROM 0 $(nof_queries)
    MCREATE $(query) 1000
    SELOUTP $(query).log
    MSEARCH 1000
  ENDFOR
```

The data was taken from the following sources and the reagents were taken from vendor catalogues:

The H3 queries molecules were taken from Celanire et al.[1]

The H3 core was taken from Lau et al.[2]

The CDK2 queries were taken from the DUD dataset[4]

The CDK2 cores were taken from Gray et al.[3]

The 5-HT2a query 1 was taken from Swain et al.[6], Ketanserin has pubmed entry 3822

The 5-HT2a core was taken from Smith et al.[5]

For generating the 2D drawings of the molecules we used 2DDraw[7]

For generating the histograms we used gnuplot 4.2(www.gnuplot.info)

2 Case studies

2.1 H3 query 1 (Sch49384)

2.1.1 Random

Randomly chosen sublibrary (starting point of optimization)

Seed 1: 0.6933

Seed 2177: 0.6359

Seed 35498: 0.6542

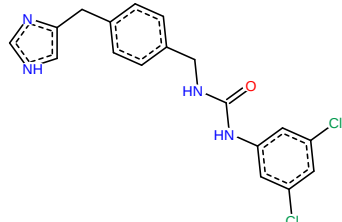
Seed 43: 0.6789

Seed 4543201: 0.6749

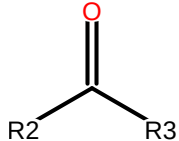
Average score: 0.6674

Solution no 1 with score 0.6933, seen after 1 step

Query:



Core:



Exceeded leadlike criteria:

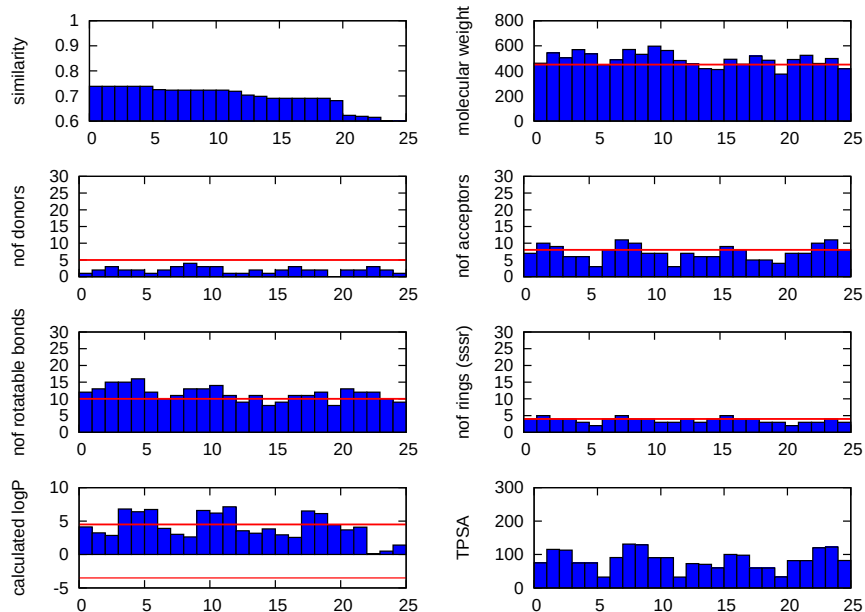
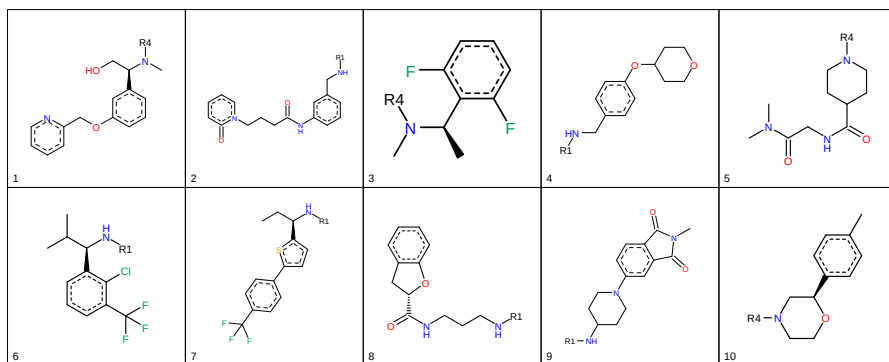
MW > 450

nof rings > 4

nof acceptors > 8

nof rotatable bonds > 10

logP > 4.5



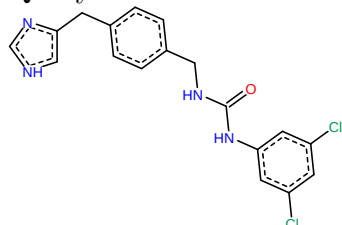
2.1.2 FTreesFS

First 5 (4 for cores with 3 links) reagents for each core link of the best 1000 products generated by FTreesFS

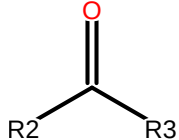
Seed 1: 0.9311

Solution no 1 with score 0.9311, seen after 107430 steps

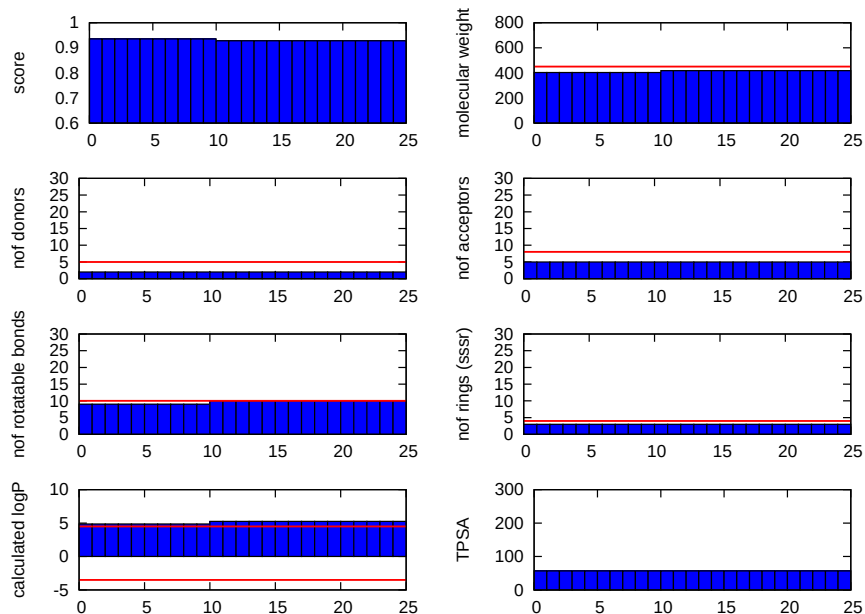
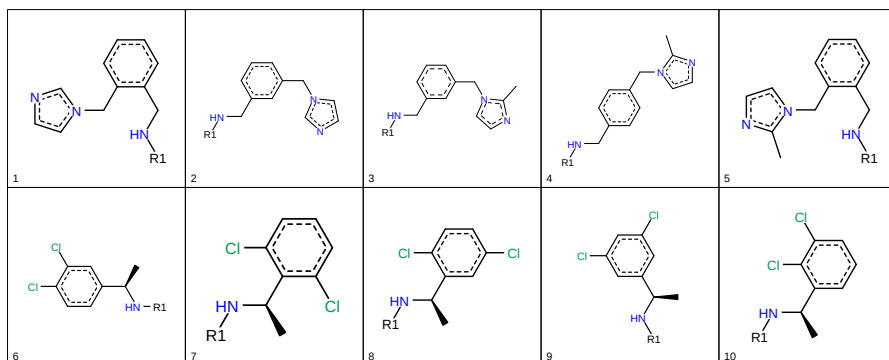
Query:



Core:



Exceeded leadlike criteria:
logP > 4.5



2.1.3 LoFT

LoFT using only ftree similarity

Seed 1: 0.9282

Seed 2177: 0.9237

Seed 35498: 0.9285

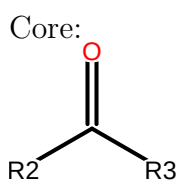
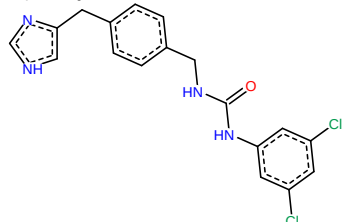
Seed 43: 0.9233

Seed 4543201: 0.9202

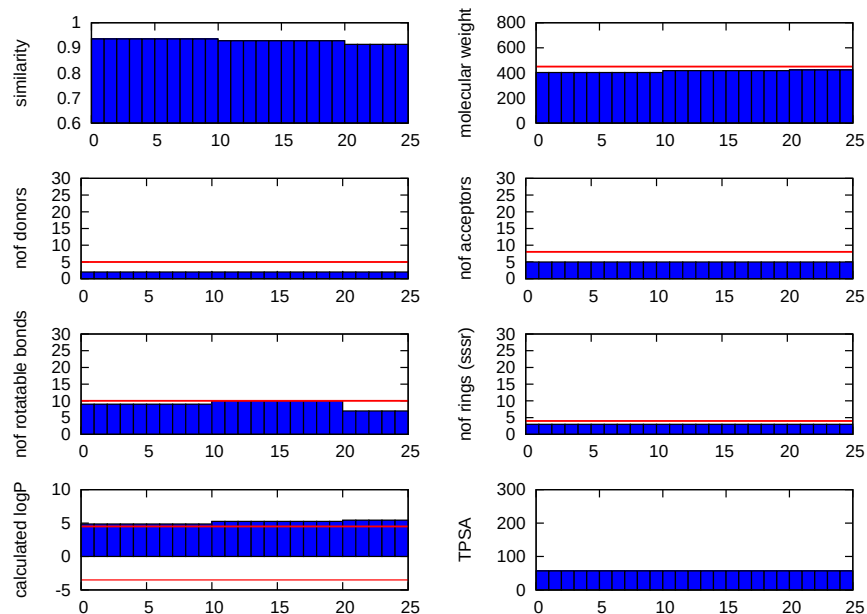
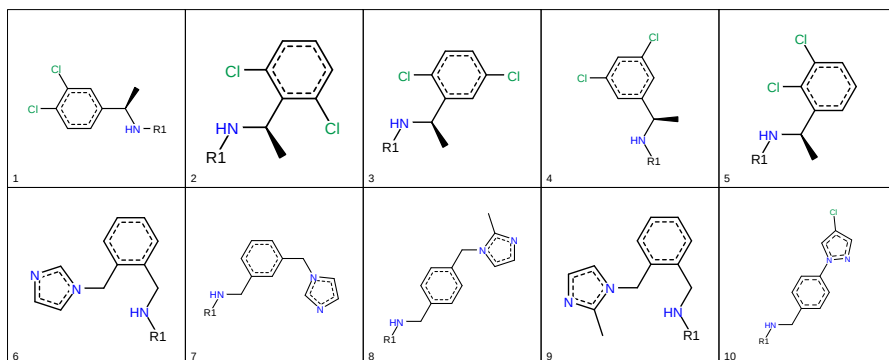
Average score: 0.9248

Solution no 1 with score 0.9282, seen after 186477 steps

Query:



Exceeded leadlike criteria:
logP > 4.5



2.1.4 Leadlike

Leadlike criteria were used as product filter.

Seed 1: 0.9120

Seed 2177: 0.9169

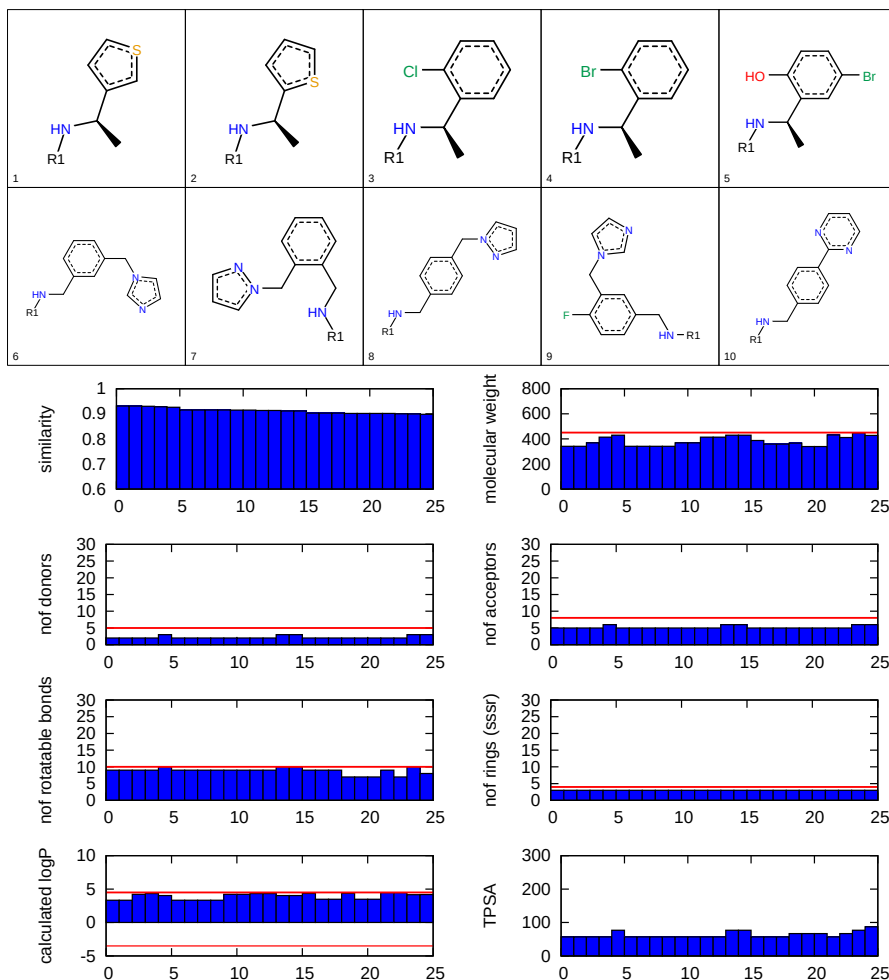
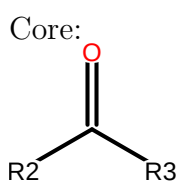
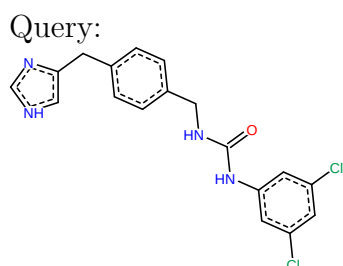
Seed 35498: 0.9034

Seed 43: 0.9177

Seed 4543201: 0.9143

Average score: 0.9129

Solution no 1 with score 0.9120, seen after 188824 steps



2.1.5 Diverse

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1.

Seed 1: 0.9110

Seed 2177: 0.9088

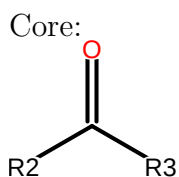
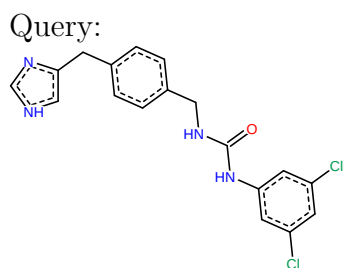
Seed 35498: 0.9129

Seed 43: 0.9130

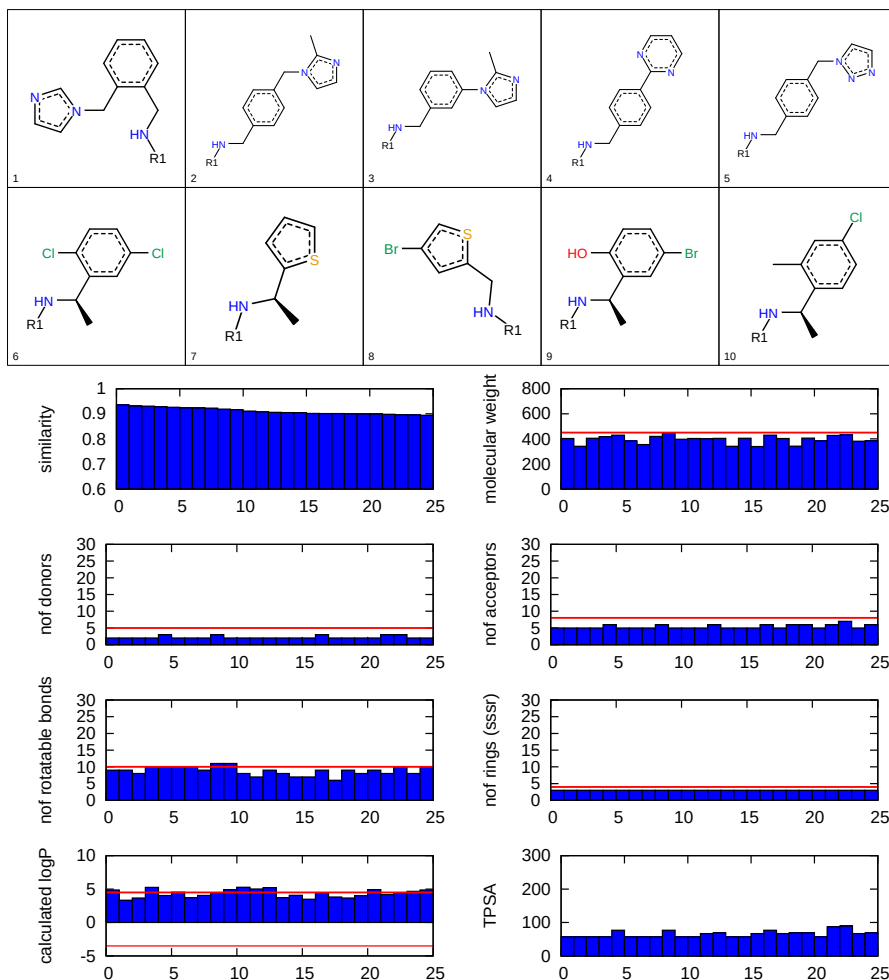
Seed 4543201: 0.9104

Average score: 0.9112

Solution no 1 with score 0.9110, seen after 195087 steps



Exceeded leadlike criteria:
 nof rotatable bonds > 10
 logP > 4.5



2.1.6 Diverse Leadlike

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1. Leadlike criteria were used as product filter.

Seed 1: 0.9012

Seed 2177: 0.9011

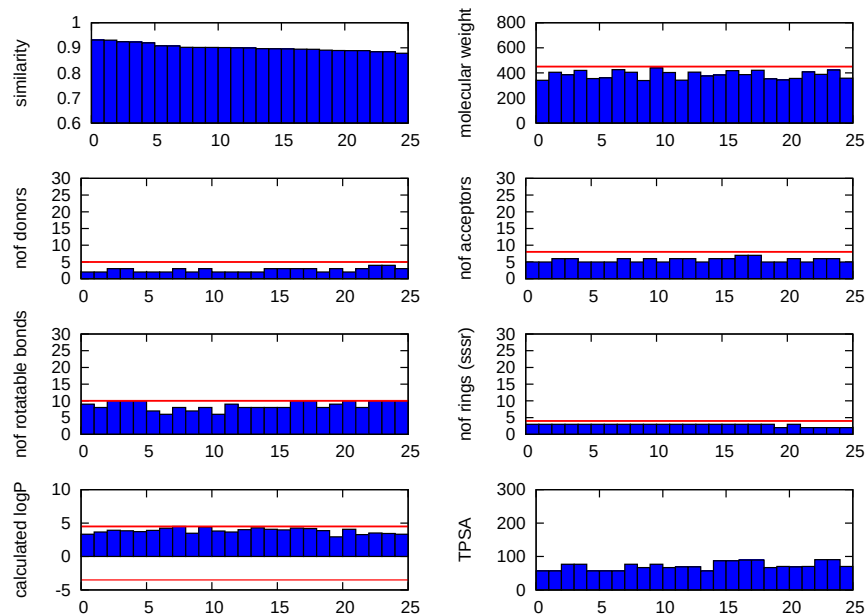
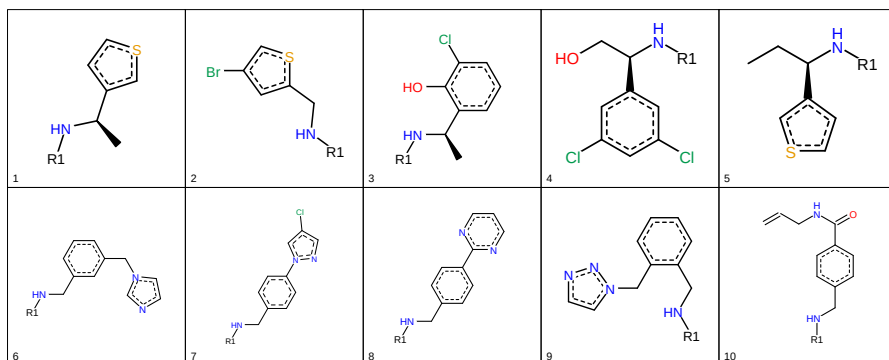
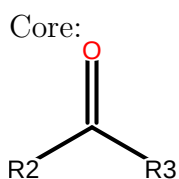
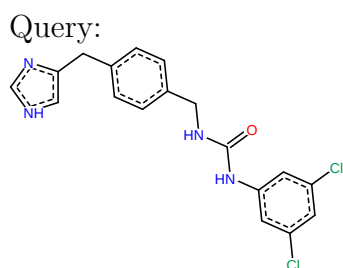
Seed 35498: 0.9009

Seed 43: 0.9041

Seed 4543201: 0.9018

Average score: 0.9018

Solution no 1 with score 0.9012, seen after 166801 step



2.2 H3 query 2

2.2.1 Random

Randomly chosen sublibrary (starting point of optimization)

Seed 1: 0.6201

Seed 2177: 0.5971

Seed 35498: 0.5979

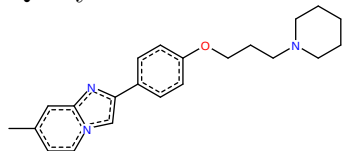
Seed 43: 0.5822

Seed 4543201: 0.6545

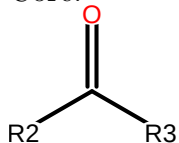
Average score: 0.6104

Solution no 1 with score 0.6201, seen after 1 step

Query:



Core:



Exceeded leadlike criteria:

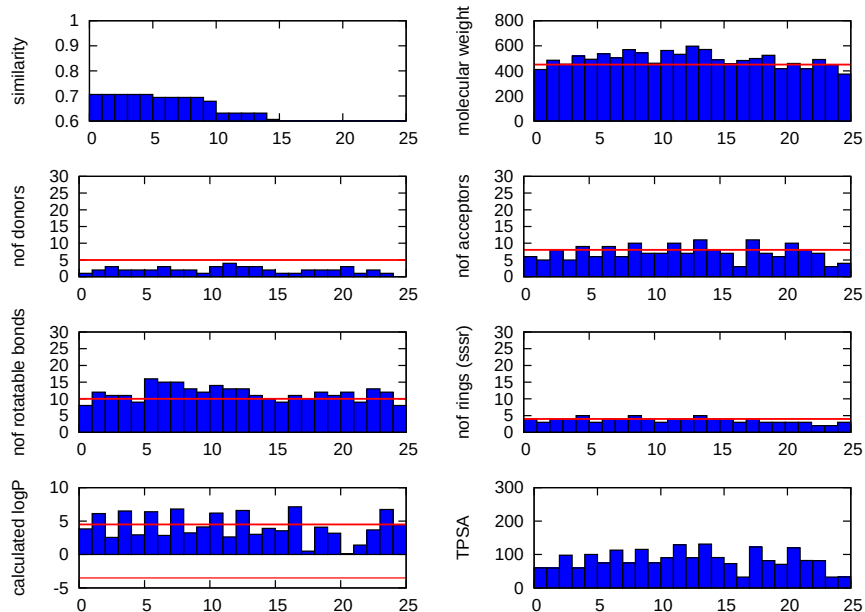
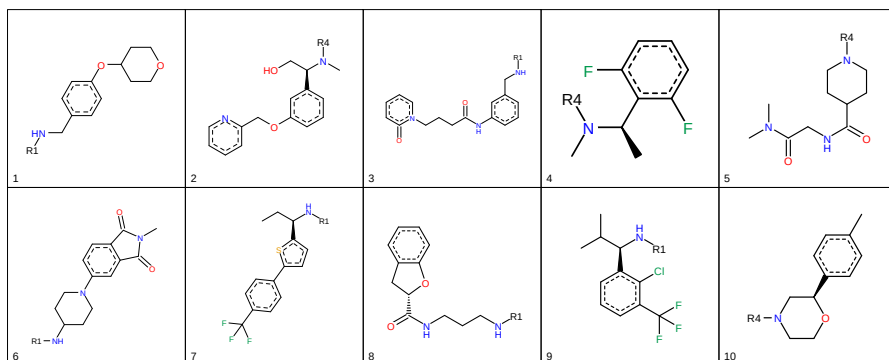
MW > 450

nof rings > 4

nof acceptors > 8

nof rotatable bonds > 10

logP > 4.5



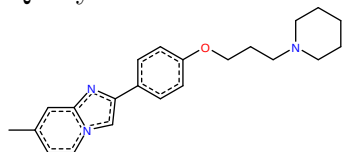
2.2.2 FTreesFS

First 5 (4 for cores with 3 links) reagents for each core link of the best 1000 products generated by FTreesFS

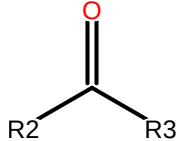
Seed 1: 0.8248

Solution no 1 with score 0.8248, seen after 52472 steps

Query:



Core:



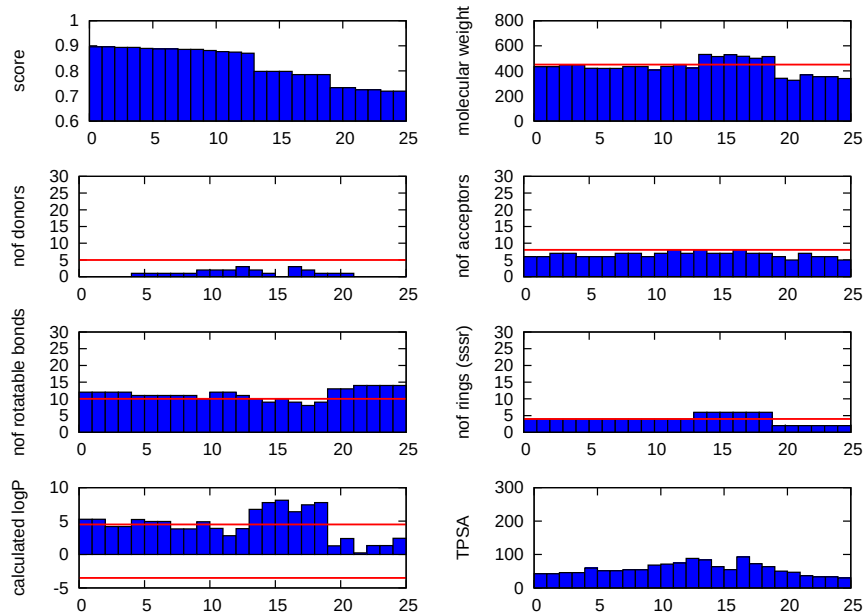
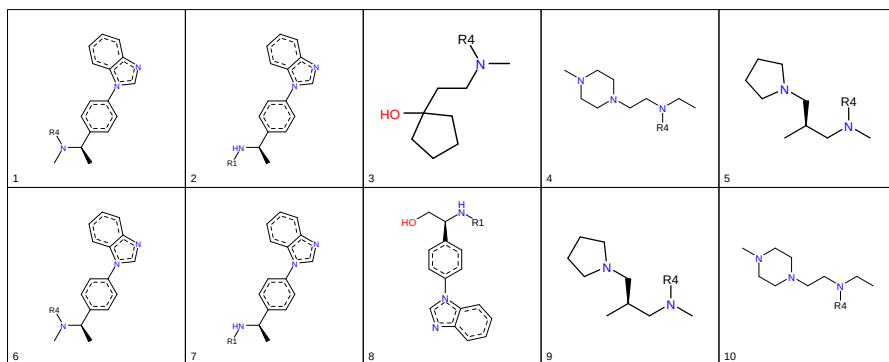
Exceeded leadlike criteria:

MW > 450

nof rings > 4

nof rotatable bonds > 10

logP > 4.5



2.2.3 LoFT

LoFT using only ftree similarity

Seed 1: 0.8857

Seed 2177: 0.8843

Seed 35498: 0.8844

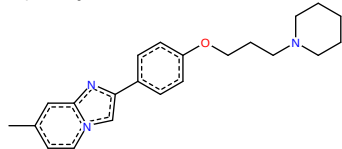
Seed 43: 0.8809

Seed 4543201: 0.8832

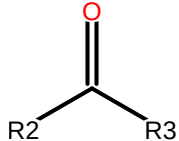
Average score: 0.8837

Solution no 1 with score 0.8857, seen after 199847 steps

Query:



Core:

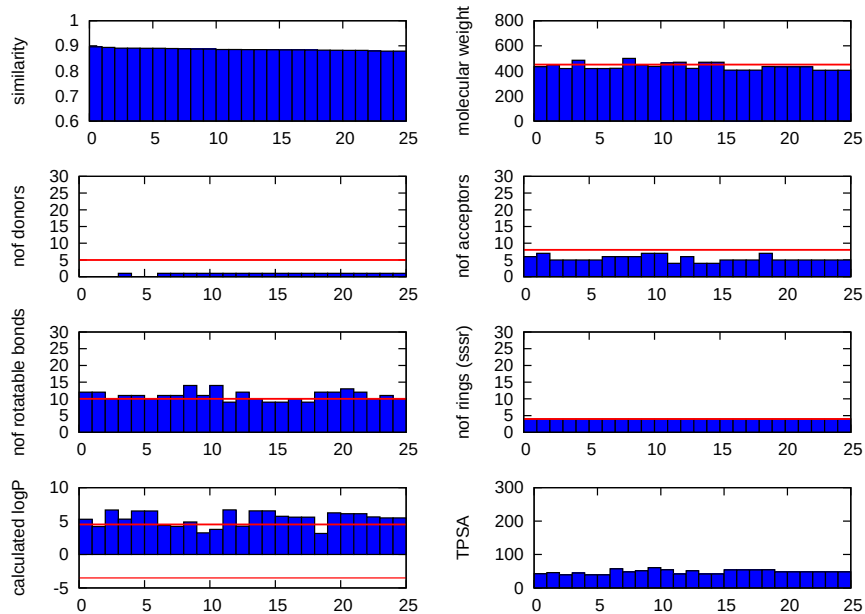
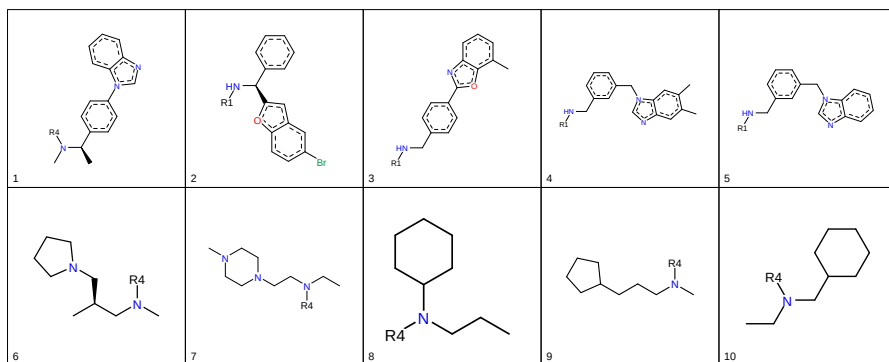


Exceeded leadlike criteria:

MW > 450

nof rotatable bonds > 10

logP > 4.5



2.2.4 Leadlike

Leadlike criteria were used as product filter.

Seed 1: 0.8574

Seed 2177: 0.8590

Seed 35498: 0.8053

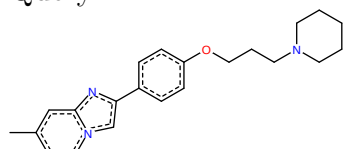
Seed 43: 0.8547

Seed 4543201: 0.8552

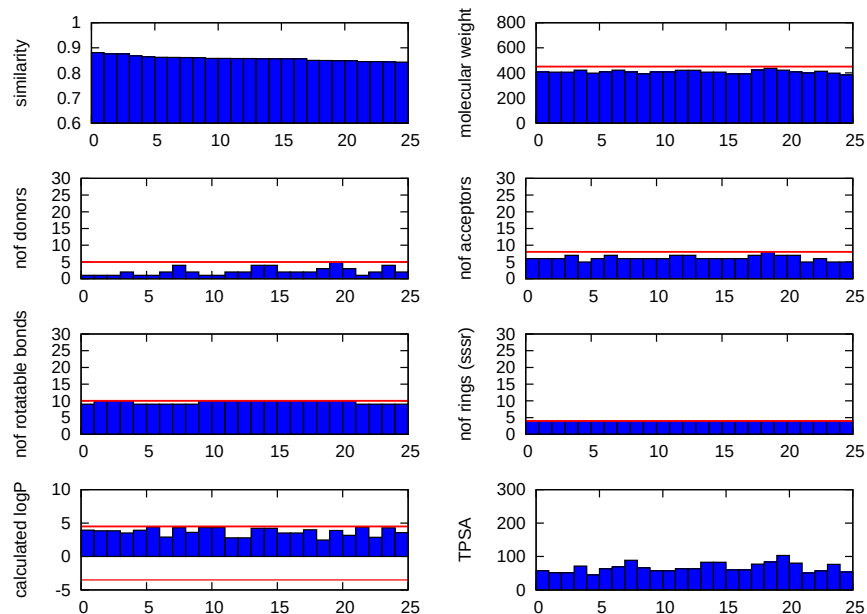
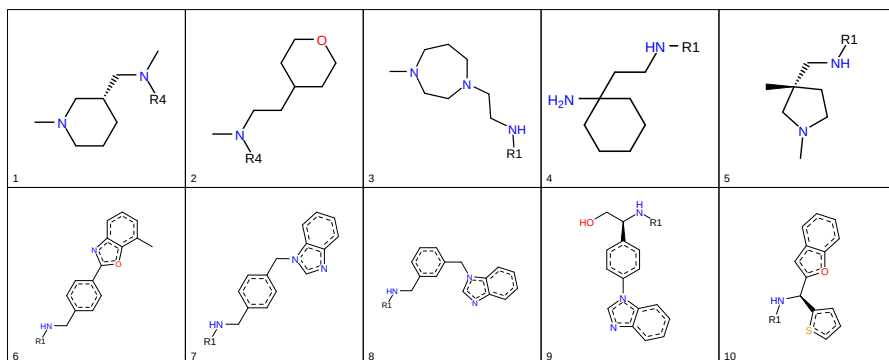
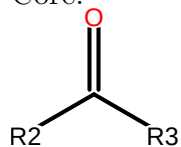
Average score: 0.8463

Solution no 1 with score 0.8574, seen after 187975 steps

Query:



Core:



2.2.5 Diverse

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1.

Seed 1: 0.8802

Seed 2177: 0.8836

Seed 35498: 0.8779

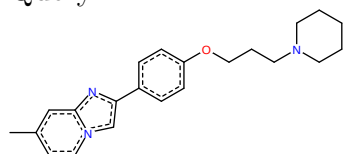
Seed 43: 0.8818

Seed 4543201: 0.8841

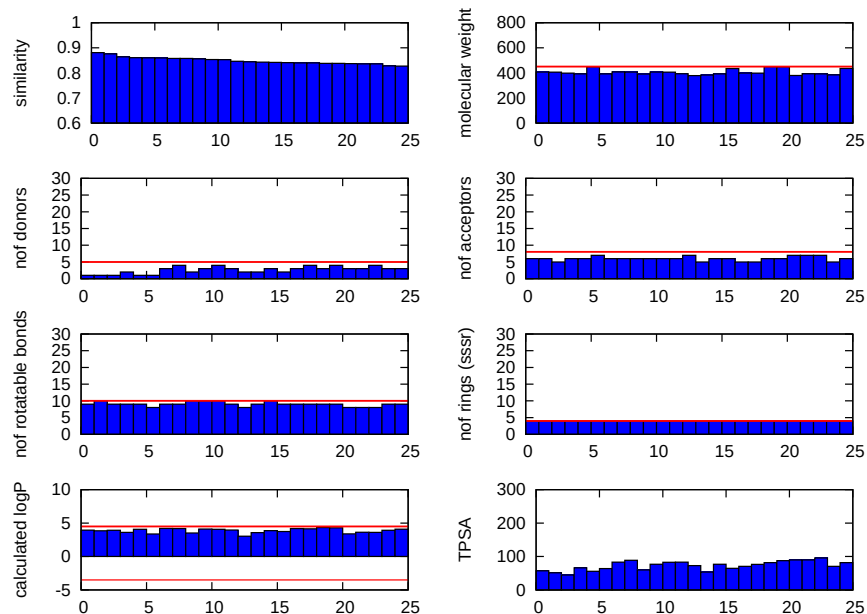
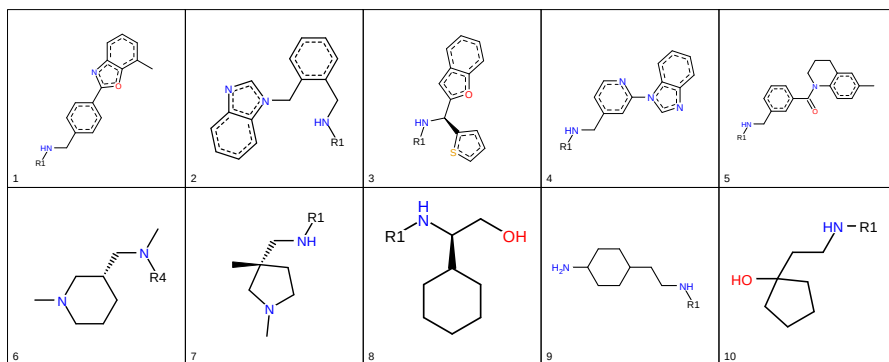
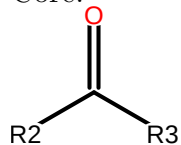
Average score: 0.8815

Solution no 1 with score 0.8802, seen after 166221 step

Query:



Core:



2.3 H3 query 3

2.3.1 Random

Randomly chosen sublibrary (starting point of optimization)

Seed 1: 0.5996

Seed 2177: 0.6220

Seed 35498: 0.5648

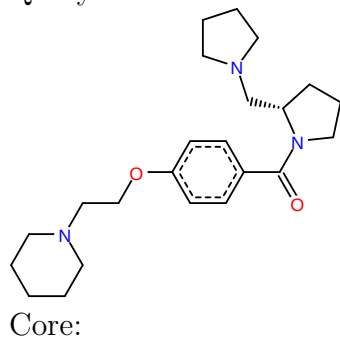
Seed 43: 0.5726

Seed 4543201: 0.6114

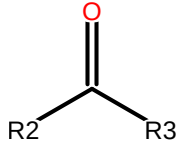
Average score: 0.5941

Solution no 1 with score 0.5996, seen after 1 step

Query:



Core:



Exceeded leadlike criteria:

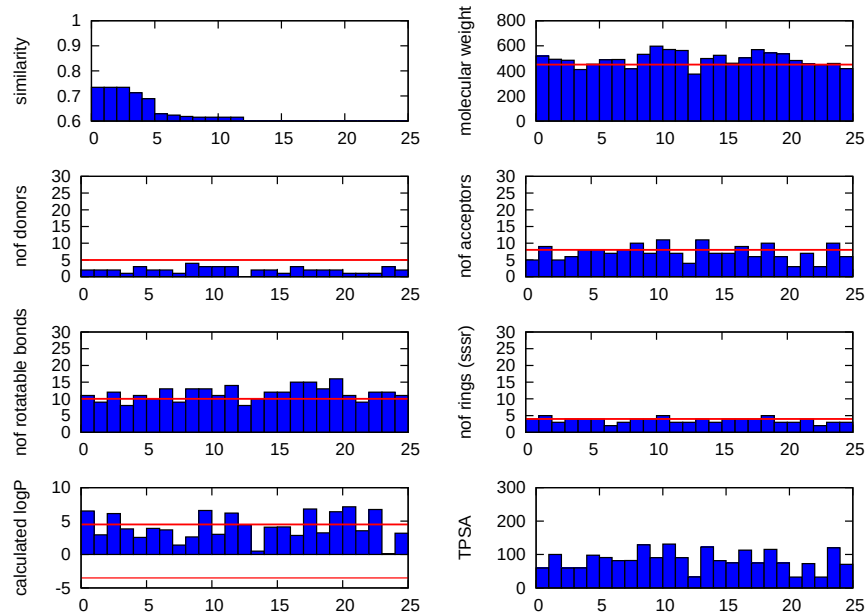
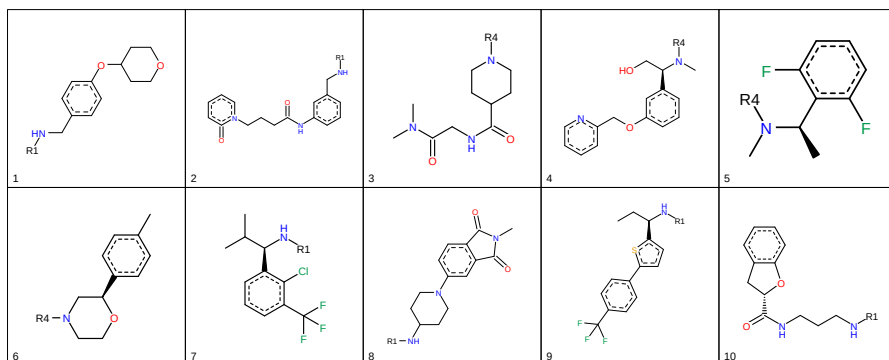
MW > 450

nof rings > 4

nof acceptors > 8

nof rotatable bonds > 10

logP > 4.5



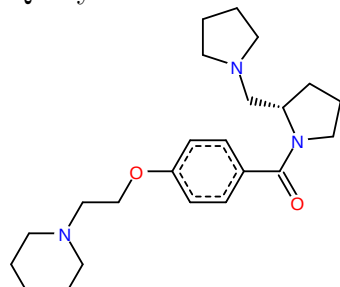
2.3.2 FTreesFS

First 5 (4 for cores with 3 links) reagents for each core link of the best 1000 products generated by FTreesFS

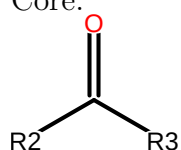
Seed 1: 0.9155

Solution no 1 with score 0.9155, seen after 105352 steps

Query:



Core:

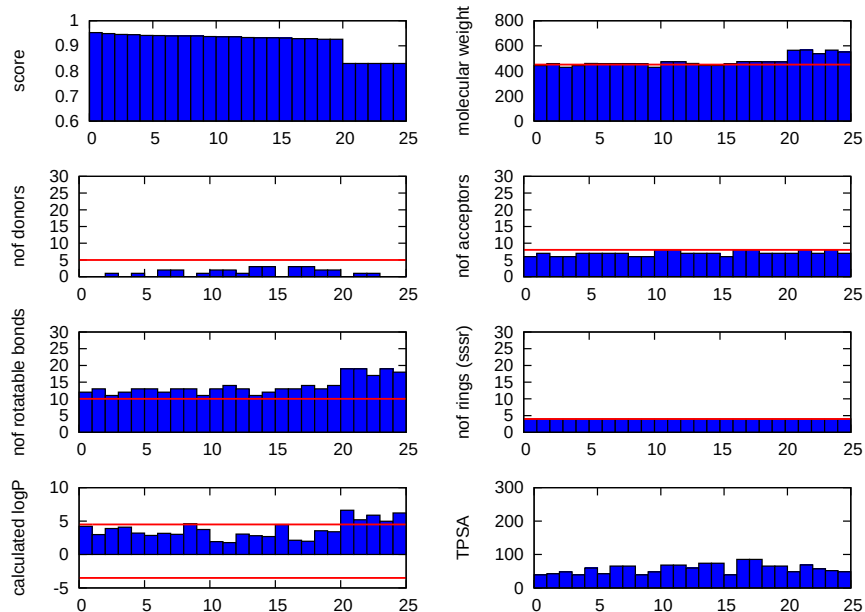
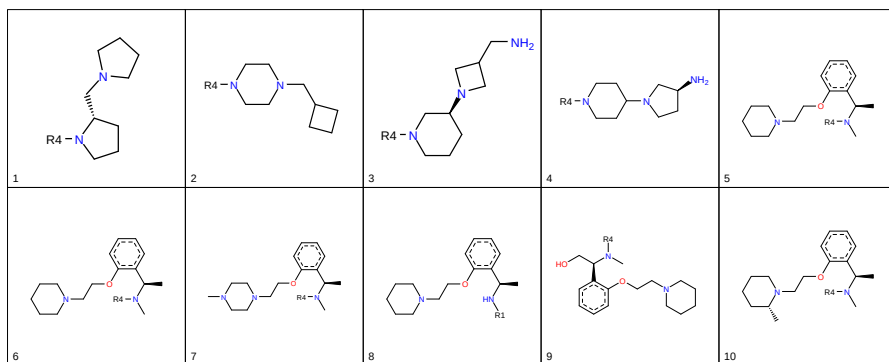


Exceeded leadlike criteria:

MW > 450

nof rotatable bonds > 10

logP > 4.5



2.3.3 LoFT

LoFT using only ftree similarity

Seed 1: 0.9392

Seed 2177: 0.9344

Seed 35498: 0.9362

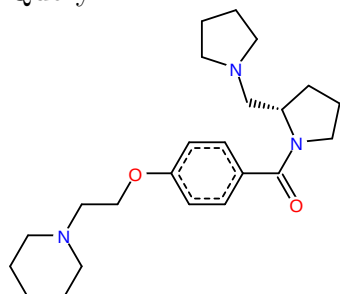
Seed 43: 0.9340

Seed 4543201: 0.9321

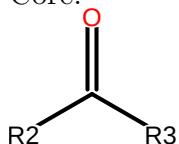
Average score: 0.9352

Solution no 1 with score 0.9392, seen after 196841 step

Query:



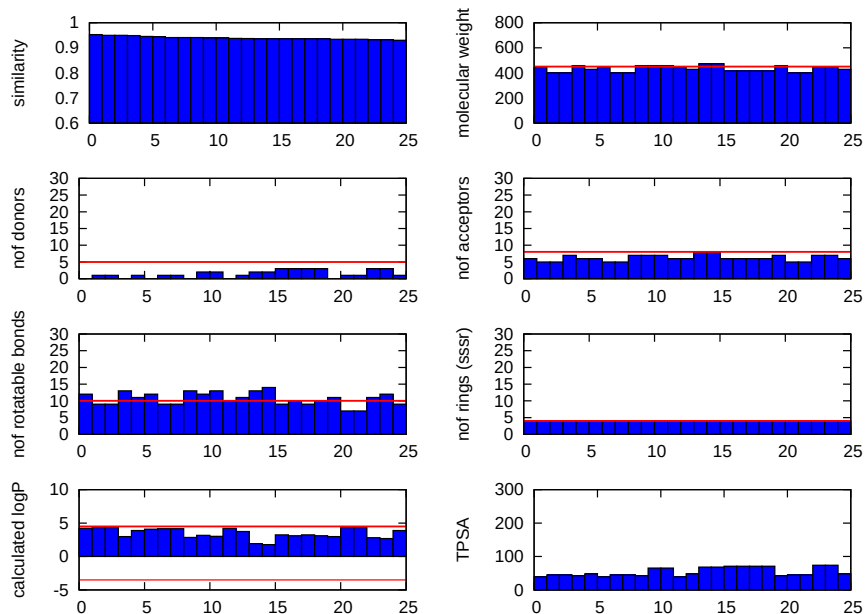
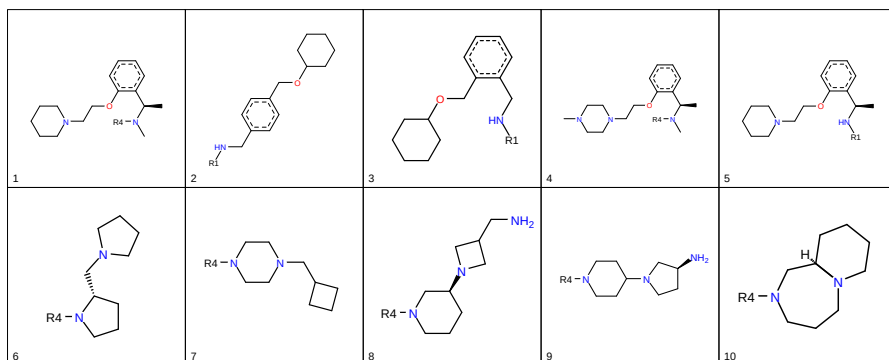
Core:



Exceeded leadlike criteria:

MW > 450

nof rotatable bonds > 10



2.3.4 Leadlike

Leadlike criteria were used as product filter.

Seed 1: 0.9223

Seed 2177: 0.9137

Seed 35498: 0.9105

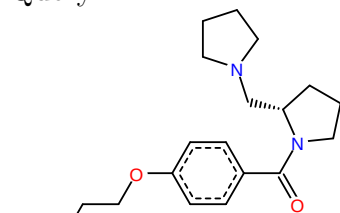
Seed 43: 0.9285

Seed 4543201: 0.9105

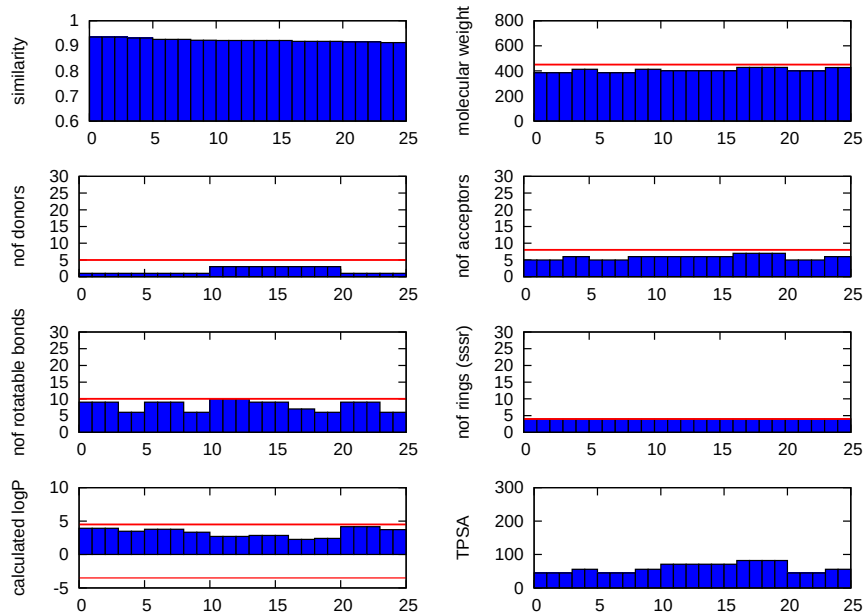
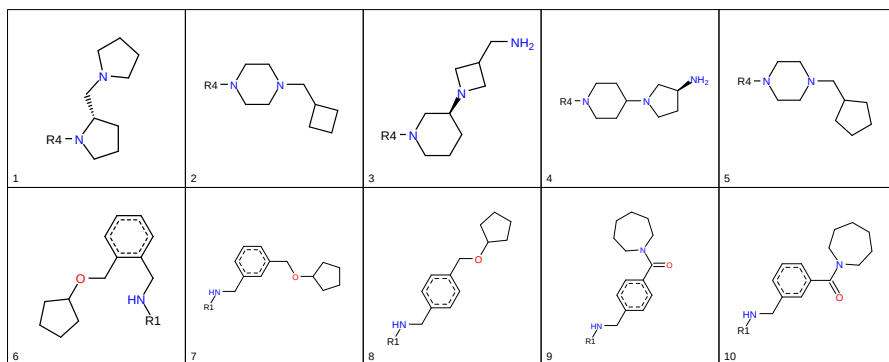
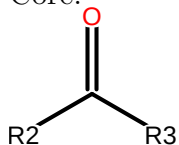
Average score: 0.9171

Solution no 1 with score 0.9223, seen after 194661 step

Query:



Core:



2.3.5 Diverse

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1.

Seed 1: 0.9247

Seed 2177: 0.9260

Seed 35498: 0.9220

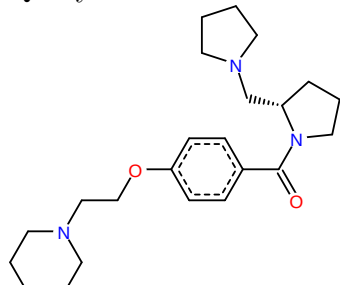
Seed 43: 0.9255

Seed 4543201: 0.9257

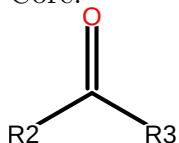
Average score: 0.9248

Solution no 1 with score 0.9247, seen after 186019 steps

Query:



Core:

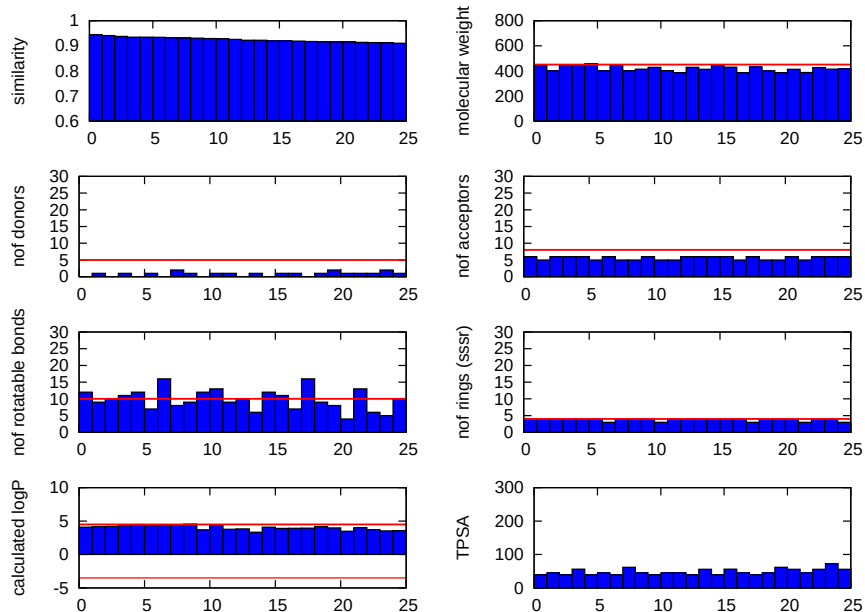
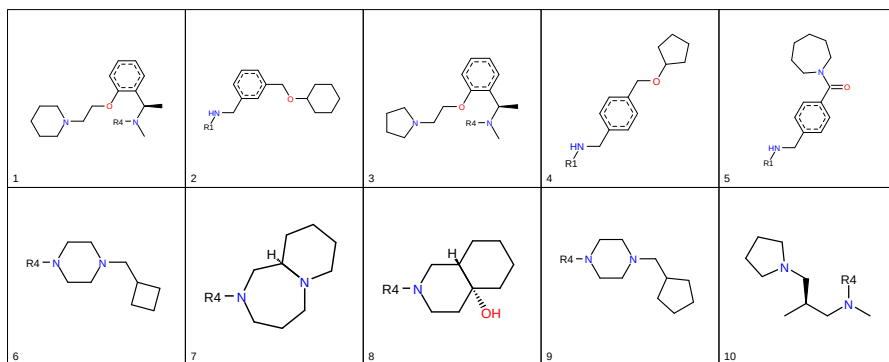


Exceeded leadlike criteria:

MW > 450

nof rotatable bonds > 10

logP > 4.5



2.3.6 Diverse Leadlike

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1. Leadlike criteria were used as product filter.

Seed 1: 0.9127

Seed 2177: 0.9041

Seed 35498: 0.9119

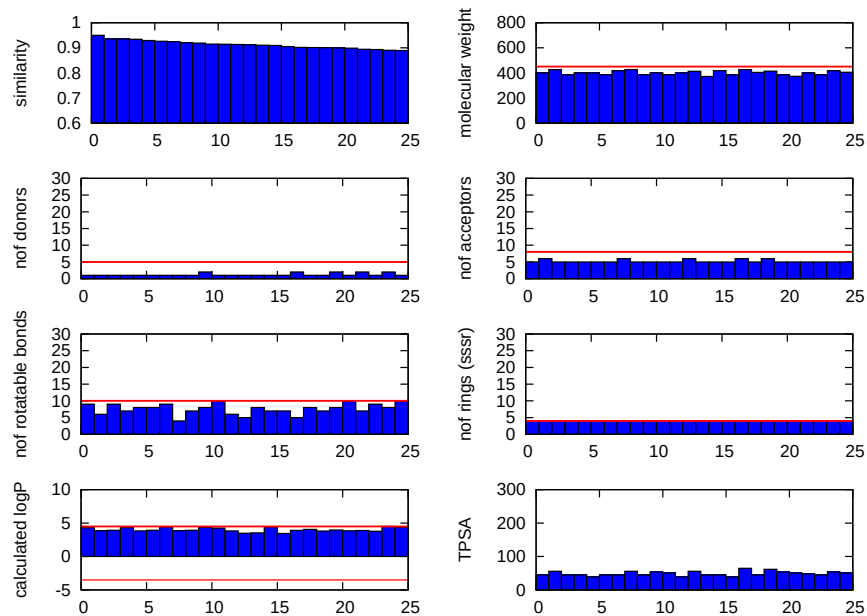
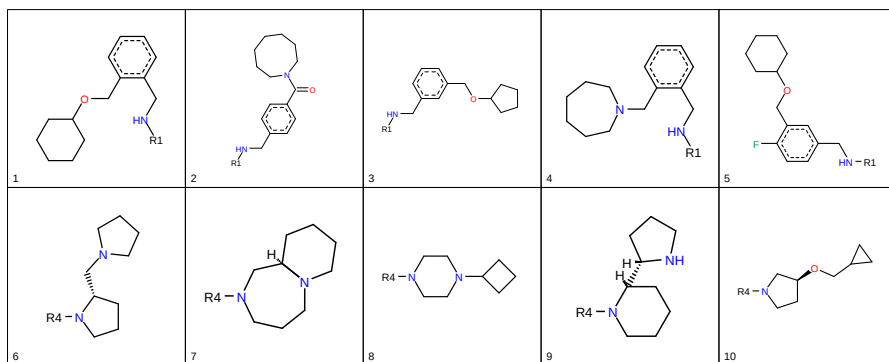
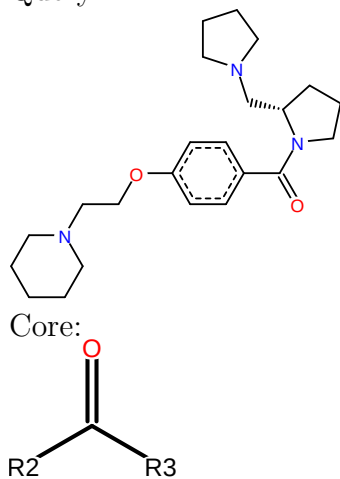
Seed 43: 0.9034

Seed 4543201: 0.9115

Average score: 0.9087

Solution no 1 with score 0.9127, seen after 164792 steps

Query:



2.4 H3 query 4

2.4.1 Random

Randomly chosen sublibrary (starting point of optimization)

Seed 1: 0.6119

Seed 2177: 0.5764

Seed 35498: 0.5969

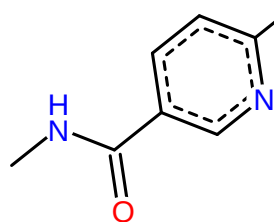
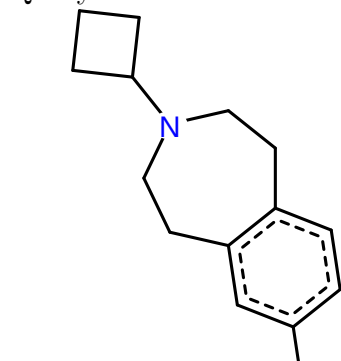
Seed 43: 0.5985

Seed 4543201: 0.6001

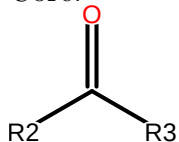
Average score: 0.5968

Solution no 1 with score 0.6119, seen after 0 steps

Query:



Core:



Exceeded leadlike criteria:

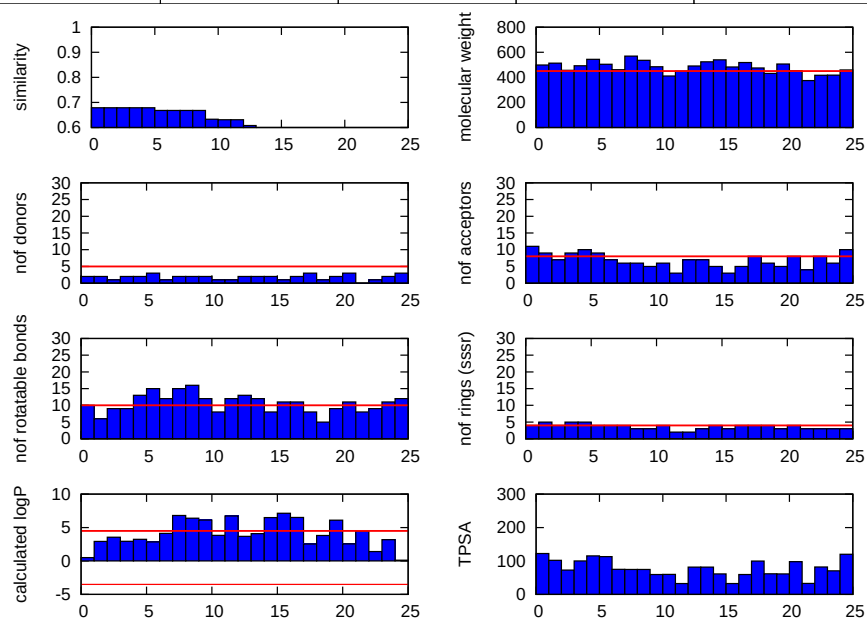
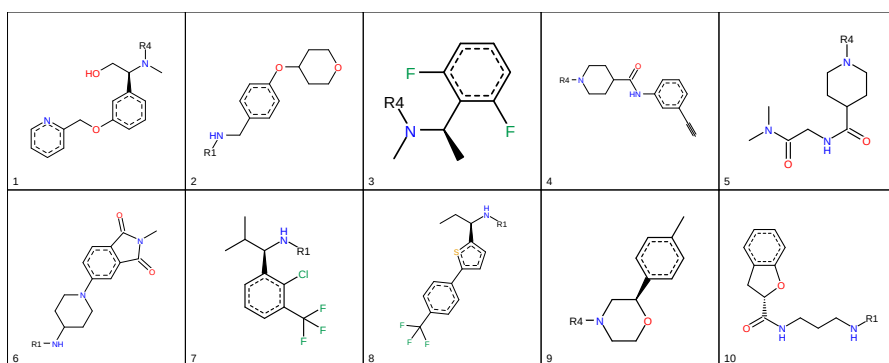
MW > 450

nof rings > 4

nof acceptors > 8

nof rotatable bonds > 10

logP > 4.5

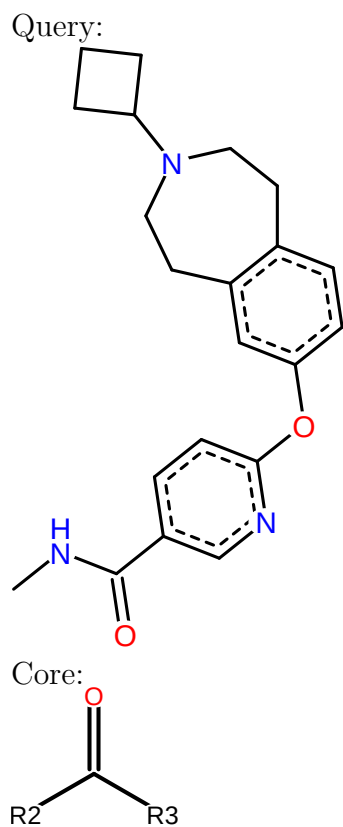


2.4.2 FTreesFS

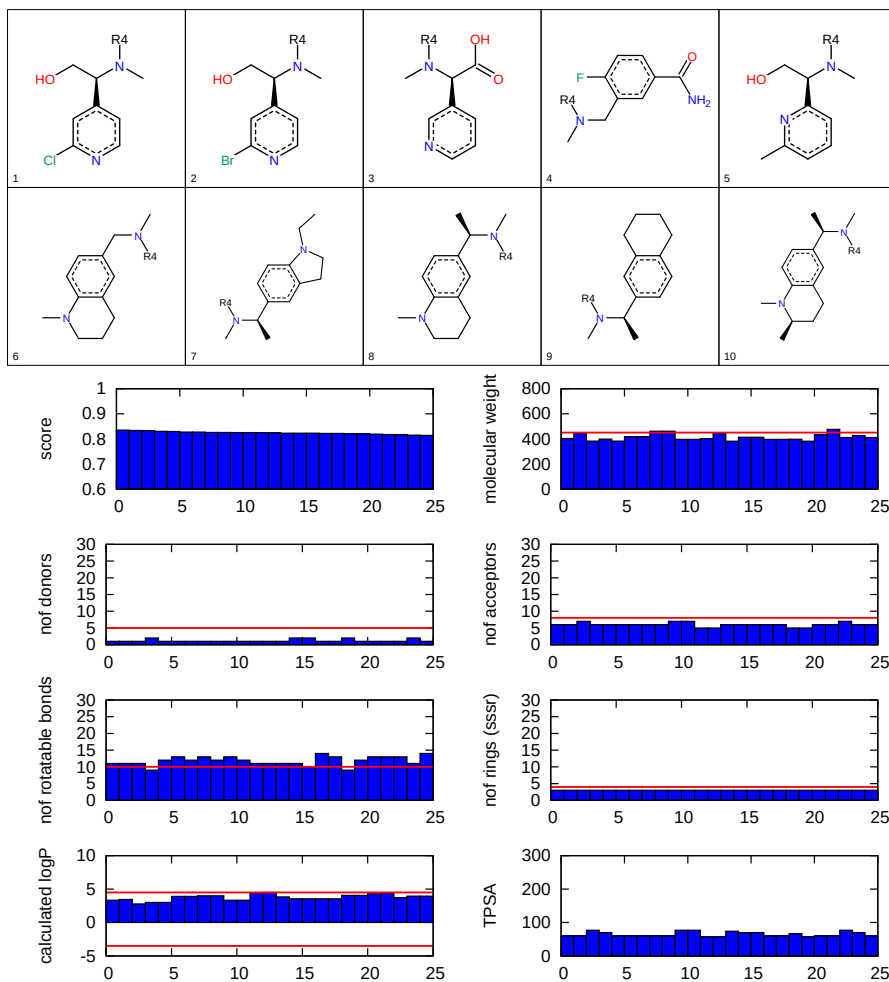
First 5 (4 for cores with 3 links) reagents for each core link of the best 1000 products generated by FTreesFS

Seed 1: 0.8238

Solution no 1 with score 0.8238, seen after 35142 steps



Exceeded leadlike criteria:
 MW > 450
 nof rotatable bonds > 10
 logP > 4.5



2.4.3 LoFT

LoFT using only ftree similarity

Seed 1: 0.8443

Seed 2177: 0.8476

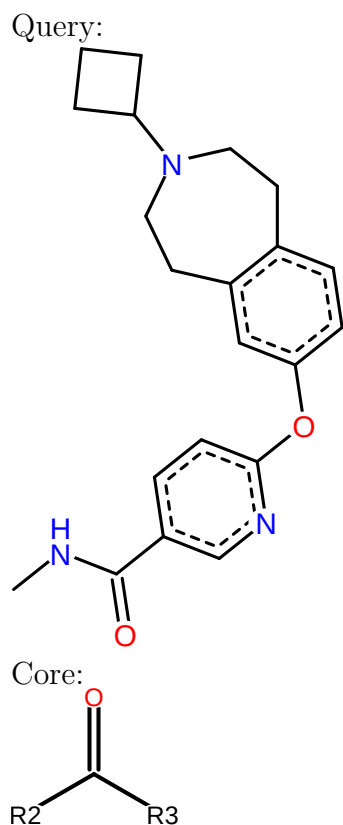
Seed 35498: 0.8467

Seed 43: 0.8462

Seed 4543201: 0.8426

Average score: 0.8455

Solution no 1 with score 0.8443, seen after 186817 steps



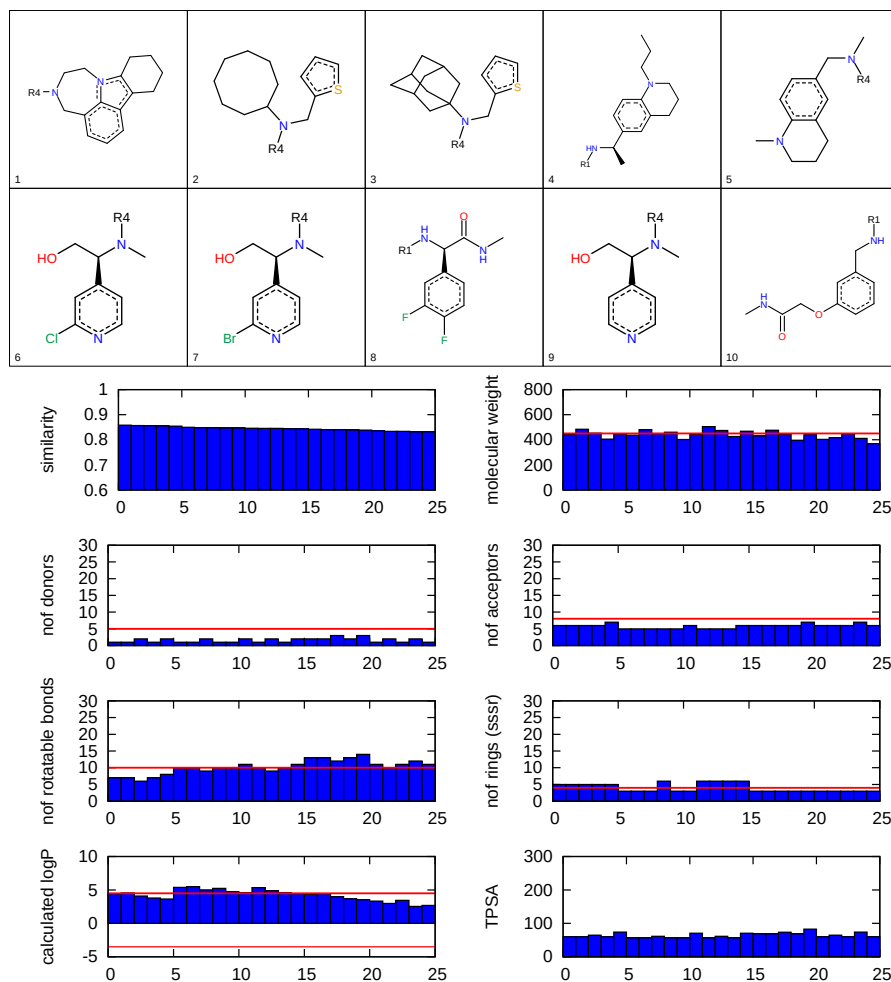
Exceeded leadlike criteria:

MW > 450

nof rings > 4

nof rotatable bonds > 10

logP > 4.5



2.4.4 Leadlike

Leadlike criteria were used as product filter.

Seed 1: 0.8148

Seed 2177: 0.8164

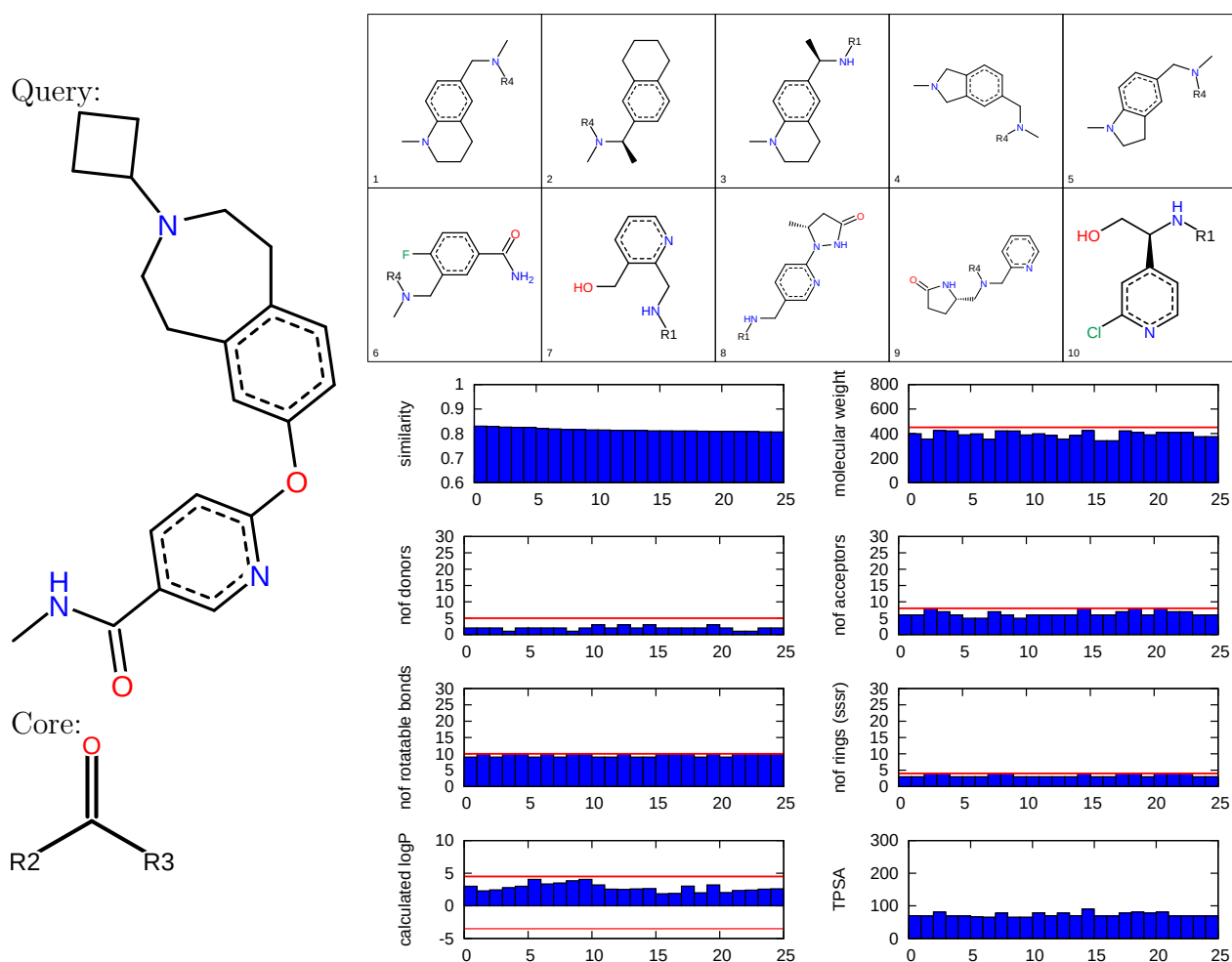
Seed 35498: 0.8092

Seed 43: 0.8175

Seed 4543201: 0.8139

Average score: 0.8144

Solution no 1 with score 0.8148, seen after 184559 steps



2.4.5 Diverse

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1.

Seed 1: 0.8470

Seed 2177: 0.8466

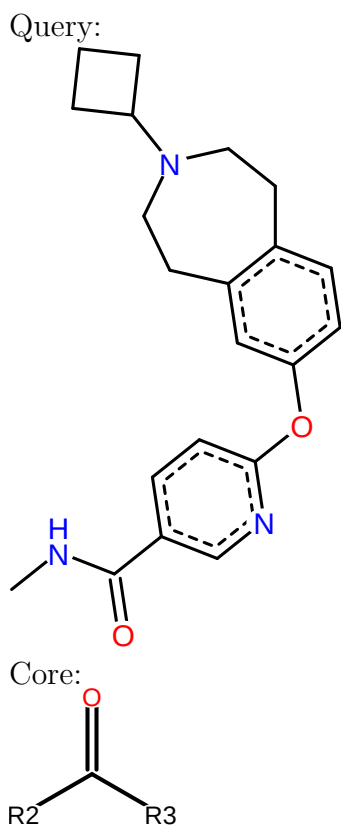
Seed 35498: 0.8450

Seed 43: 0.8446

Seed 4543201: 0.8415

Average score: 0.8449

Solution no 1 with score 0.8470, seen after 177364 steps



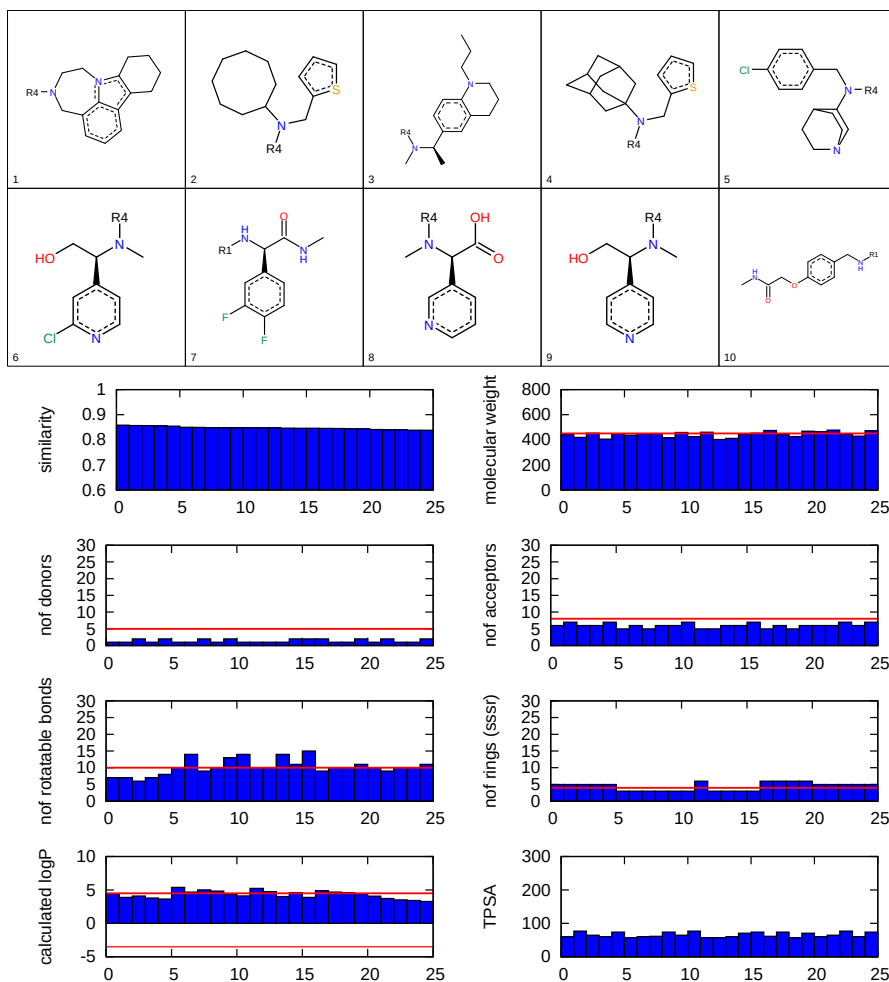
Exceeded leadlike criteria:

MW > 450

nof rings > 4

nof rotatable bonds > 10

logP > 4.5



2.4.6 Diverse Leadlike

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1. Leadlike criteria were used as product filter.

Seed 1: 0.8021

Seed 2177: 0.8061

Seed 35498: 0.8096

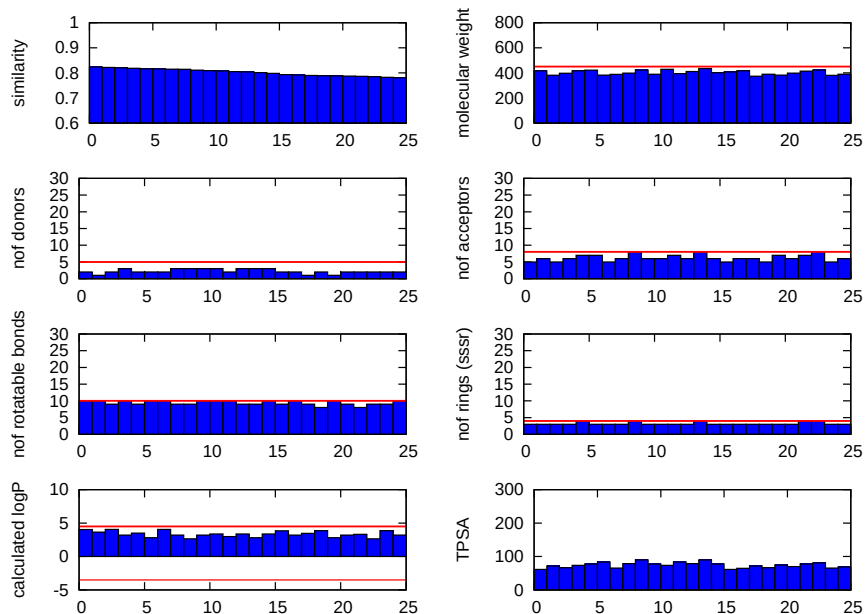
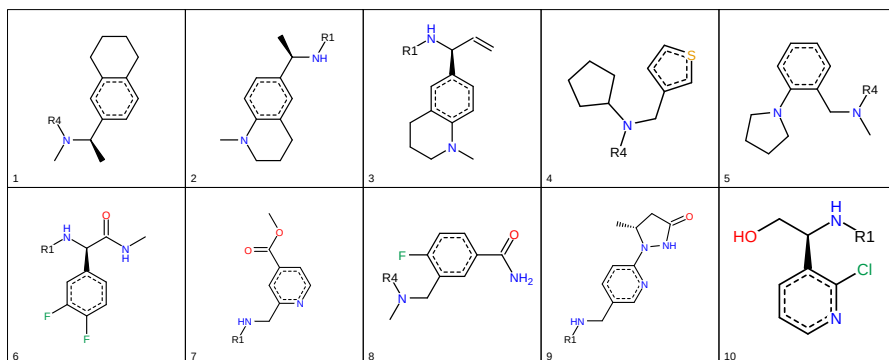
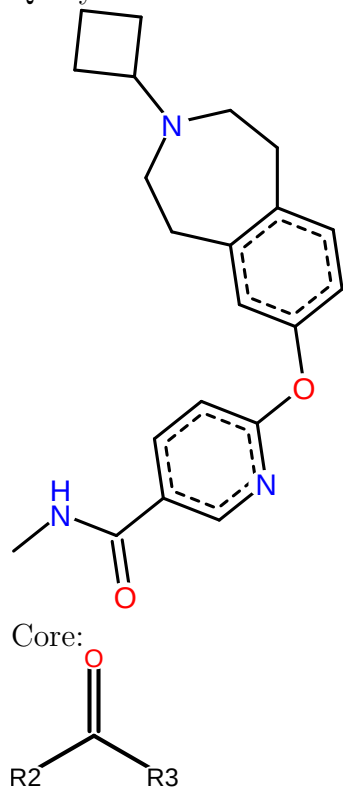
Seed 43: 0.7858

Seed 4543201: 0.8050

Average score: 0.8017

Solution no 1 with score 0.8021, seen after 184488 steps

Query:



2.5 CDK2 query ZINC00582575

2.5.1 Random

Randomly chosen sublibrary (starting point of optimization)

Seed 1: 0.6529

Seed 2177: 0.7421

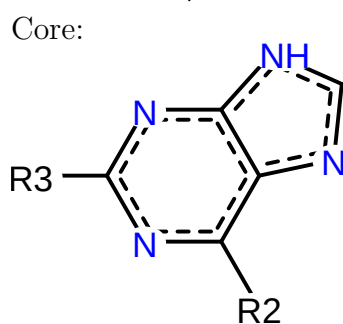
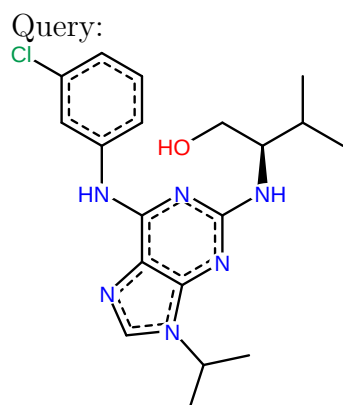
Seed 35498: 0.6826

Seed 43: 0.7006

Seed 4543201: 0.6529

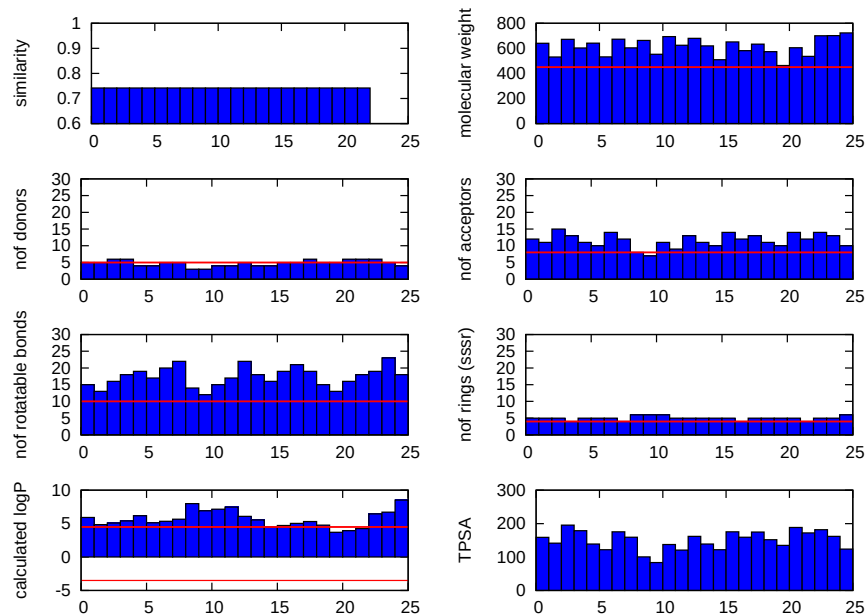
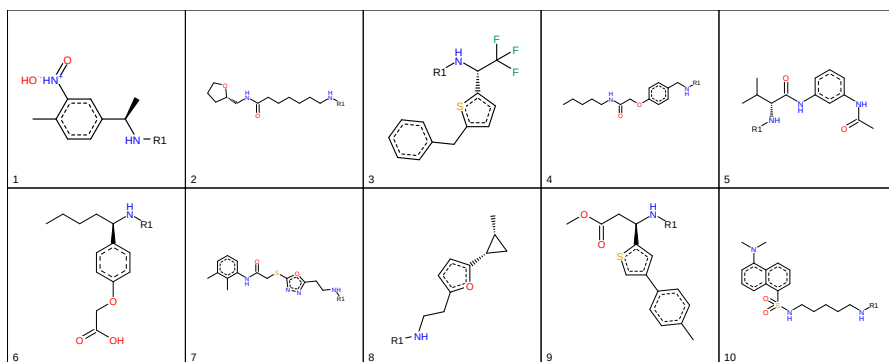
Average score: 0.6862

Solution no 1 with score 0.6529, seen after 1 step



Exceeded leadlike criteria:

- MW > 450
- nof rings > 4
- nof acceptors > 8
- nof donors > 5
- nof rotatable bonds > 10
- logP > 4.5

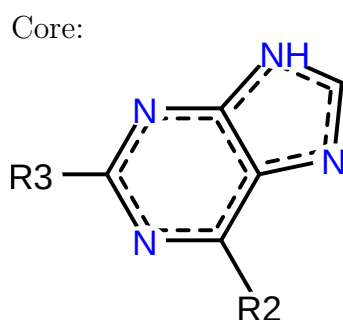
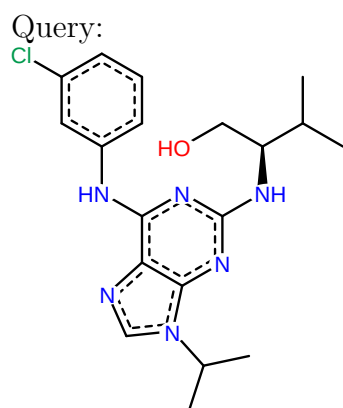


2.5.2 FTreesFS

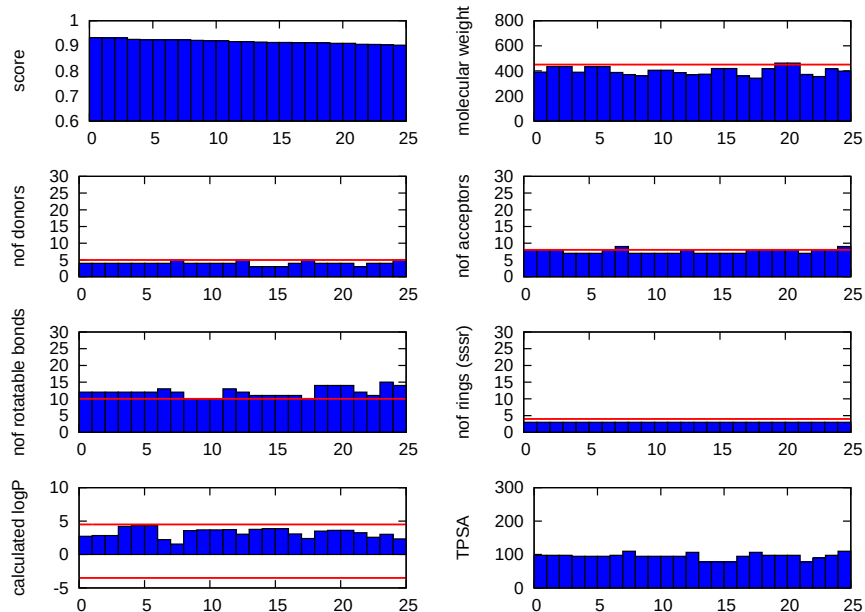
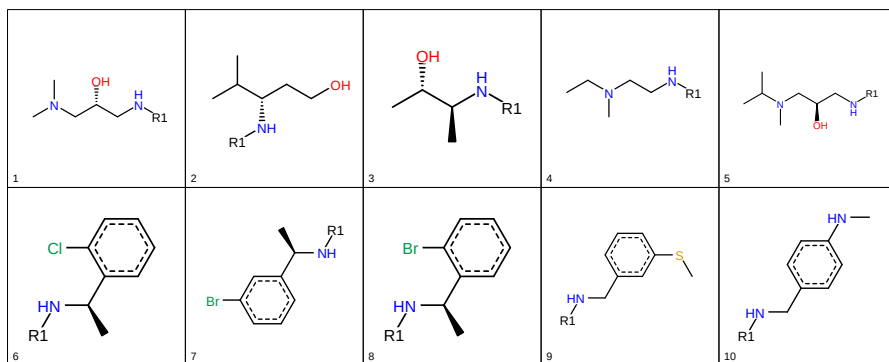
First 5 (4 for cores with 3 links) reagents for each core link of the best 1000 products generated by FTreesFS

Seed 1: 0.9169

Solution no 1 with score 0.9169, seen after 107430 steps



Exceeded leadlike criteria:
 MW > 450
 nof acceptors > 8
 nof rotatable bonds > 10



2.5.3 LoFT

LoFT using only ftree similarity

Seed 1: 0.9228

Seed 2177: 0.9256

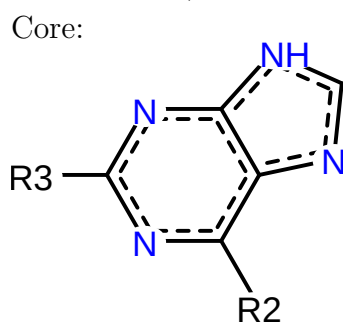
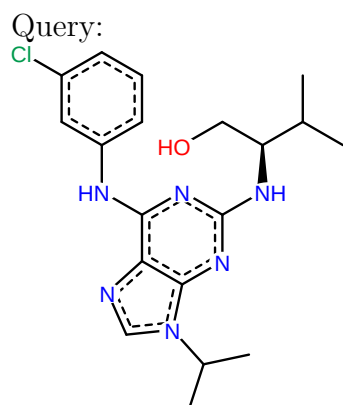
Seed 35498: 0.9242

Seed 43: 0.9238

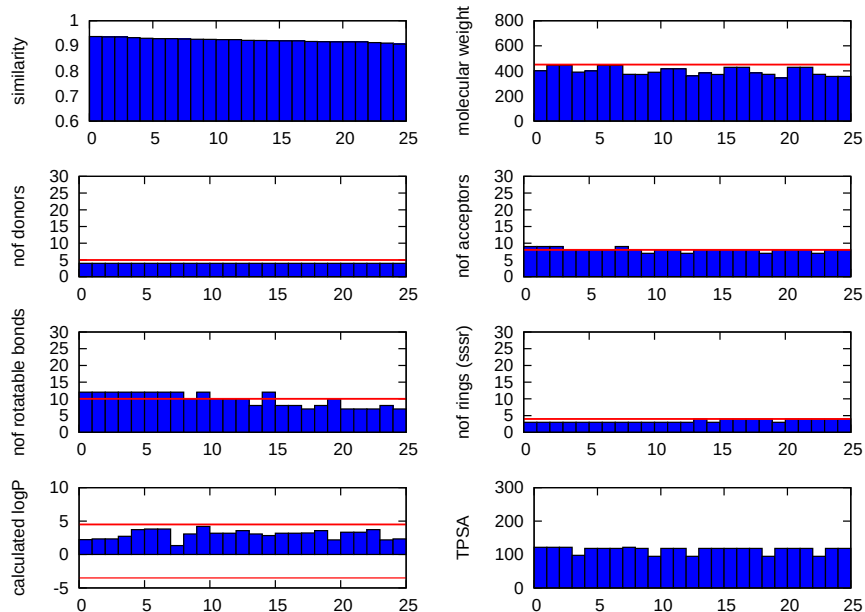
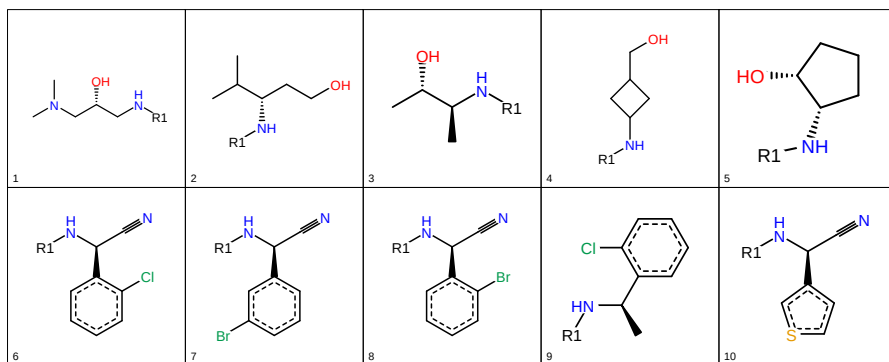
Seed 4543201: 0.9253

Average score: 0.9243

Solution no 1 with score 0.9228, seen after 184810 steps



Exceeded leadlike criteria:
 nof acceptors > 8
 nof rotatable bonds > 10



2.5.4 Leadlike

Leadlike criteria were used as product filter.

Seed 1: 0.9063

Seed 2177: 0.9112

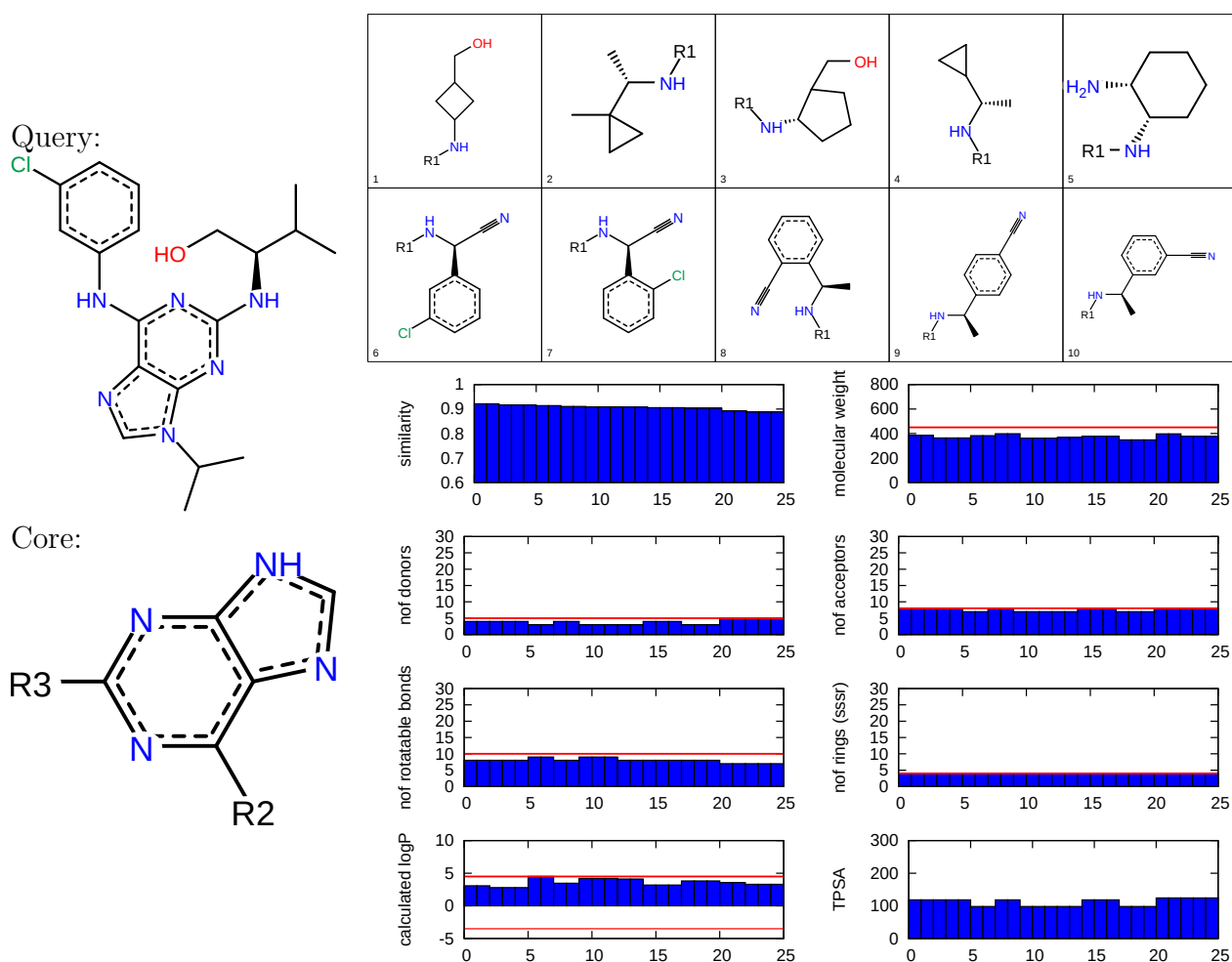
Seed 35498: 0.9140

Seed 43: 0.9070

Seed 4543201: 0.9144

Average score: 0.9106

Solution no 1 with score 0.9063, seen after 199173 steps



2.5.5 Diverse

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1.

Seed 1: 0.9176

Seed 2177: 0.9163

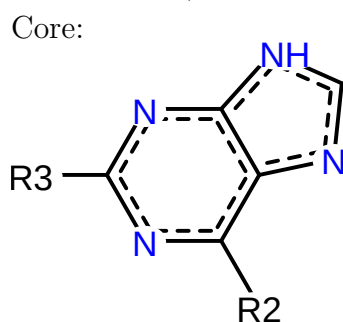
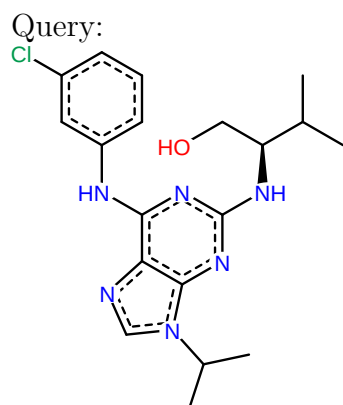
Seed 35498: 0.9186

Seed 43: 0.9171

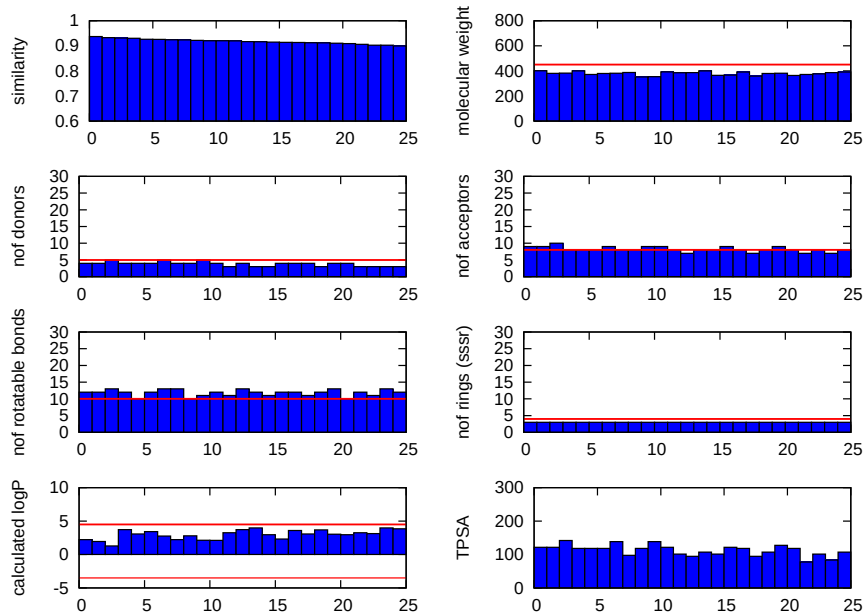
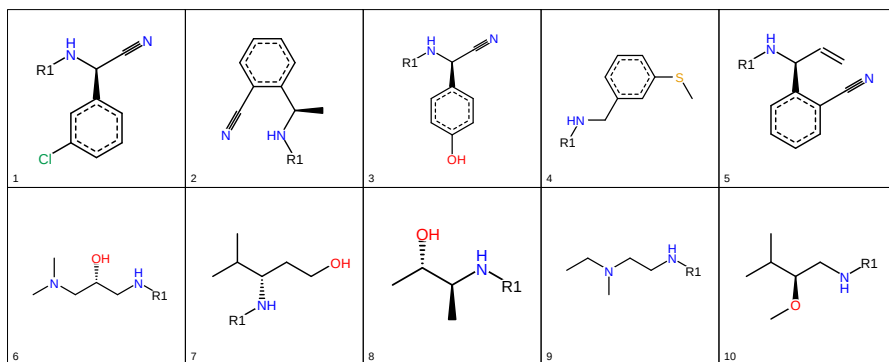
Seed 4543201: 0.9145

Average score: 0.9168

Solution no 1 with score 0.9176, seen after 177743 steps



Exceeded leadlike criteria:
 nof acceptors > 8
 nof rotatable bonds > 10



2.5.6 Diverse Leadlike

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1. Leadlike criteria were used as product filter.

Seed 1: 0.8935

Seed 2177: 0.9035

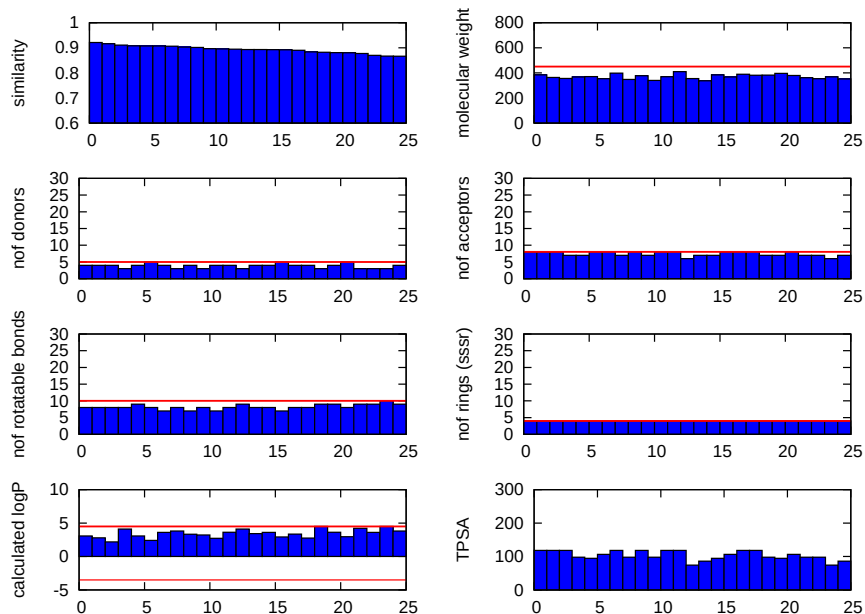
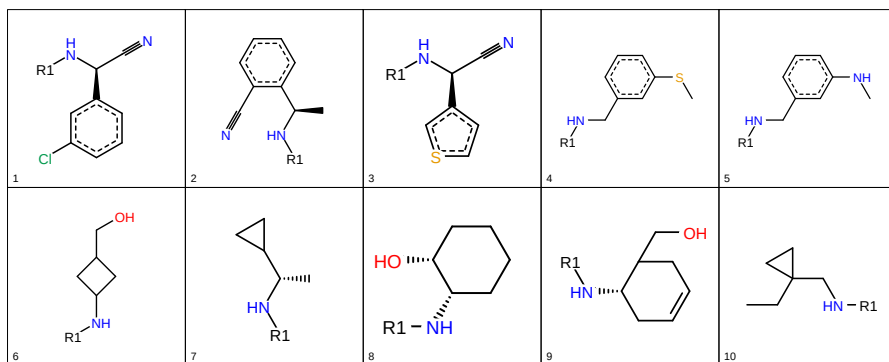
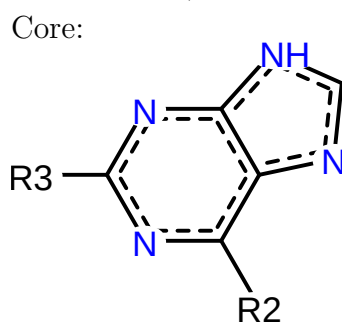
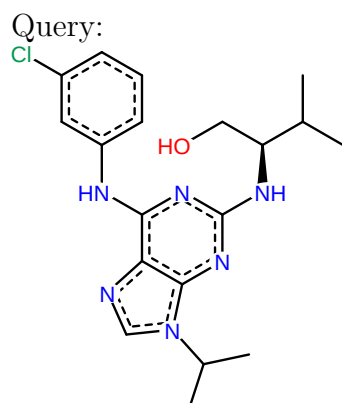
Seed 35498: 0.9024

Seed 43: 0.8935

Seed 4543201: 0.8964

Average score: 0.8979

Solution no 1 with score 0.8935, seen after 171243 steps



2.6 CDK2 query ZINC03591113

2.6.1 Random

Randomly chosen sublibrary (starting point of optimization)

Seed 1: 0.0000

Seed 2177: 0.2598

Seed 35498: 0.2598

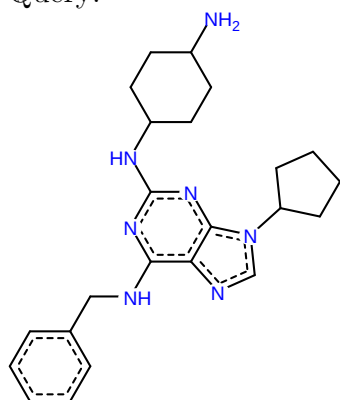
Seed 43: 0.3306

Seed 4543201: 0.2362

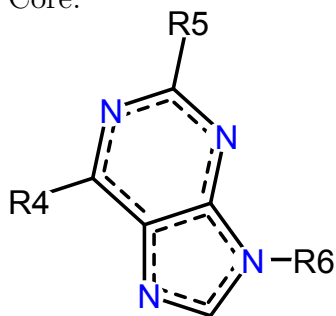
Average score: 0.2173

Solution no 1 with score 0.0000, seen after 0 steps

Query:



Core:



Exceeded leadlike criteria:

MW > 450

nof rings > 4

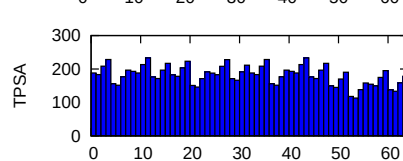
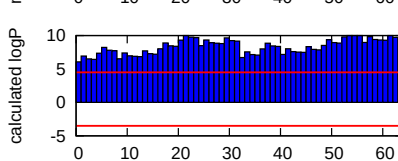
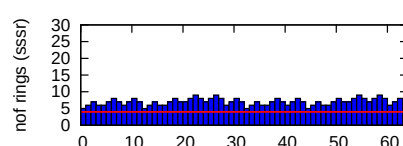
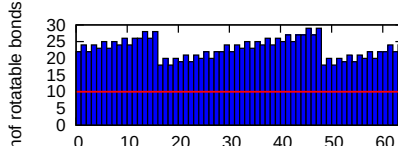
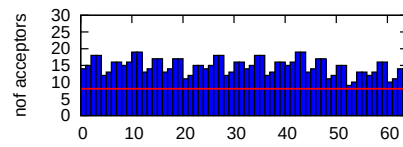
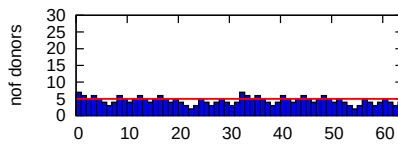
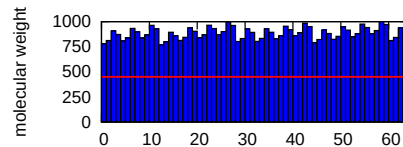
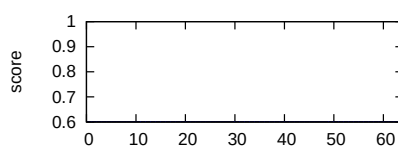
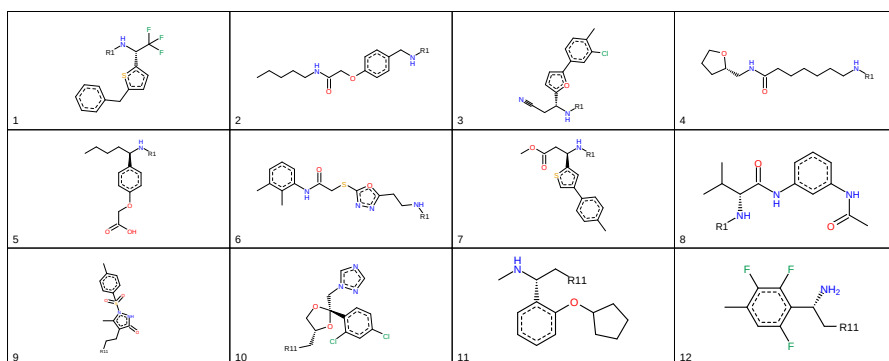
nof rings > 4

nof acceptors > 8

nof donors > 5

nof rotatable bonds > 10

logP > 4.5



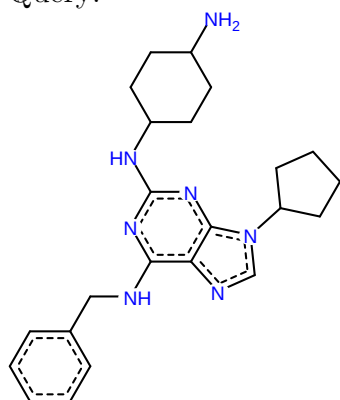
2.6.2 FTreesFS

First 5 (4 for cores with 3 links) reagents for each core link of the best 1000 products generated by FTreesFS

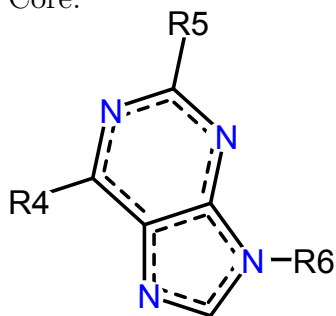
Seed 1: 0.9507

Solution no 1 with score 0.9507, seen after 9386 steps

Query:



Core:



Exceeded leadlike criteria:

MW > 450

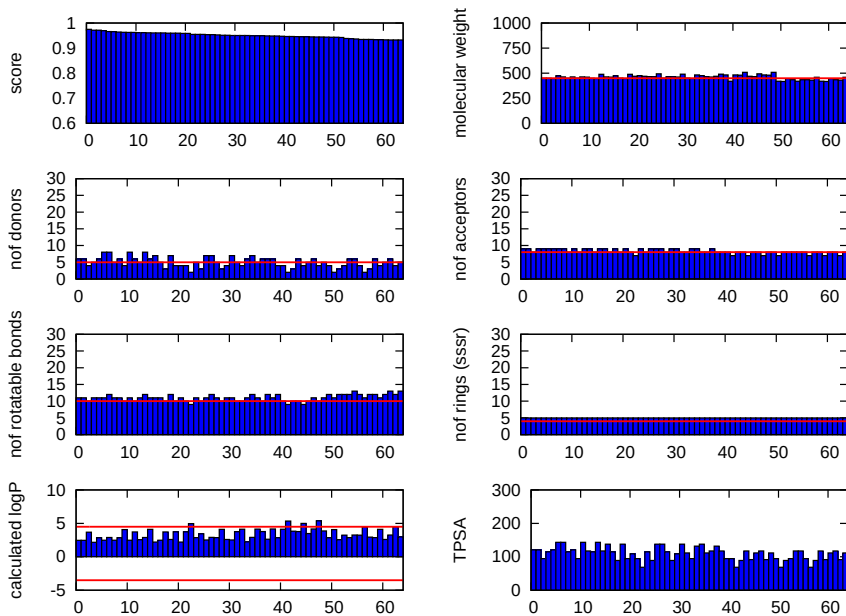
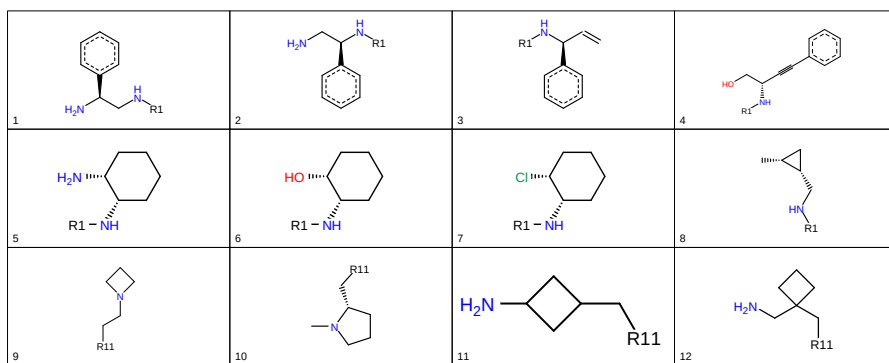
nof rings > 4

nof acceptors > 8

nof donors > 5

nof rotatable bonds > 10

logP > 4.5



2.6.3 LoFT

LoFT using only ftree's similarity

Seed 1: 0.9634

Seed 2177: 0.9611

Seed 35498: 0.9640

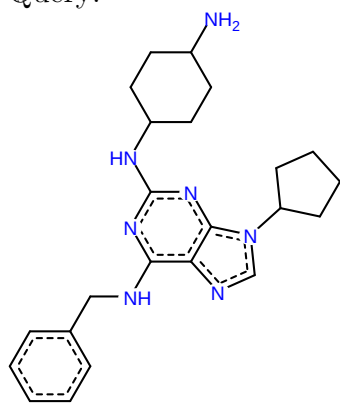
Seed 43: 0.9628

Seed 4543201: 0.9589

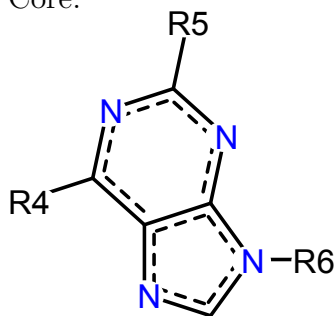
Average score: 0.9620

Solution no 1 with score 0.9634, seen after 155268 steps

Query:



Core:



Exceeded leadlike criteria:

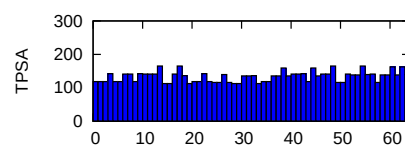
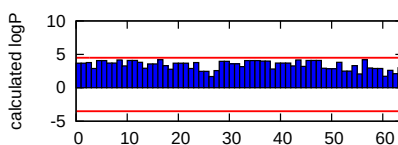
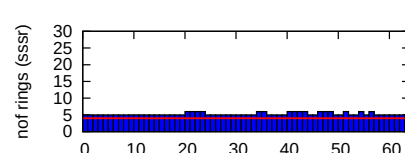
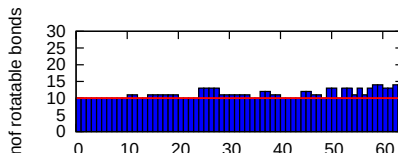
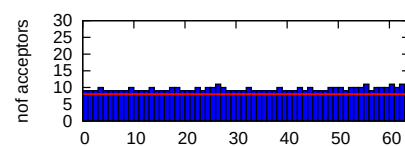
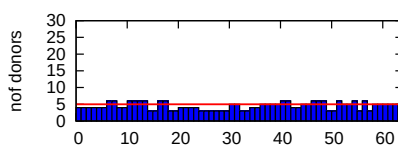
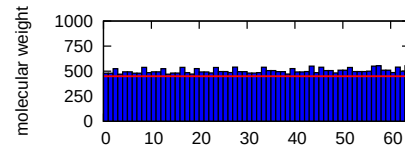
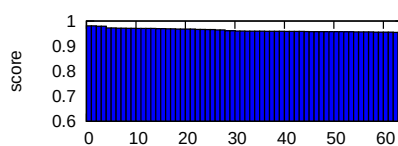
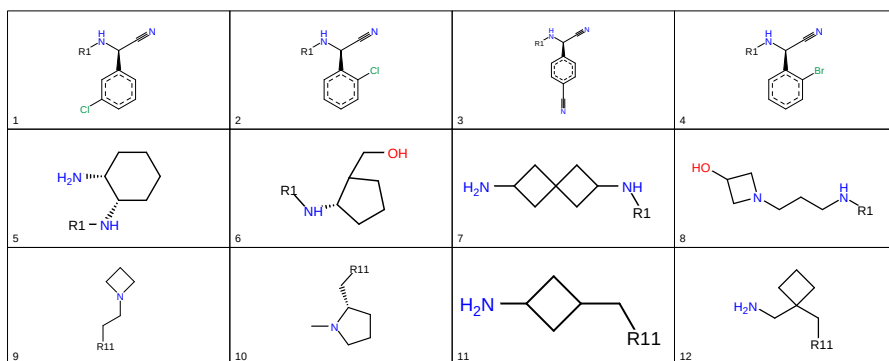
MW > 450

nof rings > 4

nof acceptors > 8

nof donors > 5

nof rotatable bonds > 10



2.6.4 Leadlike

Leadlike criteria were used as product filter.

Seed 1: 0.0000

Seed 2177: 0.0000

Seed 35498: 0.0000

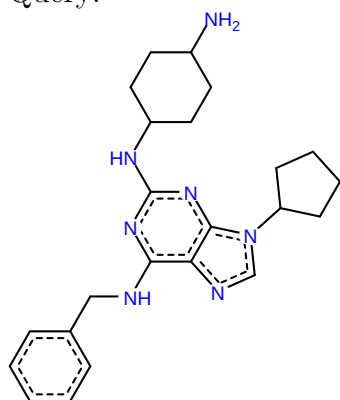
Seed 43: 0.0000

Seed 4543201: 0.0000

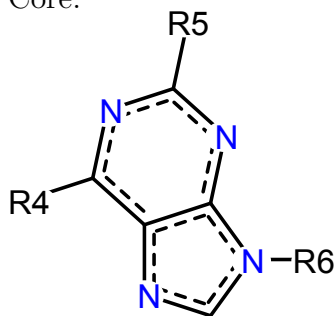
Average score: 0.0000

Solution no 1 with score 0.0000, seen after 0 steps

Query:



Core:



Exceeded leadlike criteria:

MW > 450

nof rings > 4

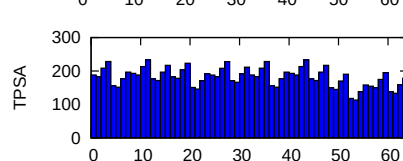
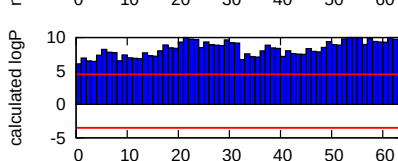
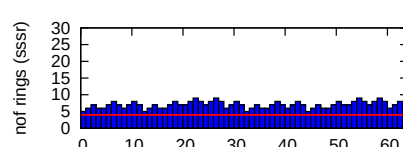
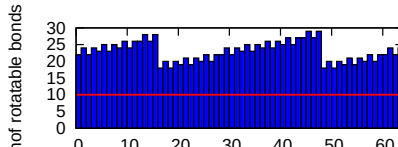
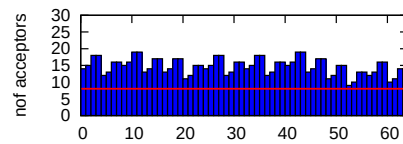
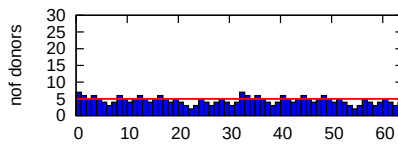
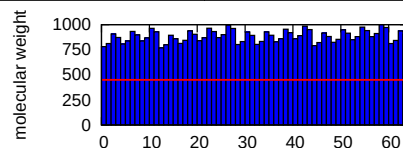
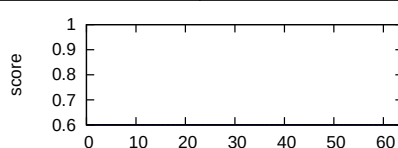
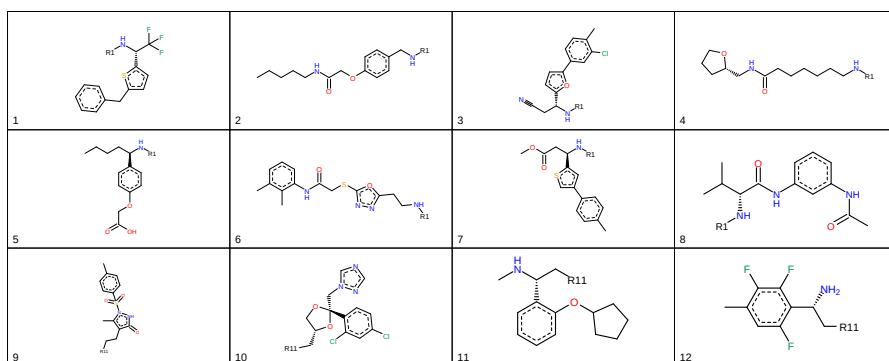
nof rings > 4

nof acceptors > 8

nof donors > 5

nof rotatable bonds > 10

logP > 4.5



2.6.5 Diverse

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1.

Seed 1: 0.9532

Seed 2177: 0.9543

Seed 35498: 0.9533

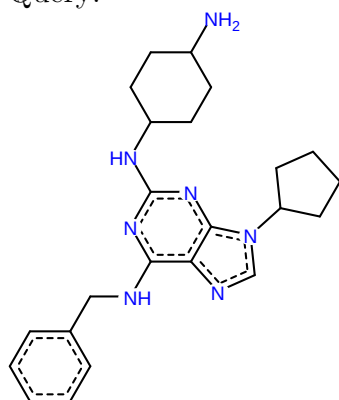
Seed 43: 0.9572

Seed 4543201: 0.9546

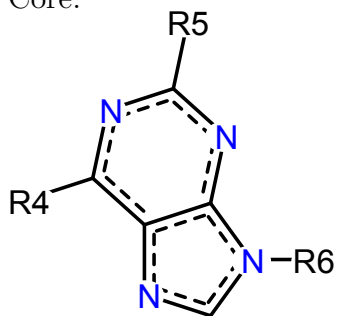
Average score: 0.9545

Solution no 1 with score 0.9532, seen after 182112 steps

Query:



Core:



Exceeded leadlike criteria:

MW > 450

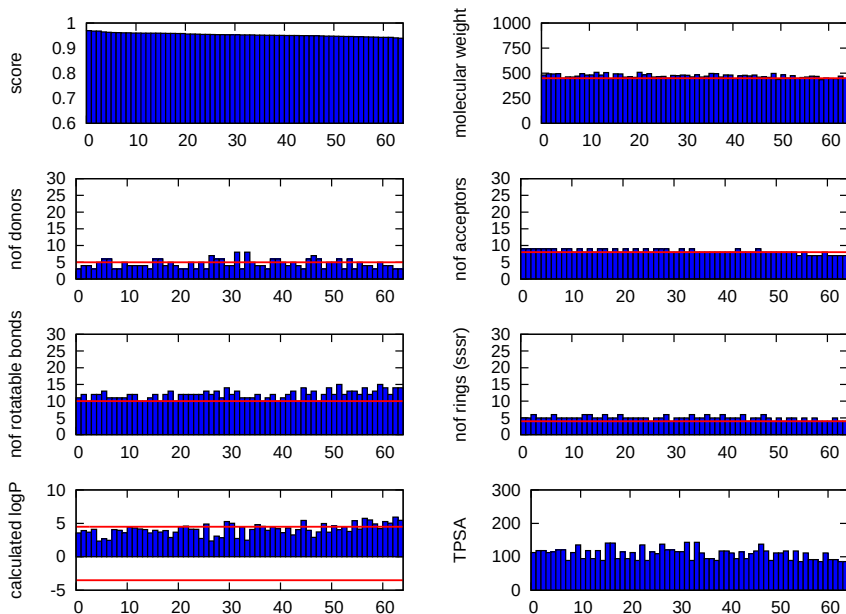
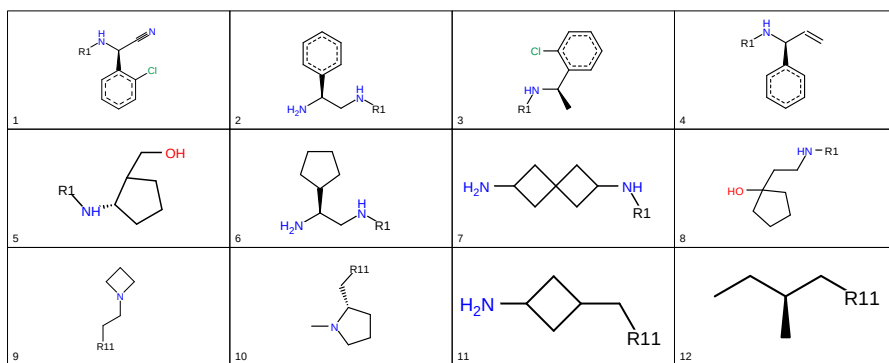
nof rings > 4

nof acceptors > 8

nof donors > 5

nof rotatable bonds > 10

logP > 4.5



2.6.6 Diverse Leadlike

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1. Leadlike criteria were used as product filter.

Seed 1: 0.0000

Seed 2177: 0.0000

Seed 35498: 0.0000

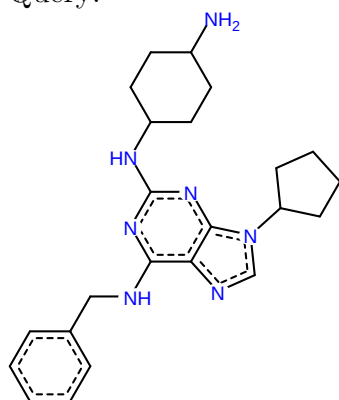
Seed 43: 0.0000

Seed 4543201: 0.0000

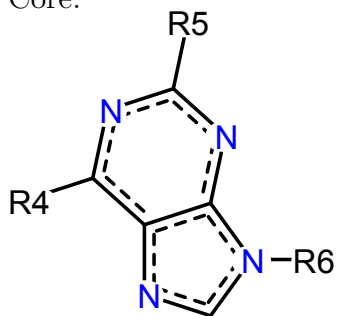
Average score: 0.0000

Solution no 1 with score 0.0000, seen after 0 steps

Query:



Core:



Exceeded leadlike criteria:

MW > 450

nof rings > 4

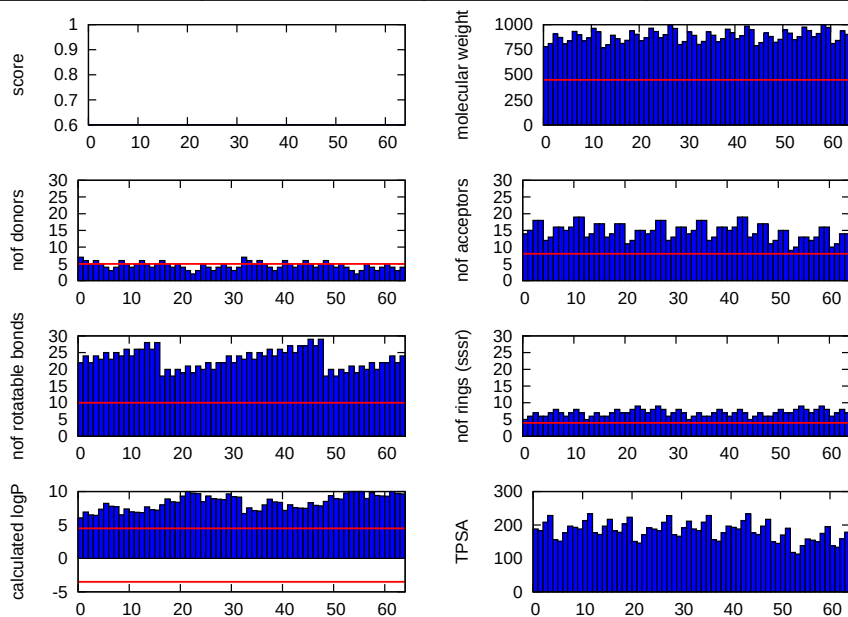
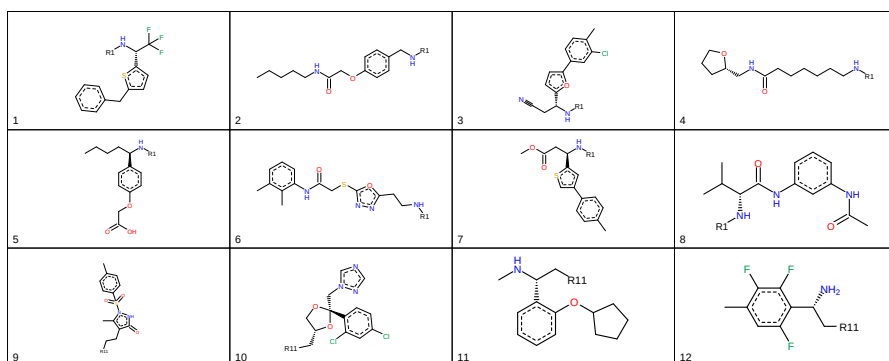
nof rings > 4

nof acceptors > 8

nof donors > 5

nof rotatable bonds > 10

logP > 4.5



2.6.7 Diverse Leadlike Scoring

Each leadlikeness criterion was used in scoring with weight 0.05. Ftree similarity had weight 0.7. Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1.

Seed 1: 0.9429

Seed 2177: 0.9412

Seed 35498: 0.9419

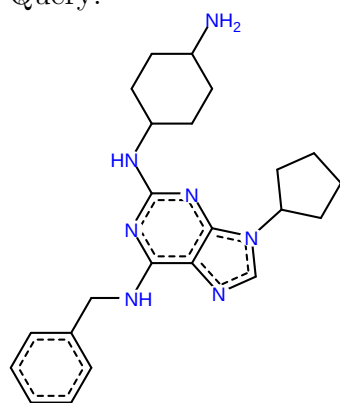
Seed 43: 0.9424

Seed 4543201: 0.9433

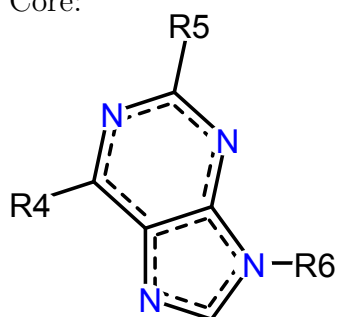
Average score: 0.9423

Solution no 1 with score 0.9429, seen after 171446 steps

Query:



Core:

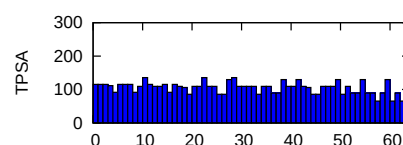
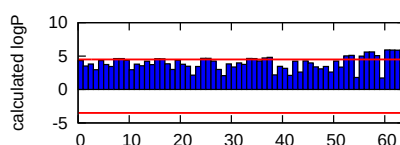
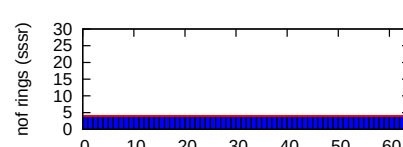
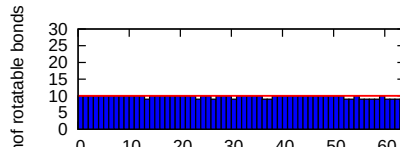
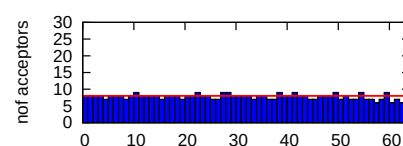
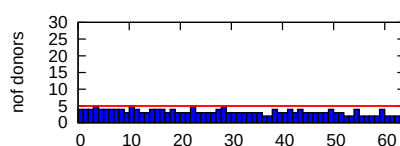
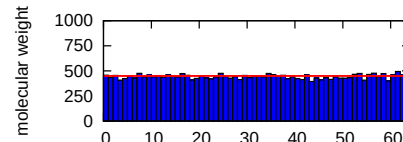
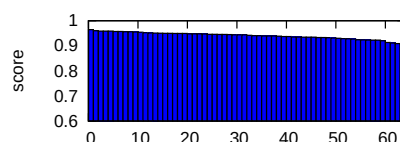
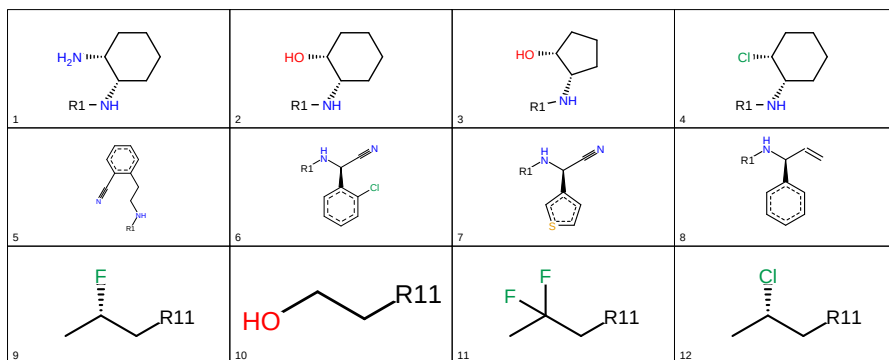


Exceeded leadlike criteria:

MW > 450

nof acceptors > 8

logP > 4.5



2.7 CDK2 query ZINC03814433

2.7.1 Random

Randomly chosen sublibrary (starting point of optimization)

Seed 1: 0.6589

Seed 2177: 0.6489

Seed 35498: 0.6541

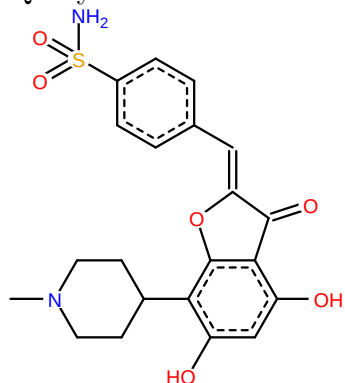
Seed 43: 0.6624

Seed 4543201: 0.6578

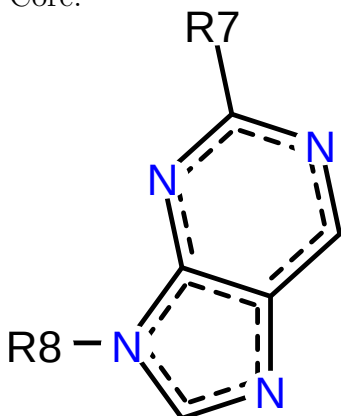
Average score: 0.6564

Solution no 1 with score 0.6589, seen after 1 step

Query:



Core:



Exceeded leadlike criteria:

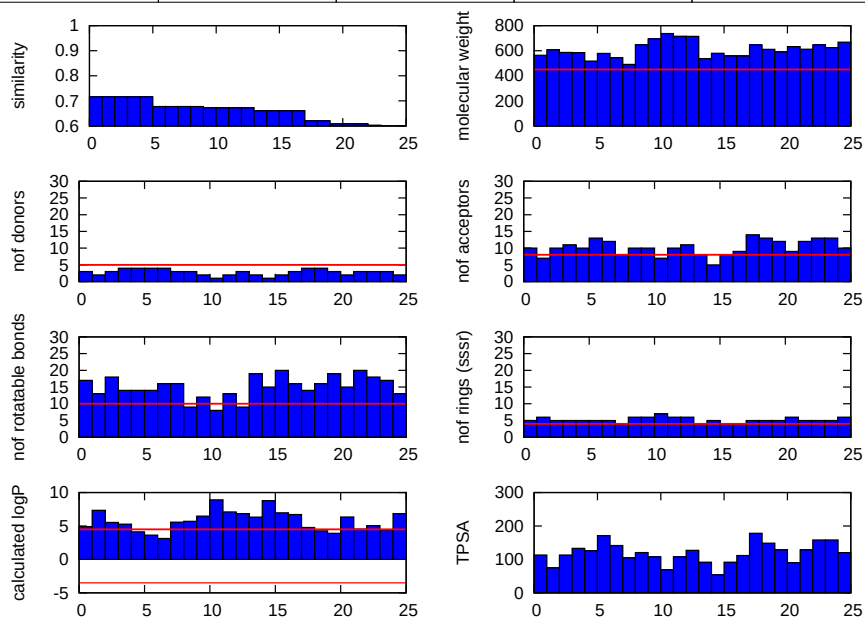
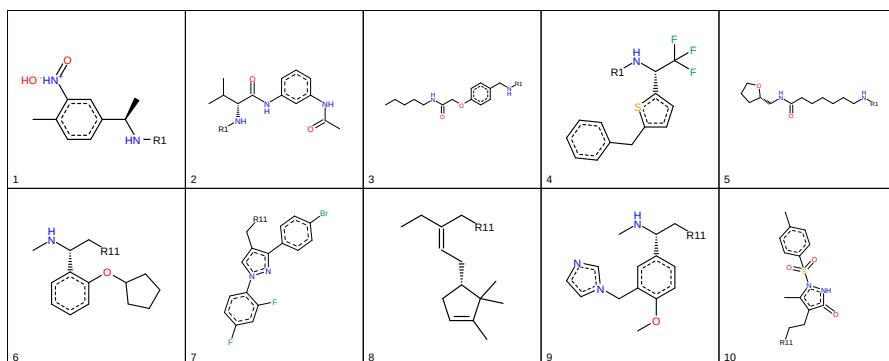
MW > 450

nof rings > 4

nof acceptors > 8

nof rotatable bonds > 10

logP > 4.5

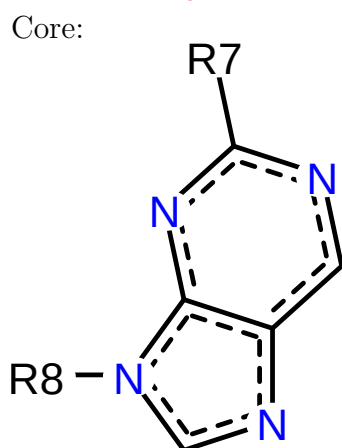
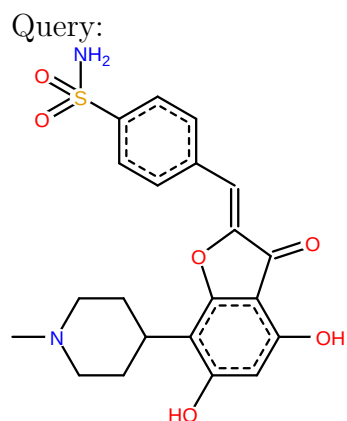


2.7.2 FTreesFS

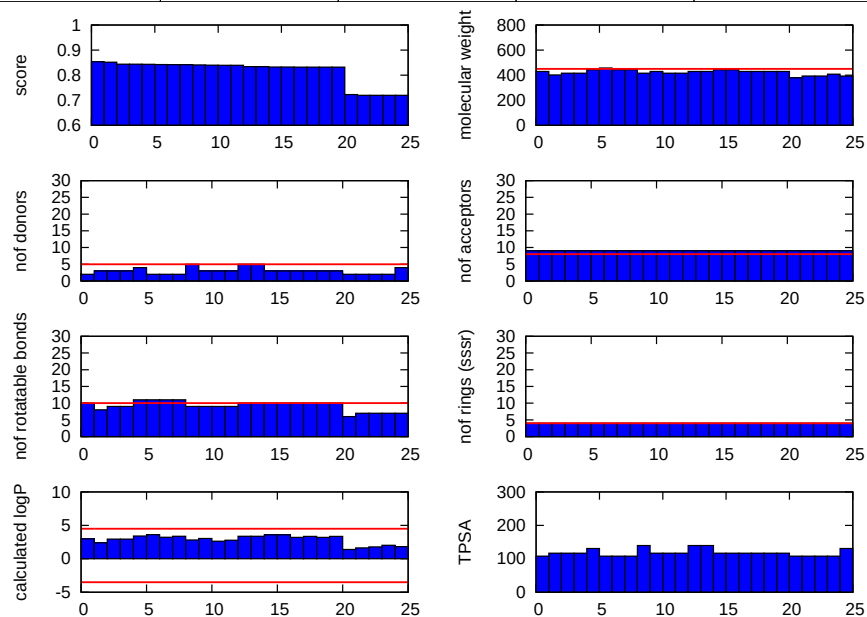
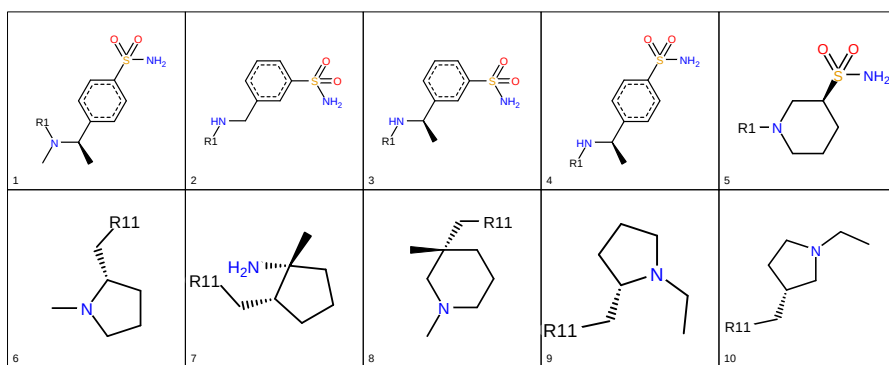
First 5 (4 for cores with 3 links) reagents for each core link of the best 1000 products generated by FTreesFS

Seed 1: 0.8155

Solution no 1 with score 0.8155, seen after 0 steps



Exceeded leadlike criteria:
 MW > 450
 nof acceptors > 8
 nof rotatable bonds > 10



2.7.3 LoFT

LoFT using only ftrees similarity

Seed 1: 0.8166

Seed 2177: 0.8174

Seed 35498: 0.8166

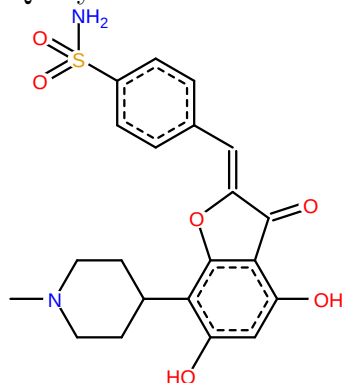
Seed 43: 0.8276

Seed 4543201: 0.8169

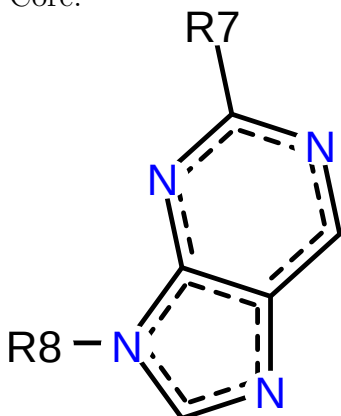
Average score: 0.8190

Solution no 1 with score 0.8166, seen after 189261 step

Query:



Core:



Exceeded leadlike criteria:

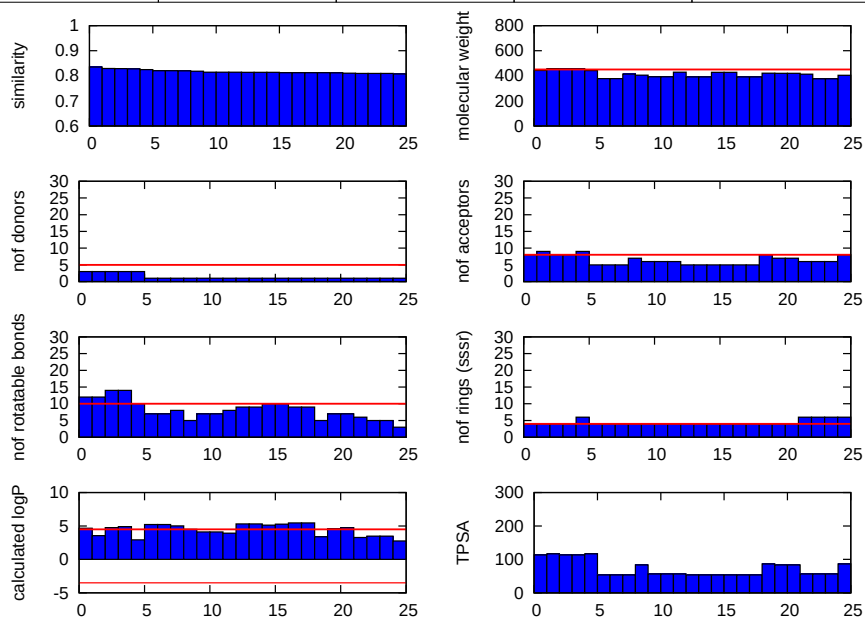
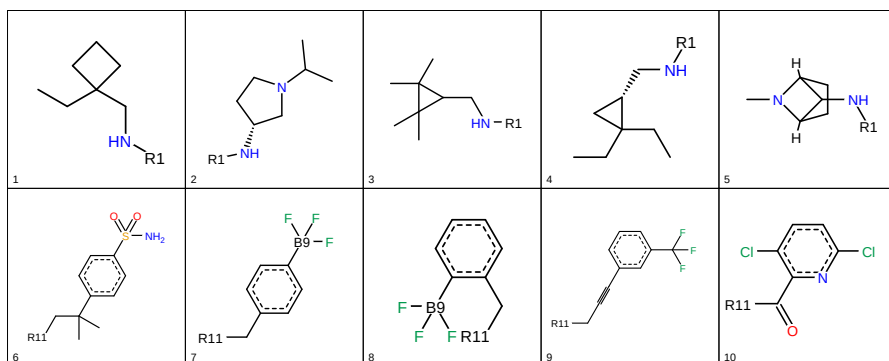
MW > 450

nof rings > 4

nof acceptors > 8

nof rotatable bonds > 10

logP > 4.5



2.7.4 Leadlike

Leadlike criteria were used as product filter.

Seed 1: 0.8086

Seed 2177: 0.8099

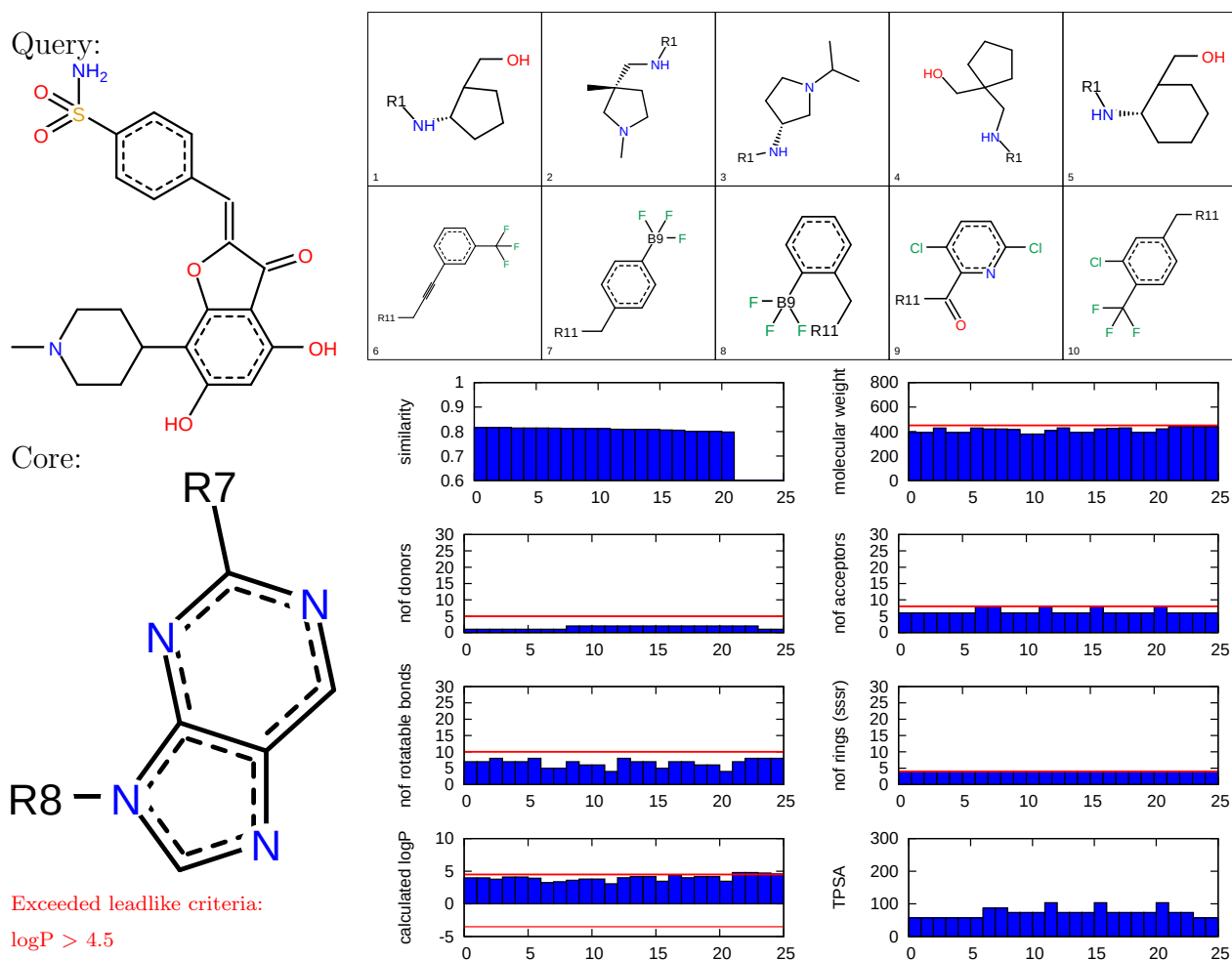
Seed 35498: 0.8079

Seed 43: 0.8092

Seed 4543201: 0.8084

Average score: 0.8088

Solution no 1 with score 0.8086, seen after 192723 steps



2.7.5 Diverse

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1.

Seed 1: 0.8103

Seed 2177: 0.8127

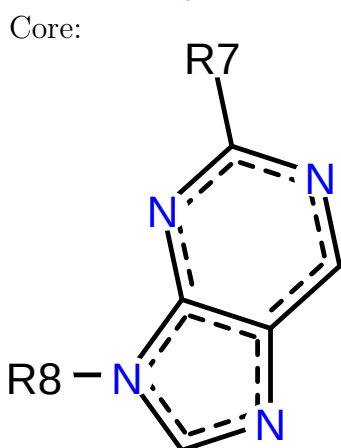
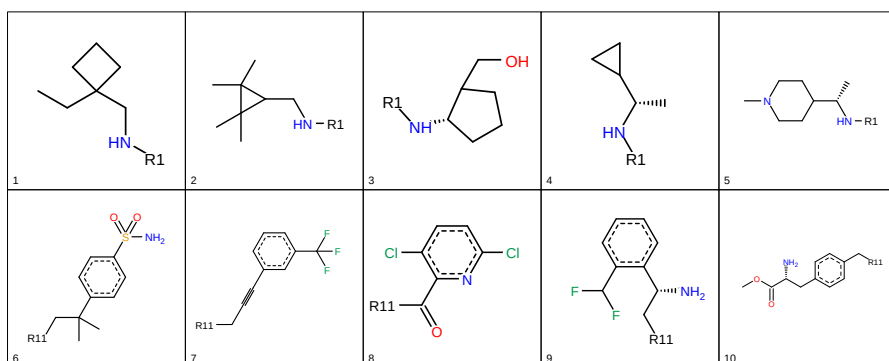
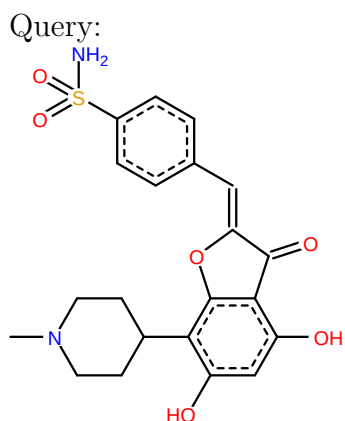
Seed 35498: 0.8124

Seed 43: 0.8129

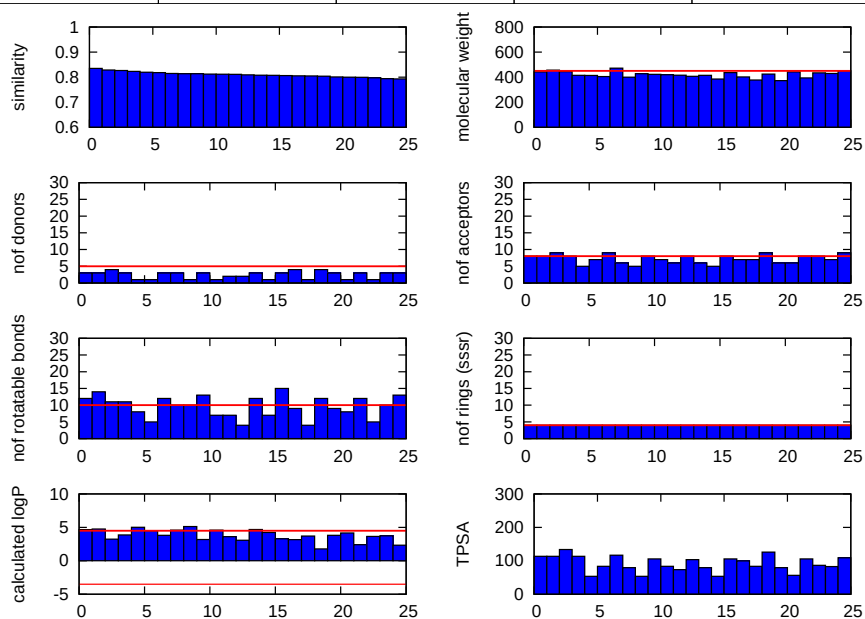
Seed 4543201: 0.8244

Average score: 0.8145

Solution no 1 with score 0.8103, seen after 139154 steps



Exceeded leadlike criteria:
 MW > 450
 nof acceptors > 8
 nof rotatable bonds > 10
 logP > 4.5



2.7.6 Diverse Leadlike

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1. Leadlike criteria were used as product filter.

Seed 1: 0.8050

Seed 2177: 0.8018

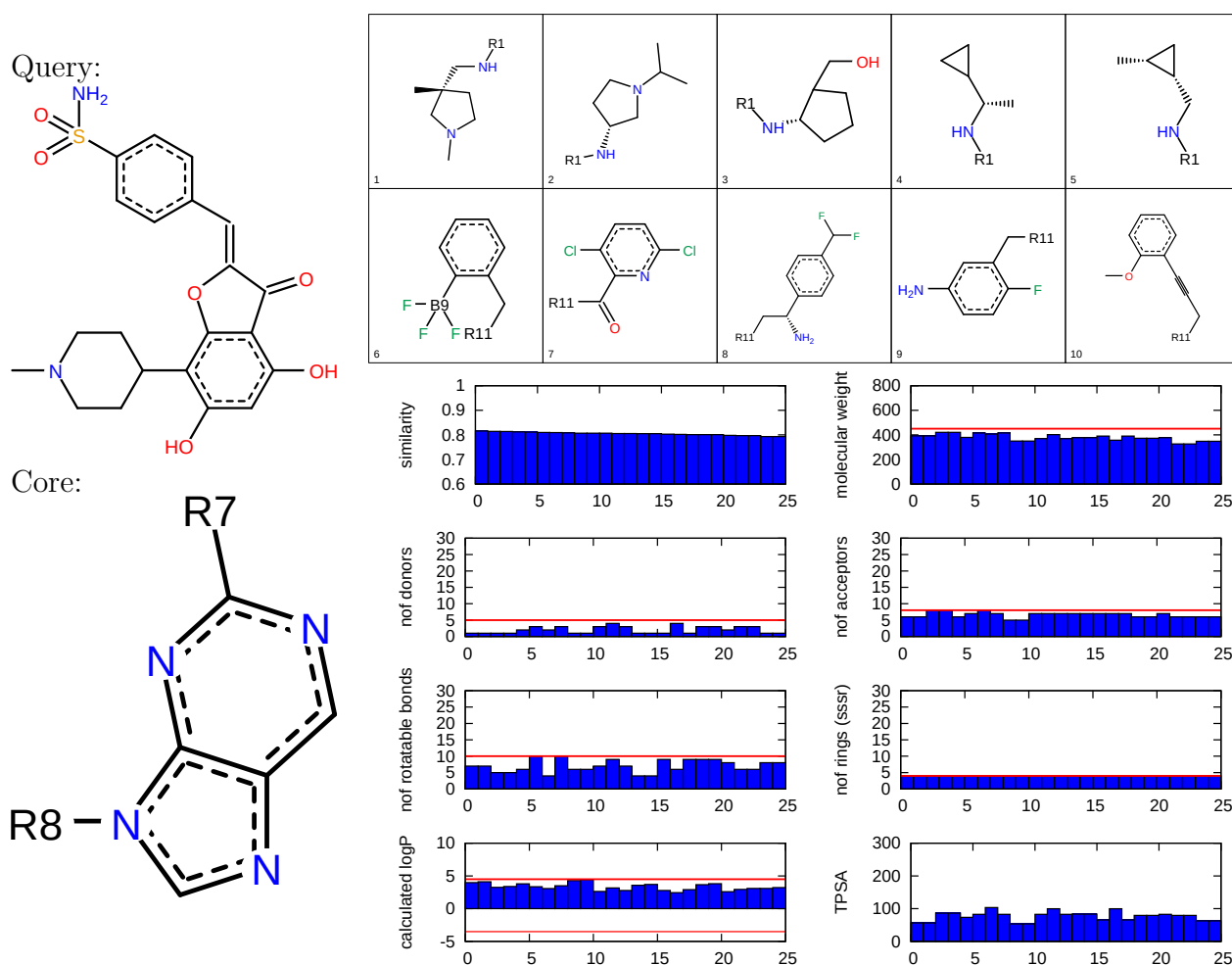
Seed 35498: 0.8018

Seed 43: 0.8018

Seed 4543201: 0.8009

Average score: 0.8023

Solution no 1 with score 0.8050, seen after 198619 steps



2.8 5HT2a query 1

2.8.1 Random

Randomly chosen sublibrary (starting point of optimization)

Seed 1: 0.5653

Seed 2177: 0.6248

Seed 35498: 0.5951

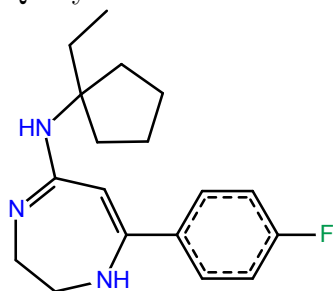
Seed 43: 0.6608

Seed 4543201: 0.4463

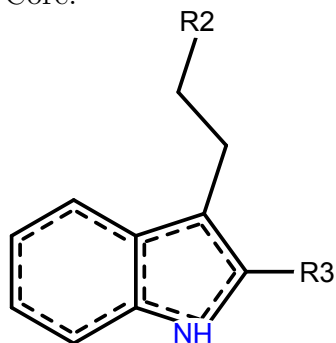
Average score: 0.5785

Solution no 1 with score 0.5653, seen after 1 step

Query:



Core:



Exceeded leadlike criteria:

MW > 450

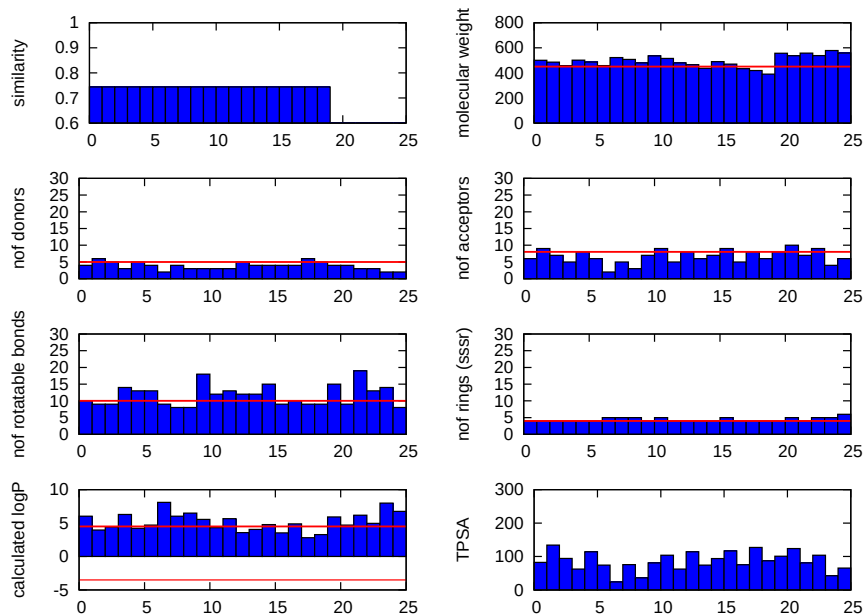
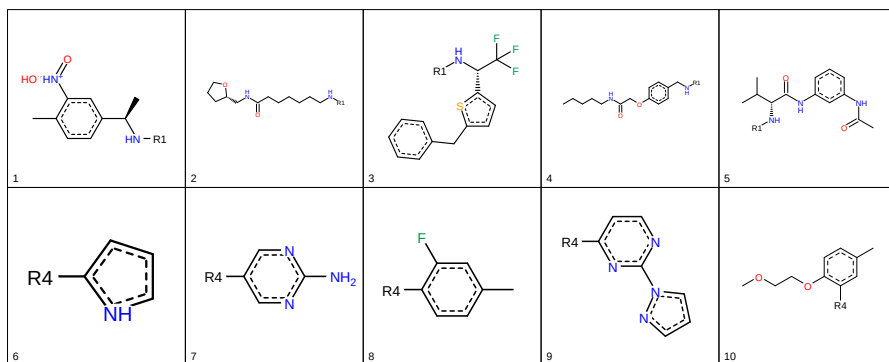
nof rings > 4

nof acceptors > 8

nof donors > 5

nof rotatable bonds > 10

logP > 4.5



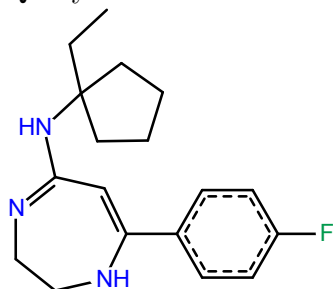
2.8.2 FTreesFS

First 5 (4 for cores with 3 links) reagents for each core link of the best 1000 products generated by FTreesFS

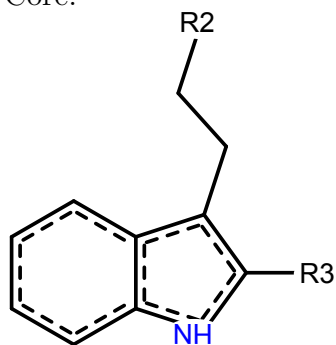
Seed 1: 0.8751

Solution no 1 with score 0.8751, seen after 0 steps

Query:

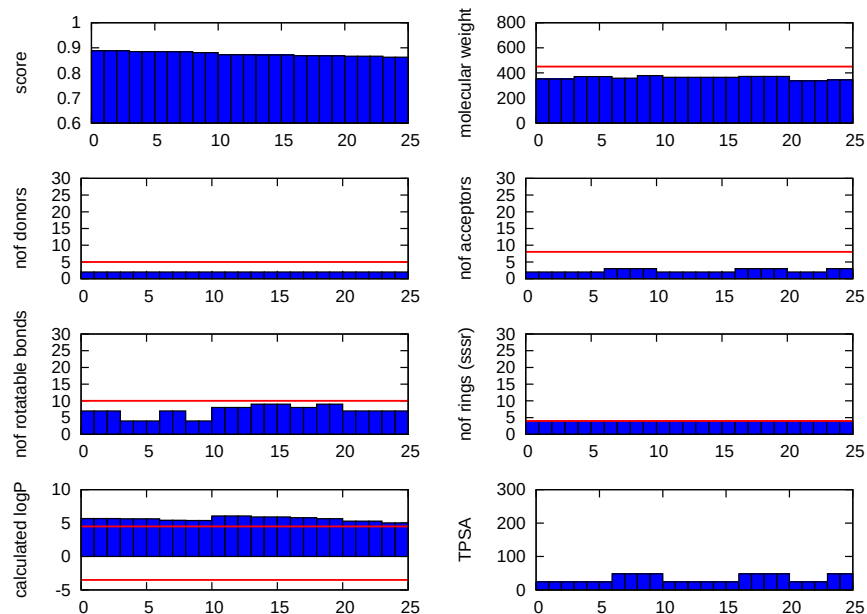
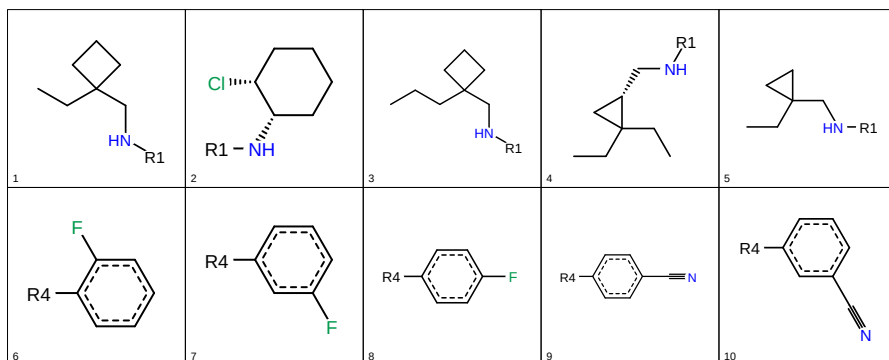


Core:



Exceeded leadlike criteria:

logP > 4.5



2.8.3 LoFT

LoFT using only ftree similarity

Seed 1: 0.8754

Seed 2177: 0.8751

Seed 35498: 0.8761

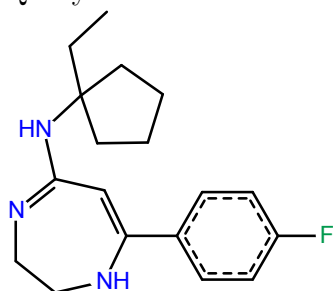
Seed 43: 0.8747

Seed 4543201: 0.8737

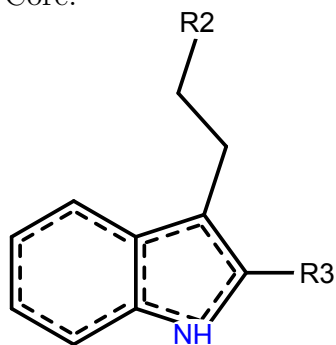
Average score: 0.8750

Solution no 1 with score 0.8754, seen after 198133 steps

Query:

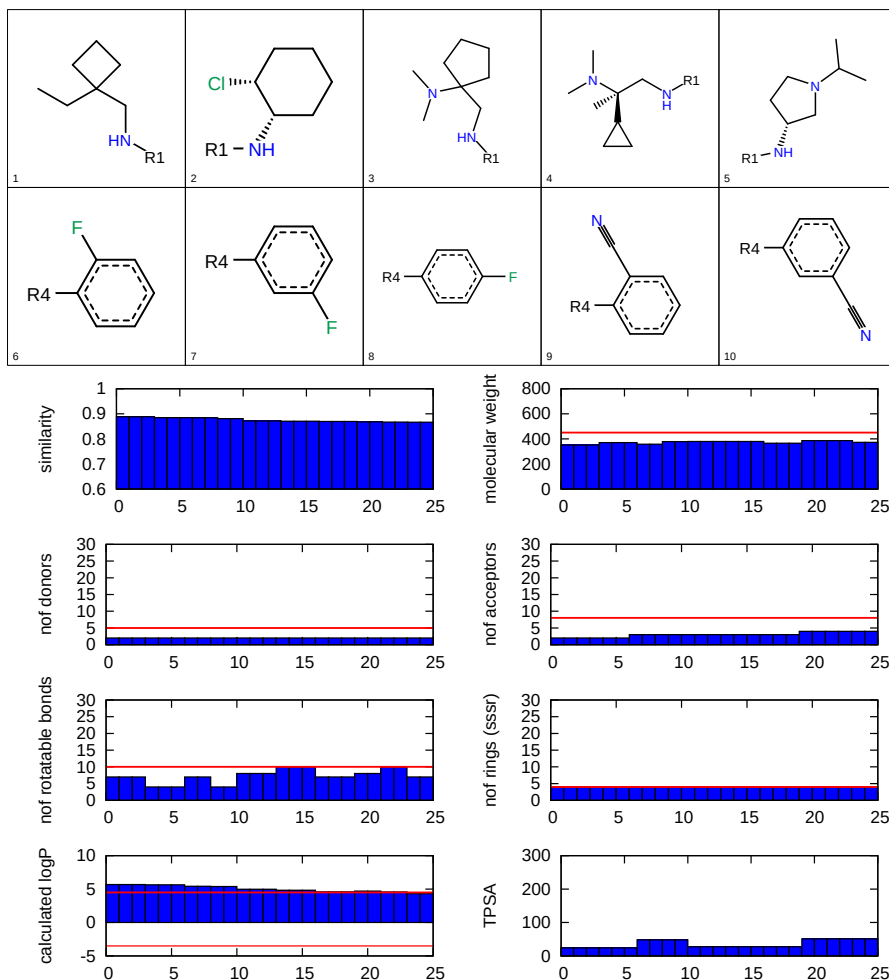


Core:



Exceeded leadlike criteria:

logP > 4.5



2.8.4 Leadlike

Leadlike criteria were used as product filter.

Seed 1: 0.8567

Seed 2177: 0.8562

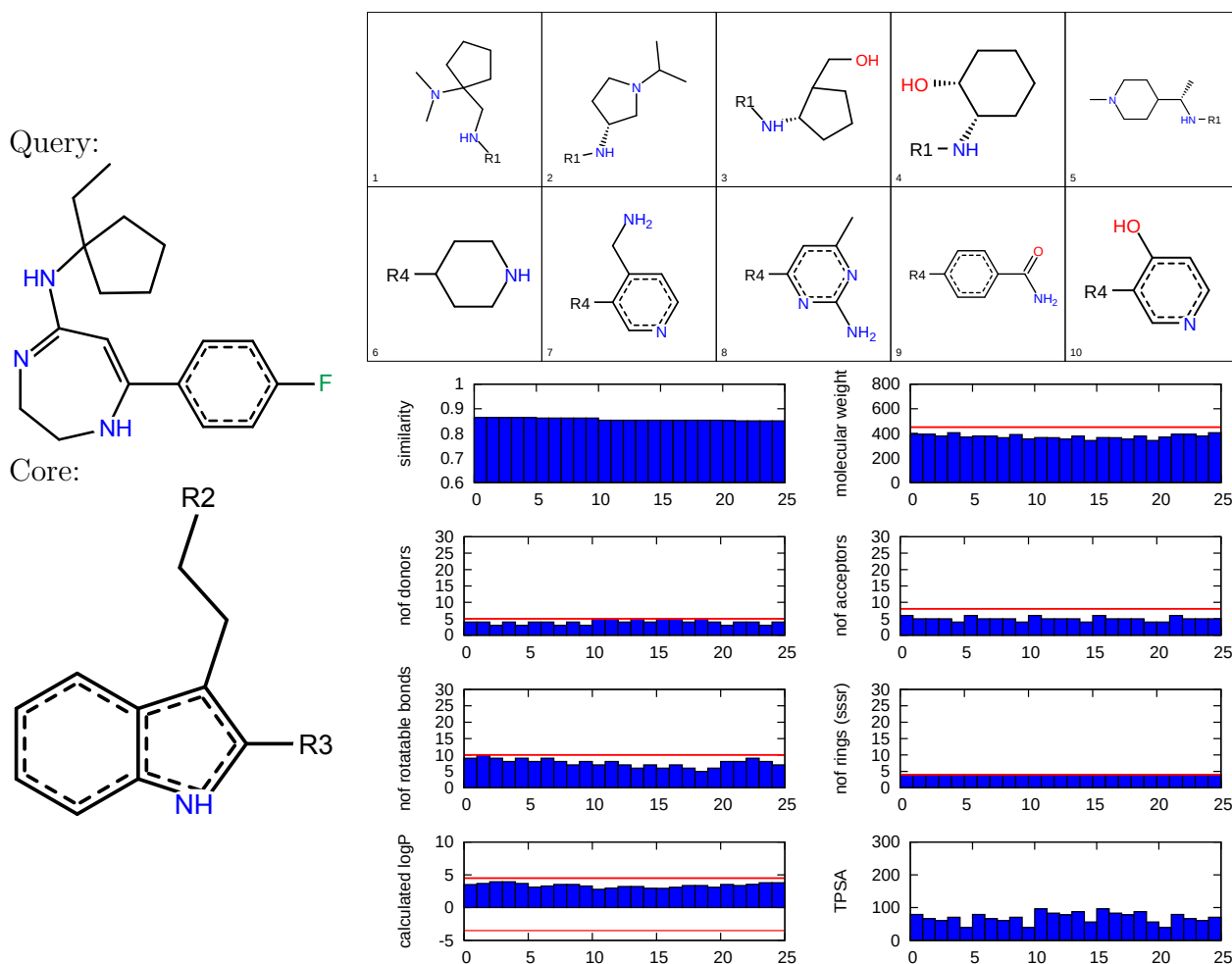
Seed 35498: 0.8564

Seed 43: 0.8561

Seed 4543201: 0.8457

Average score: 0.8542

Solution no 1 with score 0.8567, seen after 195228 steps



2.8.5 Diverse

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1.

Seed 1: 0.8659

Seed 2177: 0.8668

Seed 35498: 0.8659

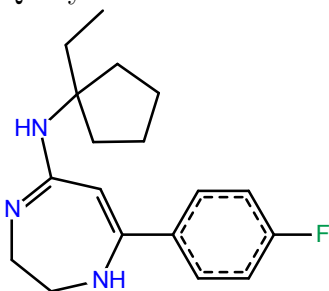
Seed 43: 0.8638

Seed 4543201: 0.8646

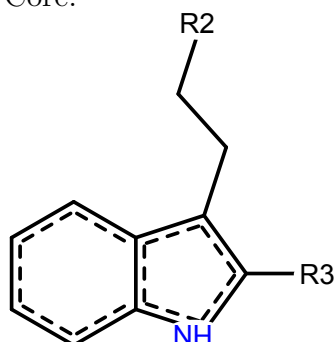
Average score: 0.8654

Solution no 1 with score 0.8659, seen after 195021 step

Query:



Core:

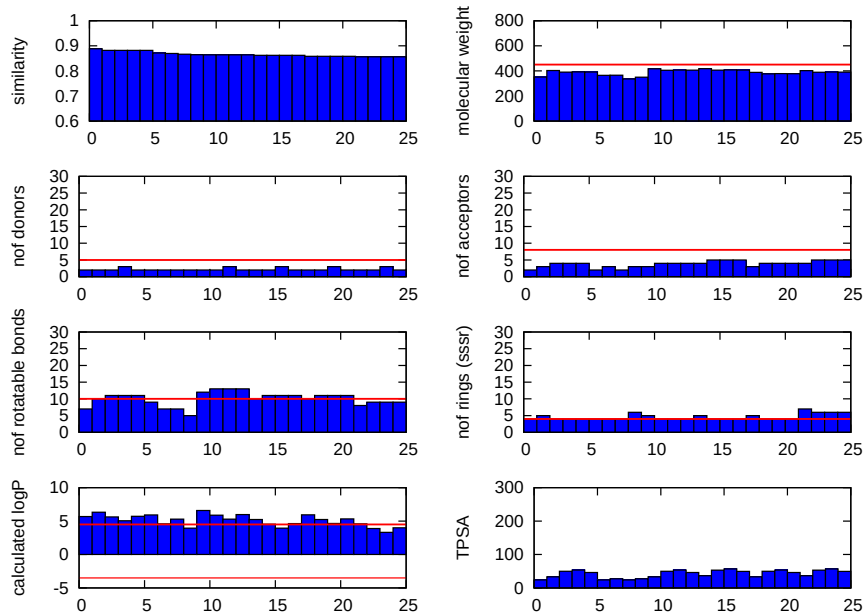
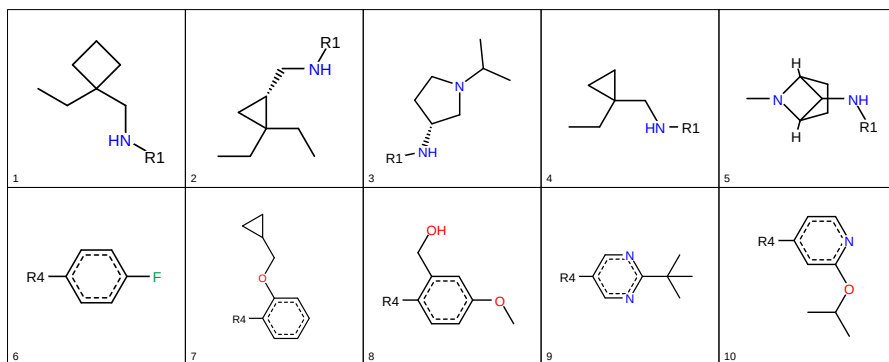


Exceeded leadlike criteria:

nof rings > 4

nof rotatable bonds > 10

logP > 4.5



2.8.6 Diverse Leadlike

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1. Leadlike criteria were used as product filter.

Seed 1: 0.8537

Seed 2177: 0.8520

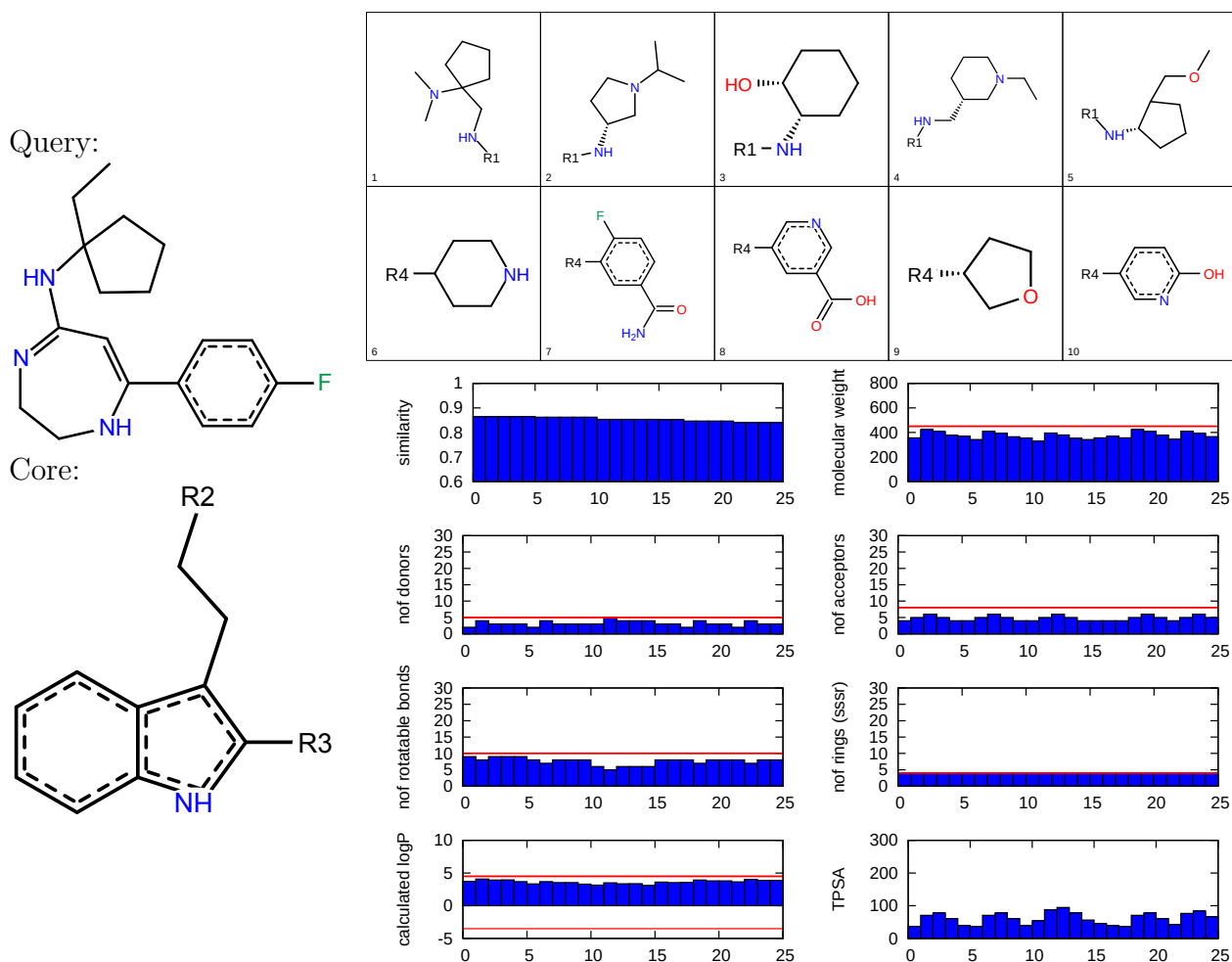
Seed 35498: 0.8559

Seed 43: 0.8547

Seed 4543201: 0.8424

Average score: 0.8517

Solution no 1 with score 0.8537, seen after 127233 steps



2.9 5HT2a query Ketanserin

2.9.1 Random

Randomly chosen sublibrary (starting point of optimization)

Seed 1: 0.7339

Seed 2177: 0.7341

Seed 35498: 0.7051

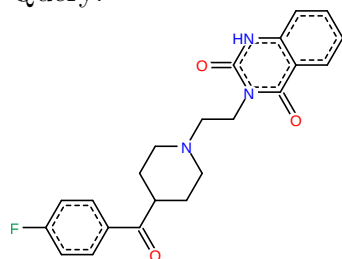
Seed 43: 0.6976

Seed 4543201: 0.7498

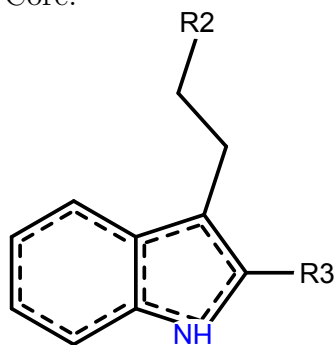
Average score: 0.7241

Solution no 1 with score 0.7339, seen after 0 steps

Query:



Core:



Exceeded leadlike criteria:

MW > 450

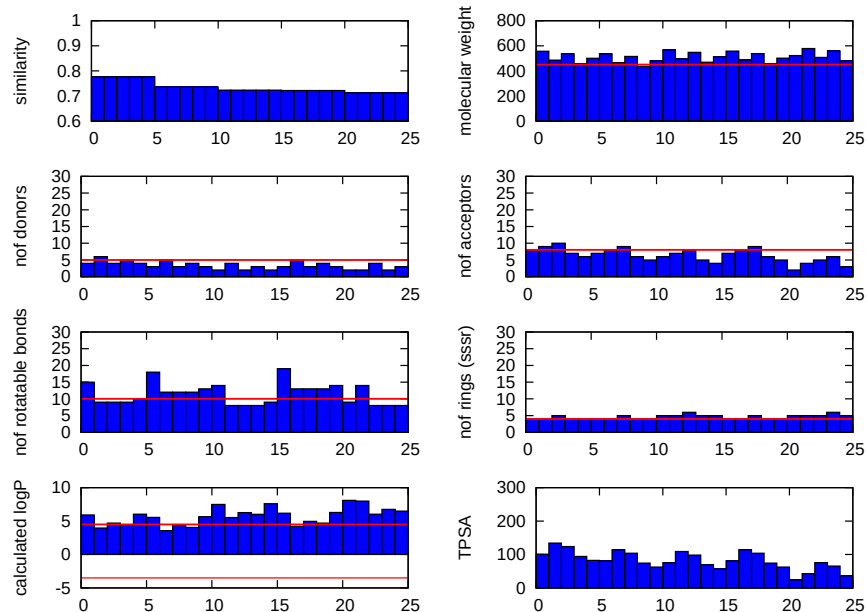
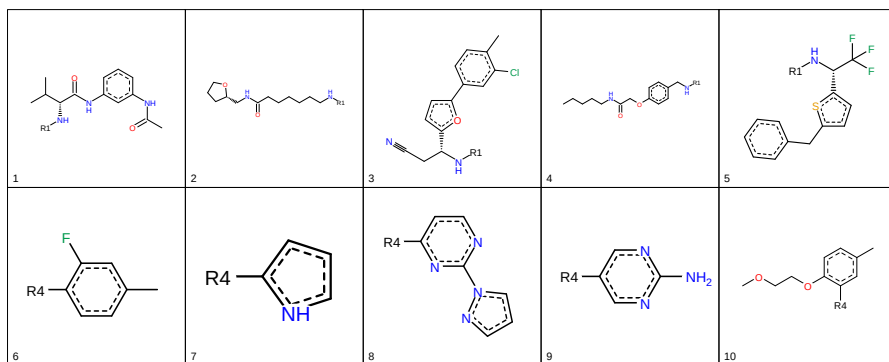
nof rings > 4

nof acceptors > 8

nof donors > 5

nof rotatable bonds > 10

logP > 4.5



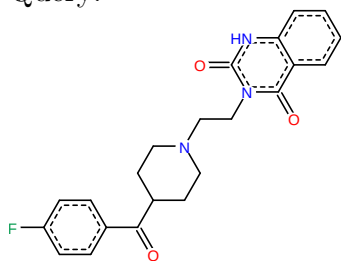
2.9.2 FTreesFS

First 5 (4 for cores with 3 links) reagents for each core link of the best 1000 products generated by FTreesFS

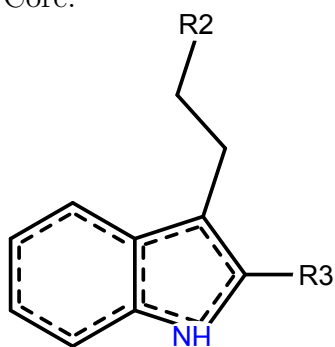
Seed 1: 0.8904

Solution no 1 with score 0.8904, seen after 0 steps

Query:



Core:

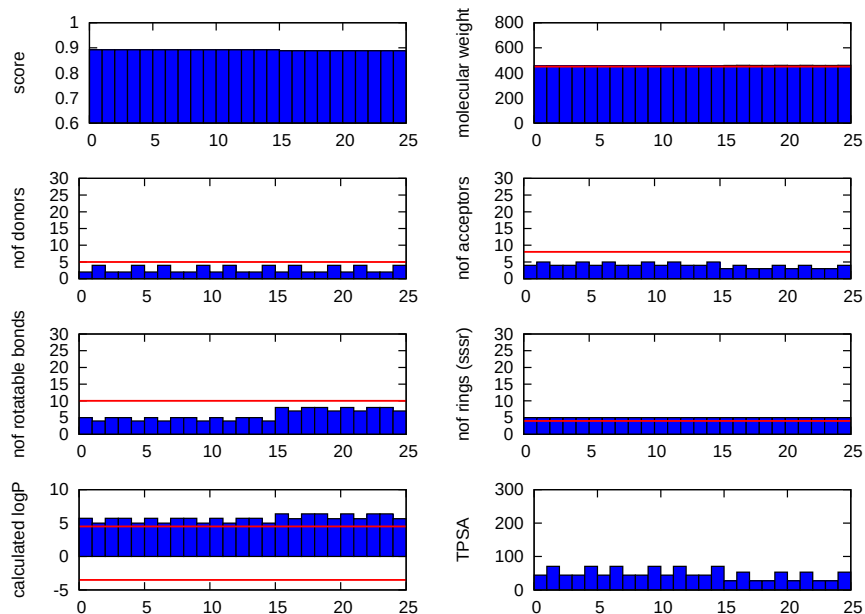
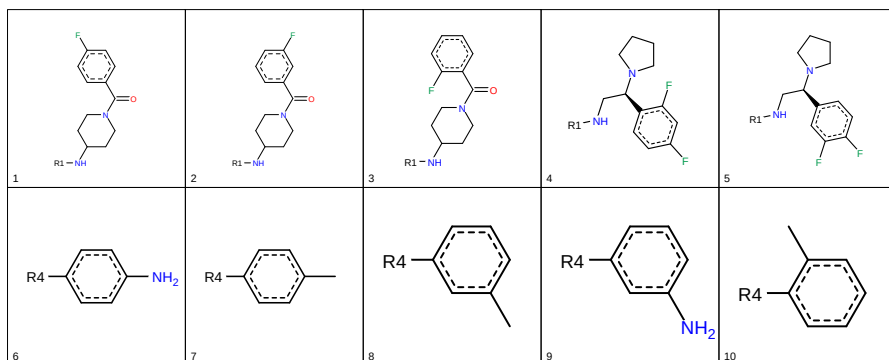


Exceeded leadlike criteria:

MW > 450

nof rings > 4

logP > 4.5



2.9.3 LoFT

LoFT using only ftree's similarity

Seed 1: 0.8891

Seed 2177: 0.8883

Seed 35498: 0.8883

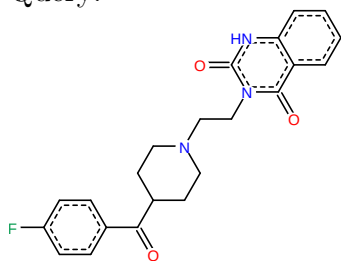
Seed 43: 0.8896

Seed 4543201: 0.8891

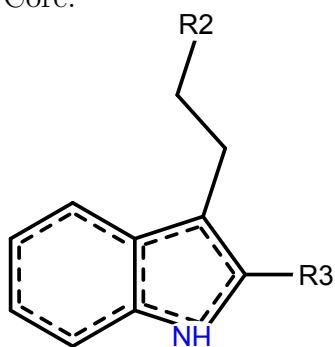
Average score: 0.8889

Solution no 1 with score 0.8891, seen after 98232 steps

Query:



Core:

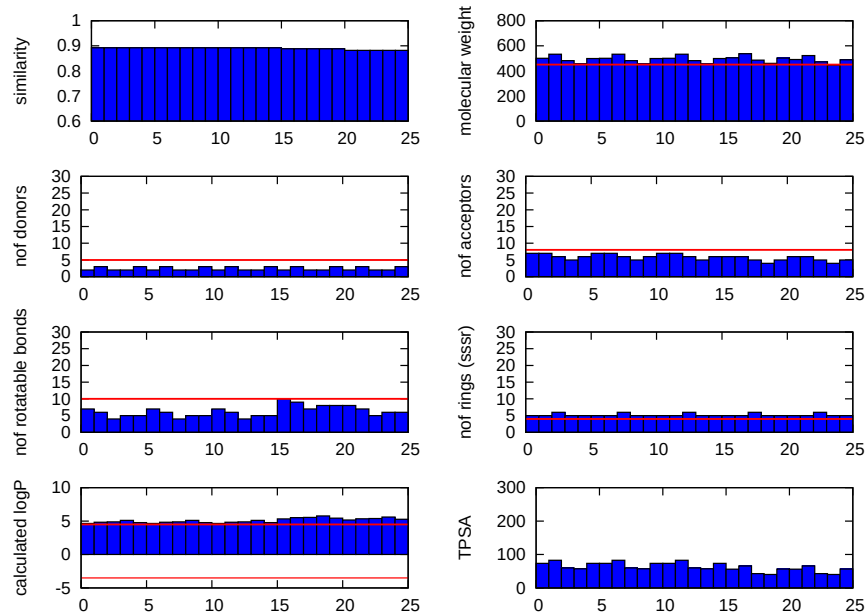
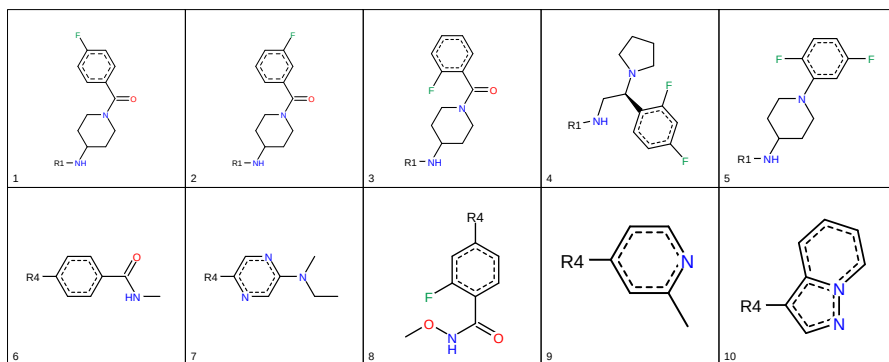


Exceeded leadlike criteria:

MW > 450

nof rings > 4

logP > 4.5



2.9.4 Leadlike

Leadlike criteria were used as product filter.

Seed 1: 0.8222

Seed 2177: 0.8411

Seed 35498: 0.8301

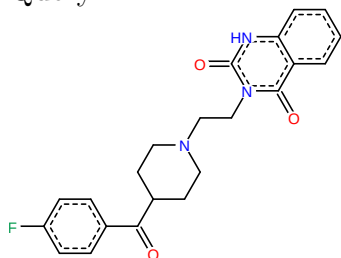
Seed 43: 0.8236

Seed 4543201: 0.8240

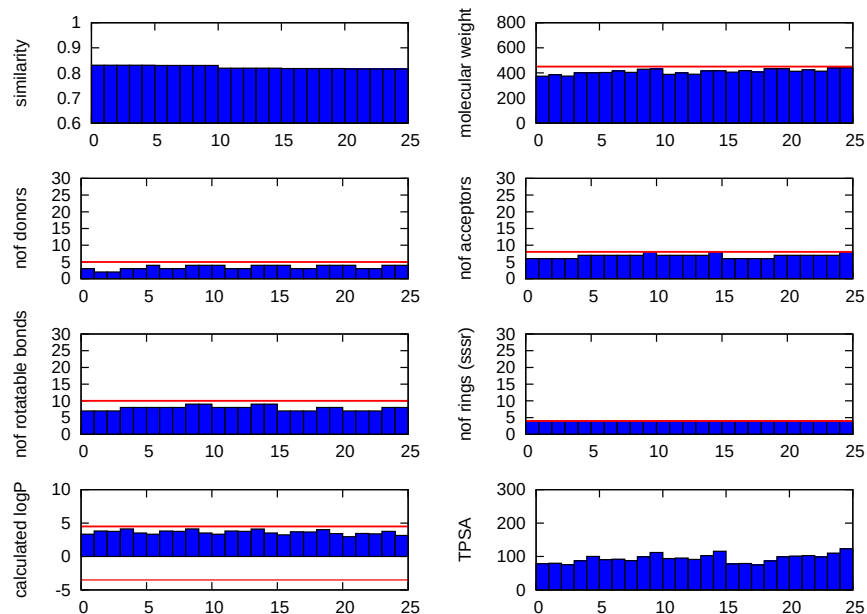
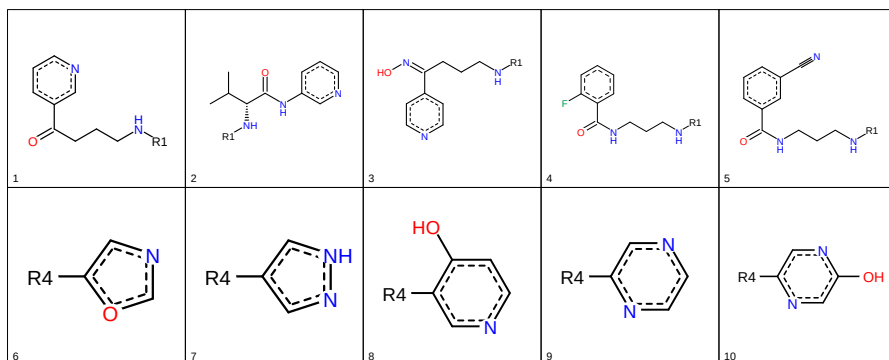
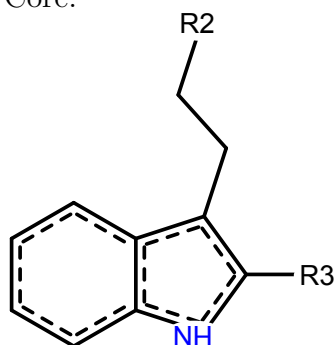
Average score: 0.8282

Solution no 1 with score 0.8222, seen after 180542 steps

Query:



Core:



2.9.5 Diverse

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1.

Seed 1: 0.8830

Seed 2177: 0.8830

Seed 35498: 0.8830

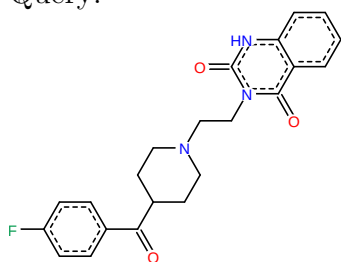
Seed 43: 0.8830

Seed 4543201: 0.8839

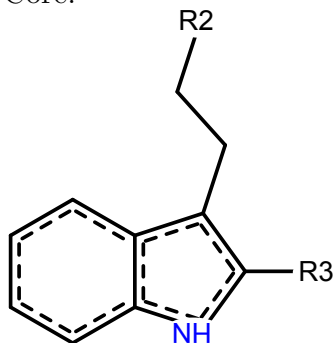
Average score: 0.8832

Solution no 1 with score 0.8830, seen after 166773 steps

Query:



Core:



Exceeded leadlike criteria:

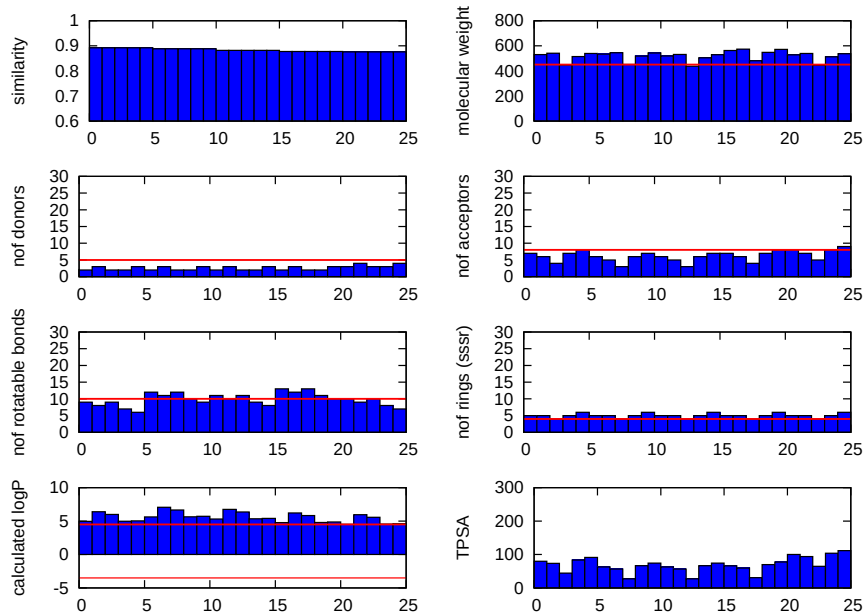
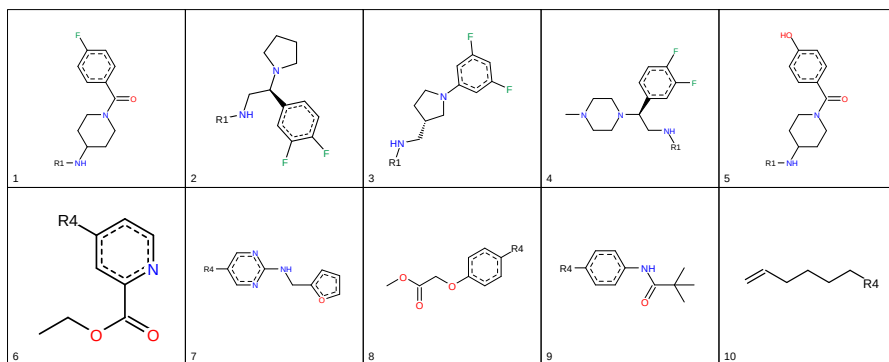
MW > 450

nof rings > 4

nof acceptors > 8

nof rotatable bonds > 10

logP > 4.5



2.9.6 Diverse Leadlike

Maximum one reagent from a cluster is allowed, reagents penalized with -0.2 otherwise. Cluster ids were generated using Complete Linkage clustering with minimum diversity of 0.1. Leadlike criteria were used as product filter.

Seed 1: 0.8223

Seed 2177: 0.8234

Seed 35498: 0.8330

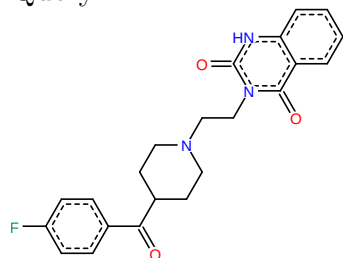
Seed 43: 0.8221

Seed 4543201: 0.8170

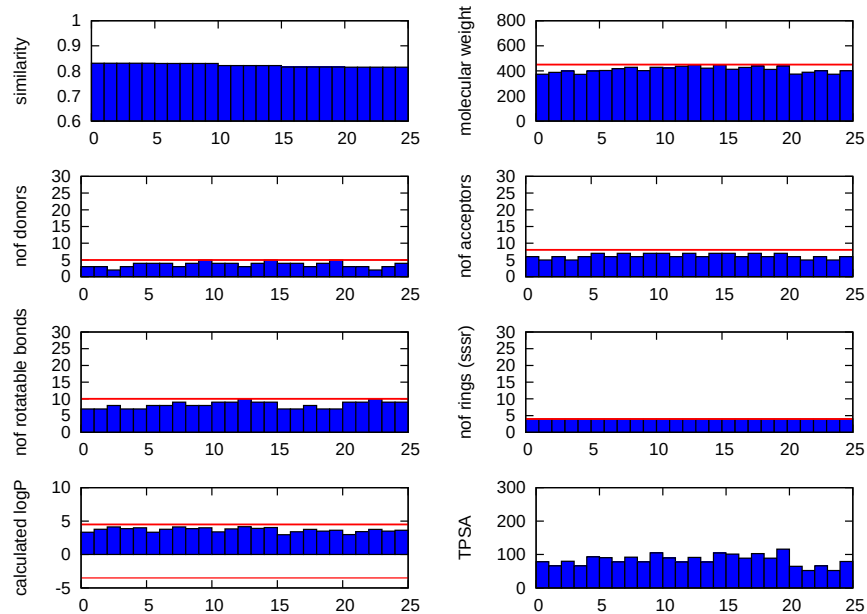
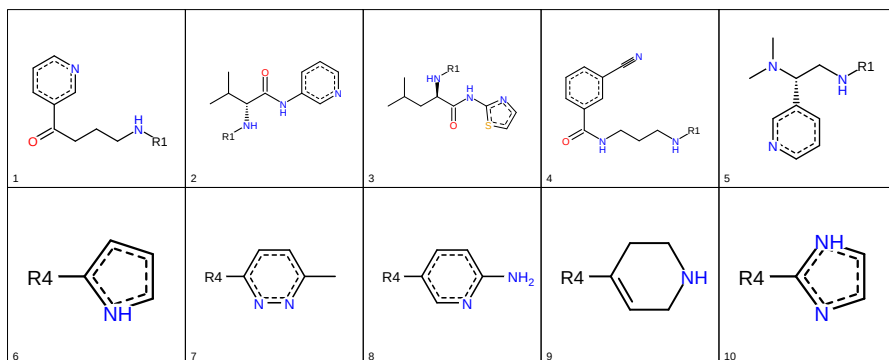
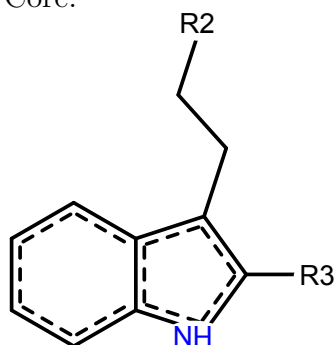
Average score: 0.8236

Solution no 1 with score 0.8223, seen after 154897 steps

Query:



Core:



Bibliography

- [1] Celanire, S., Wijtmans, M., Talaga, P., Leurs, R., & de Esch, I. J. P. (2005). Keynote review: histamine H3 receptor antagonists reach out for the clinic. *Drug Discov Today*, 10(23-24), 1613–1627.
- [2] Lau, J. F., Jeppesen, C. B., Rinvall, K., & Hohlweg, R. (2006). Ureas with histamine H3-antagonist receptor activity—a new scaffold discovered by lead-hopping from cinnamic acid amides. *Bioorg Med Chem Lett*, 16(20), 5303–5308.
- [3] Gray, N. S., Wodicka, L., Thunnissen, A. M., Norman, T. C., Kwon, S., Espinoza, F. H., et al. (1998). Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science*, 281(5376), 533–538.
- [4] Huang, N., Shoichet, B. K., & Irwin, J. J. (2006). Benchmarking sets for molecular docking. *J Med Chem*, 49(23), 6789–6801.
- [5] Smith, A. L., Stevenson, G. I., Lewis, S., Patel, S., & Castro, J. L. (2000). Solid-phase synthesis of 2,3-disubstituted indoles: discovery of a novel, high-affinity, selective h5-HT2A antagonist. *Bioorg Med Chem Lett*, 10(24), 2693–2696.
- [6] Swain, C. J., Teran, A., Maroto, M., & Cabello, A. (2006). Identification and optimisation of 5-amino-7-aryldihydro-1,4-diazepines as 5-HT2A ligands. *Bioorg Med Chem Lett*, 16(23), 6058–6062.
- [7] Fricker, P. C., Gastreich, M., & Rarey, M. (2004). Automated drawing of structural molecular formulas under constraints. *J Chem Inf Comput Sci*, 44(3), 1065–1078.