An evaluation of Catalyst's conformational search algorithm with regard to conformational diversity and conformational energy penalties

Anders Poulsen

The Royal Danish School of Pharmacy

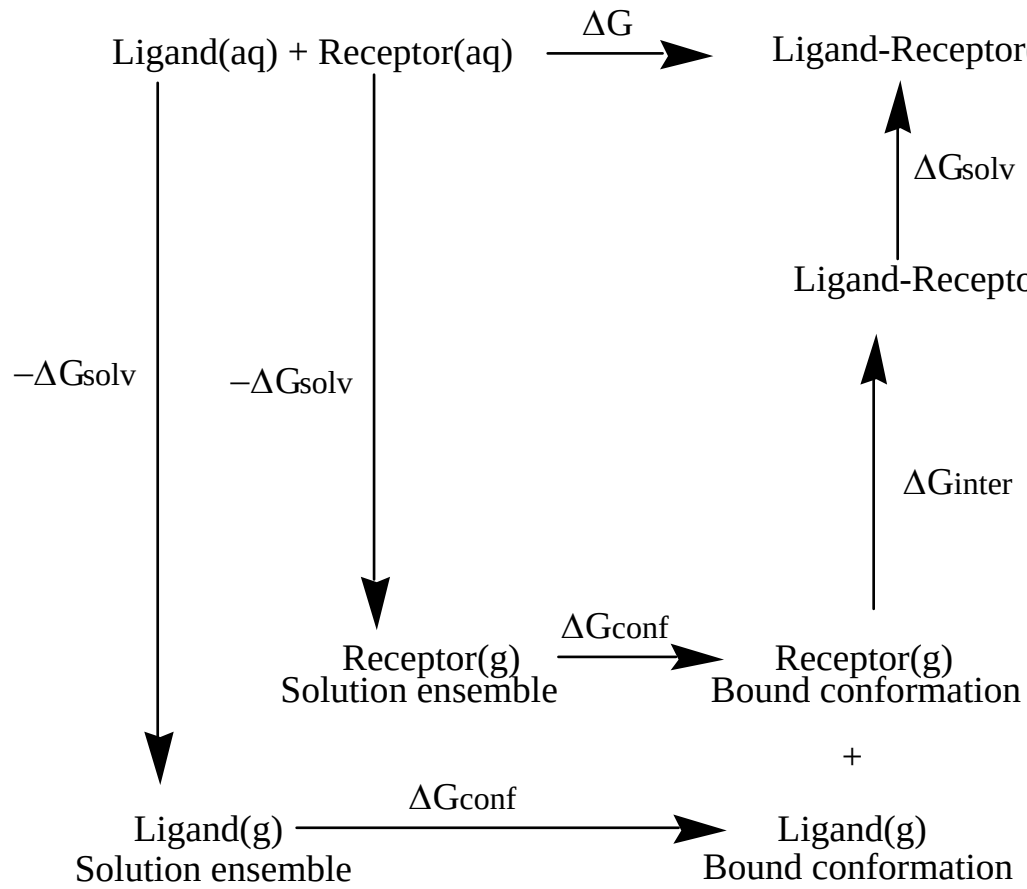
S*BIO™

H. Lundbeck A/S

ATV

Danish Academy of Technical Sciences

Conformational energy



$$\Delta G = -RT \ln K$$

ΔG_{conf}(Ligand) ~
ΔH_{conf}(Ligand Bound conf-Global min)

Bioactive Conformation < 12.6kJ/mol

Bostrom, J., Norrby, P.O., and Liljefors, T., J. Comput.-Aided Mol. Design (1998) 383.

$$\Delta G = \Delta G_{solv}(\text{Ligand-Receptor}) - \Delta G_{solv}(\text{Ligand}) - \Delta G_{solv}(\text{Receptor}) + \Delta G_{conf}(\text{Ligand}) + \Delta G_{conf}(\text{Receptor}) + \Delta G_{inter}$$

S*BIO™

Force Field Performance

Conformational energies calculated by Catalyst compared to experimental values, in kcal/mol.

Rotational Barrier	Catalyst	Experimental	Difference
Ethane	2.74	2.878	-0.1
Propene, methyl rotation	0.82[a]	1.98	-1.2
Isoprene, methyl rotation	0.96	2.71	-1.8
Ethylbenzene, ethyl rotation	1.57	1.7	-0.1
2,4,6-Trimethylisopropylbenzene, isopropyl rotation	11.43	12.8	-1.4
Styrene	7.79	1.78	6.0*
Conformational energies hydrocarbons	**Catalyst**	**Experimental**	**Difference**
Butane, g-a	1.02	0.97	0.1
2,3-Dimethylbutane, g-a	0.18	0.05	0.1
1,3,5-Trineopentylbenzene, twosyn-allsyn	0.54	1.04	-0.5
Conformational energies oxygen containing	**Catalyst**	**Experimental**	**Difference**
Methylacetate, E-Z	0.75	8	-7.3*
2-Butanone, skew-ecl	-0.11	2.0	-2.1
Ethylmethylether, g-a	1.38	1.5	-0.1
2-Methoxy-THP, eq-ax	0.87	1.0	-0.1
Ethanol, (C-O) g-a	-0.06	0.7	-0.8
Propanol, (C-C) g-a	0.68	-0.3	1.0
Conformational energies nitrogen containing	**Catalyst**	**Experimental**	**Difference**
Ethylamine, (C-N) g-a	-0.09	0.7	-0.8
N-Methylacetamide, E-Z	-2.21	2.4	-4.6*
N-Methylpiperidine, ax-eq	0.62	3.2	-2.6
2-Methylpiperidine, ax-eq	1.26	2.5	-1.2
3-Methylpiperidine, ax-eq	1.38	1.6	-0.2
4-Methylpiperidine, ax-eq	1.87	1.93	-0.1

Gundertofte et al., J. Comp. Chem. 17 (1996) 429-449

S*BIO™

Force Field Performance

Conformational energiescyclohexanes	Catalyst	Experimental	Difference
Cyclohexane, twb-ch	6.16	5.5	0.7
Phenylcyclohexane, ax-eq	2.92	2.87	0.0
Methylcyclohexane, ax-eq	1.88	1.75	0.1
Aminocyclohexane, ax-eq	0.89	1.49	-0.6
N,N-Dimethylaminocyclohexane, ax-eq	0.80	1.31	-0.5
Trans-1,2-Dimethylcyclohexane, axax-eqeq	2.42	2.58	-0.2
cic-1,2-Dimethylcyclohexane, axax-eqeq	5.08	5.5	-0.4
Conformational energies of haloalkanes	**Catalyst**	**Experimental**	**Difference**
FCH2CH2F, g-a	0.12	-0.8	0.9
PrCl, g-a	0.43	-0.36	0.8
ClCH2CH2Cl, g-a	0.78	1.05	-0.3
ClCH2CH2CH2Cl, ga-gg	-1.56	1.1	-2.7
ClCH2CH2CH2Cl, aa-gg	-1.96	1.5	-3.5
Conformational energies of halocyclohexanes	**Catalyst**	**Experimental**	**Difference**
F, ax-eq	0.19	0.16	0.0
Cl, ax-eq	0.63	0.5	0.1
Br, ax-eq	0.88	0.7	0.2
Trans-1,2-diF, axax-eqeq	0.28	0.59	-0.3
Trans-1,2-diCl, axax-eqeq	0.31	-0.93	1.2
Trans-1,2-diBr, axax-eqeq	0.07	-1.5	1.6
Trans-1,4-diF, axax-eqeq	0.37	-1.14	1.5
Trans-1,4-diCl, axax-eqeq	1.28	-0.8	2.1
Trans-1,4-diBr, axax-eqeq	1.79	-0.88	2.7
Conformational energies of conjugated compounds	**Catalyst**	**Experimental**	**Difference**
Butadiene, s-cic-s-trans	1.91	2.5	-0.6
Acrolein, s-cic-s-trans	3.31	1.7	1.6

Gundertofte et al., J. Comp. Chem. 17 (1996) 429-449

S*BIO™

Force Field Performance

Average absolute errors, in kcal/mol.

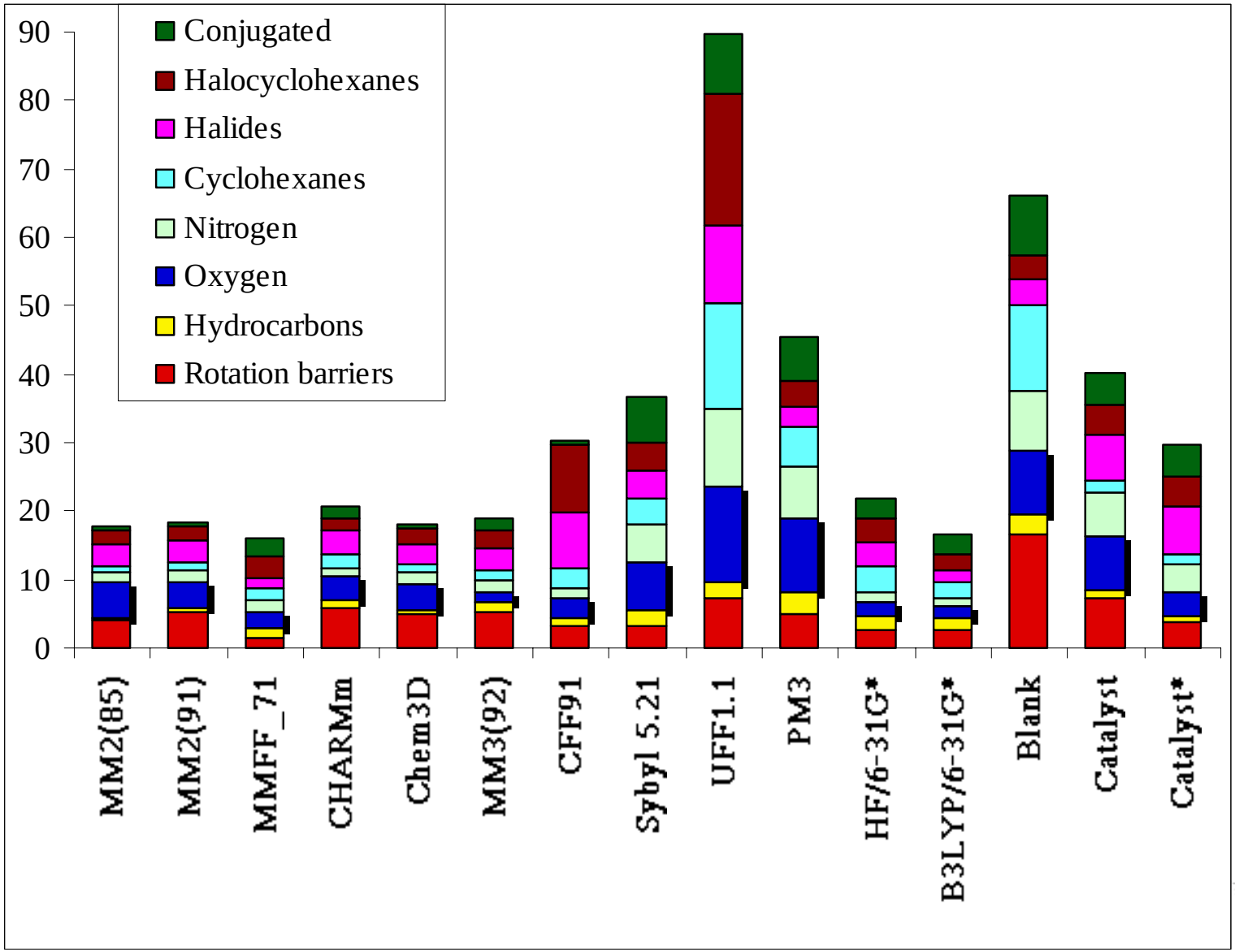
Compound class	Catalyst	MM3* Macromodel 4.0	MMFFs 93 Ceriuc 2	Blank
Rotational barriers	1.76	1.00	0.35	3.97
Hydrocarbons	0.23	0.36	0.39	0.69
Oxygen compounds	1.89	0.50	0.52	2.34
Nitrogen compounds	1.58	0.38	0.43	2.05
Cyclohexanes	0.36	0.43	0.71	3.00
Haloalkanes	1.62	0.72	0.39	0.96
Halocyclohexanes	1.08	0.69	0.68	0.80
Conjugated compounds	1.10	0.46	0.19	2.10
Total	1.20	0.60	0.51	2.03
Total excluding Rotational barriers	1.12	0.53	0.53	1.73
Rotational barriers*	0.91	1.00	0.35	3.97
Oxygen compounds*	0.82	0.50	0.52	2.34
Nitrogen compounds*	0.98	0.38	0.43	2.05
Total*	0.89	0.60	0.51	2.03
Total excluding Rotational barriers*	0.88	0.53	0.53	1.73

* Excluding largest difference from catalyst result

Gundertofte et al., J. Comp. Chem. 17 (1996) 429-449

S*BIO™

Force Field Performance



Mean Absolute Error/class in kJ/mol

Gundertofte et al., J. Comp. Chem. 17 (1996) 429-449

S*BIO™

Catalyst

The primary aims of Catalysts conformation generation module are 1: Speed. 2: To explore compounds in terms of all the energetically accessible conformations available under physiological conditions http://www.accelrys.com/

Catalyst searches feature space. More diverse sampling than other conformational search algorithms.

S*BIO™

Overview

Catalyst

Best
Fast

Generate Ensemble

Ensemble
Global E. Minima

Export

MacroModel

MM3* ± GB/SA
MMFFs ± GB/SA

Minimise

Ensemble
Global E. Minima

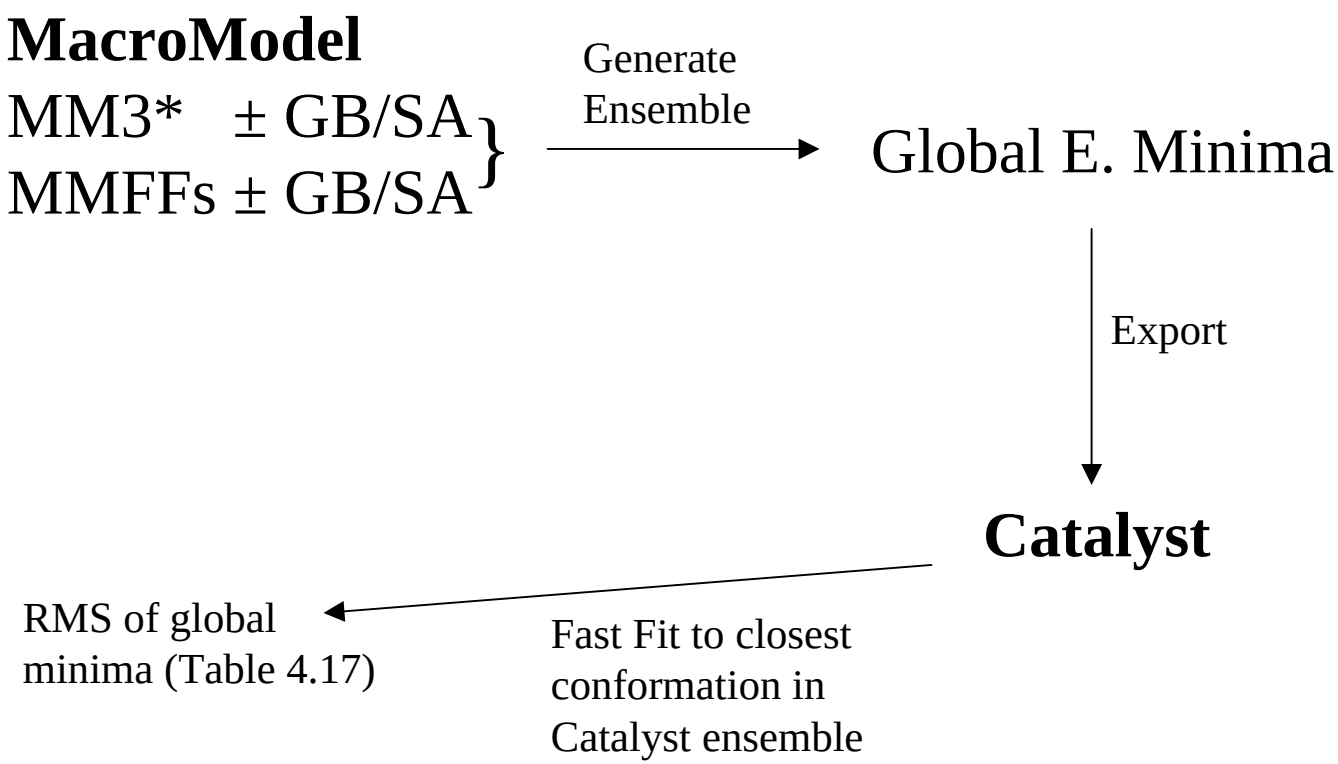
Average energy of ensemble generated by various energy limits

Average energy of ensemble

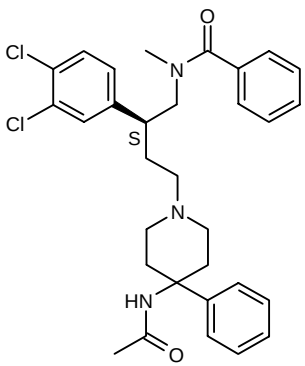
Energy and RMS of global minima (

S*BIO™

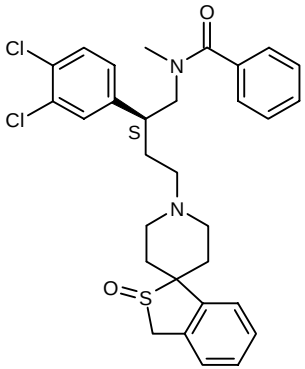
Overview


MacroModel
MM3* ± GB/SA
MMFFs ± GB/SA
Generate Ensemble
Global E. Minima
Export
Catalyst
Fast Fit to closest conformation in Catalyst ensemble
RMS of global minima (Table 4.17)

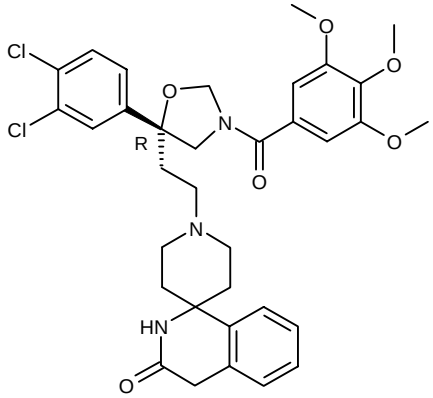
Compounds



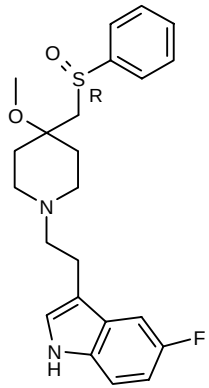
1



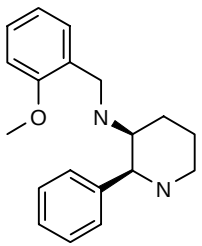
2



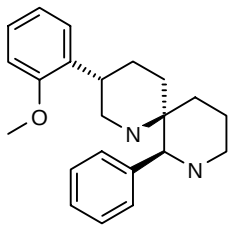
3



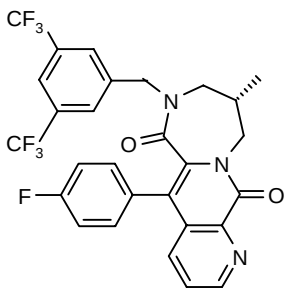
4



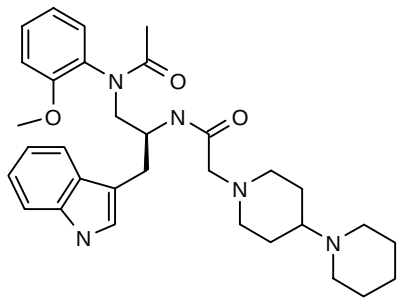
5



6



7



8

Selective NK1 and NK2 antagonists. Dual NK1 and NK2 antagonists. Varying degree of flexibility

Energy of Global energy minima calculated by other force fields

Catalyst —Conf. Model→ Global Minima —Export→ MacroModel —Minimise→ Energy

Compound	Catalyst Method	MM3*	MM3* GB/SA	MMFFs	MMFFs GB/SA	MM3* Constr.	MM3* GB/SA Constr.	MMFFs Constr.	MMFFs GB/SA Constr.
1	BEST	37.5	24.7	39.1	20.2	41.0	33.0	50.6	39.0
1	FAST	26.3	9.6	37.6	4.8	29.9	16.8	39.7	14.6
2	BEST	13.6	8.2	22.3	10.0	14.1	10.5	22.6	17.9
2	FAST	28.6	13.4	37.9	4.5	30.6	16.8	40.5	8.8
3	BEST	16.7	6.1	15.1	3.2	28.4	18.4	44.8	33.4
3	FAST	34.3	24.8	57.6	24.6	61.9	60.5	70.0	55.4
4	BEST	11.4	4.5	33.5	14.7	12.3	6.3	39.4	18.1
4	FAST	18.1	10.7	21.3	5.6	32.5	18.0	51.2	21.9
5	BEST	28.5	23.0	13.7	1.5	30.8	25.0	16.0	3.2
5	FAST	28.2	20.6	24.2	6.7	33.5	25.8	31.9	14.5
6	BEST	4.9	9.0	17.1	10.3	8.6	13.8	27.6	22.8
6	FAST	24.6	37.4	61.4	63.0	71.1	81.3	86.7	90.3
7	BEST	0.3	1.0	0.1	0.3	48.0	34.2	24.4	4.9
7	FAST	0.1	1.0	4.5	0.3	26.9	18.9	17.1	6.6
8	BEST	46.0	33.0	49.0	32.5	40.0	32.9	103.9	81.1
8	FAST	31.7	14.8	28.6	10.1	18.9	13.2	33.2	20.6
Average	BEST	19.9	13.7	23.7	11.6	27.9	21.8	41.2	27.6
Average	FAST	24.0	16.5	34.1	15.0	38.2	31.4	46.3	29.1

	Within 4.2 kJ
	Within 8.4 kJ
	Within 12.6 kJ

Best lower in energy than Fast

Only fully minimised structures low in energy

S*BIO™

Structure of Global energy minima minimised by other force fields

Catalyst —Conf. Model→ Global Minima —Export→ MacroModel —Minimise→ RMS

Compound	Catalyst Method	MM3*	MM3* GB/SA	MMFFs	MMFFs GB/SA	MM3* Constr.	MM3* GB/SA Constr.	MMFFs Constr.	MMFFs GB/SA Constr.
1	BEST	1.259	1.016	0.899	0.747	0.139	0.140	0.139	0.139
1	FAST	0.539	0.506	0.473	0.413	0.124	0.124	0.131	0.131
2	BEST	0.355	0.398	0.405	0.417	0.129	0.128	0.134	0.128
2	FAST	0.368	0.338	0.484	0.373	0.130	0.126	0.130	0.131
3	BEST	0.800	0.886	1.077	1.086	0.147	0.147	0.143	0.144
3	FAST	1.371	1.107	0.581	1.018	0.130	0.132	0.146	0.147
4	BEST	0.409	0.407	0.594	0.476	0.127	0.139	0.142	0.139
4	FAST	0.648	0.699	0.596	0.560	0.139	0.138	0.139	0.138
5	BEST	0.580	0.523	0.579	0.468	0.126	0.125	0.127	0.124
5	FAST	0.736	0.660	0.362	0.339	0.139	0.135	0.132	0.134
6	BEST	0.404	0.454	0.541	0.569	0.129	0.130	0.142	0.141
6	FAST	1.797	1.797	0.680	0.723	0.143	0.144	0.149	0.150
7	BEST	0.905	0.875	0.920	0.774	0.104	0.106	0.124	0.125
7	FAST	0.607	0.567	0.613	0.504	0.122	0.121	0.123	0.123
8	BEST	0.519	0.579	2.598	2.547	0.148	0.145	0.155	0.155
8	FAST	0.527	0.841	0.652	0.716	0.131	0.131	0.148	0.147
Average	BEST	0.654	0.642	0.952	0.886	0.131	0.133	0.138	0.137
Average	FAST	0.824	0.814	0.555	0.581	0.132	0.131	0.137	0.138

Most of catalysts global energy minima changes conformation upon minimization

Far from local minima
High conformational energy

Within 0.5 Å

S*BIO

Average energy of ensemble calculated by other force fields

Catalyst —Conf. Model→ Ensemble —Export→ MacroModel —Minimise→ Energy

Compound	Catalyst Method	MM3*	MM3* GB/SA	MMFFs	MMFFs GB/SA	MM3* Constr.	MM3* GB/SA Constr.	MMFFs Constr.	MMFFs GB/SA Constr.	Catalyst	Number of Conf.
1	BEST	40.4	35.2	39.0	28.9	66.8	62.1	77.9	68.8	47.8	192
1	FAST	37.2	29.2	40.7	24.8	53.7	45.5	66.3	51.1	48.3	199
2	BEST	27.7	25.0	33.5	20.4	53.1	50.0	73.7	62.3	45.3	183
2	FAST	33.8	29.0	42.0	24.8	49.7	45.1	65.0	49.1	55.3	100
3	BEST	39.1	35.0	39.6	25.6	56.4	52.1	74.7	65.7	44.5	107
3	FAST	45.3	38.3	54.0	32.3	78.1	71.6	89.8	70.4	49.8	225
4	BEST	25.7	22.5	41.9	33.9	49.0	45.3	81.6	70.7	43.7	230
4	FAST	36.1	32.0	44.3	36.7	55.4	51.8	77.0	65.0	41.6	225
5	BEST	31.5	26.7	27.5	15.2	54.9	48.8	60.7	46.8	49.7	71
5	FAST	27.6	22.7	15.8	7.0	40.4	35.4	38.9	29.3	25.3	10
6	BEST	30.5	38.4	45.1	44.4	59.3	66.6	79.0	75.4	54.2	56
6	FAST	24.7	35.8	52.6	58.1	71.1	81.3	93.0	97.4	44.2	5
7	BEST	10.8	3.3	9.7	2.8	41.9	33.1	47.1	27.7	32.8	140
7	FAST	11.5	4.7	15.2	3.6	42.3	36.2	40.7	25.3	29.7	67
8	BEST	54.0	41.7	66.2	47.3	76.9	66.2	125.6	96.6	41.1	202
8	FAST	55.8	42.4	72.6	46.7	65.3	53.8	110.8	77.0	47.2	216
Average	BEST	32.5	28.5	37.8	27.3	57.3	53.0	77.5	64.3	44.9	147.6
Average	FAST	34.0	29.3	42.2	29.2	57.0	52.6	72.7	58.1	42.7	130.9

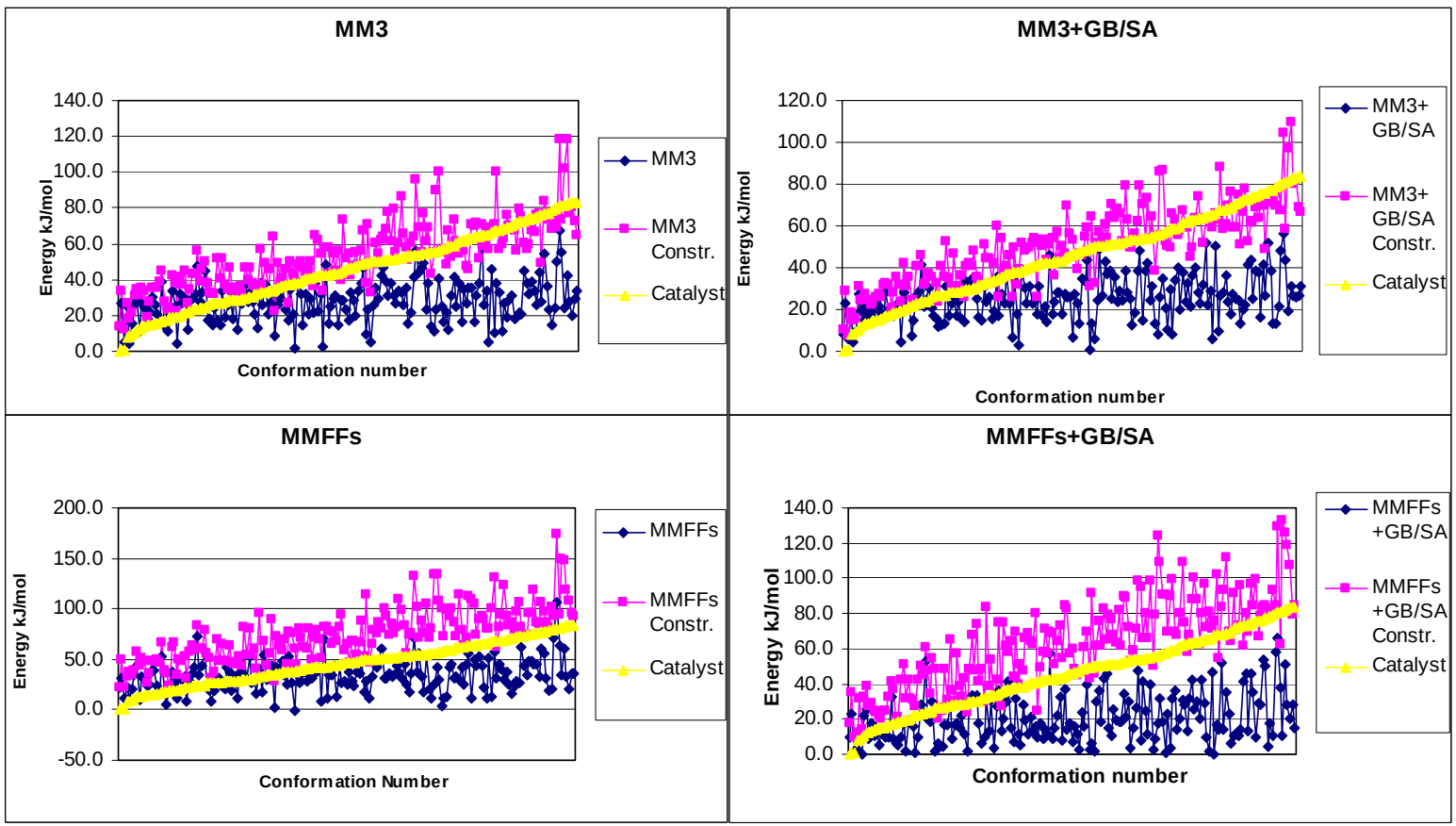
	Below Glob. min.
	Within 4.2 kJ/mol of glob. min.
	Within 8.4 kJ/mol of glob. min.

Generally within the standard 84kJ energy limit. Fast and Best equal in energy.

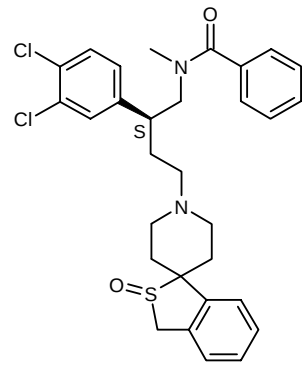
Average energy comparable to Global minima

S*BIO™

Energy of ensemble calculated by various force fields



Energy penalty calculated by Catalyst is almost a random number



2

Structure of Global energy minima found by other force fields

MacroModel —Conf. Model→ Global Minima —Export→ Catalyst —Fast Fit→ RMS

	Catalyst Best search					Catalyst Fast search				
Compound	MM3*	MM3* GB/SA	MMFFs	MMFFs GB/SA	Catalyst Fast	MM3*	MM3* GB/SA	MMFFs	MMFFs GB/SA	Catalyst Best
1	1.440	1.657	1.323	1.694	1.700	1.203	1.633	1.358	1.633	1.977
2	1.494	1.594	1.121	1.360	1.287	1.871	1.970	0.938	1.358	1.505
3	1.945	1.935	1.820	1.985	1.060	2.477	1.231	1.697	1.773	1.346
4	1.395	0.751	0.723	1.005	1.601	1.553	1.364	1.648	1.716	1.606
5	0.767	0.754	1.306	1.307	0.987	1.289	1.295	1.260	1.237	1.361
6	0.736	0.409	0.577	0.592	0.824	1.657	1.551	1.714	1.709	1.559
7	0.547	1.065	0.482	1.201	0.197	0.759	0.877	0.597	0.666	0.483
8	1.942	1.909	1.548	1.889	1.952	1.685	1.776	1.182	1.802	2.446
Average	1.283	1.259	1.113	1.379	1.201	1.562	1.462	1.299	1.487	1.535

	Within 0.5 Å
	Within 1 Å

Trend: Fast search gives higher RMS than Best search

Global energy minima found by other methods generally not precent in catalysts ensemble

S*BIO™

Energy limits for conformational search

The mean energies of the poled set were somewhat higher than those of the unpoled set, consistent with our goal of covering conformational space with respect to a user-defined energy threshold rather than just elucidating local minima.

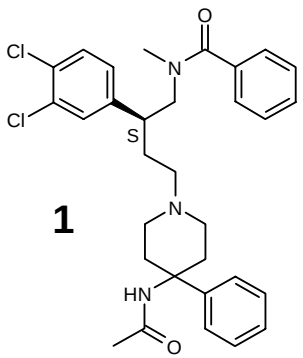
A. Smellie et al., J. Comp. Chem., Vol. 16, No. 2, 171-187 (1995)

S*BIO™

Energy limits for conformational search

Catalyst —Conf. Model→ Ensemble —Export→ MacroModel —Minimise→ Energy

	Catalyst Fast search				Catalyst Best search			
Energy limmit	MMFFs	MMSFFs Constr	Catalyst	Number of conf.	MMFFs	MMSFFs Constr	Catalyst	Number of conf.
2.1	15.0	31.2	0.2	2	20.4	31.6	0.6	2
4.2	14.7	32.5	1.3	2	26.1	32.7	1.3	3
8.4	41.1	51.6	2.8	2	34.9	40.9	4.2	7
12.6	41.1	43.2	1.4	2	26.7	34.6	7.7	9
16.8	30.2	44.0	9.7	19	29.7	36.9	8.2	10
21	31.7	45.4	14.5	77	27.6	36	9.6	12
42	32.0	51.8	22.6	74	39.2	54.3	19.3	44
63	32.1	55.1	37.7	185	39.5	69.5	34.1	134
84	34.4	60.1	43.6	236	34.9	74.9	43.8	197



1

Low E limit insufficient coverage of conf. space

S*BIO™

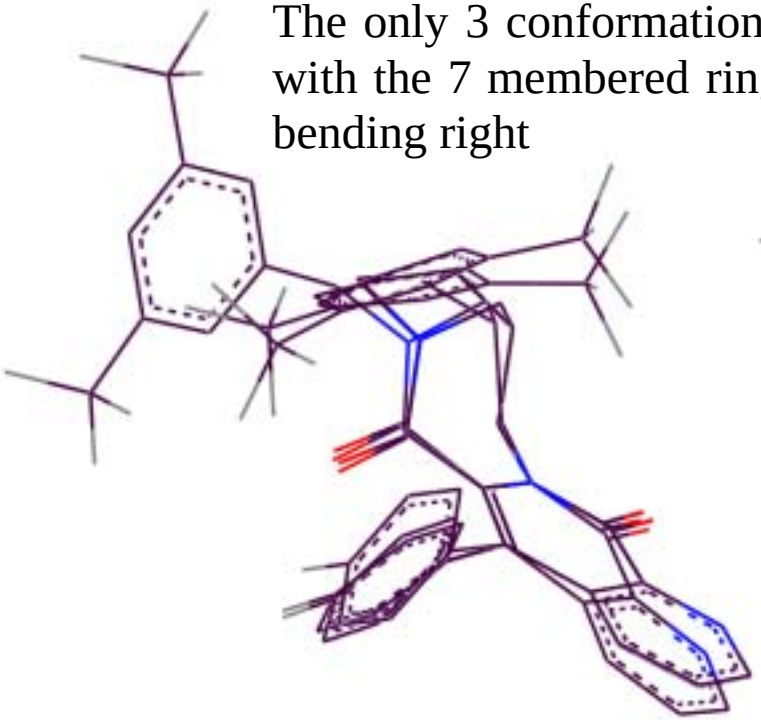
Conformational diversity

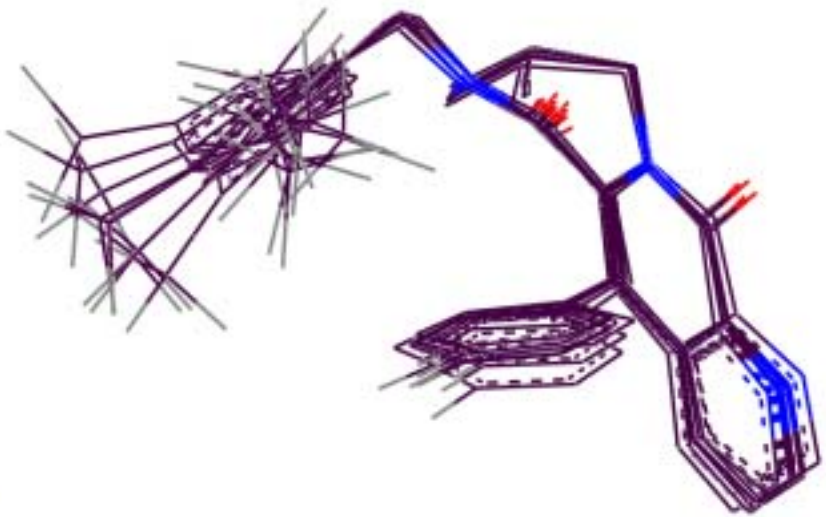
The 7 membered ring can bend to the left or right of the tricyclic ring system

Catalyst Best Search

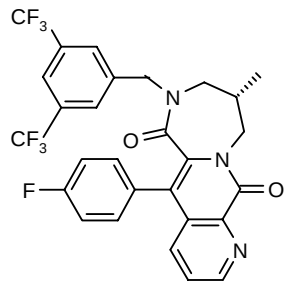
Method	Right	Left
Catalyst Best	3	137
Catalyst Fast	24	43
MMFFs	29	24
MMFFs+GB/SA	30	47
MM3*	9	8
MM3*+GB/SA	9	7

The only 3 conformations with the 7 membered ring bending right





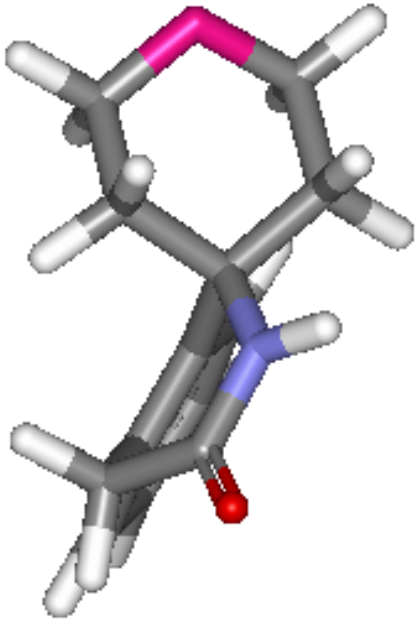
10 first conf. in ensemble

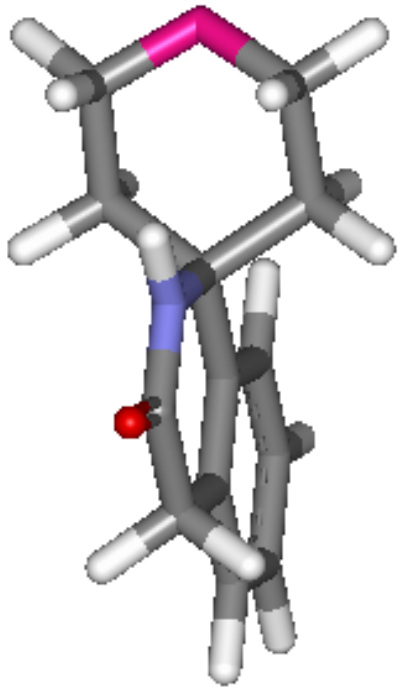


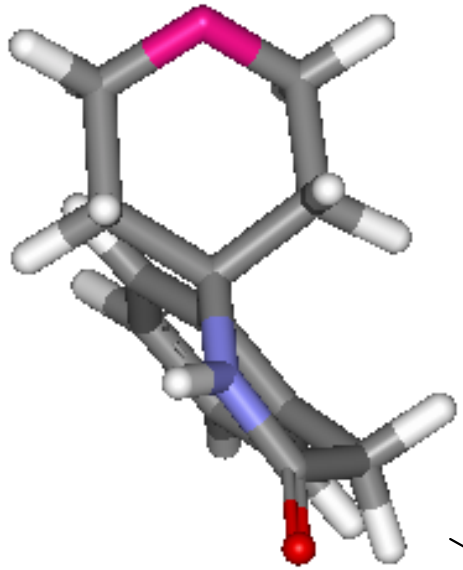
7

S*BIO™

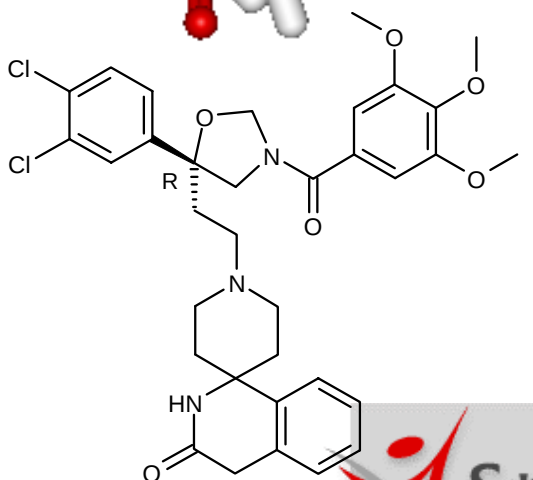
Conformational diversity







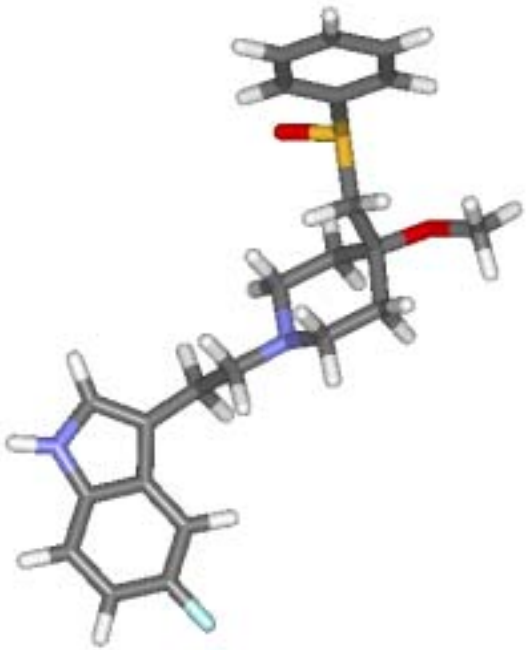
Method	Left	Center	Right
Catalyst Best	94%	2.6%	3.5%
Catalyst Fast	15%	85%	0%
MMFFs MMFFs+GB/SA MM3* MM3*+GB/SA	~50%	0%	~50%

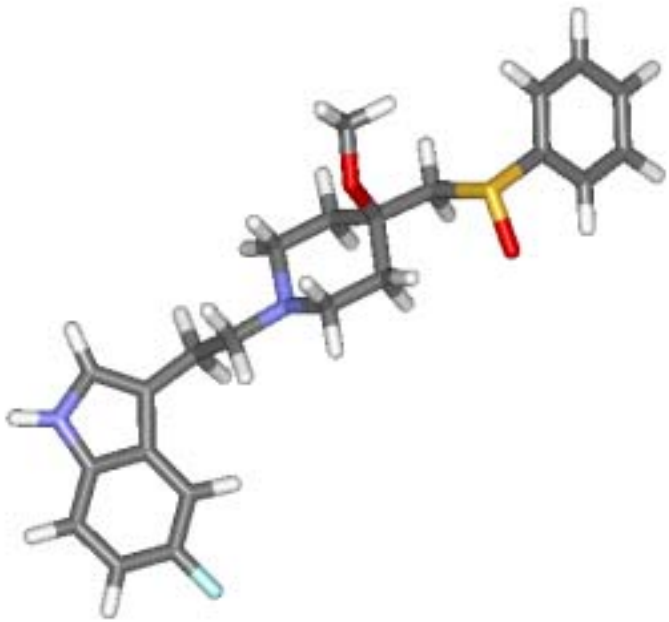


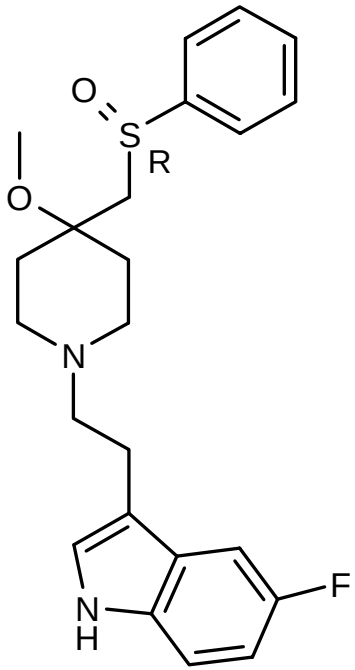
S*BIO™

Conformational diversity

Catalyst Method	Fast	Best
Axial Chair	0 (0%)	2 (0.8%)
Axial Twist	0 (0%)	40 (17%)
Equatorial	49 (100%)	188 (81%)
No. Conformations	49	230



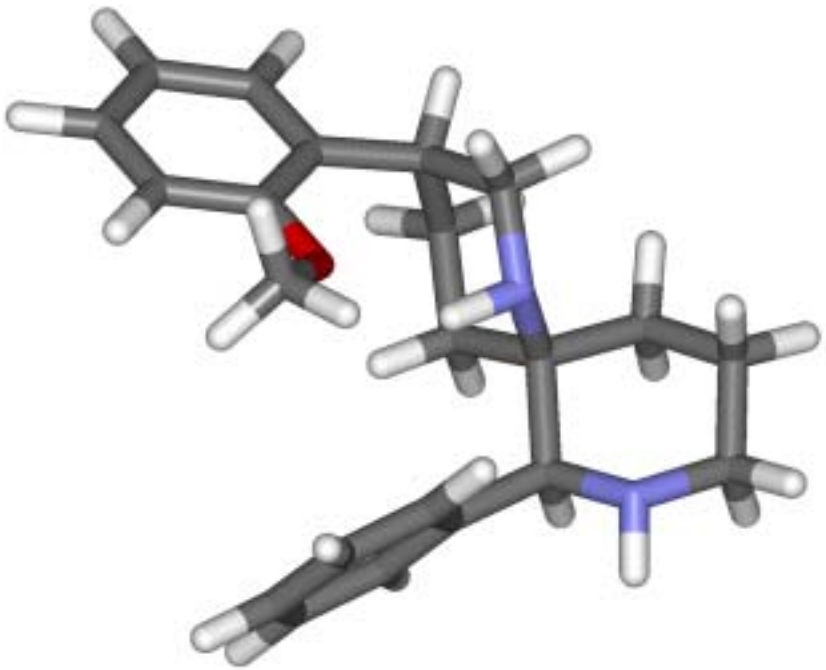


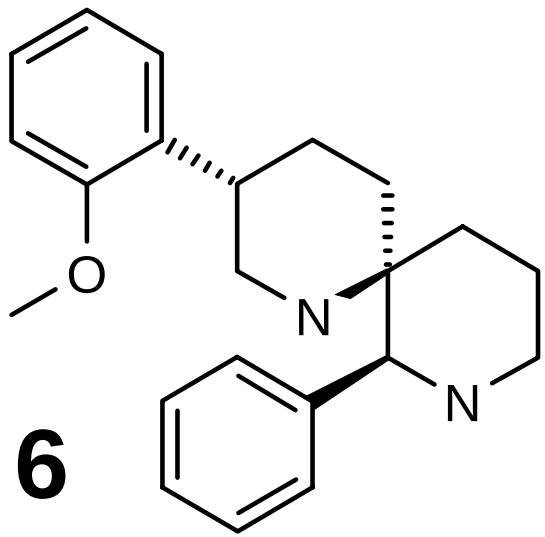


4

Conformational diversity

Catalyst Method	Fast	Best
Chair Chair	0 (0%)	3 (5.3%)
Unexpected	5 (100%)	53 (94.6%)
No. Conformations	5	56

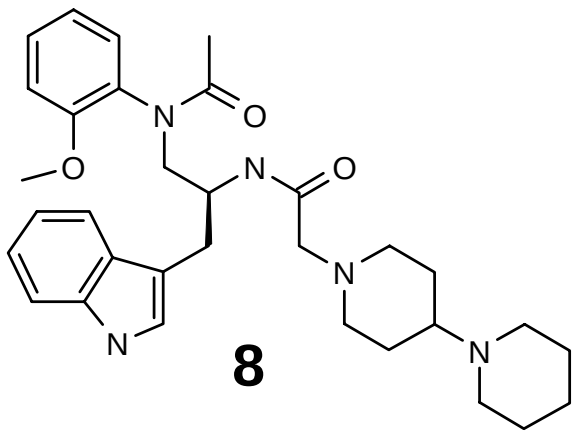




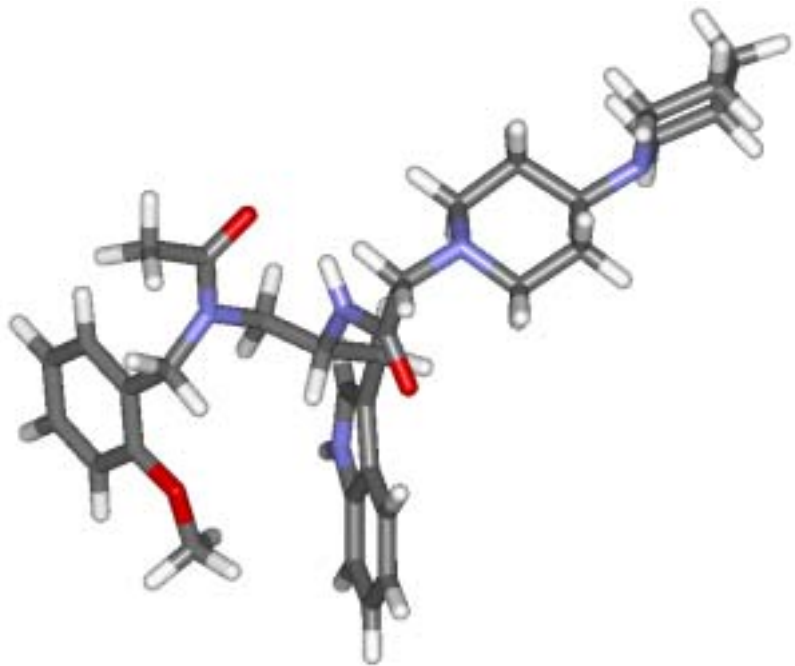
6

Conformational diversity

Catalyst Method	Fast	Best
Chair Chair Equatorial substituents	36 (17%)	4 (1.8%)
Chair Twist Equatorial substituents	9 (4.2%)	6 (2.7%)
Non Extended Conformations	171 (79.2%)	212 (95.5%)
No. conformations	216	222

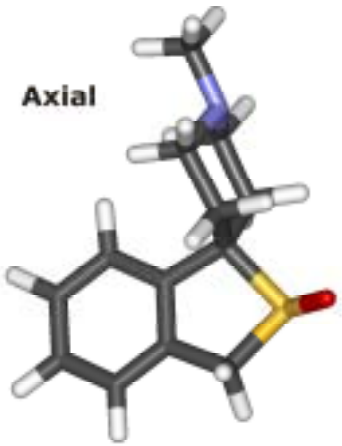


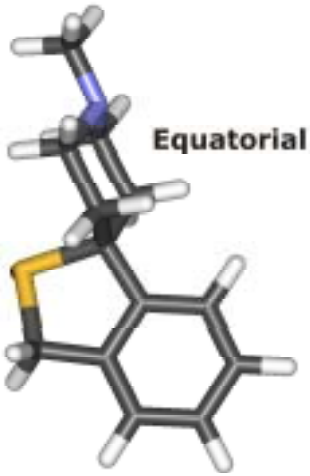
8



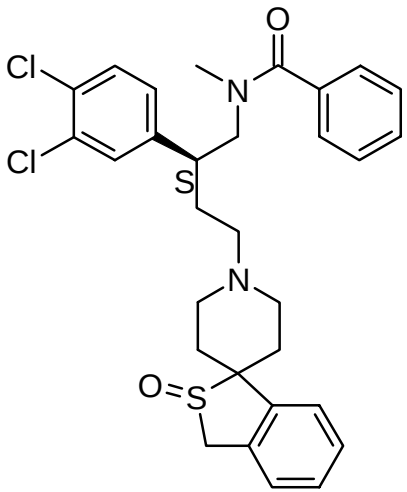
S*BIO™

Conformational diversity





Conformation	Axial	Equatorial	Number
Catalyst Best	7	176	183
Catalyst Fast	22	78	100
MMFFs			
MMFFs+GB/SA	~50%	~50%	
MM3*			
MM3*+GB/SA			



Conformational diversity

6 out of 8 conformational models generated by Catalyst are not diverse

Diversity in feature space is not the same as diversity in conformational space

Poling: Promoting Conformational Variation
A. Smellie et al., J. Comp. Chem., Vol. 16, No. 2, 171-187 (1995)

Poling: Preventing Conformational Variation

S*BIO™

Conclusion

The primary aim of Catalysts conformation generation module is to explore compounds in terms of all the energetically accessible conformations available under physiological conditions.

The present work clearly demonstrates that this is not achieved, unless physiological conditions includes Dante's version of Hell and crematorium ovens

Conclusion

The force Field performance are comparable to Tripos 3.0 and better than UFF 1.01, Dreiding 2.21 and CVFF

Significantly better if problem with Cis/Trans and halocarbons are fixed

Most conformations generated by Catalyst are far from a local minima and high in energy.

The rank ordering of the conformational model is more or less arbitrary. (Conformations are two high in energy that it makes sense to use a force field to calculate a energy)

S*BIO™

Conclusion

Diverse sampling of feature space does not always result in a diverse conformational model. (The diversity that we want to see are not observed)

The conformational diversity is often seen in distortions of low energy conformations. (The diversity that we observe are not the diversity that we want to see)

Especially for rings, high energy conformations are over represented in the conformational model. Sometimes so much that no low energy conformation is found for the ring system.

S*BIO™

Conclusion

The high energy conformations are just noise that results in false positives when doing a database search.

The missing low energy conformations results in false negatives.

This noise and missing low energy conformations are dangerous when doing automatic hypothesis generation, since it leads to wrong models.

Catalyst is a good tool for 3D database search.

Substituting the conformational search algorithm for a fragmented algorithm without poling would greatly enhance the quality of the program.

S*BIO™

Acknowledgement

Financial support for my Ph.D. Project:

Lundbeck

ATV

Danish Academy of
Technical Sciences

Ardent support of free speech:

accelrys

S*BIO™