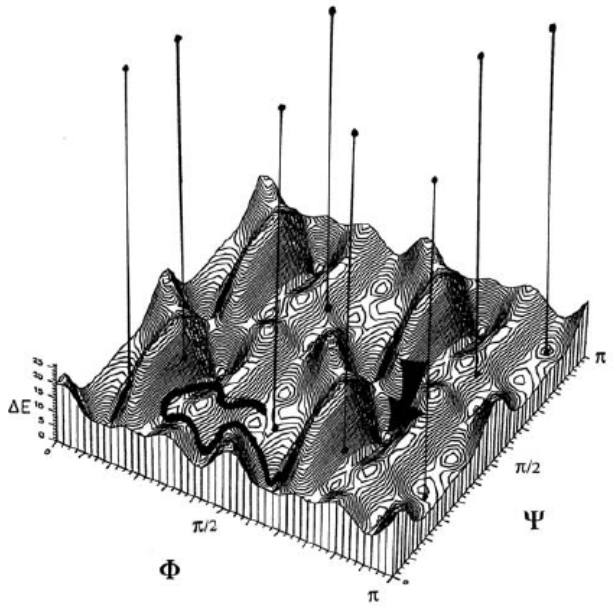
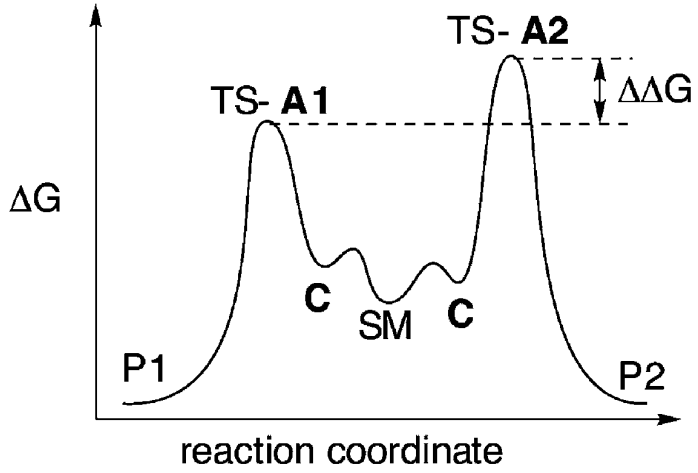
Predicative Design of Stereoinduction in Catalytic, Asymmetric Reactions



Structure
To
Mechanism
?
Mechanism
To
Structure

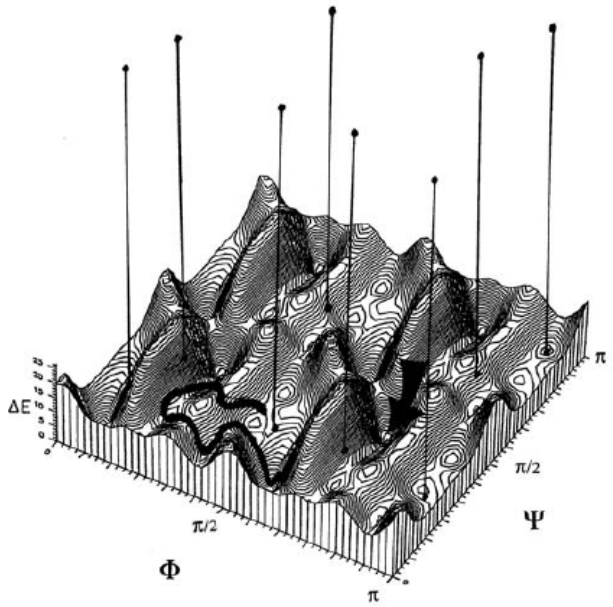


William Collins

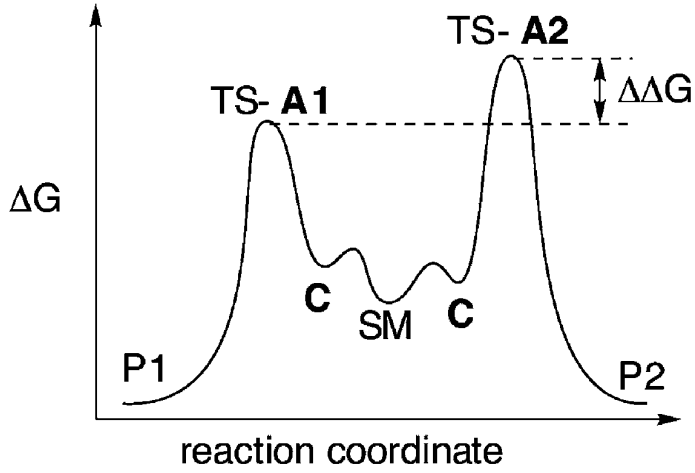
Denmark Group Meeting

(08-28-07)

Predicative Design of Stereoinduction in Catalytic, Asymmetric Reactions



Structure
To
Mechanism
?
Mechanism
To
Structure

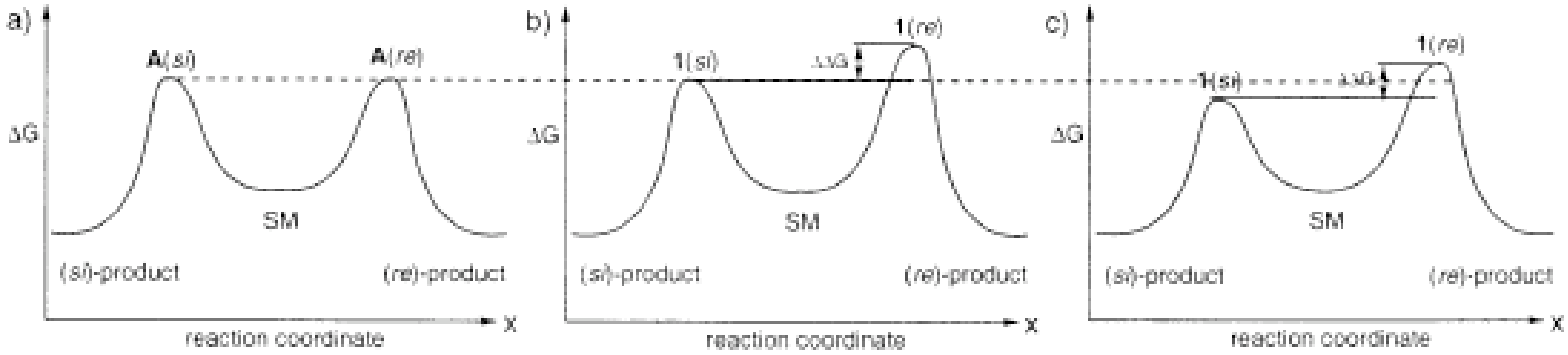


"…in spite of literally thousands of chiral ligands that have been reported in the past, there is no unique rationale for the efficiency of all catalysts not even for those that have been applied to the same reaction. Therefore the design of new catalysts is based on trial and error."

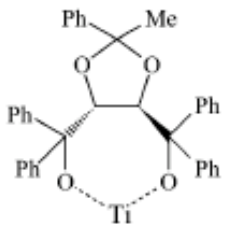
Holz, J. *et. al. Eur. J. Org. Chem.* (2001), 4615.

Chiral Catalysts

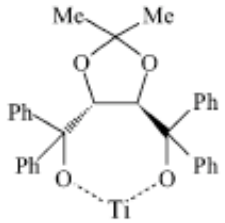
What are the requirements for successful enantiodiscrimination (or diastereo)?



For the Diels-Alder reaction of 3-acryloyl-1,3-oxazolidininone with cyclopentadiene:



Highly Selective



Very low Selectivity

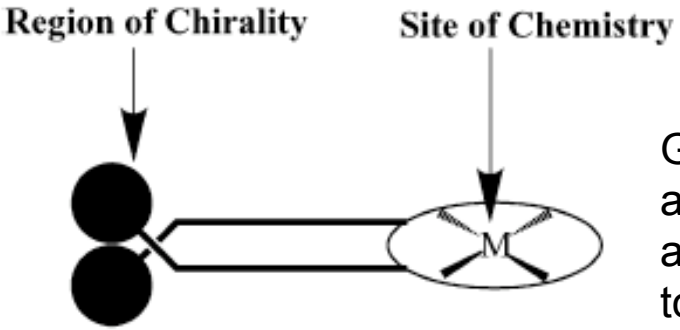
DG (kcal/mol)	X:Y	%X : %Y
0.41	1:2	33 : 67
0.65	1:3	25 : 75
0.95	1:5	17 : 83
1.36	1:10	9 : 91
2.72	1:100	1 : 99
4.08	1:1000	0.1 : 99.1

There is a clear lack of guidelines or rules for chiral catalyst design…

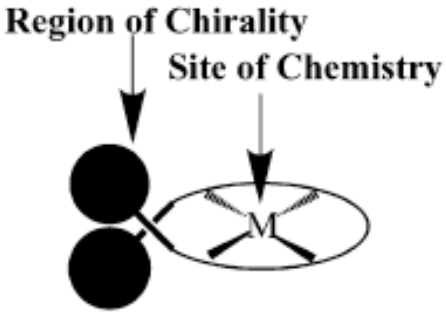
- Kozlowski, M.; Panda, M. *J. Org. Chem.* (2003), **68**, 2061.

Stereocartography

Prove or Disprove: Maximum asymmetric induction by a catalyst is achieved when the region of greatest stereoinduction is spatially congruent with the site of chemistry



Generalization for the intermolecular interactions between a chiral ligand and the reagents undergoing a chemical trans. at the M center fall off as $1/r$ (ion-ion) to $1/r^2$ (ion-dipole) to $1/r^3$ (dipole-dipole) to $1/r^6$ (dispersion)

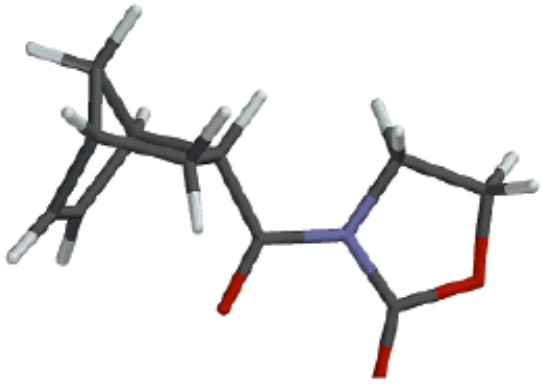


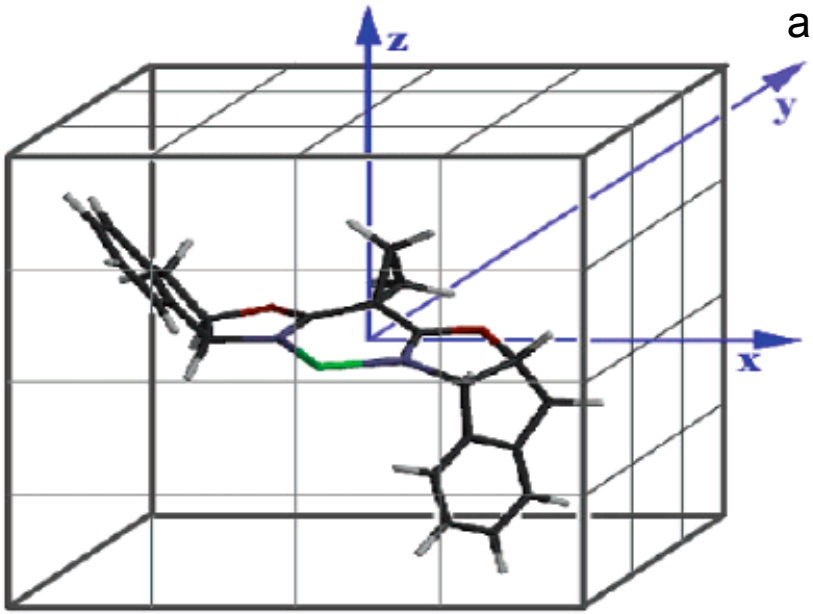
Having “chirality” at the reaction site will “desymmetrize” an otherwise degenerate set of reaction coordinates

How do we prove this?

- Lipkowitz, K. *et. al. J. Amer. Chem. Soc.* (2002), **124,** 14255.

Stereocartography





"Mapping" of stereodiscriminating regions around a chiral catalyst:

1) Place the catalysts center of mass @ origin of a Cartesian coordinate system

2) Chiral "probe" molecule = TS of molecules reacting In the presence of the catalyst

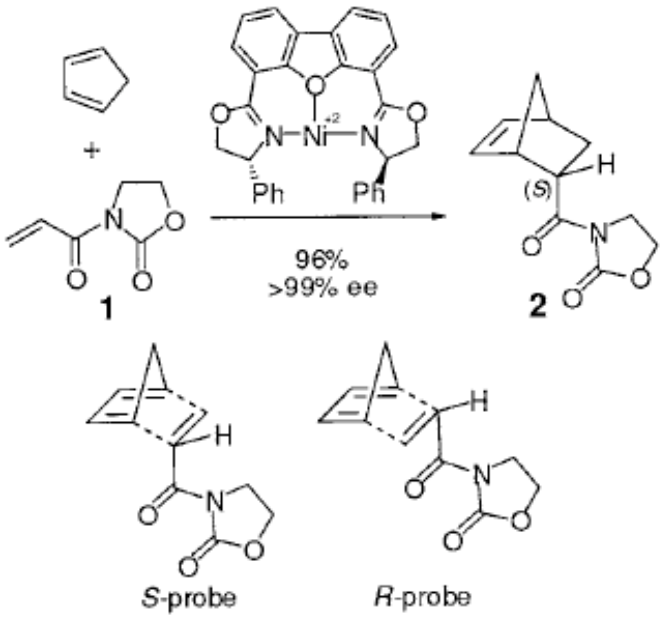
3) At each grid point, a Boltzmann weighted energy is determined between the probe and the catalyst for a large number of probe orientations

-Both (R) and (S) antipodes of probe are considered in the calculations

-where interpentration of molecules occur, the data is ignored

- Lipkowitz, K. *et. al. J. Amer. Chem. Soc.* (2002), **124,** 14255.

Stereocartography example:



Plotting an isodesmic map:

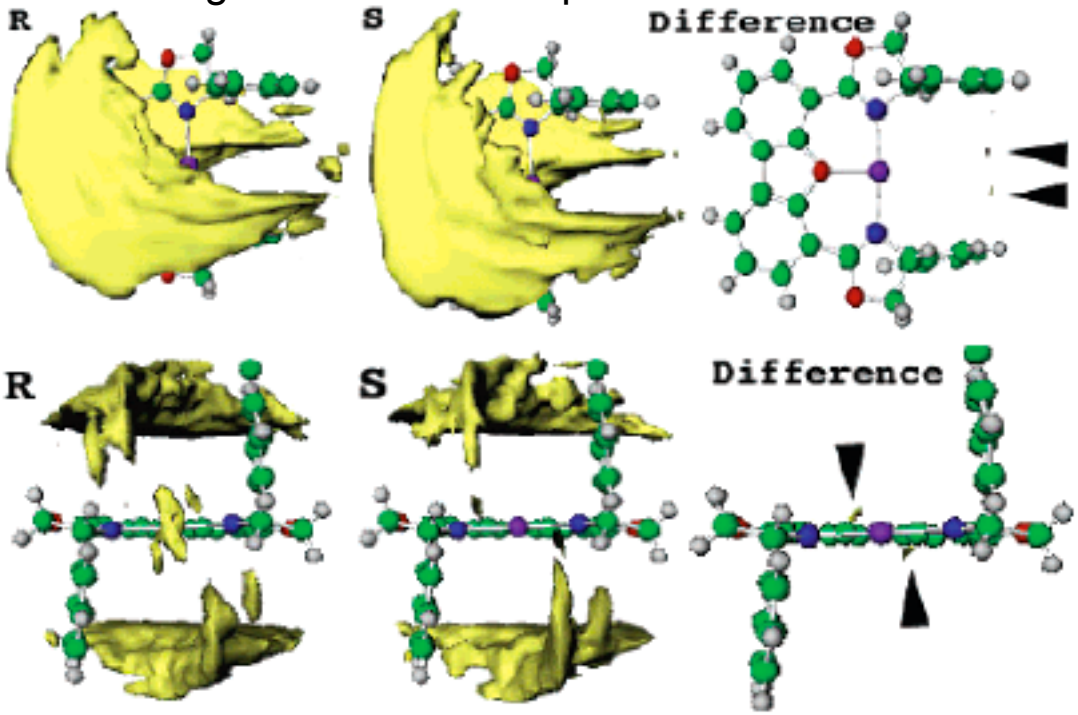


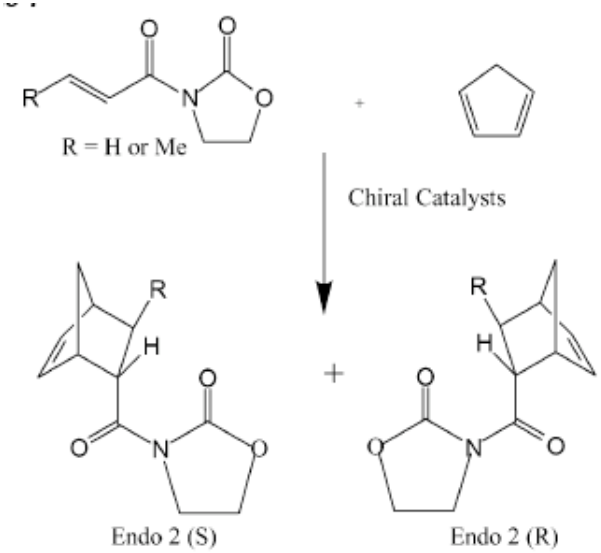
Figure 4. Top view (top panel) and front view looking down the Ni—O bond (bottom panel) of the preferred binding domain of the (R) transition-state probe around the catalyst, the (S) transition-state probe around the catalyst, and the difference map defining the region of greatest stereo-induction.

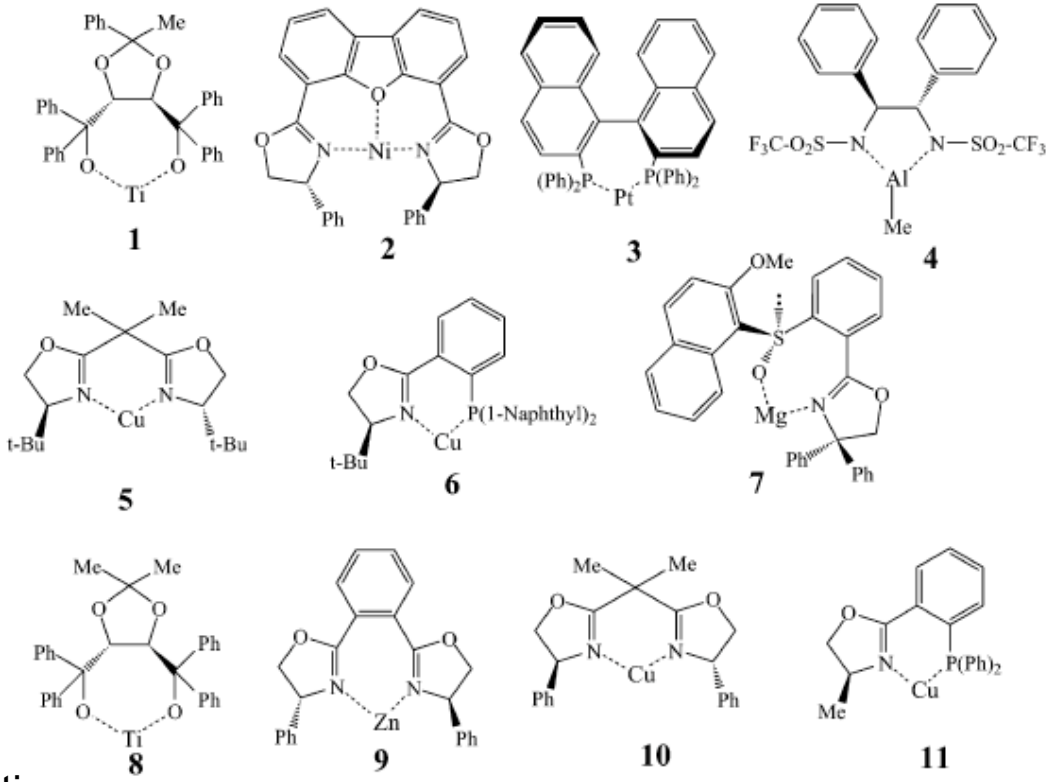
Most enantiodiscriminating region "in front" of the catalyst, near the metal center.

Grid spacing 0.25 angstroms, at each point the TS is rotated 30 deg. about x,y,z axis giving 1728 intermol. calc. Typical run of 17 angstroms with ~275,000 grid points (475 million configs for each probe)

- Lipkowitz, K. *et. al. J. Amer. Chem. Soc.* (2002), **124,** 14255.
- Lipkowitz, K.; Kozlowski, M. *Synlett* (2003), 1547.

Stereocartography Examples:





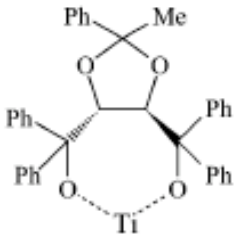
1) Many examples known with reliable ee (also need complete rxn conditions)
2) Rxn should be as simple as possible
3) Catalyst ligands have as few conformations as possible

Catalysts 1-7 give excellent selectivities (above 90% ee)
Catalysts 8-11 give poor selectivites (all below 60% ee)

Is there a spatial congruence for 1-7 but not for 8-11?

- Lipkowitz, K. *et. al. J. Amer. Chem. Soc.* (2002), **124,** 14255.

Stereocartography Examples:



-Crystal structure of dimethyl acetal in CSD. Structure modified and used as a starting point for optimization using PM3.

-Distance between the site of maximum stereoinduction and TS is 1.57 A

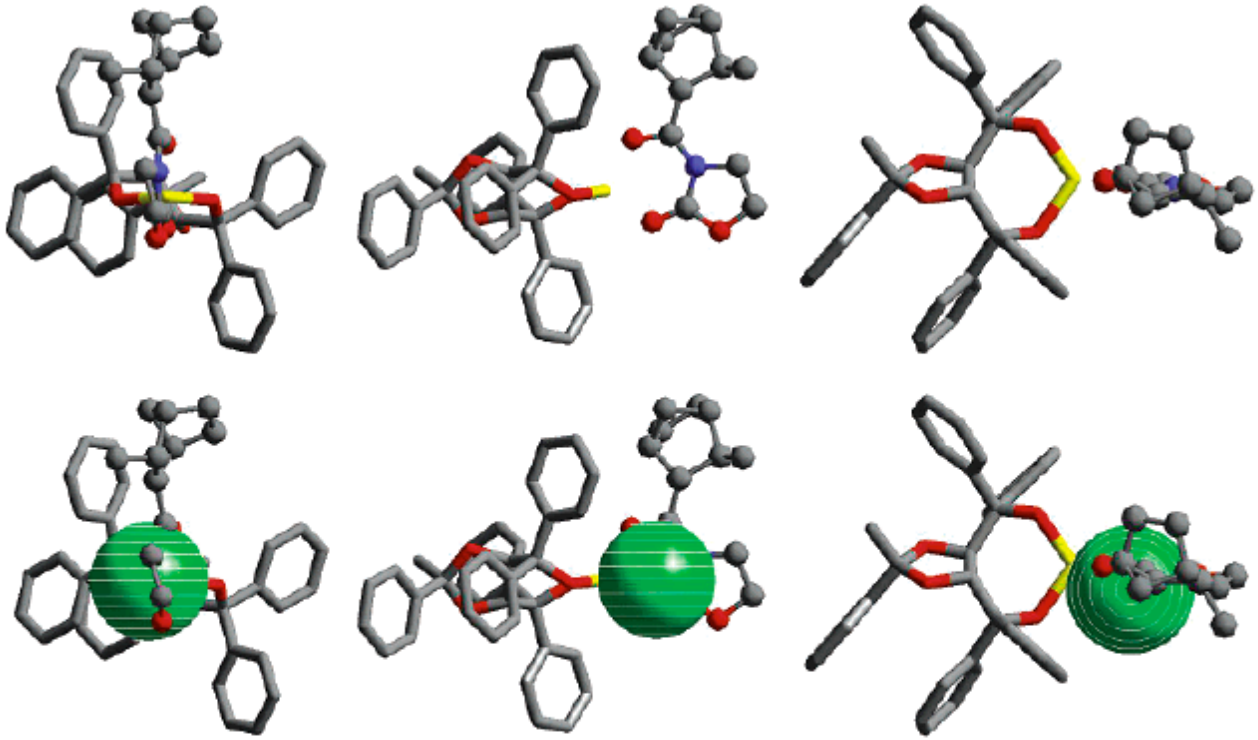
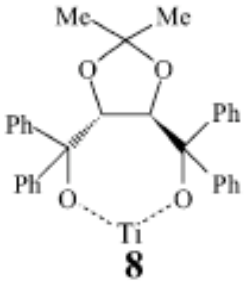


Figure 9. Location of transition-state probe and region of maximum stereoinduction for catalyst **1**. The frames in the columns from left to right show, respectively, front, side, and top views of the docked transition state with the catalyst. In all frames, the catalyst is depicted in a cylinder (tube) format while the transition-state probe is shown in a ball-and-stick format. All hydrogen atoms have been removed for clarity. The top row of frames illustrates only the docked transition-state structure for this catalyst system. The bottom row shows the most enantiodiscriminating region encapsulated within a sphere of 1 Å radius (centered about the most enantiodiscriminating grid point).

- Lipkowitz, K. *et. al. J. Amer. Chem. Soc.* (2002), **124,** 14255.

Stereocartography Examples:



8

-The site of chemistry is “in front” of the Ti-ligand complex whereas the site of maximum stereoinduction is 5.33 A removed from the optimal TS

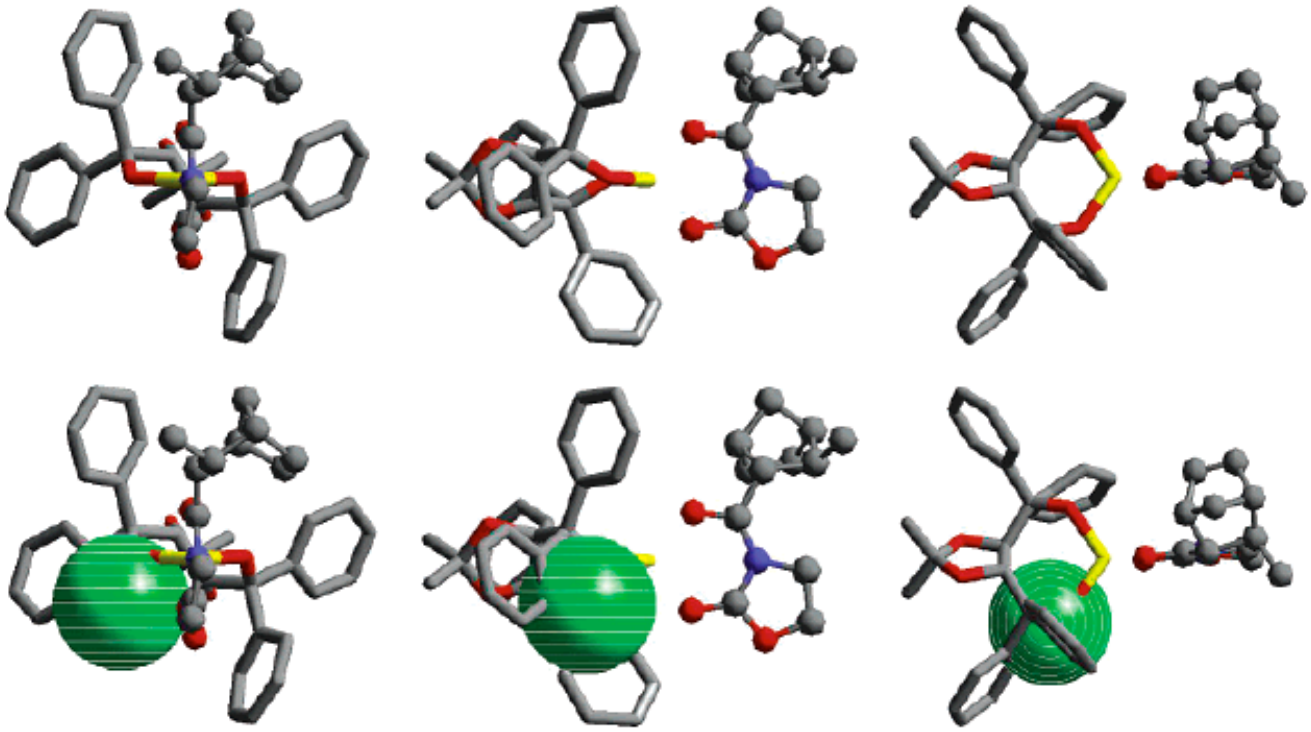


Figure 16. Location of transition-state probe and region of maximum stereoinduction for catalyst **8**. See caption of Figure 9 for details concerning the graphic.

- Lipkowitz, K. *et. al. J. Amer. Chem. Soc.* (2002), **124,** 14255.

Stereocartography

-Of the 18 examples catalysts studies 17 conformed to the hypothesis. They could not rationalize why the single catalyst did not conform.

"**What this means is that one can now predict, in a yes/no way, whether a given catalyst will be effective as asymmetric induction for a particular reaction with 94% probability of being right**."

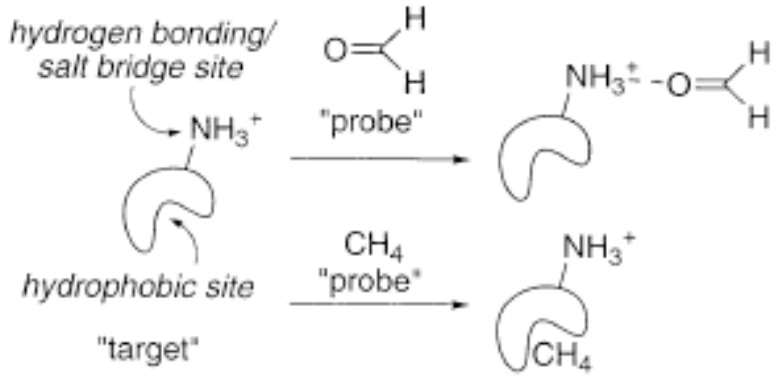
-Caveats: Not all structures were reoptimized from the CSD. Counterions were not involved directly in the calculations (this was done by moving them infinite distances away from the TS)

-Additionally, a linear relationship between enantioselectivity and distance from the most enantiodiscriminating space could not be found

- Lipkowitz, K. *et. al. J. Amer. Chem. Soc.* (2002), **124,** 14255.

Functionality Mapping

A program designed to determine the most energetically favorable positions of functionality in the binding domains of proteins:

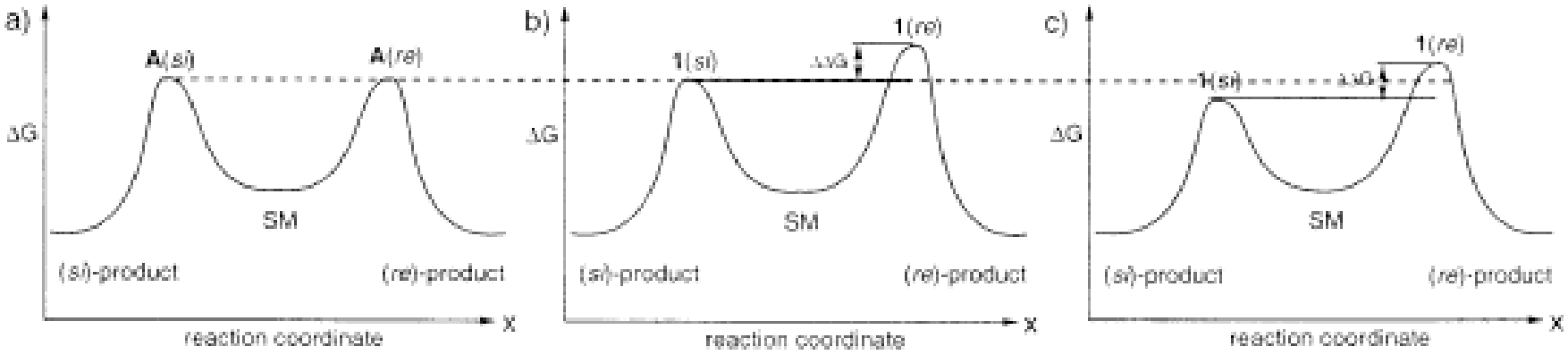


The program operates by:
1) rigidly defining a target molecule
2) Sequential, independent energy minimizations (MM2*) of a series of the same functional group (probes) placed “randomly” around the target structure.
3) Only nonbonding parameters are incorporated into MM2* and thus only van der Waals, and electrostatics are responsible for driving the probes to minimized locations around the target (pi-pi and CH-pi are incorporated / approximated in the van der Waals parameters)
4) The probes that converge (RMSD) are computationally clustered into groups.
5) The clusters are sorted by energy and a representative probe is written to an output file for each cluster

- Kozlowski, M.; Panda, M. *J. Org. Chem. (2003)*, **68**, 2061.

Functionality Mapping to Determine Stereocontrol Elements

What if TS structures are used instead of macromolecules / proteins?



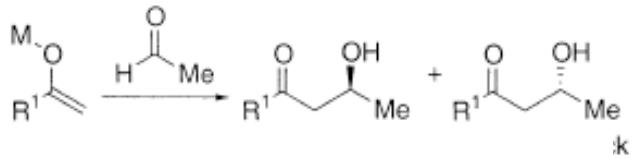
Functionality mapping to the TS of an organic reaction would permit identification of the optimal interactions between defined functional groups and the TS. Knowledge of these interactions could give rise to ligand design!

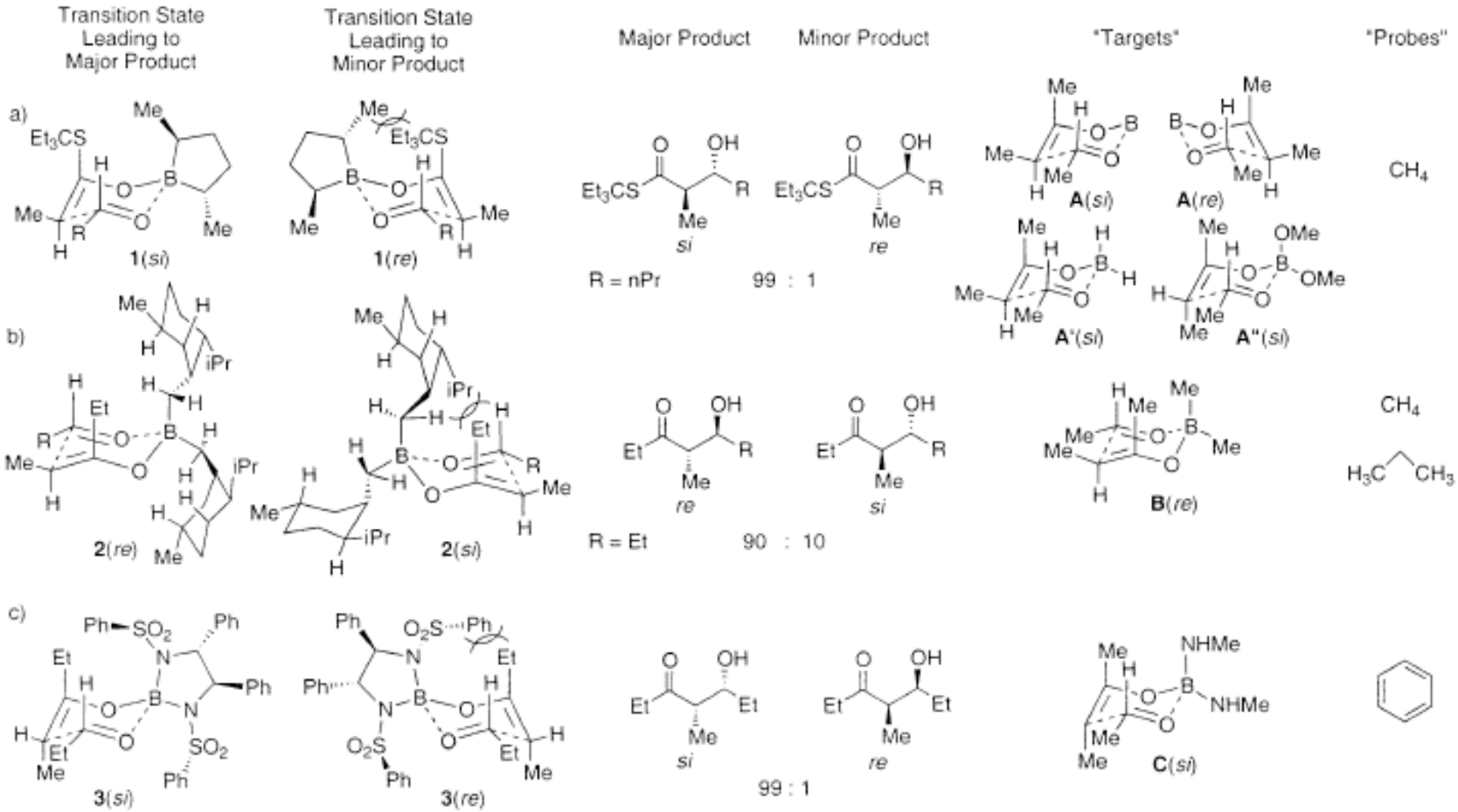
- Kozlowski, M.; Panda, M. *J. Org. Chem. (2003)*, **68**, 2061.

Proof of Principle: Functional Mapping of the Boron Aldol

Can functionality mapping provide a realistic and usable assesment of the optimal positions of a given functional group with respect to a TS structure?

Stereochemical control of the aldol reaction:





- Kozlowski, M.; Panda, M. *J. Org. Chem. (2003)*, **68**, 2061.

Masamune dimethylborolane Functional Mapping

Beginning by removing the chiral ligand from the lowest energy TS. The new structure was then minimized providing the “core structures” **A**

A series of methane molecules were randomly placed around the surface of the chair TS. Theses methanes were independently minimized to the surface of the TS.

TABLE 1. Relationship between the Starting Number of Methane Probes and the Output Clusters Found with the PRCG and TNCG Gradients in the Functionality Mapping of A

entry	TS	gradient[a]	no. of probes	no. of clusters[b]	population of clusters 1st	2nd	3rd	4th	5th
1	A(*si*)	PRCG	200	10	29	43	33	26	13
2[c]	A(*si*)	PRCG	200	10	32	24	34	31	26
3	A(*si*)	PRCG	500	11	81	91	75	73	61
4	A(*si*)	PRCG	1000	12	148	196	189	147	119
5	A(*re*)	PRCG	200	10	33	31	45	30	28
6	A(*re*)	PRCG	500	10	74	89	88	77	72
7	A(*si*)	TNCG	200	9	23	47	37	25	25
8[c]	A(*si*)	TNCG	200	9	22	40	49	25	29
9	A(*si*)	TNCG	500	9	89	82	91	72	62
10	A(*si*)	TNCG	1000	9	168	181	188	151	120
11	A(*re*)	TNCG	200	9	43	25	28	39	29
12	A(*re*)	TNCG	500	9	90	99	89	66	58

[a] Convergence criteria: PRCG 0.001 kJ/(mol·Å); TNCG 0.01 kJ/(mol·Å). [b] A 1.0 Å RMSD (heavy atoms) was used for clustering. [c] Repeat of the run in the prior table entry.

Methanes are nonpolar and tend to only have weak interactions (-5 to -8 kJ/mol) w/ the target.

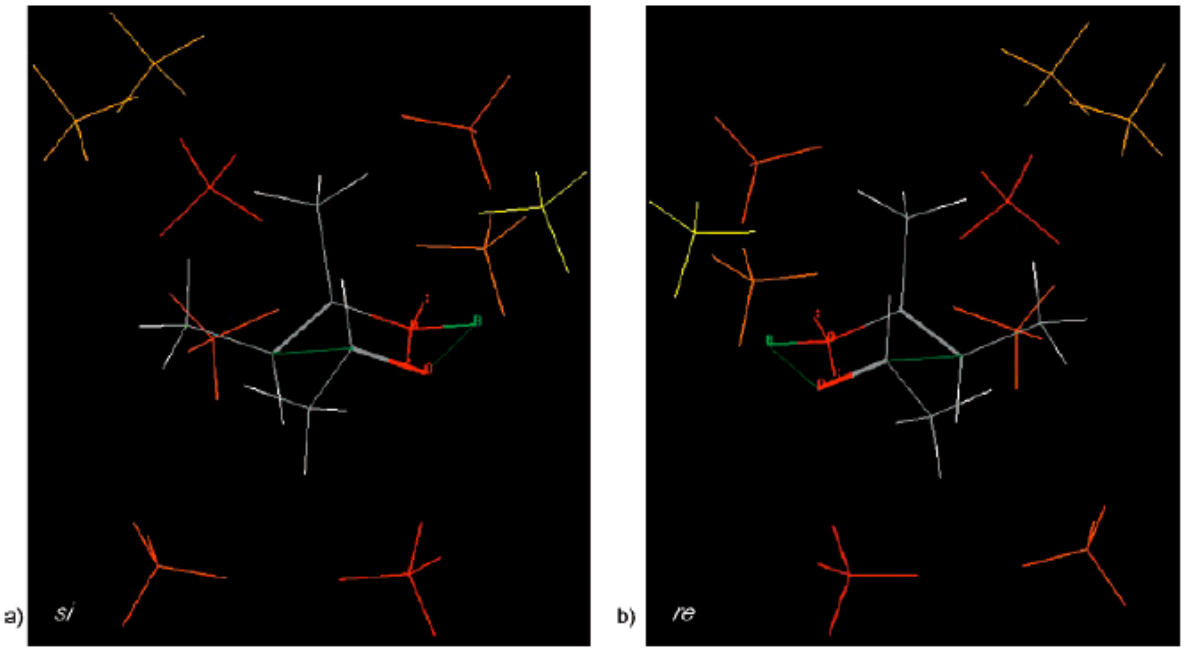
The methane molecules were found to distribute over large portions of the TS, but discrete clusters did form. The clusters favored Impressions in the TS structure.

- Kozlowski, M.; Panda, M. *J. Org. Chem. (2003)*, **68**, 2061.

Masamune dimethylborolane Functional Mapping

entry	target	relevant TS[b]	probe	no. of clusters[c]	populations; energies (kJ/mol) of the top five clusters 1st	2nd	3rd	4th	5th	energy range (kJ/mol)	distance in Å (energy in kJ/mol, cluster no.)[d]
1	A(*si*)	**1**(*si*)	CH_4	9	23; 7.84	47; −7.78	37; −7.26	25; −7.09	25; −7.05	−7.84 to −5.19	0.95 (−7.05, 5) 3.11 (−7.78, 2)
2	A(*re*)	**1**(*re*)	CH_4	9	28; −7.84	31; −7.78	47; −7.26	32; −7.09	25; −7.05	−7.84 to −5.19	1.80 (−7.05, 5) 2.58 (−7.78, 2)
3	A′(*si*)	**1**(*si*)	CH_4	9	28; −8.50	35; −7.92	34; −7.31	23; −7.13	11; −7.13	−8.50 to −5.81	1.66 (−7.13, 4) 2.59 (−8.50, 1)
4	A″(*si*)		CH_4	8	39; −10.88	29; −10.51	38; −9.45	8; −9.21	20; −8.99	−10.88 to −6.97	
5	A″(*re*)		CH_4	8	38; −9.82	41; −9.61	33; −9.63	20; −9.52	29; −9.47	−9.82 to −6.92	

[a] 200 probes with a TNCG gradient and a convergence criterion of 0.01 kJ/(mol·Å) were used in all cases. [b] Corresponds to the transition state incorporating the relevant chiral auxiliary (see Figure 2), which is being compared to the target and its functionality map. [c] A 1.0 Å RMSD (heavy atoms) was used for clustering. [d] Distances between stereodiscriminating groups of the transition state from the third column and their closest clusters. The interaction energy of these clusters and their rank, in terms of favorable interaction energy, are



Most favorable points of interactions were above and below plane of chair (lesser interactions w/ oxygen heteroatoms).

Consistent w/ optimization of van der Waals interactions.

FIGURE 3. Functionality mapping of target **A** with 200 methane probes. The reddest clusters have the largest favorable interactions and the most yellow clusters have the smallest favorable interactions: (a) target **A**(*si*), 9 clusters were found; (b) target **A**(*re*), 9 clusters were found.

- Kozlowski, M.; Panda, M. *J. Org. Chem. (2003)*, **68**, 2061.

Masamune dimethylborolane Functional Mapping

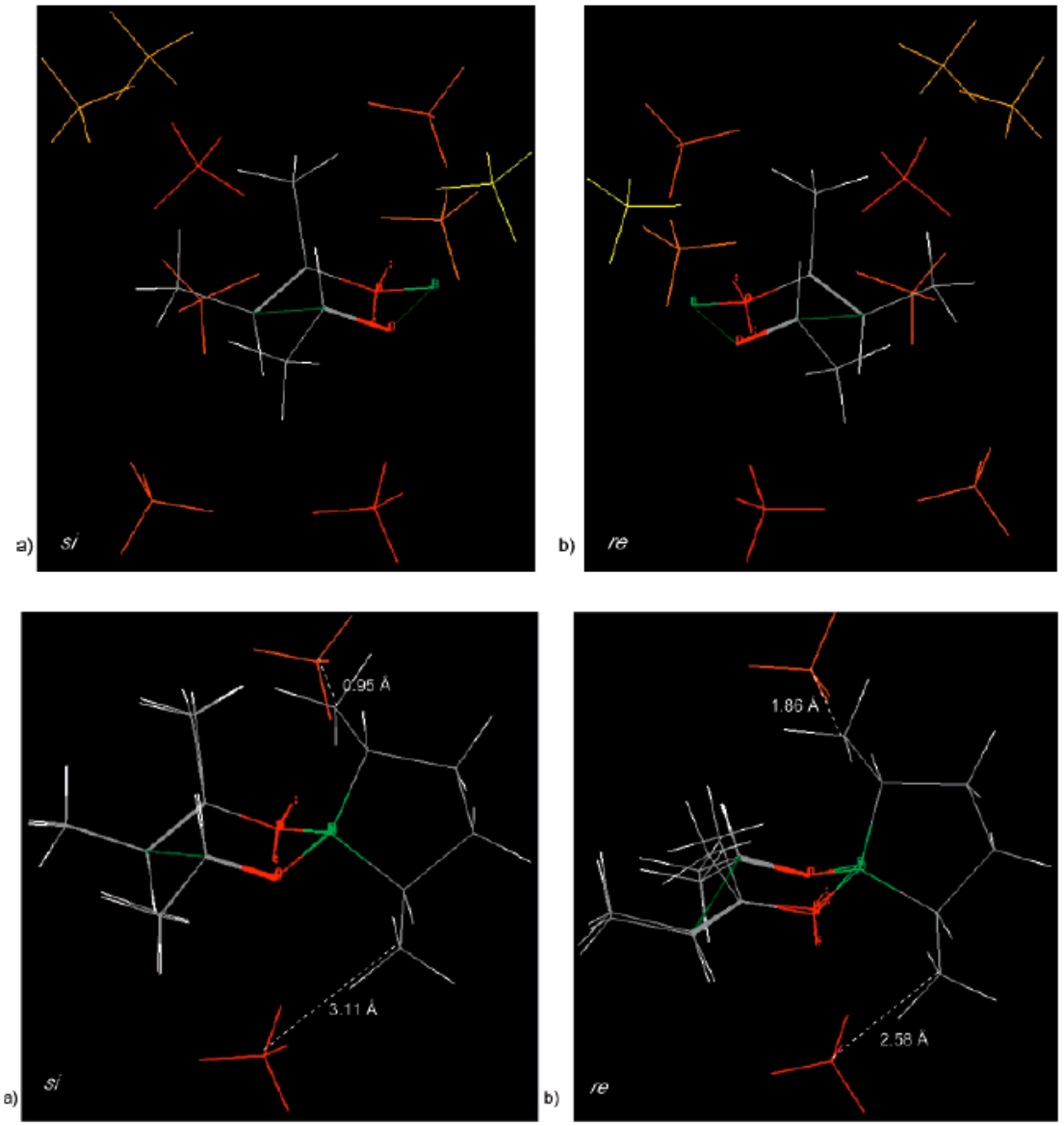
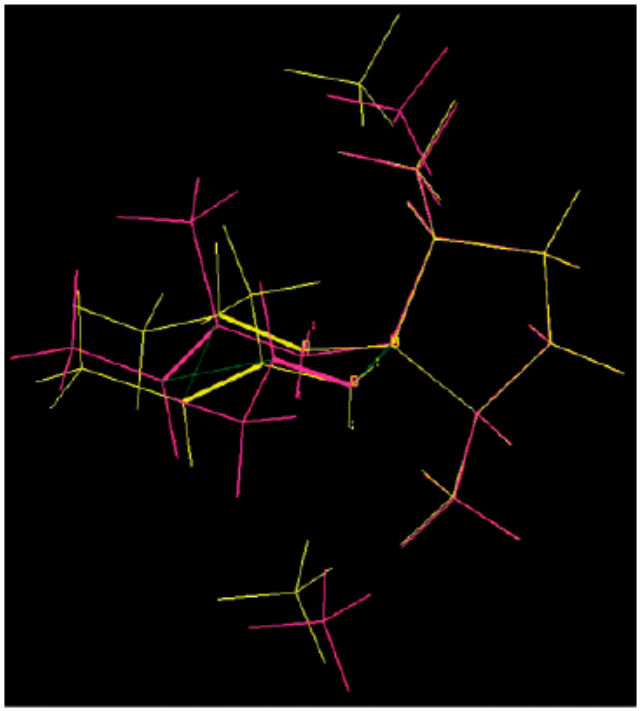


FIGURE 4. Functionality mapping of target **A** with 200 methane probes. The reddest clusters have the largest favorable interactions and the most yellow clusters have the smallest favorable interactions. Only the probes closest to the stereodiscriminating groups of the chiral ligand are illustrated. (a) Target **A**(*si*) is overlaid with **1**(*si*). (b) Target **A**(*re*) is overlaid with **1**(*re*).

- Kozlowski, M.; Panda, M. *J. Org. Chem. (2003)*, **68**, 2061.

Masamune dimethylborolane Functional Mapping



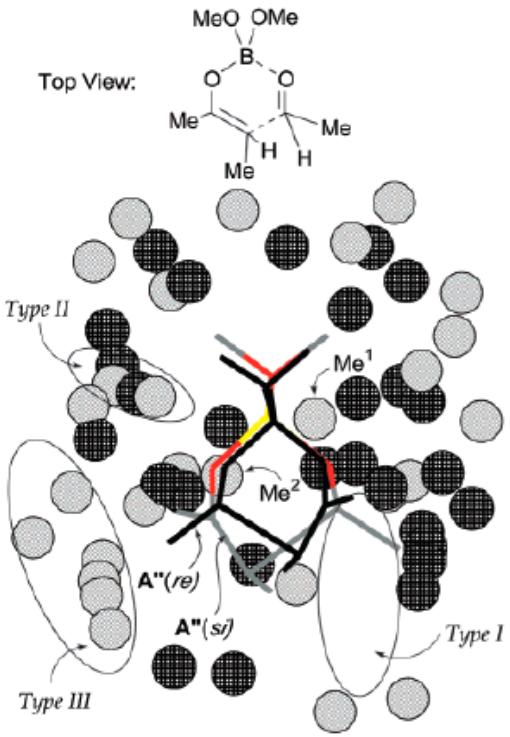
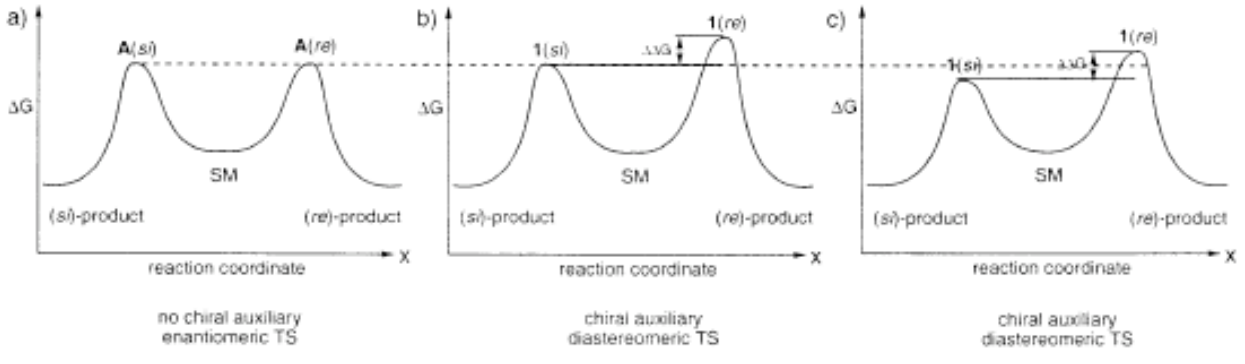
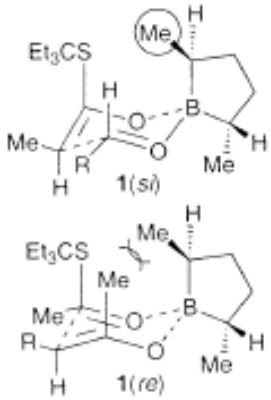


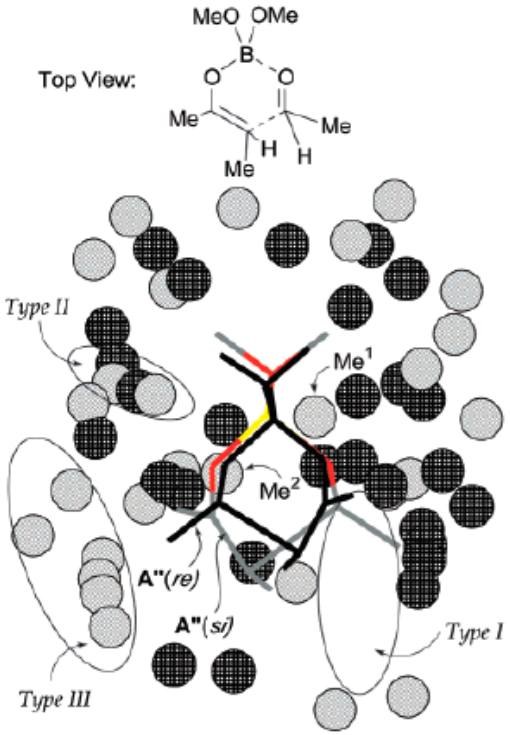
FIGURE 5. Overlay of **1**(*si*) (purple) and **1**(*re*) (yellow). The methane probes from the corresponding functionality maps of **A**(*si*) and **A**(*re*) which are closest to the methyl substituents of the chiral ligands are also shown.



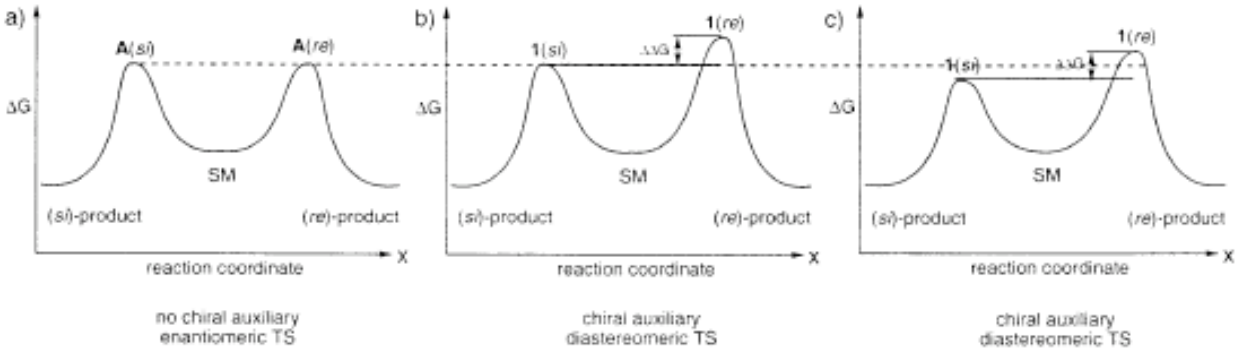


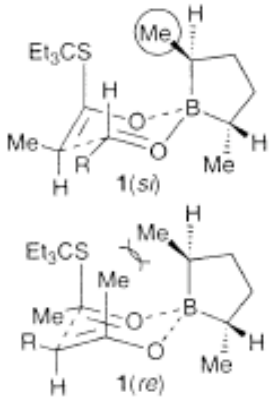
Masamune dimethylborolane Functional Mapping

Type 1: neither enantiomorphic TS undergoes favorable interactions (neg. energy)
-absence of interactions (zero energy)
-disfavorable inter. (positive energy)
Type 2: Both TS favor probes at these positions
Type 3: Only one TS prefers probes at this location. There is no favorable interaction with enantiomeric TS

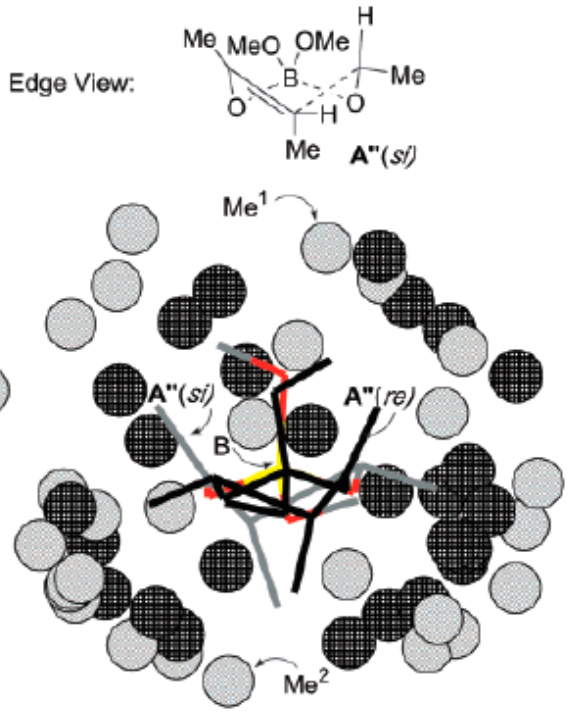


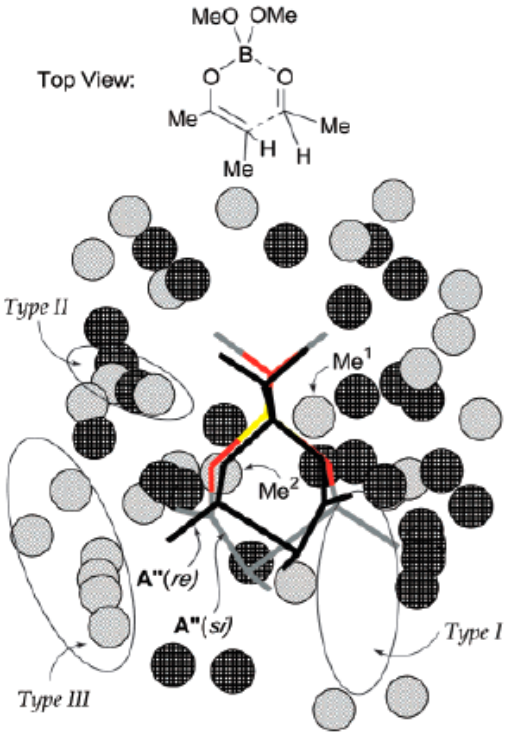
Carbons of methane shown at 15% van der Waals radii - space filling



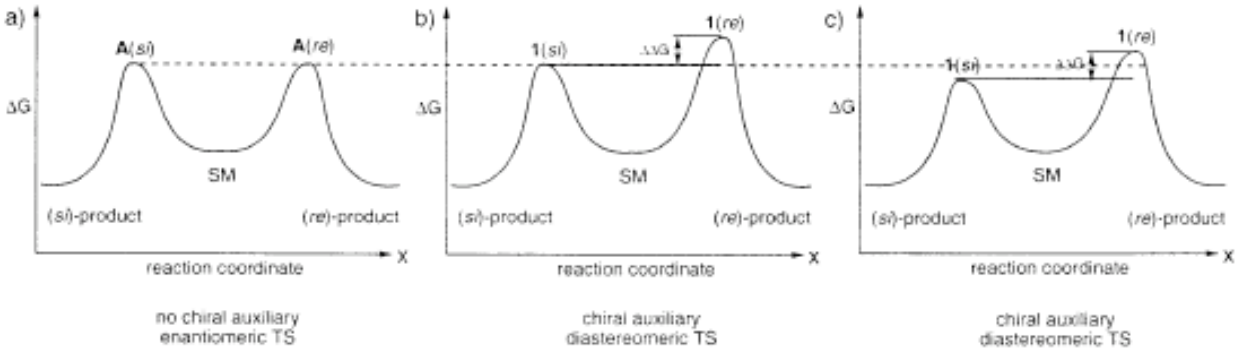


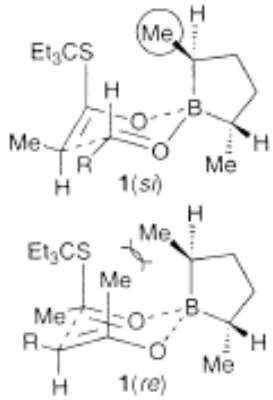
Masamune dimethylborolane Functional Mapping



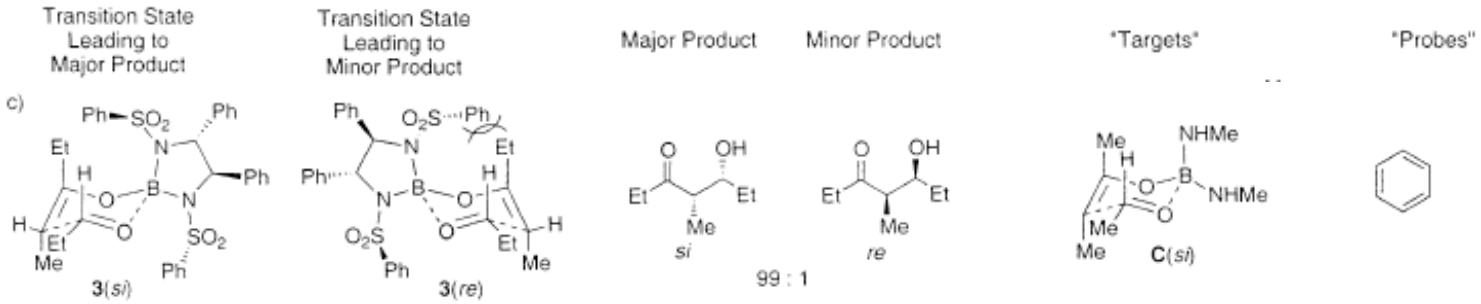


Carbons of methane shown at 15% van der Waals radii - space filling





Example #2: Corey (Stein) Functional Mapping



entry	target	relevant TS[b]	probe	no. of clusters[c]	populations; energies (kJ/mol) of the top five clusters: 1st	2nd	3rd	4th	5th	energy range (kJ/mol)	distance in Å (energy in kJ/mol, cluster no.)[d]
8	C(*si*)	3(*si*)	PhH	74[e]	9; −21.61	6; −21.61	7; −21.61	5; −21.61	5; −21.61	−21.61 to −11.86	0.90[f] (−19.29, 21) 1.33[f] (−21.61, 1)

[a] 200 probes with a TNCG gradient and a convergence criterion of 0.01 kJ/(mol·Å) were used in all cases. [b] Corresponds to the transition state incorporating the relevant chiral auxiliary (see Figure 2), which is being compared to the target and its functionality map. [c] A 1.0 Å RMSD (heavy atoms) was used for clustering. [d] Distances between stereodiscriminating groups of the transition state from the third column and their closest clusters. The interaction energy of these clusters and their rank, in terms of favorable interaction energy, are given in parentheses. [e] The high symmetry of benzene leads to mapping of up to 6 identical clusters having the same interaction energy. [f] RMSD values comparing the 6 carbons of the benzene probe with those of the phenyl group.

Interactions much stronger than that of methane (-4 to -11 kJ/mol vs. -12 to -22 kJ/mol)

- Kozlowski, M.; Panda, M. *J. Org. Chem. (2003)*, **68**, 2061.

Corey (Stein) Functional Mapping

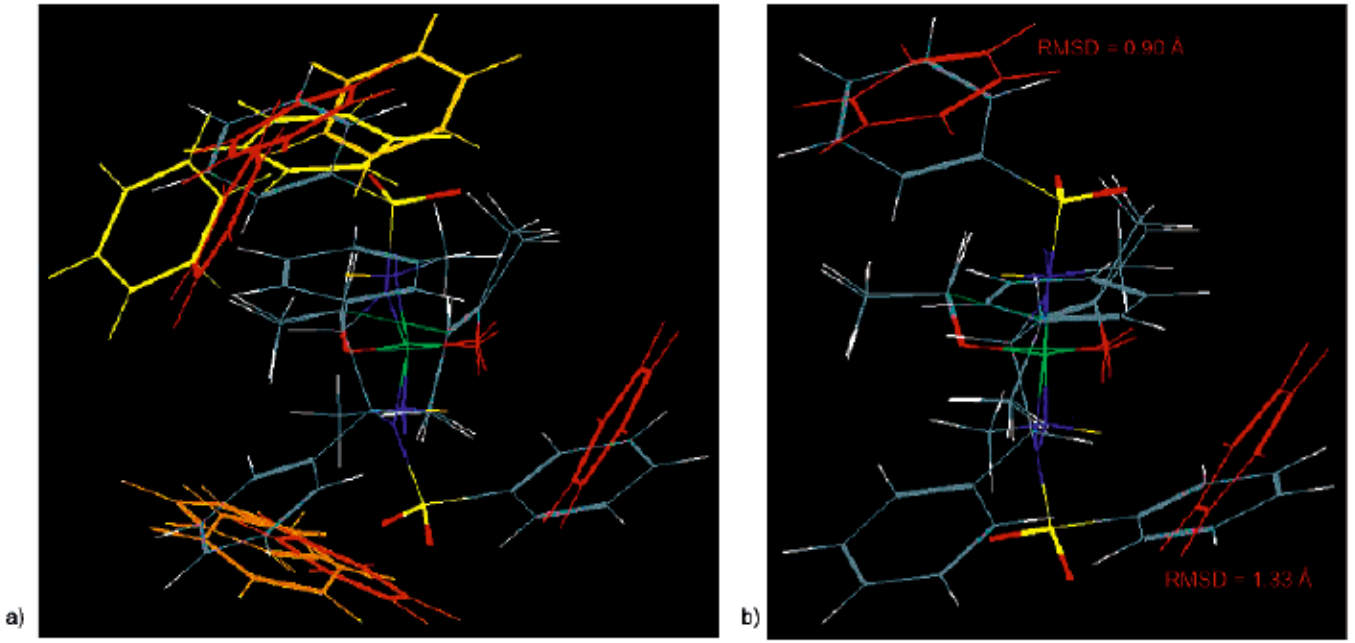


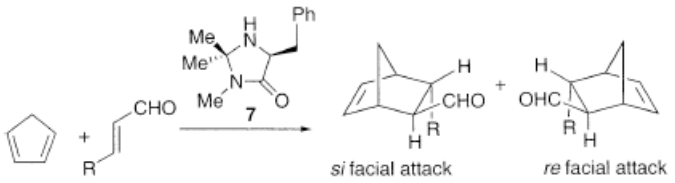
FIGURE 10. Functionality mapping of target **C**(*si*) with benzene probes. Target **C**(*si*) is overlaid with **3**(*si*). (a) The first 40 clusters out of the 74 found are shown. (b) The clusters closest to the stereodiscriminating phenyl groups, and the RMSD between them, are shown.

-Locations of two dense clusters of benzene probes with the phenyl rings of the arylsulfonyl groups of the stein reagent. Additionally, the orientation of the phenyl rings is similar to the stein reagent.

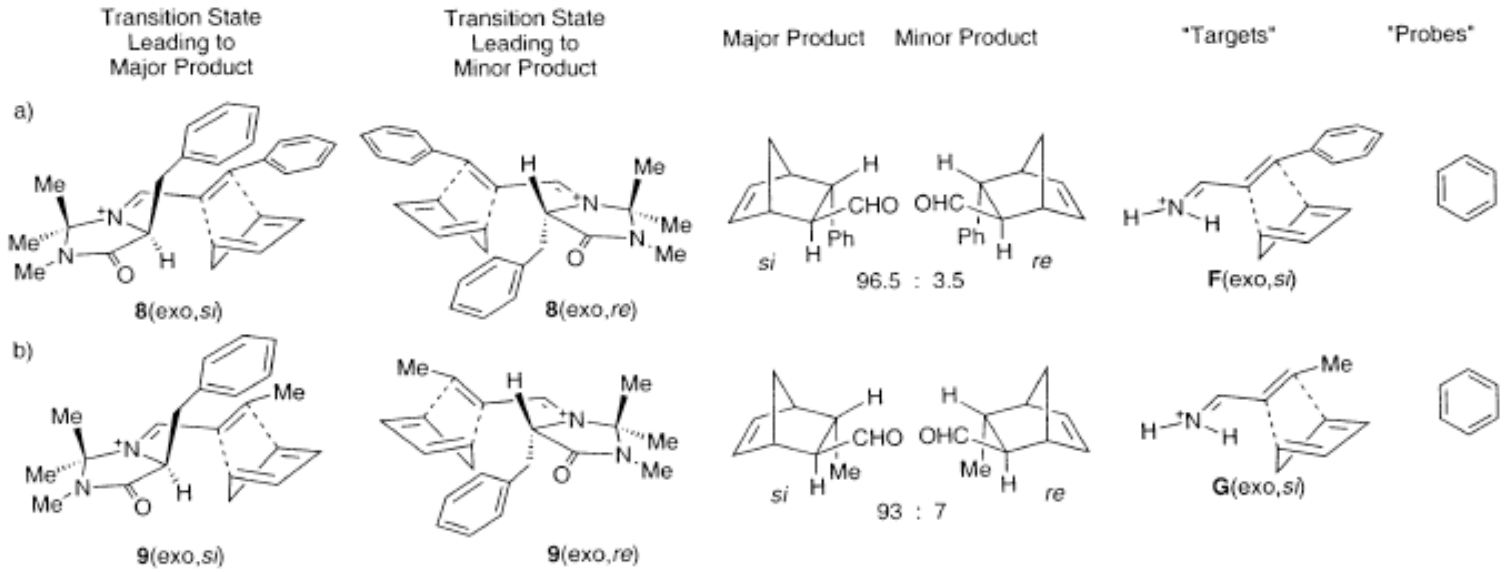
-Mapping of the benzenes on both sides of the TS w/ little change in the interaction energy suggests the importance of both phenyl substituents in the C2 symmetric ligand.

- Kozlowski, M.; Panda, M. *J. Org. Chem. (2003)*, **68**, 2061.

Example #3: MacMillan's Iminium Diels-Alder Functional Mapping



Based on ground state MM2 force field calc. MacMillan suggested that the benzyl group shields the *re* face of the dienophile.



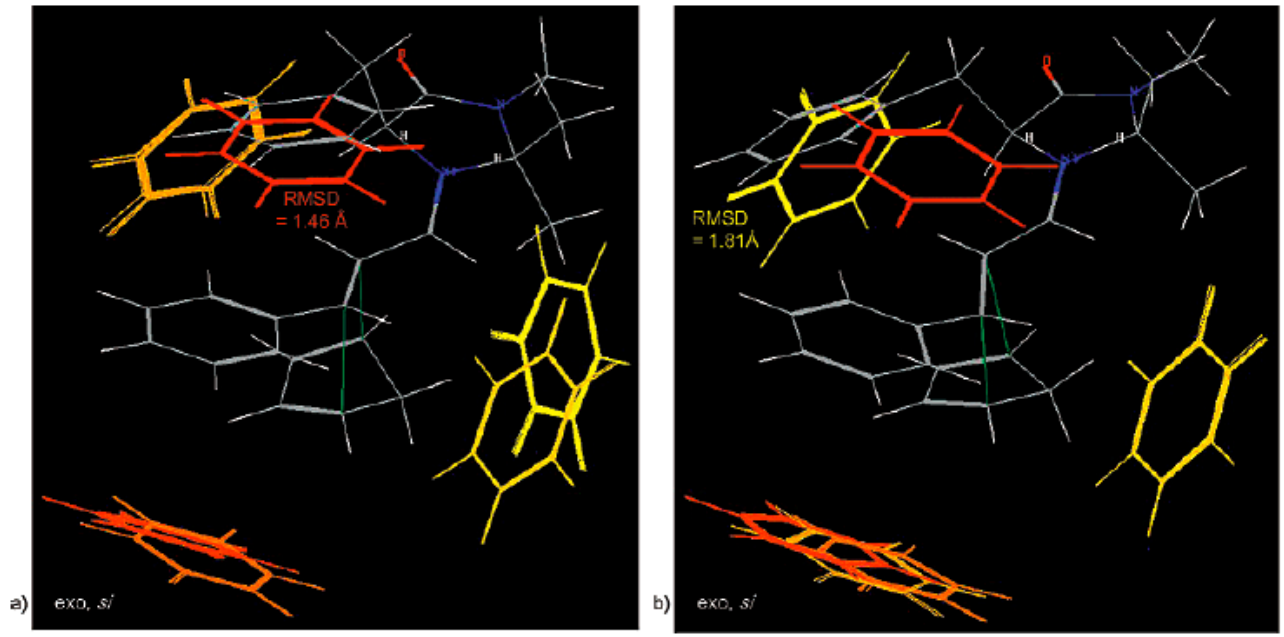
First step involved a Monte-Carlo TS search (MM2 force field). Boltzmann distribution of these TS At 298 K gives 87% ee (exo S), which is in agreement with experimental 93% ee. Exo/endo ratio Calc at 1.7 : 1 was also in agreement exp = 1.3 : 1.

- Kozlowski, M.; Panda, M. *J. Org. Chem. (2003)*, **68**, 2061.
- MacMillan, D. *et al J. Am. Chem. Soc.* (2000), **122**, 4243.

Iminium Diels-Alder Functional Mapping

entry	target	relevant TS[b]	probe	no. of clusters[c]	populations; energies (kJ/mol) of the top five clusters					energy range (kJ/mol)	RMSD in Å (energy in kJ/mol, cluster #)[d]
					1st	2nd	3rd	4th	5th		
1	**F**(exo,*si*) MM2*	**8**(exo,*si*)	PhH	71	7; −22.27	9; −22.27	7; −22.27	10; −22.27	11; −22.27	−22.27 to −12.96	1.46 (−22.27, 1)
2	**F**(exo,*re*) MM2*	**8**(exo,*re*)	PhH	67	20; −22.03	13; −22.03	3; −22.03	7; −22.03	5; −22.03	−22.03 to −13.29	6.95 (−22.03, 1)
3	**F**(endo,*si*) MM2*	**8**(endo,*si*)	PhH	74	10; −21.94	7; −21.94	8; −21.94	1; −21.94	12; −21.94	−21.94 to −12.51	2.15 (−21.94, 1)
4	**F**(exo,*si*) HF/3-21G	**8**(exo,*si*)	PhH	64	13; −22.72	8; −22.72	6; −22.72	4; −22.72	11; −22.72	−22.72 to −13.83	1.81 (−16.55, 33)

[a] 200 probes with a TNCG gradient and a convergence criterion of 0.01 kJ/(mol·Å) were used in all cases. [b] Corresponds to the transition state incorporating the relevant chiral auxiliary (see Figure 16) which is being compared to the target and its functionality map. [c] The high symmetry of benzene leads to mapping of up to 6 identical clusters having the same interaction energy. A 1.0 Å RMSD (heavy atoms) was used for clustering. [d] RMSD values comparing the 6 carbons of the closest benzene clusters with those of the phenyl stereodiscriminating group of the transition state from the third column. The interaction energy of these clusters and their rank, in terms of favorable interaction energy, are given in parentheses.



Benzene probes Concentrated to 3 general areas. Most favorable over the surface of the target (-22 kJ/mol)

FIGURE 17. Functionality mapping of target **F**(exo,*si*) with benzene probes. Target **F**(exo,*si*) is overlaid with **8**(exo, *si*). (a) With the MM2* target, FUNMAP found 71 clusters (the first 40 are shown). (b) With the HF/3-21G target, FUNMAP found 64 clusters (the first 40 are shown).

Iminium Diels-Alder Functional Mapping

Observed (exo, si):(exo, re): 96.5 : 3.5 (exo : endo) 1.3 : 1

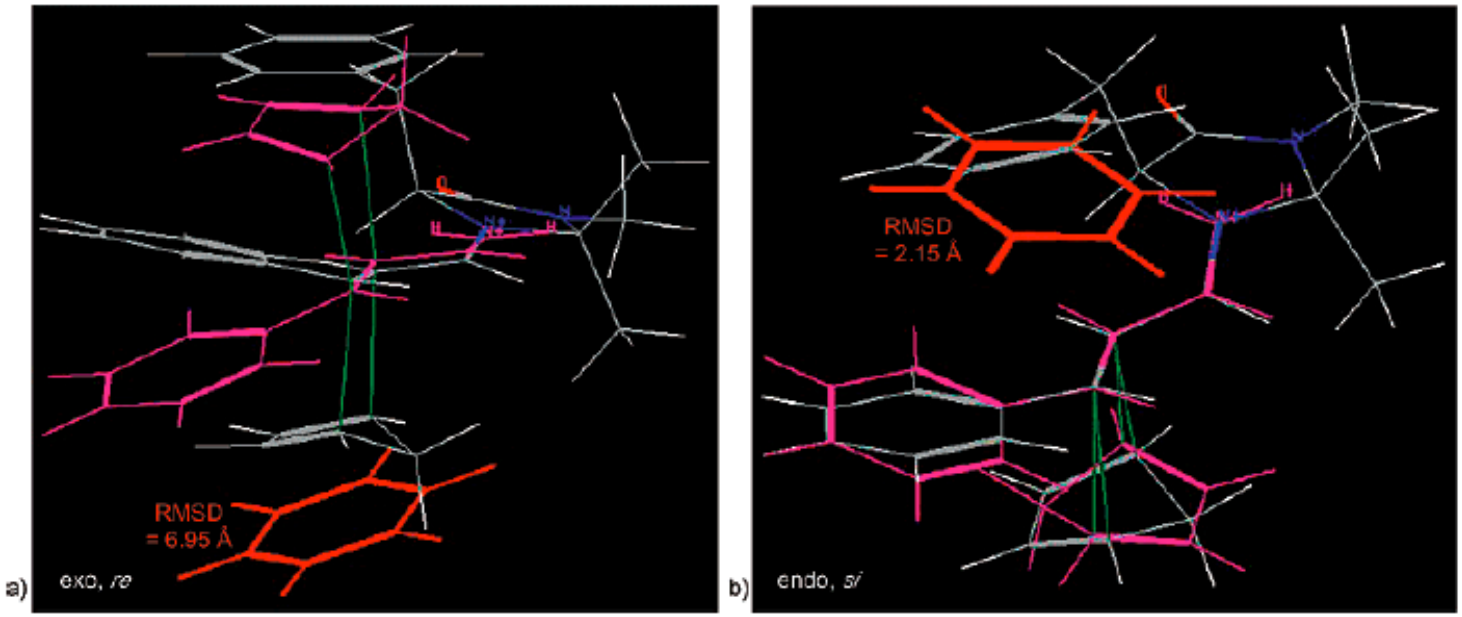


FIGURE 18. (a) Functionality mapping of MM2* target **F**(exo,*re*), which is shown in purple, with benzene probes (the 12 overlapping clusters with the best interaction energies are shown in red). The α,β-unsaturated iminium of **F**(exo,*re*) is overlaid with that of **8**(exo,*si*), which is shown in half-bond colors. (b) Functionality mapping of MM2* target **F**(endo,*si*), which is shown in purple, with benzene probes (the 11 overlapping clusters with the best interaction energies are shown in red). The α,β-unsaturated iminium of **F**(endo,*si*) is overlaid with that of **8**(exo,*si*), which is shown in half-bond colors.

-Functionality mapping tools are a way to better understand the governing non-bonding interactions of ligand functional groups in TS leading to chiral products.
-Orientation as well as positional information about potential groups is obtained.
-It is possible to determine if a chiral ligand imparts the observed selectivity through stabilization of a pathway, destabilization of a pathway or both by the quantification of the positions and energies afforded by non-bonding interactions.

Can this be used for De-Novo design?

- Kozlowski, M.; Panda, M. *J. Org. Chem. (2003)*, **68**, 2061.

Database Search Methods to Identify Chiral Ligands

Conventional approach towards ligand design / identification relies on intuition or random screening approachs.

Is there a more systematic method that could lead to ligand design?

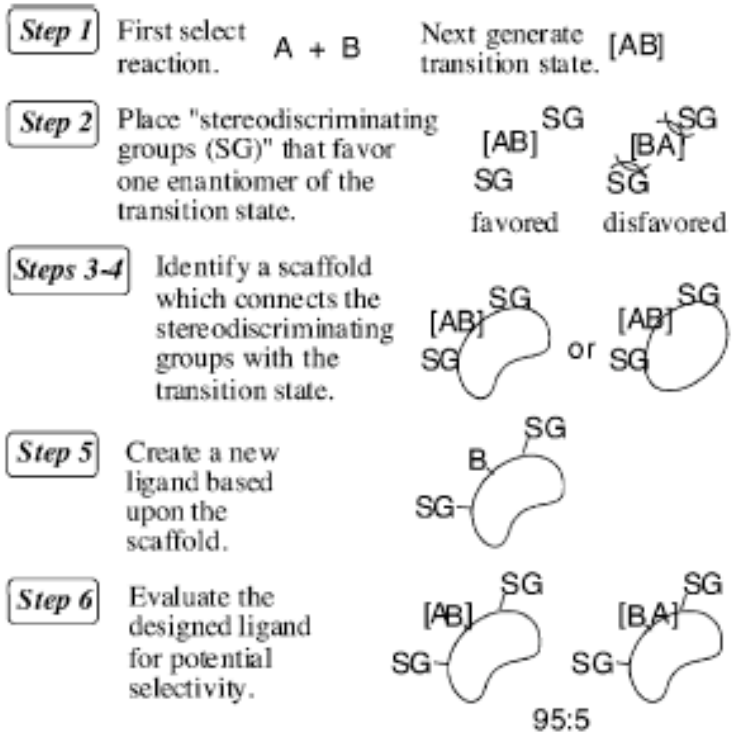


Fig. 1. Generalized schematic of the ligand design protocol.

Can be done computationally

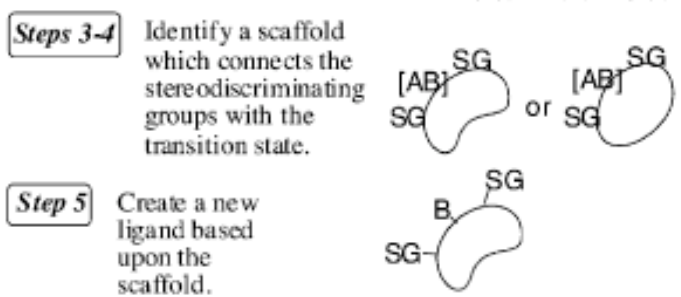
Done either emperically / combinatorially, computationally or by models (or serendipity)

Not obvious how this could happen…

Done in a flask / computationally

Kozlowski, M.; Panda, M. *J. Mol. Graph.* (2002), **20**, 399.

Database Search Methods to Identify Chiral Ligands



For the most part, the identification of a scaffold is generally based upon previously successful chiral ligands (e.g. BINOL, pybox, salen, cinchona ect….)

Therefore, a new method to identify ligands based on their ability to act as stereodiscriminating scaffolds would be immensely important.

Computational method to reduce 3-dimensional topographical information to a series of vectors:

CAVEAT: program designed to facilitate the structure-based design of enzyme inhibitors. Devised By Paul Bartlett of UC Berkeley as an interactive “idea generator”

-Realized that the unifying problem encountered with structure-based searches was their reliance upon the *location of atoms*. The focus should be on the *orientation of the bonds.*

Therefore, the CAVEAT program treats molecules as a set of defined vector relationships. These Vectors are then placed into a search engine (CSD, TRIAD (tricyclic structures), ILIAD (acyclic Structures), CAS-3D (chemical abstracts services three-dimensional database)

Kozlowski, M.; Panda, M. *J. Mol. Graph.* (2002), **20**, 399.
Lauri, G.; Bartlett, P. *J. Comp. Mol. Des.* (1994), **8**, 51.

CAVEAT

In a typical CAVEAT search a bond is treated as a vector; one of the atoms in the bond, the Base, defines the location of the vector and the tip atom defines the orientation.

The relative orientation of several bonds is defined by the combination of several vectors (taken in pairwise combination). These vector pairs are defined by: 1) distance between the two atoms, 2) the dihedral angle between the vectors and 3) the two exterior angles for each vector.

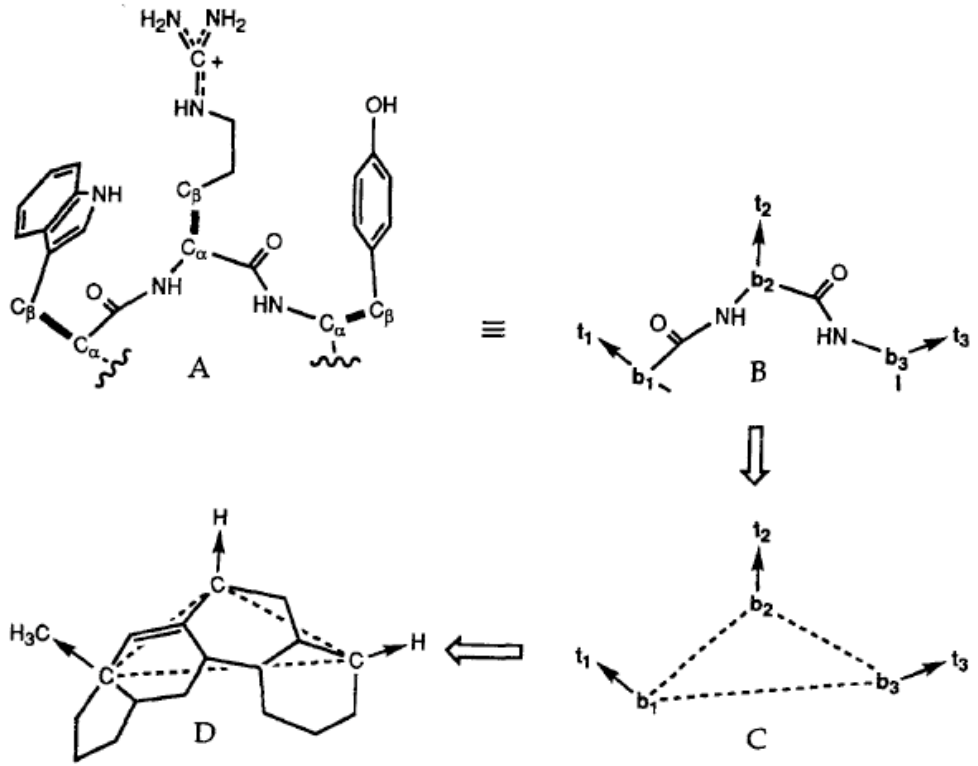
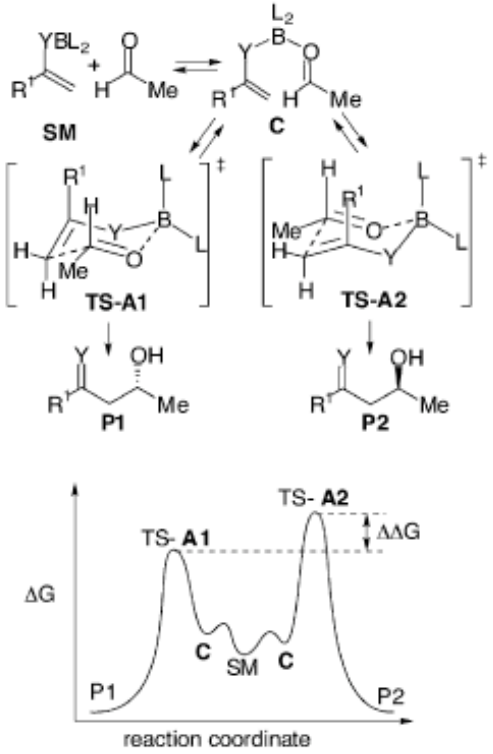


Fig. 1. To devise a peptide mimic, the C_α-C_β bonds of the peptide (A) are represented as vectors (B); the relationship between these vectors is determined by the three vector-pair combinations (C), which must also be present in a hit structure (D) if it is to serve as a template for the design.

- Lauri, G.; Bartlett, P. *J. Comp. Mol. Des.* (1994), **8**, 51.

Database Mining Example #1: Boron Aldol

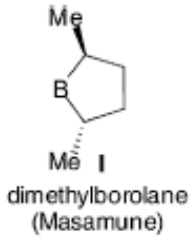
Stereochemical analysis:

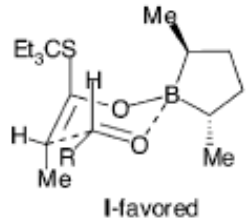


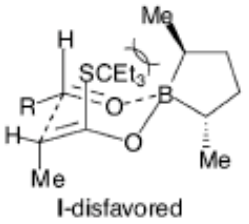
Computational methods have thus far been restricted to the *prediction* of stereoselectivity for given substrates and chiral ligands.

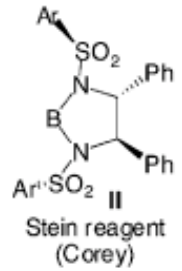
The use of computational methods for the *generation* of chiral ligands has not been documented.

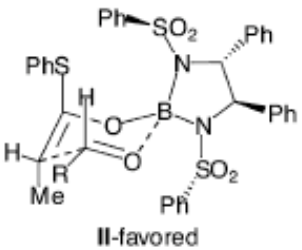
Known highly enantioselective boron auxillaries:

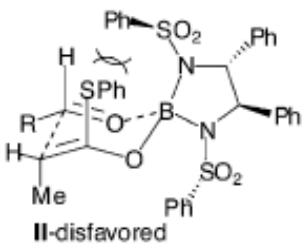










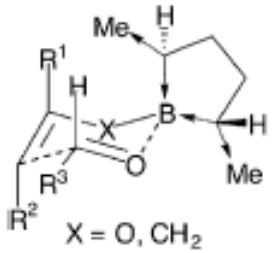


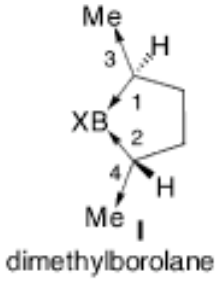
Can separate the “stereodiscriminating” groups from the rest of the scaffold (which rigidifies the ligand)

- Kozlowski, M.; Panda, M. *J. Mol. Graph.* (2002), **20**, 399.

Masamune Dimethylborolane Database mining

Vectors for the database were taken from minimized TS structures:



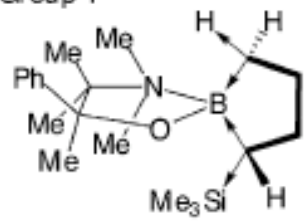


Vectors are the C-B bonds in the 5-membered ring and the Stereodiscriminating C-CH3 bonds.

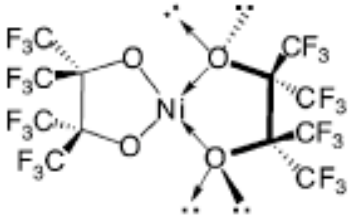
Searching CSD: with "linked" backbone

With 0.5 angstrom vector pair distance tolerance, and 0.4 rad angle tolerance 17,035 structures were retrieved. These hits were divided into 444 clusters and 60 Main groups (384 subgroups)

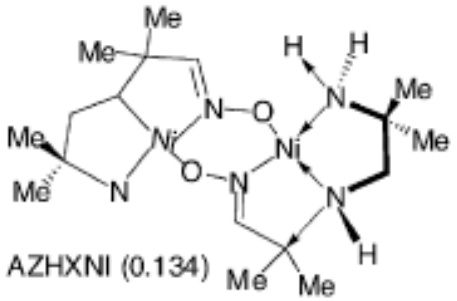
Group 1



VAMNOM (0.132)

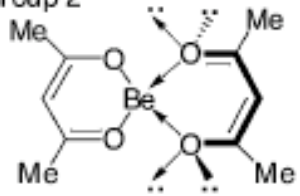


BIBSAG (0.139)

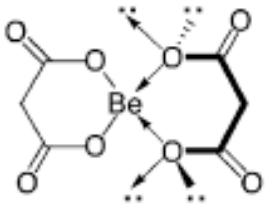


AZHXNI (0.134)

Group 2

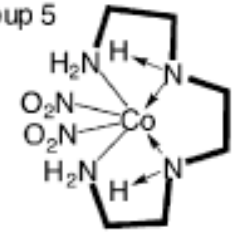


ACACBE02 (0.186)



MALBEK (0.193)

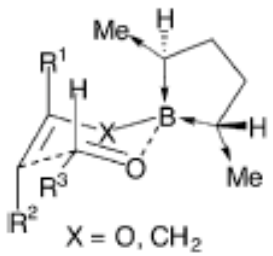
Group 5

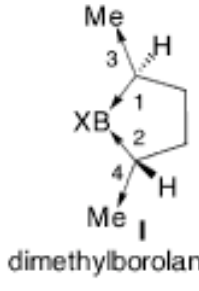


CUNKEB (0.191)

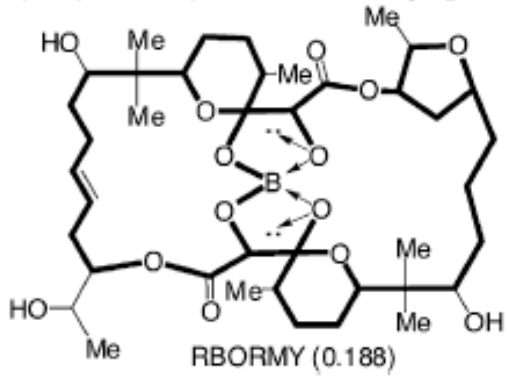
- Kozlowski, M.; Panda, M. *J. Mol. Graph.* (2002), **20**, 399.

Masamune Dimethylborolane Database mining

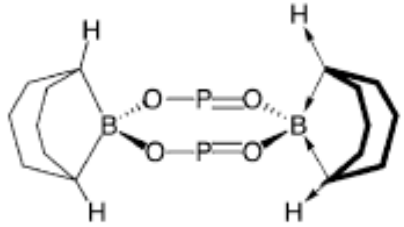




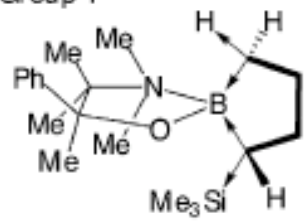
Group 3 (and Group 4 when rotated by C_2-axis)

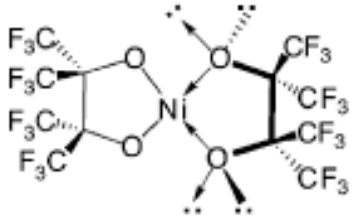


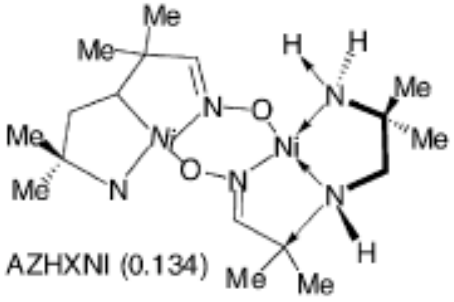
Group 60



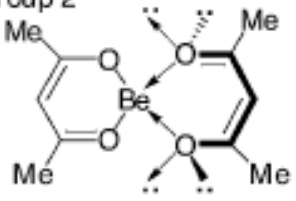
Group 1

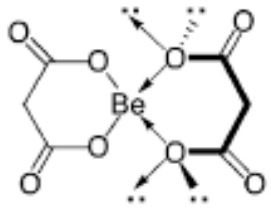




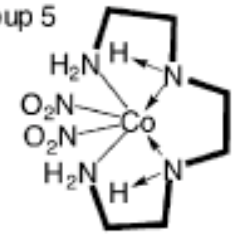


Group 2



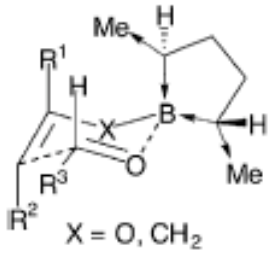


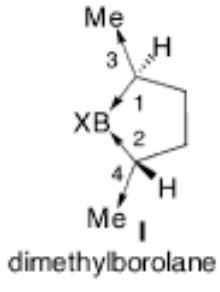
Group 5

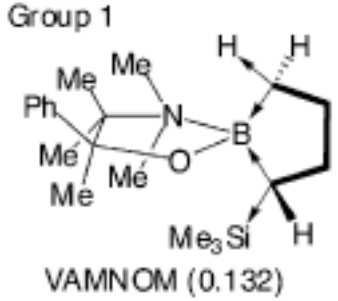


- Kozlowski, M.; Panda, M. *J. Mol. Graph.* (2002), **20**, 399.

Masamune Dimethylborolane Database mining

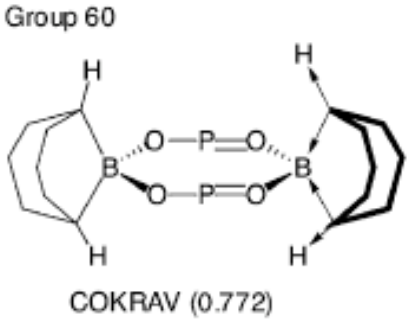






Has been shown in allylborane Reactions to give higher Selectivities than dimethyl Borolane variant!

"Best hit" reported by Masamune. The identification of this closely related compound validates (1) the use of vectors from calculated structures and (2) the ability of the database to retrieve usable structures



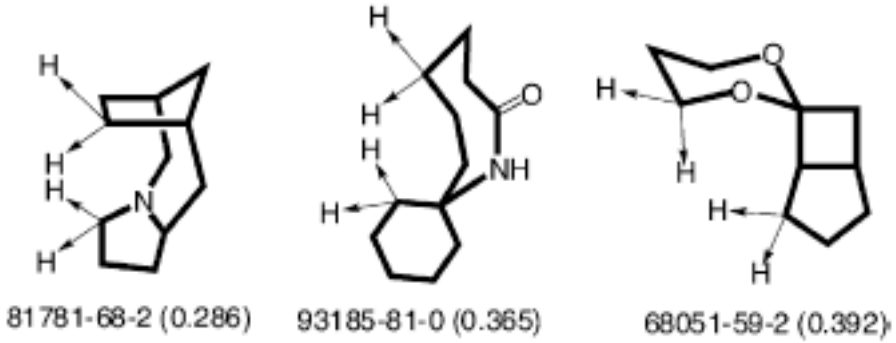
Akin to the 9-borabicyclodecane chiral auxillaries developed by Soderquist!

- Kozlowski, M.; Panda, M. *J. Mol. Graph.* (2002), **20**, 399.
- Short, R.; Masamune, T. *J. Am. Chem. Soc.* (1989), **111**, 1892.
- Soderquist, J. *et. Al.* Organoboranes for Synthesis, ACS symp ser. (2001), **783**, 176.

Masamune Dimethylborolane Database Mining

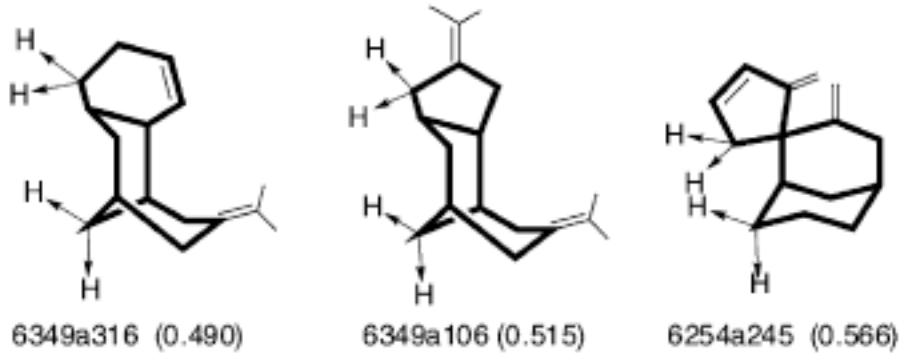
Searching the "unlinked" CSD and TRIAD

Unlinked CSD:



5406 hits (15 main groups, 280 Subgroups)

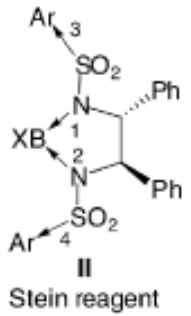
Unlinked TRIAD:



13, 806 hits (98 main groups, 139 Subgroups)

- Kozlowski, M.; Panda, M. *J. Mol. Graph.* (2002), **20**, 399.

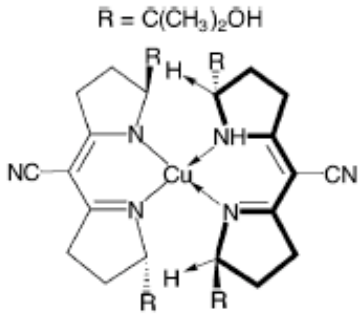
Example #2: Corey Stein Database mining



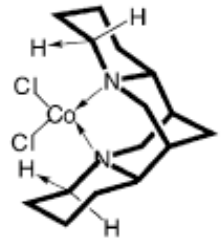
II

Stein reagent

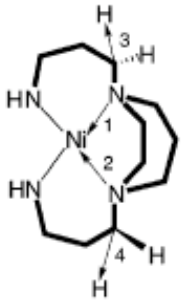
Linked CSD: 2391 hits



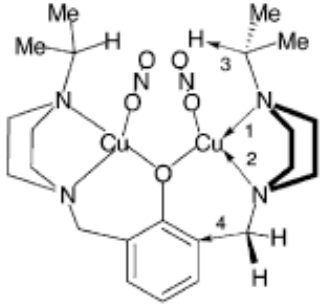
KAFSEP (0.313)



ISPCCO (0.325)

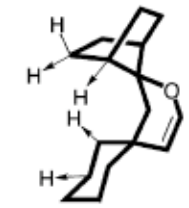


SELSEH (0.345)

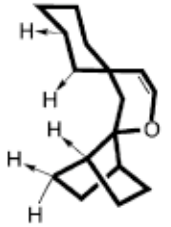


SAFHEM (0.356)

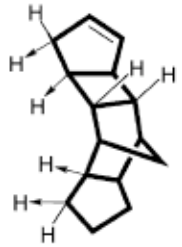
Unlinked CSD: 3299 hits



114447-04-0 (0.244)

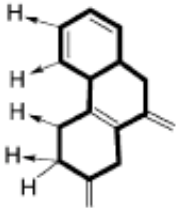


114447-04-0 (0.308)

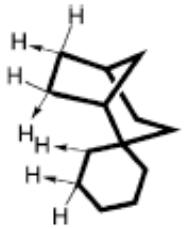


95313-58-9 (0.386)

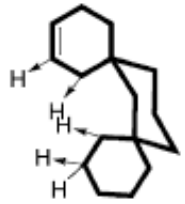
Unlinked TRIAD: 3279 hits



6312a448 (0.410)



6254a503 (0.512)

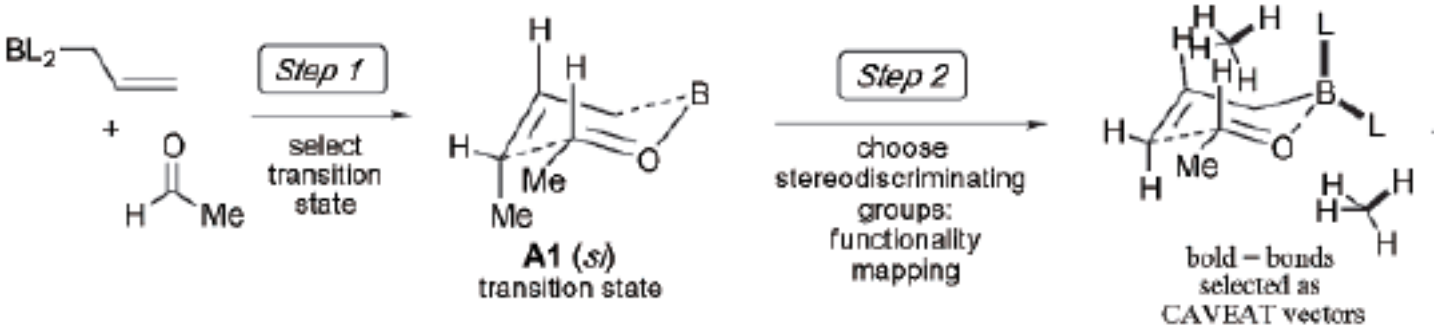


45a701 (0.520)

From the connecting elements and the stereo-discriminating elements a series of bond vectors can be selected which “uniquely define the chiral auxillary”. Using CAVEAT a number of scaffolds can be found for starting points for chiral ligand design.

- Kozlowski, M.; Panda, M. *J. Mol. Graph.* (2002), **20**, 399

De-novo design of a chiral ligand for asymmetric allylation



Product ratios controlled by the relative TS Energies shown below:

Functionality mapping: Targets A1/A2
Methanes as the probes

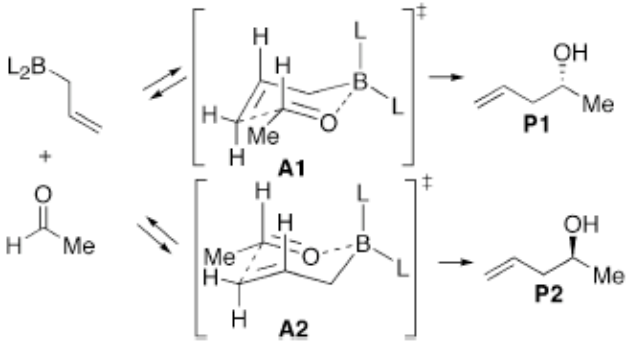
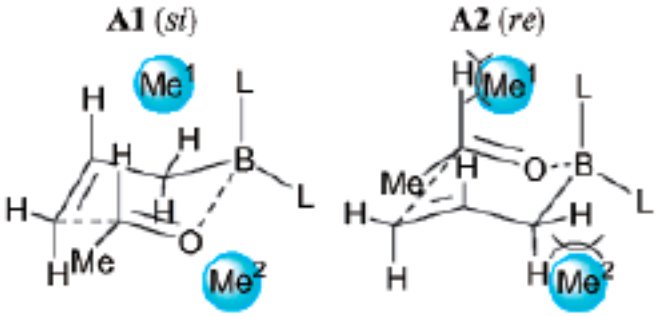


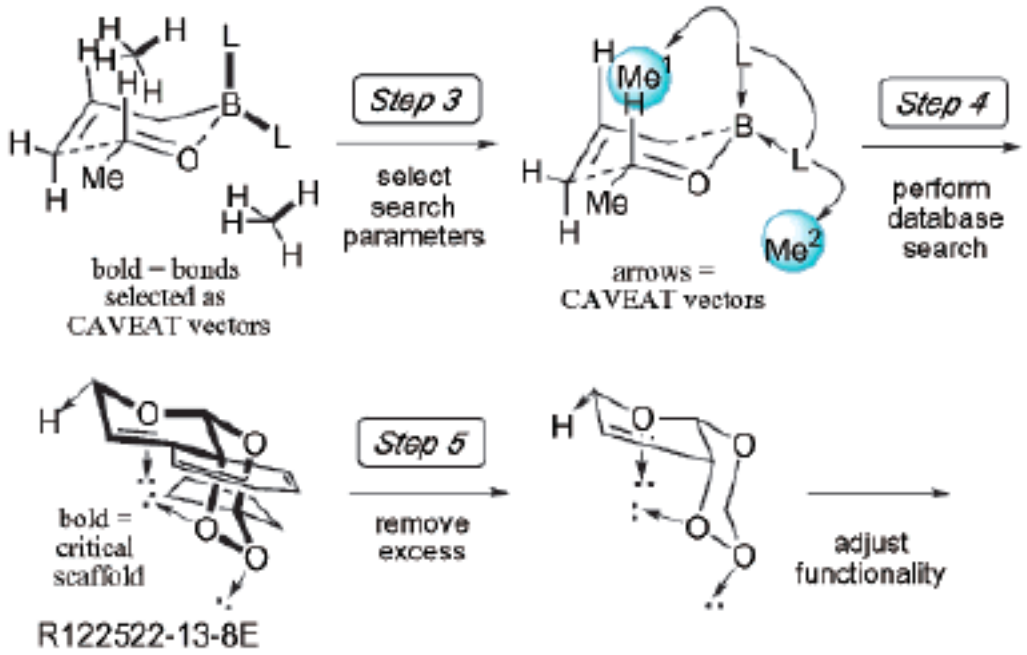
Figure 1. Boron allylation reaction.



MM2-force field calc.

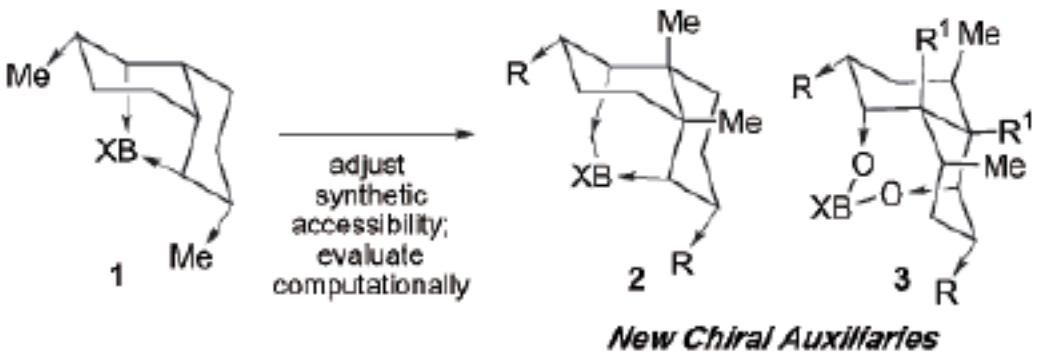
- Kozlowski, M. *et. al. Org. Lett. (2002),* **4,** 4391.

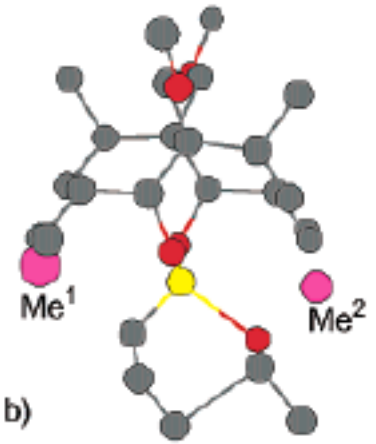
De-novo design of a chiral ligand for asymmetric allylation



Searched both linked CSD and CAS-3D

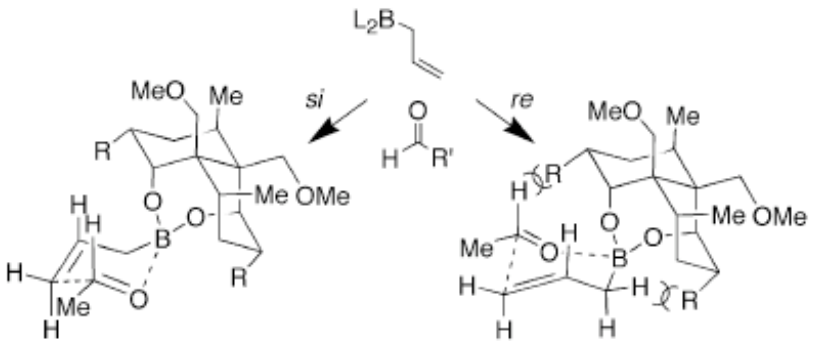
Cis-decalin structure!!!



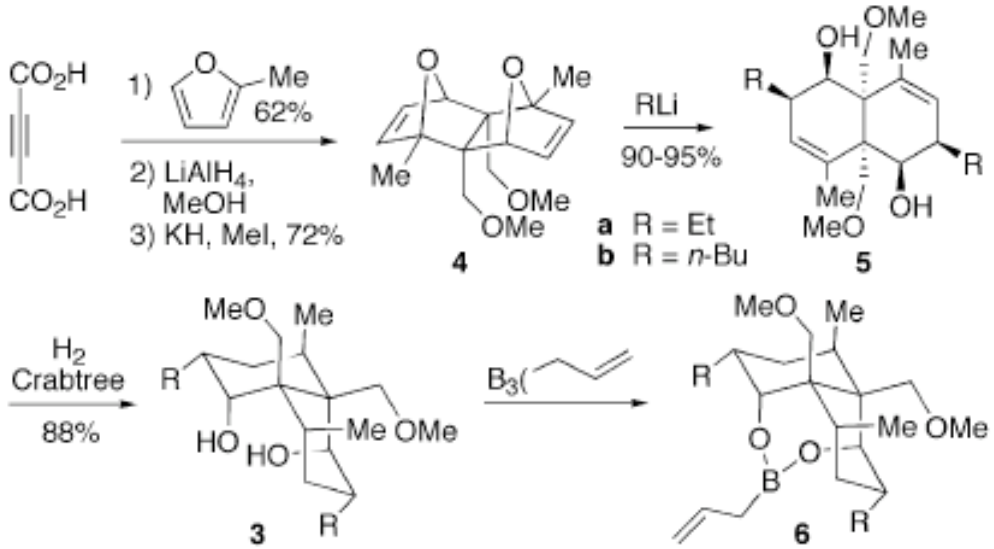


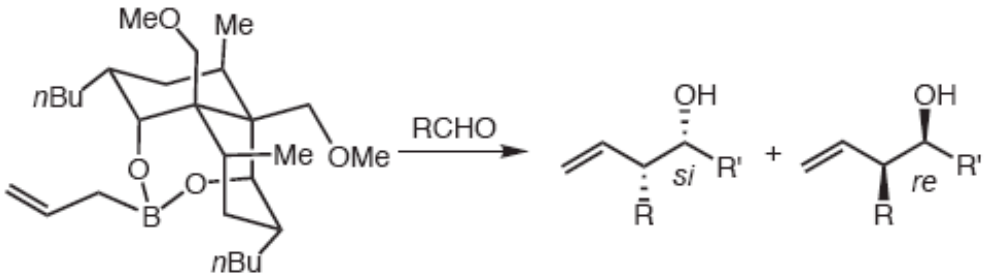
- Kozlowski, M. *et. al. Org. Lett. (2002),* **4,** 4391.

De-novo design of a chiral ligand for asymmetric allylation



TS structure calculations
(85 : 15 *si:re*, $\Delta\Delta G$ = 0.68 kcal/mol)
Based on MM2*, AM1, HF)

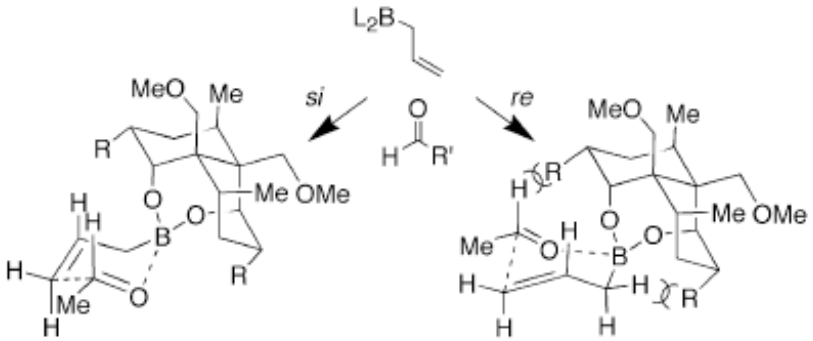




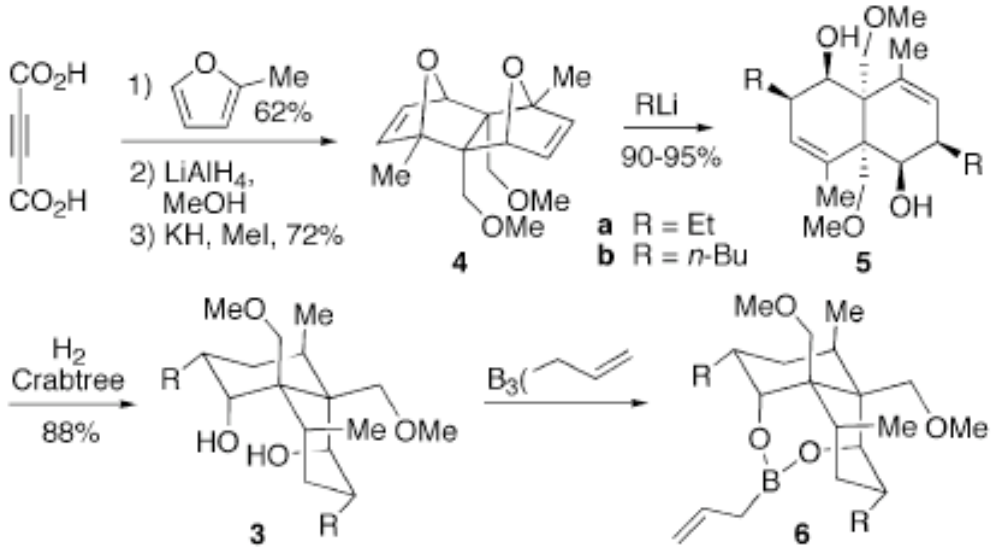
Observed selectivity
(71:29 $\Delta\Delta G$ = 0.36 kcal/mol)
Is within error (0.2-0.5 kcal/mol) of the TS calculations

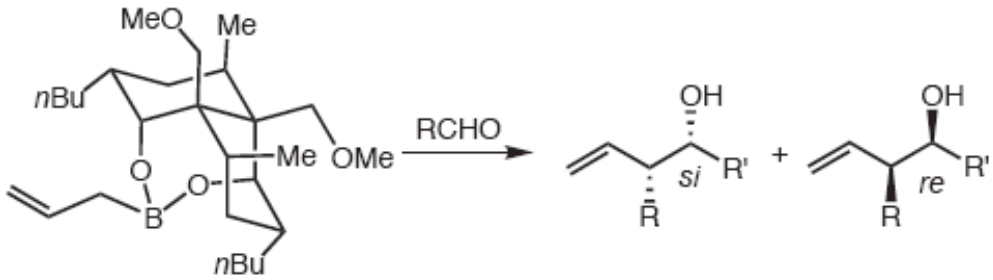
- Kozlowski, M. *et. al. Org. Lett. (2002),* **4,** 4391.

De-novo design of a chiral ligand for asymmetric allylation



TS structure calculations
(85 : 15 *si:re*, ΔΔG = 0.68 kcal/mol)
Based on MM2*, AM1, HF)





Observed selectivity
(71:29 ΔΔG = 0.36 kcal/mol)
Is within error (0.2-0.5 kcal/mol) of the TS calculations

"The decalin ligand could have been envisioned as a ligand for allylation w/o the use of computational protocols. The benefit of this method is that a large number of potential scaffolds can be rapidly screened. Additionally, the databases allow the user to identify structure types which the user may not be familiar with."

- Kozlowski, M. *et. al. Org. Lett. (2002),* **4,** 4391.

Quantum Mechanical Models Correlating Structure w/ Selectivity:

What if you want to improve an existing catalyst structure? What if the mechanism is not entirely understood?

Can you still use computational methods to predict structures with superior enantiodiscrimination?

QSSR (quantitative structure selectivity relationship): computational (mathmatical) correlation / relationship between the structure of chiral ligands to enantioselectivity in a given transformation.

Importantly, the QSSR can be predictive if satisfactory cross-validation is demonstrated

Physical descriptors: obtained by placing each catalyst, oriented in the same way, at the center of a cartesian coordinate. At each grid point the interaction between the aligned molecule and a probe (a positron) is evaluated computationally.

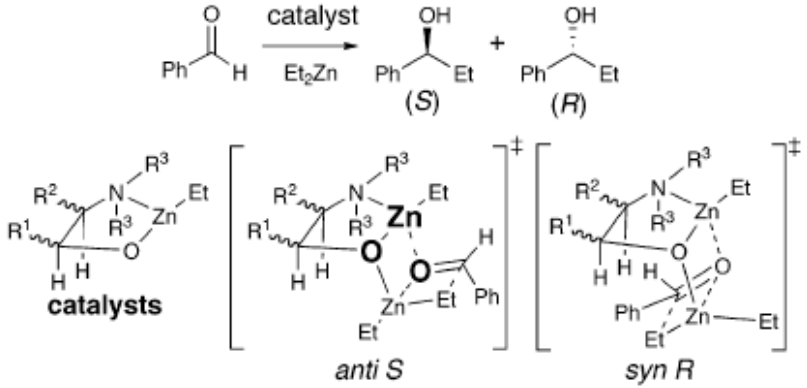
These interaction energies are regressed onto the enantioselectivity data to generate a mathmatical model that can be used to predict the selectivity of new analogues.

***Two sets of catalysts are employed in generating predictions using this approach:
1) A parameterization (testing) set to generate a model, and 2) a prediction set to which this model is applied

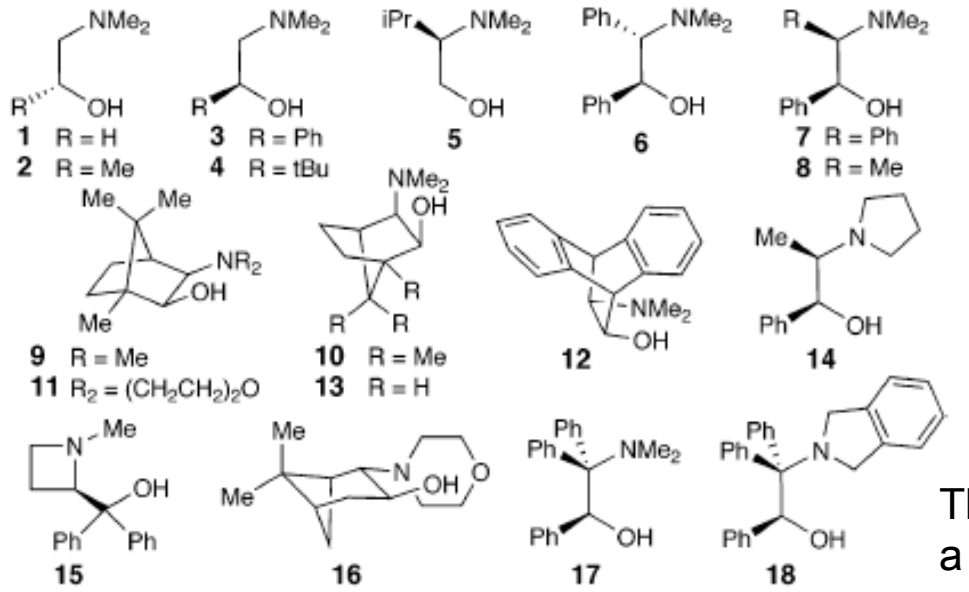
- Kozlowski, M. *et. al. J. Amer. Chem. Soc. (2004),* **126**, 15473

QSSR: Predicting the enantioselectivity of aldehyde alkylation

Start simple: can QSSR predict known selectivities of diethyl zinc to the structures of the β-aminoalkoxide zinc catalysts?



anti S (located using PM3 methods) the structures were aligned using the O, Zn, and O atoms



The structures were then split into a *training set* and *predictive set*

The QSSR models were aligned on a grid and a set of PM3 *probe interaction energies* (PIEs) were computed separately for each structure
The diethylzinc and aldehyde were ignored

- Kozlowski, M. *et. Al. J. Amer. Chem. Soc.* (2003), **125**, 6614.

QSSR: Predicting the enantioselectivity of aldehyde alkylation

The PIE data are the pool of independent variables from which the mutli-linear regression models are built.

The regression models that provided cross-validated results of the parametrization set (in the aldehyde alkylation) were comprised of two PIE points per catalyst. (aka, most of the variance in the er was restricted to two regions of the catalyst; groups of different sizes or electronic aspects at these positions modify the enantioselection)

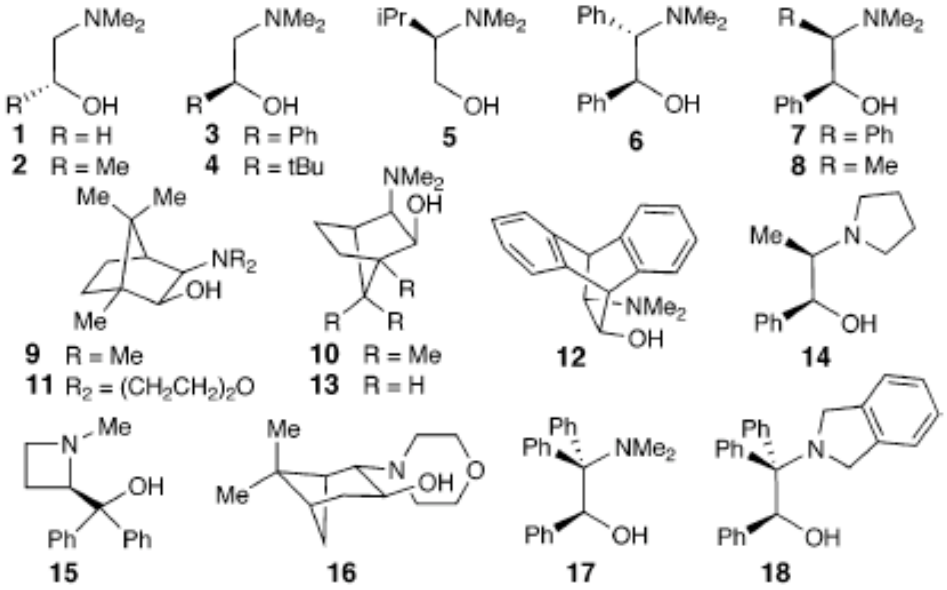


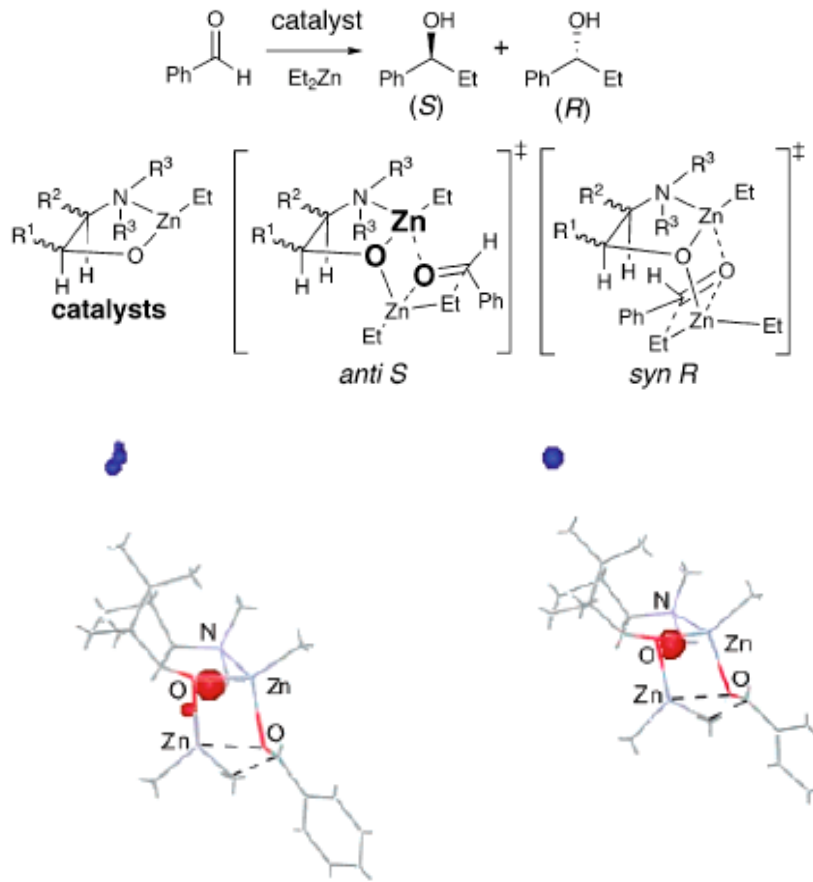
Table 1. QSSR Calculations Using Catalysts from **1**–**18**[a]

	expt.		*anti S*1 best				*anti S*1 avg	
cmpd	% ee[b]	ΔG[c]	% ee_{fit}	ΔG_{fit}[d]	PIE_1[e]	PIE_2[e]	% ee_{fit}	ΔG_{fit}
			Training Set[f]					
1	0	0.00	28	0.32	25.92	4.76	26	0.30
3	59	0.76	70	0.97	21.79	3.54	68	0.92
4	93	1.85	89	1.61	20.62	4.44	90	1.63
5	49	0.60	24	0.27	26.25	4.87	34	0.39
6	66	0.88	72	1.00	21.92	3.72	70	0.96
7	73	1.04	72	1.00	24.28	5.44	76	1.10
8	81	1.26	79	1.18	23.15	5.10	81	1.26
9	98	2.56	98	2.44	18.39	5.11	98	2.47
10	95	2.04	97	2.35	18.53	4.95	97	2.38
11	98	2.56	97	2.35	19.14	5.39	97	2.37
12	96	2.17	97	2.26	21.55	6.94	96	2.18
13	94	1.94	91	1.69	20.92	4.88	93	1.83
17	94	1.94	95	2.02	21.35	6.12	93	1.89
18	97	2.33	97	2.43	20.97	6.97	96	2.23
			Prediction Set[g]					
2	3	0.03	11	0.12	27.91	5.67	5	0.06
14	86	1.43	81	1.25	21.45	4.05	76	1.10
15	98	2.56	99	3.36	15.29	5.36	99	2.82
16	63	0.83	83	1.31	23.66	5.85	75	1.09

[a] Catalyst geometries taken from *anti S* transition structures. Grid1 orientation, 0.7 Å grid spacing. [b] (*S*)-product. [c] The % ee is converted to ΔG (kcal/mol) using $\Delta G = RT \ln K$, K is ratio of the (*R*) and (*S*) enantiomers. [d] $\Delta G_{fit} = a + c_1(PIE_1) + c_2(PIE_2)$; $a = 5.48$ kcal/mol, $c_1 = -0.27$, $c_2 = 0.36$. [e] Probe interaction energies (kcal/mol) at the two grid points identified in the QSSR analysis. [f] best,avg: SD = 0.23, 0.17 kcal/mol; $R^2 = 0.93, 0.95$. [g] best,avg: RMSE = 0.49, 0.29 kcal/mol; $R^2 = 0.72, 0.90$; CC = 0.95, 0.96.

- Kozlowski, M. *et. Al. J. Amer. Chem. Soc.* (2003), **125**, 6614.

QSSR: Predicting the enantioselectivity of aldehyde alkylation



-Average model consistantly gave better results
-Important structural features (for stereoinduction):

a) The first point was found near the bridging dimethylmethylene of DAIB
(TS analysis shows that bulk at this position is more Readily accepted in anti S than anti R, syn R, or syn S)

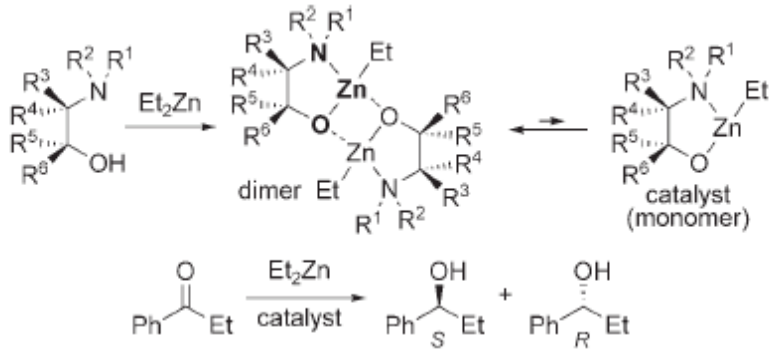
b) The bottom face of the 5 membered zinc chelate (destabilization by steric interactions, potential carbonyl dipole interactions)

Figure 3. Models from QSSR calculations (grid1, 0.7 Å) superimposed on the *anti S* transition structure from the DAIB ligand (**9**). Left = *average* model; Right = *best individual* model. Blue = ee increases with increasing PIE values; Red = ee decreases.

- Kozlowski, M. *et. Al. J. Amer. Chem. Soc.* (2003), **125**, 6614.

QSSR: a priori prediction of new catalysts from ground state

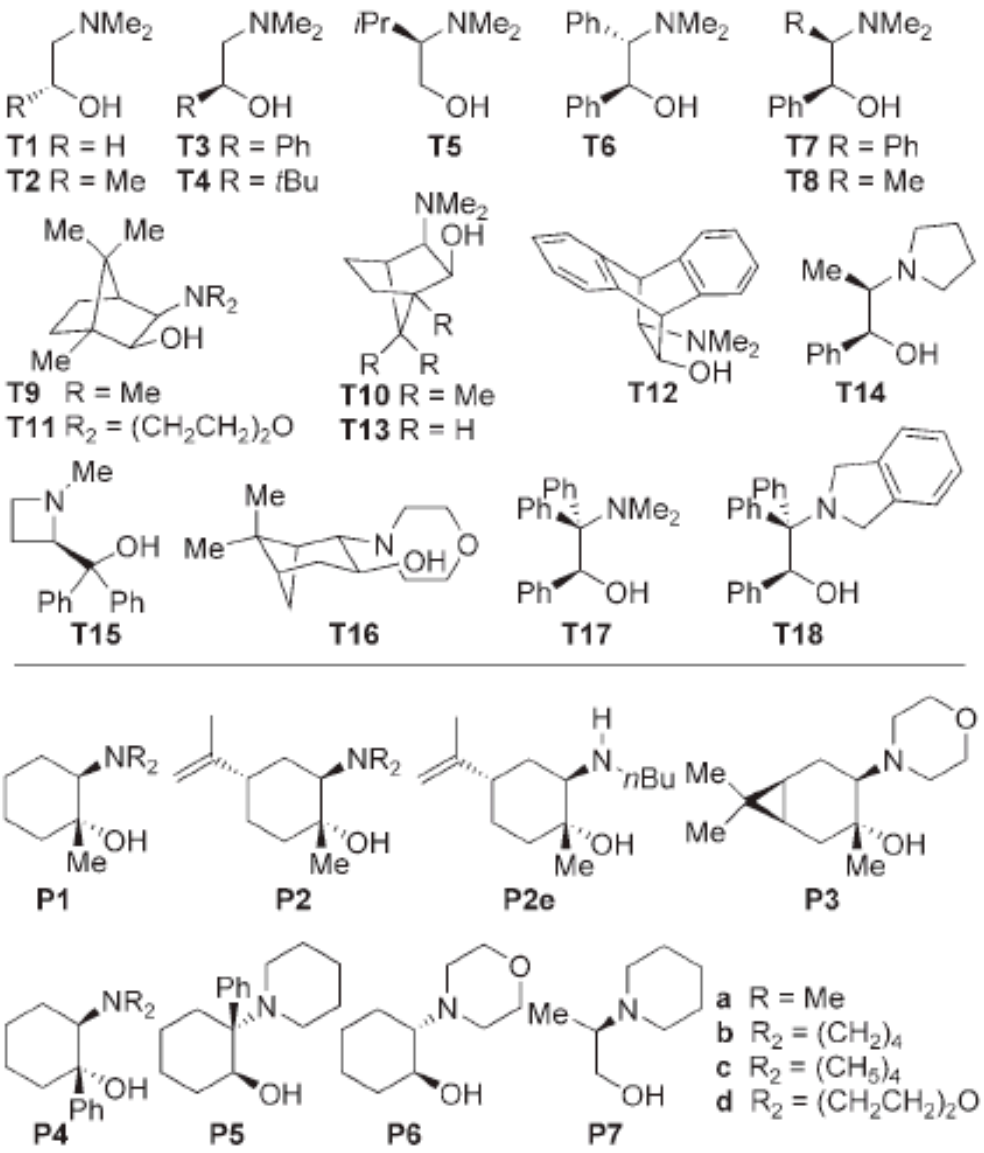
What if the TS structure is not known?



Zinc dimers (known cryst and in soln) are tetrahedral at Zn, which is likely to reflect the interactions encountered during rxn. (calculated using PM3)

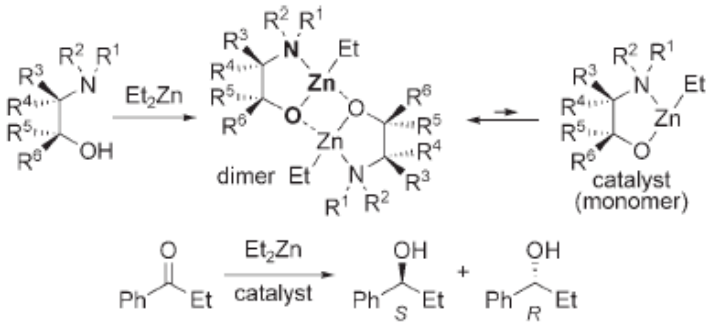
-Aligned by N, Zn, O atoms

-Prediction set of **T1-T18**
-Test set **P1-P7**



- Kozlowski, M. *et. al. Angew. Chem.* (2006), **118**, 5628

QSSR: a priori prediction of new catalysts from ground state



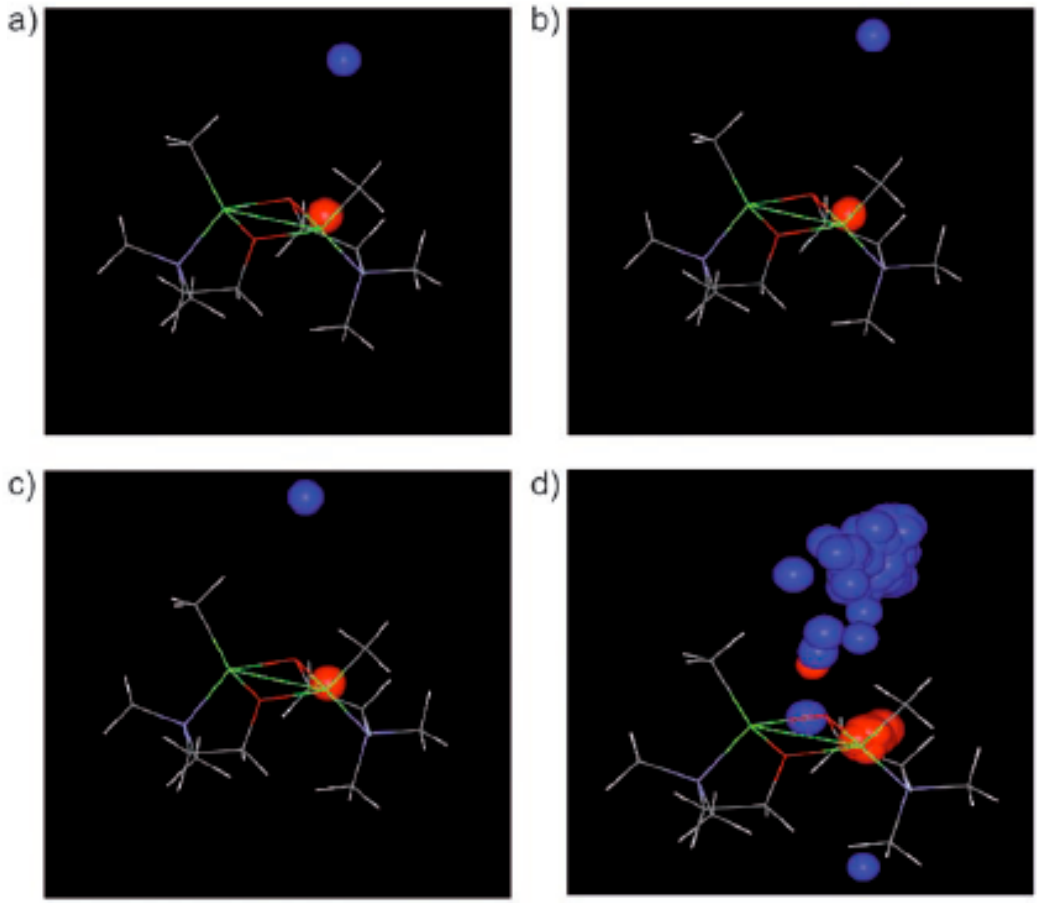


Figure 2. a–c) Several of the models from the "leave-two-out" analyses. d) Superimposed models from all 153 trials. Blue: *ee* value increases with increasing grid-point energy; red: *ee value* decreases.

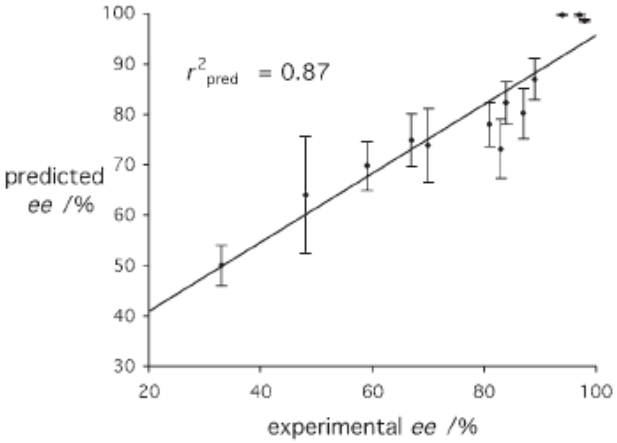
- Kozlowski, M. *et. al. Angew. Chem.* (2006), **118**, 5628

QSSR: a priori prediction of new catalysts from ground state

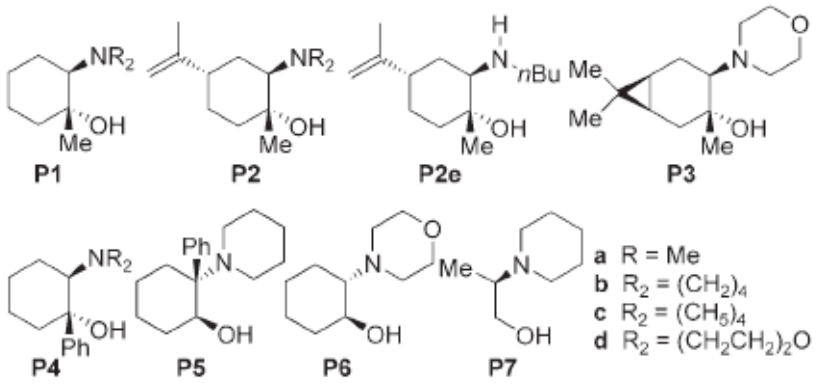
Table 1: QSSR a priori predictions for catalysts derived from **P1–P7**.

Ligand	Single run[a] ΔG[c]	Single run[a] % *ee*[d]	"Leave-two-out"[b] Mean ΔG[c]	"Leave-two-out"[b] SD ΔG[c]	"Leave-two-out"[b] % *ee*[d]	"Leave-two-out"[b] CI % *ee*[e]	Expt. % *ee*[f]
P1a	1.74	92.3	1.68	0.22	91.5	3.2	–
P1b	0.81	63.5	0.91	0.33	68.8	15.9	–
P1c	1.16	78.9	1.45	0.38	87.2	8.2	89
P1d	0.80	63.0	1.01	0.28	73.2	11.8	83
P2a	1.01	73.1	1.05	0.26	74.9	10.3	67
P2b	0.92	69.3	1.03	0.35	73.9	14.6	70
P2c	1.14	78.5	1.19	0.31	80.2	10.0	87
P2d	1.00	72.7	1.13	0.24	78.0	8.7	81
P2e	0.74	59.2	0.82	0.43	64.1	23.1	48
P3	1.11	77.4	1.26	0.29	82.3	8.4	84
P4a	2.88	99.0	2.68	0.31	98.6	0.8	–
P4b	2.10	96.0	2.01	0.38	95.2	3.2	–
P4c	3.61	99.7	3.77	0.77	99.8	0.3	97
P4d	3.67	99.8	3.78	0.75	99.8	0.2	94
P5	2.71	98.7	2.69	0.16	98.6	0.4	98
P6	0.89	67.6	0.93	0.21	69.8	9.8	59
P7	0.47	41.2	0.55	0.24	50.1	8.0	33

[a] Compounds **T1–T18** served as the parameterization set. [b] "Leave-two-out" cross-validating analysis using 16 compounds from **T1–T18** as the parameterization set (all 153 combinations). [c] Calculated ΔG in kcal mol^{-1}. [d] Generated from ΔG at 273 K; *S* enantiomer product. [e] 95% confidence interval. [f] From reactions performed at 273 K. *S* enantiomer product.

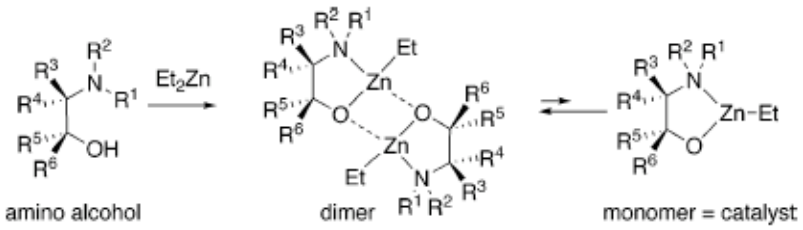


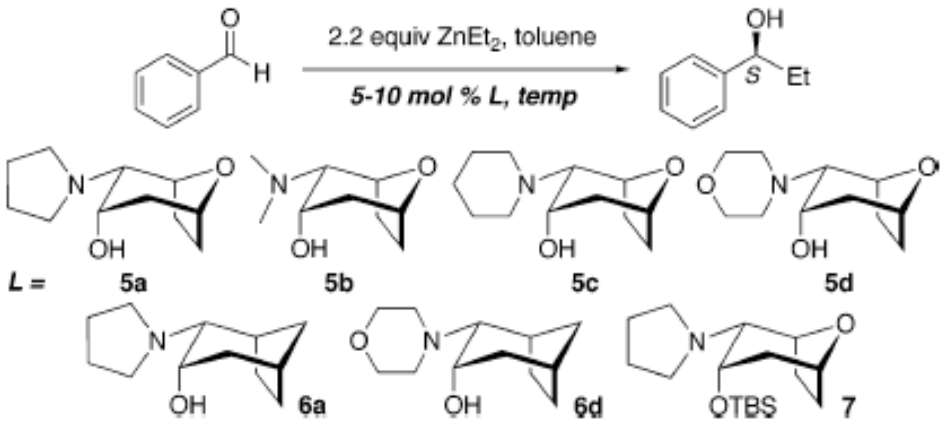
Notably: **P1** to **P4** was predicted, additionally
The low selectivity of **P7** was fairly accurately identified



- Kozlowski, M. *et. al. Angew. Chem.* (2006), **118**, 5628

QSSR: De Novo Chiral Amino Alcohols



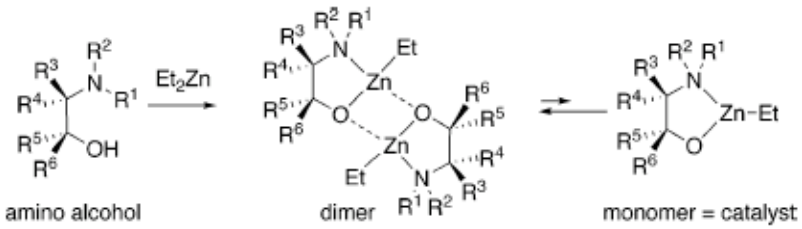


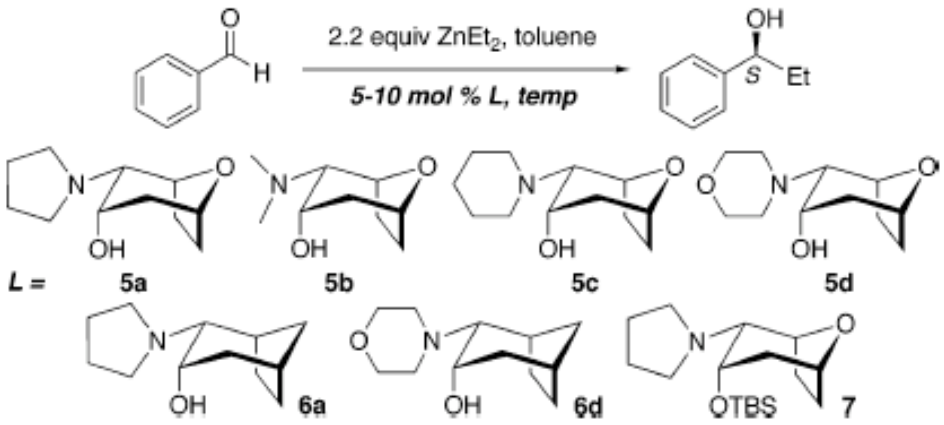
-Using the same parameterization set as Before.
-The test set was the previous 18 amino alkoxide zinc catalysts.

-2 programs were utilized to minimize the Catalyst conformers (PM3 in spartan and HF/3-21G*) thus two analyses were Completed and compared QMQSAR, and GQSAR

- Kozlowski, M. Hsung, R. *et. al. Org. Lett.* (2006), **8**, 1565.

QSSR: De Novo Chiral Amino Alcohols



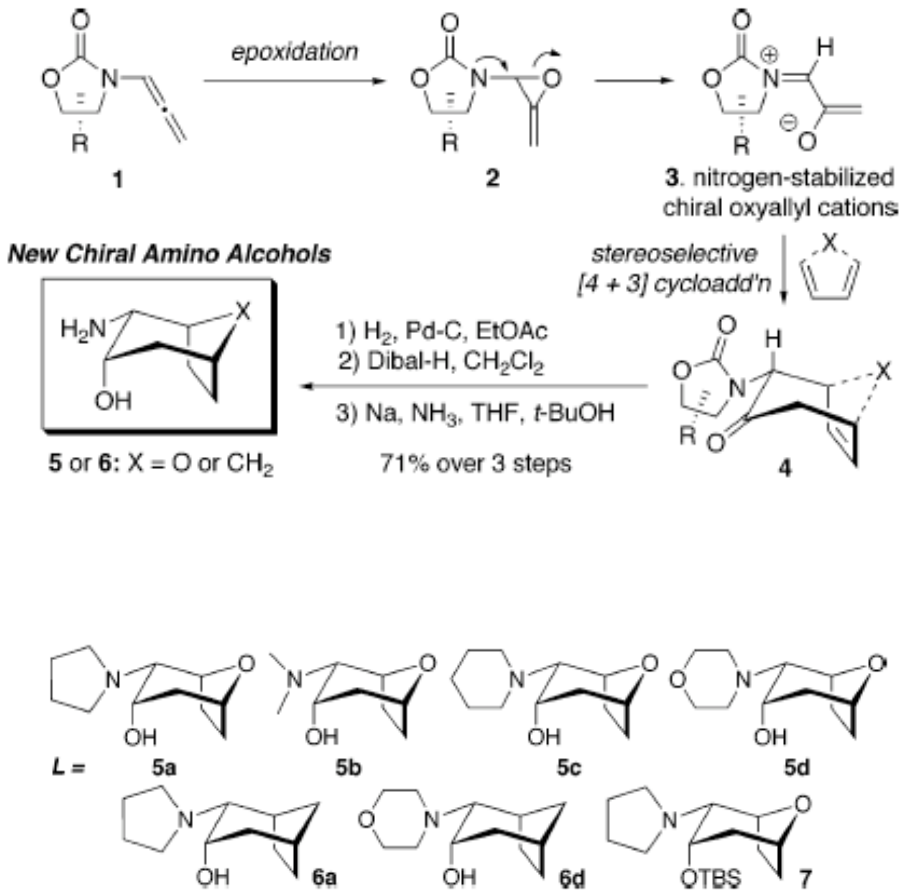


	average predictions (18 Models Total)			
	QMQSAR		GQSAR	
L	mean ee (%)	SD ee (%)	mean ee (%)	SD ee (%)
	228 K = −45 °C			
5a	99.2	0.1	98.1	0.7
5b	99.0	0.2	98.0	0.6
5c	99.2	0.1	98.0	0.6
5d	99.1	0.2	97.9	0.7
6a	99.0	0.1	97.8	0.5
6d	98.8	0.1	97.6	0.5
	278 K = 0 °C			
5a	98.0	0.2	95.9	1.2
5b	97.6	0.3	95.7	1.0
5c	98.1	0.2	95.8	1.0
5d	97.9	0.4	95.6	1.1
6a	97.6	0.2	95.5	0.9
6d	97.3	0.2	95.1	0.9
	298 K = 25 °C			
5a	97.0	0.2	94.3	1.5
5b	96.5	0.4	94.1	1.3
5c	97.2	0.3	94.3	1.2
5d	96.9	0.5	94.0	1.4
6a	96.6	0.2	93.8	1.1
6d	96.1	0.2	93.3	1.1

- Kozlowski, M. Hsung, R. *et. al. Org. Lett.* (2006), **8**, 1565.

Synthesis and evaluation

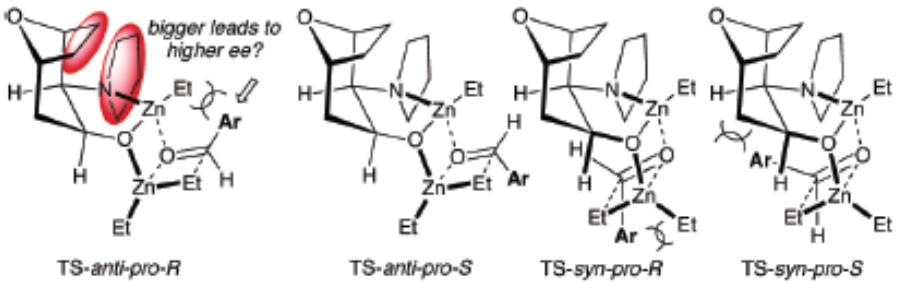
Scheme 1. Oxyallyl Cation [4 + 3] Cycloadditions



	average predictions (18 Models Total)					
	QMQSAR		GQSAR		expt. values	
L	mean ee (%)	SD ee (%)	mean ee (%)	SD ee (%)	ee (%)	yield (%)
			228 K = −45 °C			
5a	99.2	0.1	98.1	0.7	95	81
5b	99.0	0.2	98.0	0.6		
5c	99.2	0.1	98.0	0.6	97	81
5d	99.1	0.2	97.9	0.7	97	81
6a	99.0	0.1	97.8	0.5		
6d	98.8	0.1	97.6	0.5	95	73
			278 K = 0 °C			
5a	98.0	0.2	95.9	1.2	90	90
5b	97.6	0.3	95.7	1.0		
5c	98.1	0.2	95.8	1.0	93	92
5d	97.9	0.4	95.6	1.1	91	90
6a	97.6	0.2	95.5	0.9	93	88
6d	97.3	0.2	95.1	0.9	79	86
			298 K = 25 °C			
5a	97.0	0.2	94.3	1.5	89	93
5b	96.5	0.4	94.1	1.3		
5c	97.2	0.3	94.3	1.2	89	87
5d	96.9	0.5	94.0	1.4	89	93
6a	96.6	0.2	93.8	1.1		
6d	96.1	0.2	93.3	1.1	44	53

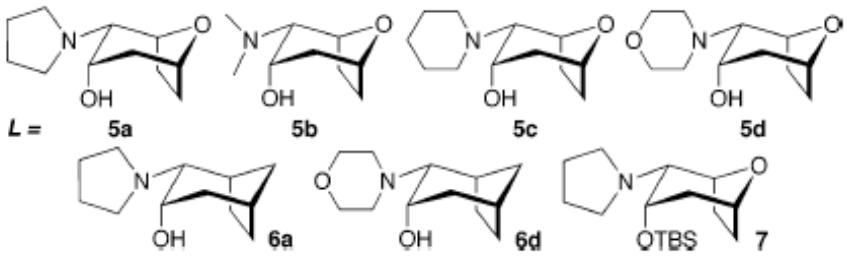
- Kozlowski, M. Hsung, R. *et. al. Org. Lett.* (2006), **8**, 1565.

Synthesis and evaluation



The two anti-TS more favorable than syn-TS
The anti-TS-pro-S most favorable.

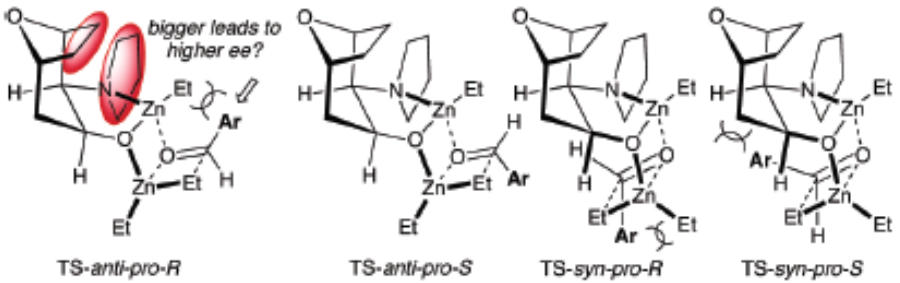
Speculated from this model that bulky
Aldehydes would further the selectivity



	average predictions (18 Models Total)					
	QMQSAR		GQSAR		expt. values	
L	mean ee (%)	SD ee (%)	mean ee (%)	SD ee (%)	ee (%)	yield (%)
			228 K = −45 °C			
5a	99.2	0.1	98.1	0.7	95	81
5b	99.0	0.2	98.0	0.6		
5c	99.2	0.1	98.0	0.6	97	81
5d	99.1	0.2	97.9	0.7	97	81
6a	99.0	0.1	97.8	0.5		
6d	98.8	0.1	97.6	0.5	95	73
			278 K = 0 °C			
5a	98.0	0.2	95.9	1.2	90	90
5b	97.6	0.3	95.7	1.0		
5c	98.1	0.2	95.8	1.0	93	92
5d	97.9	0.4	95.6	1.1	91	90
6a	97.6	0.2	95.5	0.9	93	88
6d	97.3	0.2	95.1	0.9	79	86
			298 K = 25 °C			
5a	97.0	0.2	94.3	1.5	89	93
5b	96.5	0.4	94.1	1.3		
5c	97.2	0.3	94.3	1.2	89	87
5d	96.9	0.5	94.0	1.4	89	93
6a	96.6	0.2	93.8	1.1		
6d	96.1	0.2	93.3	1.1	44	53

- Kozlowski, M. Hsung, R. *et. al. Org. Lett.* (2006), **8**, 1565.

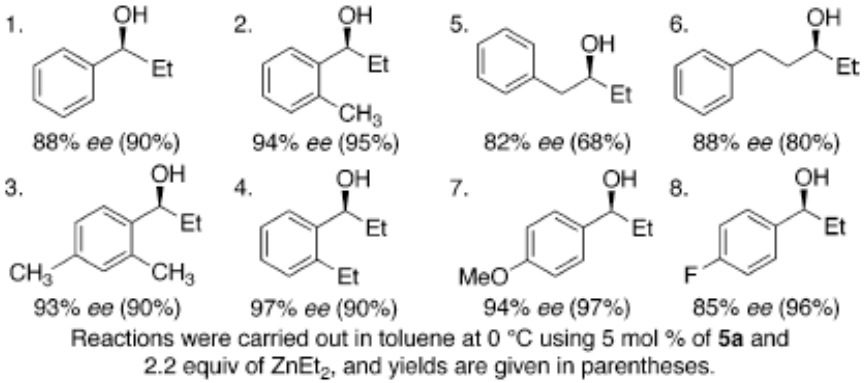
Synthesis and evaluation



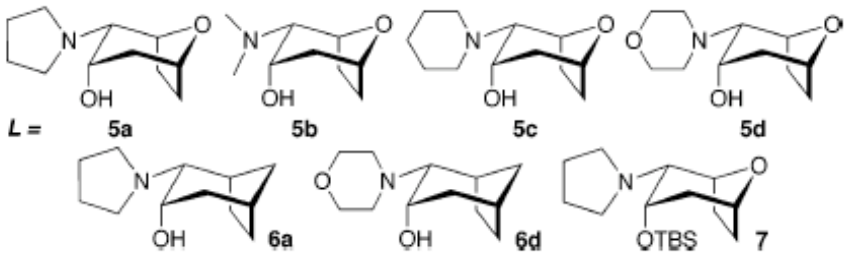
The two anti-TS more favorable than syn-TS
The anti-TS-pro-S most favorable.

Speculated from this model that bulky Aldehydes would further the selectivity

Tighter coordination of the aldehyde to Zn seems to further the selectivity.



Reactions were carried out in toluene at 0 °C using 5 mol % of **5a** and 2.2 equiv of $ZnEt_2$, and yields are given in parentheses.

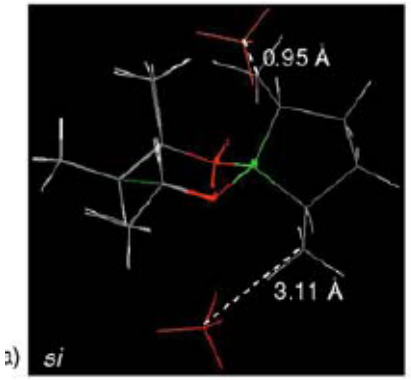


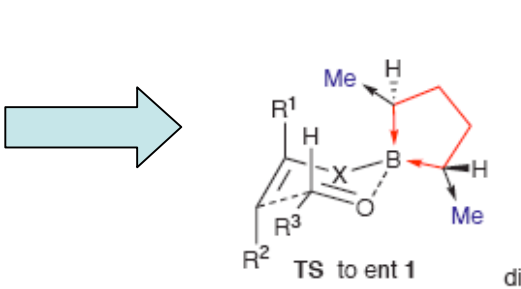
The implementation of a completely a prioi designed chiral ligand has been demonstrated using only the knowledge of the ground state conformation and the selectivities of a pool of known ligands

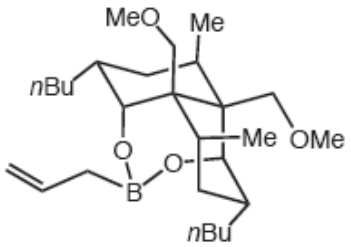
- Kozlowski, M. Hsung, R. *et. al. Org. Lett.* (2006), **8**, 1565.

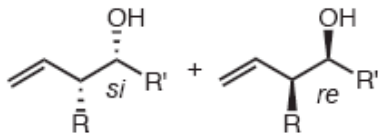
Conclusions

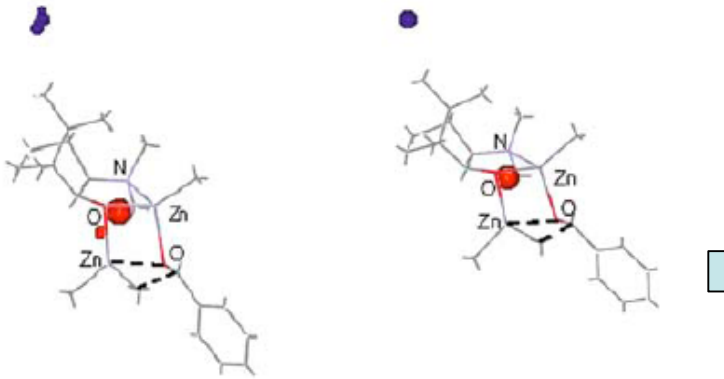
Many levels of success already exist for prediction-based discovery of chiral ligands:



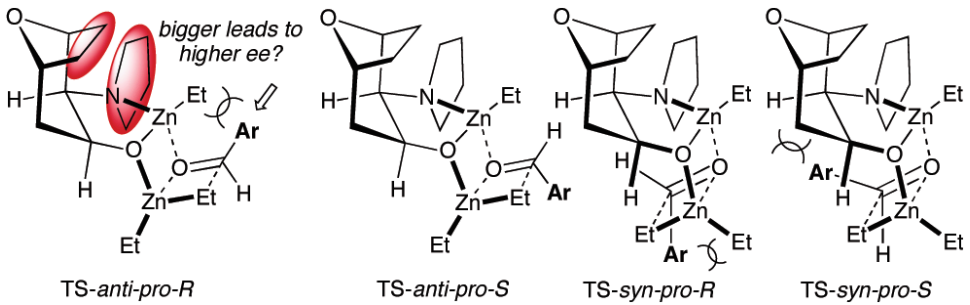








Average Model (6 grid points) *Best Individual* Model (2 grid points)



Ultimately, though, it will not be the power of the computations (MM2, HF, ect..) that bridge the divide between synthetic organic chemistry and computational chemistry.It will be the enginuity of the program developers to identify algorithms that accurately portray chemical systems. Even more importantly than the programmers will be the role of the synthetic chemist to adopt computational chemistry as a daily routine and not an exotic method to be used only at post-chemistry rationalization.

- Lipkowitz, K.; Kozlowski, M. *Synlett* (2003), 1547.

Bibliography

General Reading:

1) Lipkowitz, K. “Atomistic Modeling of Enantioselective Binding” *Acc. Chem. Res.* (2000), **33**, 555.
2) Lipkowitz, K.; Kozlowski, M. “Understanding Stereoinduction in Catalysis via Computer: New Tools for Asymmetric Synthesis” *Synlett* (2003), 1547.

Stereocartography:

1) Lipkowitz, K.; *et. al.* “Stereocartography: A Computational Mapping Technique That Can Locate Regions of Maximum Stereoinduction around Chiral Catalysts” *J. Amer. Chem. Soc.* (2002), **124**, 14255.

3d-QSSR/QSAR Studies:

1) Kozlowski, M. *et. al.* “Quantum Mechanical Models Correlating Structure with Selectivity…” *J. Amer. Chem. Soc.* (2003), **125**, 6614.
2) Kozlowski, M. *et. al.* “A Priori Theoretical Prediction of Selectivity…” *Angew. Chem.* (2006), **118**, 5628.
3) Kozlowski, M. *et. al. “De Novo* Chiral Amino Alcohols in Catalyzing Asymmetric Additions….” *Org. Lett.* (2006), **8**, 1565.
4) Kozlowski, M. *et. al.* “Is the A-ring of Spareine Essential for High Enantioselectivity in the Asymmetric Lithiation-substitution of N-Boc pyrrolidines?” *J. Amer. Chem. Soc.* (2004), **126**, 15473.

Functional Mapping/Database Mining/CAVEAT:

1) Kozlowski, M. *et. al.* “Computer Aided design of Chiral Ligands Part 1. Database search Methods…” *J. Mol. Graph. Mod.* (2002), **20**, 399.
2) Kozlowski, M. *et. al.* “Computer Aided design of Chiral Ligands Part 2. Functionality Mapping…” *J. Org. Chem.* (2003), **68**, 2061.
3) Kozlowski, M. *et. al.* “Computer Aided design of Chiral Ligands Part 3. A Novel Ligand for…” *Org. Lett.* (2002), **4**, 4391.
4) Bartlett, P. *et. al.* “ CAVEAT: A program to facilitate the design of organic molecules” *J. Comp. Aided Mol. Des. (*1994), **8**, 51.