Calculation on enantiomeric excess of catalytic asymmetric reactions of diethylzinc addition to aldehydes with topological indices and artificial neural network

Jiang Chen, Wen Jiwu, Li Mingzong, Tianpa You [*]

Department of Chemistry, University of Science and Technology of China, China

Received 21 January 2006; received in revised form 10 May 2006; accepted 11 May 2006
Available online 30 June 2006

Abstract

The relationships between the enantiomeric excess of products in several types of asymmetric diethylzinc addition reactions and the structures of the catalysts and reagents were studied by using the BP neural network with topological indices. Moderate cross-validated correlation coefficients were yielded by the models generated from the neural network. Then the models were further used to validate reaction mechanisms, determine the proper reaction environment, and predict the activities of new catalysts.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Asymmetric catalytic reaction; Enantiomer excess; Topological indices; Artificial neural network; QSAR

1. Introduction

Ever since the initial report of Oguni and Omi on the addition reaction of diethylzinc to benzaldehyde in the presence of a catalytic amount of (*S*)-leucinol with moderate enantioselectivity (49% e.e.) in 1984 [1], researches on asymmetric organozinc additions to carbonyl compounds have grown dramatically as this type of reaction is able to form carbon–carbon bond and asymmetric configuration synchronously as shown in Scheme 1. The first highly enantioselective ligand for the dialkylzinc addition to aldehydes is (−)-3-exo-dimethylaminoisoborneol [(−)-DAIB] discovered by Noyori and his co-workers in 1986 [2]. Over the past decades, a large number of chiral catalysts have been developed and high enantioselectivities have been achieved. In addition, the reaction of diethylzinc with aldehydes has also become a classical test in the design of new ligands for catalytic enantioselective syntheses [3].

As the growing number of the catalysts has been applied in this field, new ligands must be carefully designed and screened to achieve improved enantioselectivities. In this article, QSAR (quantitative structure–activity relationship) calculations were performed on some typical diethylzinc addition reactions with topological indices and artificial neural networks. The calculation results were useful to refine ligand structures, optimize catalysis environment, and validate reaction mechanisms.

2. Method of calculation and experimental data

2.1. Molecular descriptor

Because of the easiness of calculation and the flexibility in varieties of situations, topological descriptors including Randic index [4], Kier and Hall index [5] and Kier shape index [6] were used in our calculations.

The atom partial charge was also introduced as one of the descriptors of certain molecules after structure being optimized by the AM1 method [7] with the lowest energy.

Since the structures of catalysts varied a lot from each other in different types of reactions, a "universal" descriptor set could not be applied. Different descriptors were screened for different reactions to produce the models with acceptable predictive abilities.

* Corresponding author. Tel.: +86 551 3606944.
  E-mail address: ytp@ustc.edu.cn (T. You).
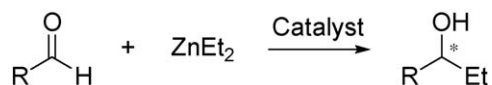
Scheme 1. Diethylzinc addition to aldehyde: the typical type of organozinc additions to carbonyl compounds.

2.2. ANN calculation

The ANN calculations were carried out in ANN-analyser, a generalized program that we have developed. Feed-forward networks with three layers were trained with the back-propagation of errors algorithm [8]. In all calculations, r is the correlation coefficient, and q is the cross-validated correlation coefficient.

2.3. Experimental data

All experimental data were obtained from papers published [9,10,13] or the work of our laboratory.

3. Results and discussion

3.1. Diethylzinc additions to aldehydes catalyzed by amino alcohols

Amino alcohols have been used as classic ligands for the asymmetric diethylzinc additions since the discovery of the first effective catalyst to this type of reactions by Noyori and his co-workers. Here a set of data from the work done by Pericàs et al. [9] was calculated. Amino alcohol ligand shown in Scheme 2 derived from fluorene was found to be an excellent ligand to a number of aldehydes to produce preponderant S configuration products. Topological indices Randic order 2, Kier and Hall
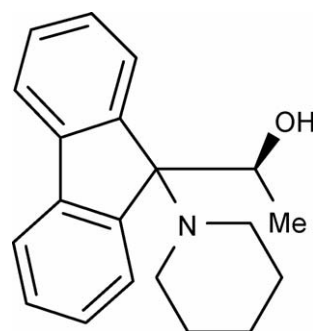
Scheme 2. Amino alcohol catalyst derived from fluorene.

order 2 and Kier shape index order 3 were selected to represent the aldehyde structures as the input variables among topological candidates after screening. The enantiomeric excess (e.e.%) was chosen as the dependent variable. A BP network consisted of three input neurons, one output neuron and three hidden neurons was used to produce the model. All 19 entries were divided into five groups to perform cross-validation. Prediction results in each validation group are shown in Table 1 and Fig. 1. Acceptable predictions were generated from the calculation.

3.2. Diethylzinc additions to aldehydes catalyzed by amino thiols

The amino thiols with cyclic and acyclic amino substituents shown in Scheme 3 were used by Kang et al. to catalyze diethylzinc addition reactions with 70–100% e.e. of R configuration products [10]. This set of reactions was also inspected by us.

Topological indices Kier and Hall (order 2) were used to represent the aldehydes and the substituent R^2 in different

Table 1
Observed [9] and predicted e.e.% of diethylzinc additions to aldehyde catalyzed by amino alcohols[a]

Entry	Aldehyde	Configuration	e.e.% (experimental)	e.e.% (predicted)
1	Benzaldehyde	S	96	96
2	o-Chlorobenzaldehyde	S	91	94
3	o-Methylbenzaldehyde	S	96	95
4	o-Methoxybenzaldehyde	S	82	82
5	m-Chlorobenzaldehyde	S	97	95
6	m-Methylbenzaldehyde	S	96	96
7	m-Methoxybenzaldehyde	S	96	97
8	p-Fluorobenzaldehyde	S	97	96
9	p-Methylbenzaldehyde	S	97	95
10	1-Naphthaldehyde	S	90	95
11	2-Naphthaldehyde	S	96	95
12	Hexanal	S	91	91
13	Heptanal	S	91	90
14	Nonyl aldehyde	S	90	90
15	Isovaleraldehyde	S	95	93
16	3-Phenylpropionaldehyde	S	91	95
17	Cyclohexanecarboxaldehyde	S	98	96
18	2-Ethylbutyraldehyde	S	94	94
19	E-α-Methylcinnamaldehyde	S	94	93
Correlation coefficient (r)				0.8575
Squared cross-validated correlation coefficient (q^2)				0.5245

[a] The reactions were carried out at 0 °C in toluene in the presence of 3 mol% of catalyst and 2.2 equiv. of diethylzinc (1 M in hexane).

Table 2
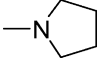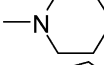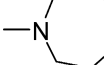Observed [10] and predicted e.e.% of diethylzinc additions to aldehyde catalyzed by amino thiols[a]
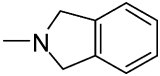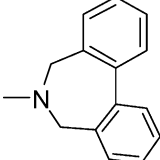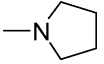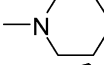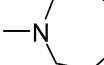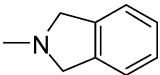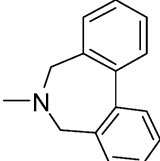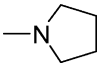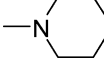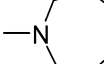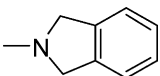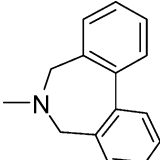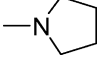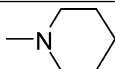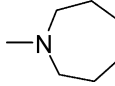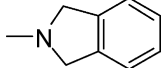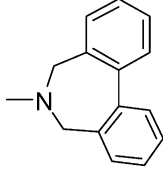
Entry	Aldehyde	R^2 in catalysts	Configuration	e.e.% (experimental)	e.e.% (predicted)
1	Benzaldehyde	—N (pyrrolidine)	R	99	97
2	Benzaldehyde	—N (piperidine)	R	100	100
3	Benzaldehyde	—N (azepane)	R	86	96
4	Benzaldehyde	—N (isoindoline)	R	97	100
5	Benzaldehyde	—N (dibenzazepine)	R	73	93
6	Benzaldehyde	$-N^nBu_2$	R	99	99
7	Benzaldehyde	$-NBn_2$	R	72	71
8	Cinnamaldehyde	—N (pyrrolidine)	R	78	85
9	Cinnamaldehyde	—N (piperidine)	R	77	87
10	Cinnamaldehyde	—N (azepane)	R	74	74
11	Cinnamaldehyde	—N (isoindoline)	R	72	87
12	Cinnamaldehyde	—N (dibenzazepine)	R	66	68
13	Cinnamaldehyde	$-N^nBu_2$	R	81	84
14	Cinnamaldehyde	$-NBn_2$	R	72	70
15	Cyclohexanecarboxaldehyde	—N (pyrrolidine)	R	99	100
16	Cyclohexanecarboxaldehyde	—N (piperidine)	R	100	96
17	Cyclohexanecarboxaldehyde	—N (azepane)	R	97	100
18	Cyclohexanecarboxaldehyde	—N (isoindoline)	R	96	100
19	Cyclohexanecarboxaldehyde	—N (dibenzazepine)	R	97	99
20	Cyclohexanecarboxaldehyde	$-N^nBu_2$	R	100	99
21	Cyclohexanecarboxaldehyde	$-NBn_2$	R	97	100
22	2-Naphthaldehyde	—N (pyrrolidine)	R	99	94

Table 2 (*Continued*)

Entry	Aldehyde	R^2 in catalysts	Configuration	e.e.% (experimental)	e.e.% (predicted)
23	2-Naphthaldehyde		R	99	100
24	2-Naphthaldehyde		R	96	95
25	2-Naphthaldehyde		R	97	93
26	2-Naphthaldehyde		R	96	91
27	2-Naphthaldehyde	$-N^nBu_2$	R	98	97
28	2-Naphthaldehyde	$-NBn_2$	R	95	91
Correlation coefficient (r)					0.8580
Squared cross-validated correlation coefficient (q^2)					0.6376

[a] The reactions were carried out in toluene over 12 h using 1 M hexane solution of diethylzinc.
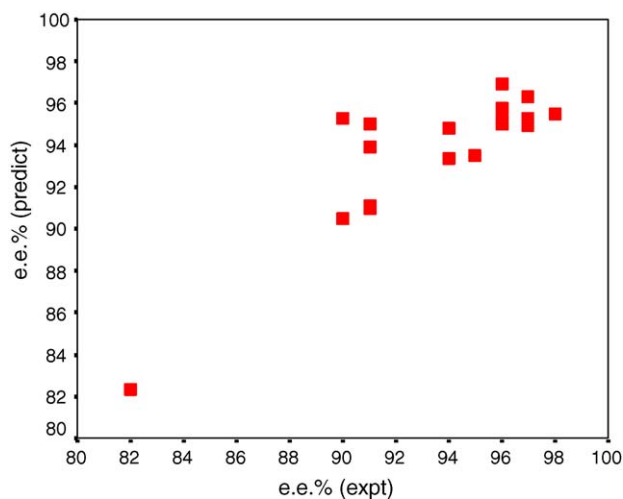

Fig. 1. Plot of experimental vs. ANN predicted e.e.% of diethylzinc additions to aldehyde catalyzed by amino alcohols ligands.
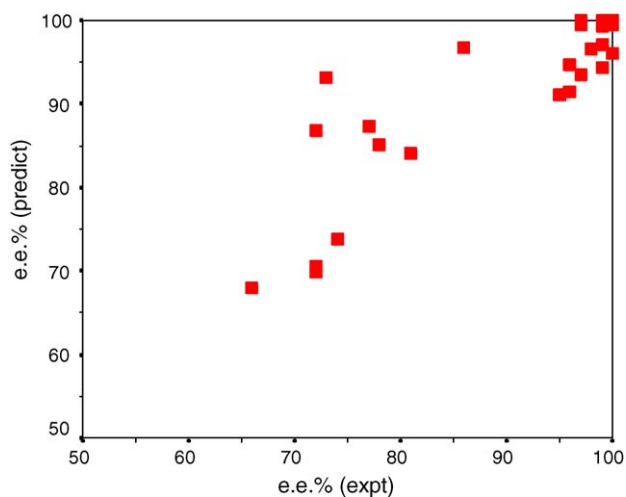

Fig. 2. Plot of experimental vs. ANN predicted e.e.% of diethylzinc additions to aldehyde catalyzed by amino thiols.

ligands. The partial charge on the N atom of the ligands was also introduced as the independent variable after structure being optimized by the AM1 method. Enantiomeric excess was chosen as the dependent variable. The BP network consisted of three input neurons, one output neuron and three hidden neurons. All 28 entries were divided into seven groups to perform cross-validation. Results are shown in Table 2 and Fig. 2.

As Noyori and co-workers have proposed the mechanism of dimethylzinc addition to benzaldehyde catalyzed by amino alcohols [11,12], shown in Scheme 4, we checked whether the amino thiols catalysts follow a similar path.

From Scheme 4, it is obvious that the steric hindrance of R plays a dominant role in the formation of stereoselective
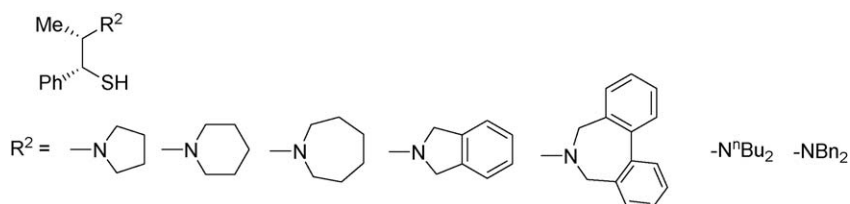
Scheme 3. Amino thiol catalysts used by Kang et al.
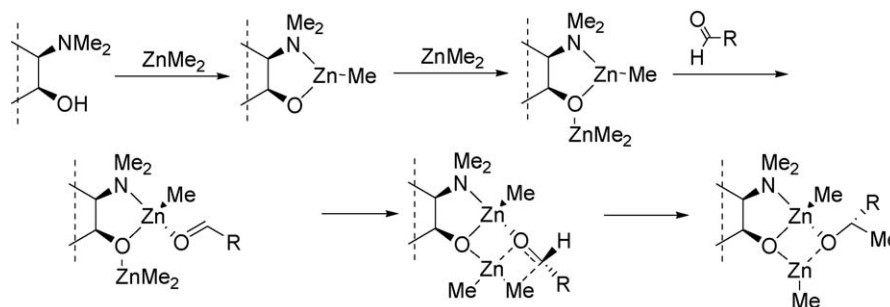
Scheme 4. Proposed mechanism for the dimethylzinc addition to benzaldehyde catalyzed by amino alcohol by Noyori and Co-workers.

products. Fixing the catalyst to , a set of test struc-

tures of aldehydes from PhCH=CHCHO to Ph(CH=CH)$_3$CHO, together with benzaldehyde, were inputted into the trained network. Predicted values of e.e.% are shown in Fig. 3. The hindrance as well as the e.e.% value predicted decreases when the carbon chain increases between carbonyl and phenyl gradually. This phenomenon is consistent with the Noyori's mechanism. It may imply that the mechanism of the reaction catalyzed by amino thiols is similar to the one by amino alcohols.

3.3. Diethylzinc additions to aldehydes catalyzed by amines

Chiral amines containing no hydroxyl or thiol groups but having multiple nitrogen atoms capable of chelate coordination are also found to be effective in the dialkylzinc additions to aromatic aldehydes. (S)-2-(2′-pyrrolidinyl)pyridine as shown in Scheme 5 was utilized by Falorni et al. to catalyze the addition reaction with up to 100% e.e.[13].

Fig. 3. Relation of structure of aldehyde vs. ANN predicted e.e.% of diethylzinc additions to aldehyde catalyzed by amino thiols.

With e.e.% as the usual output variable, topological indices Randic order 3, Kier and Hall order 3 for the aldehydes, the partial charge on the O atom of the aldehyde, temperature, and reaction time were chosen as the input variables. Cross-

Table 3
Observed [13] and predicted e.e.% of diethylzinc additions to aldehyde catalyzed by amines[a]

Entry	Aldehyde	Temperature (°C)	Time (h)	Configuration	e.e.% (experimental)	e.e.% (predicted)
1	Benzaldehyde	20	20	S	93	90.18
2	Benzaldehyde	−10	28	S	100	92.32
3	p-Methoxybenzaldehyde	−10	20	S	98	100.00
4	p-Chlorobenzaldehyde	−10	16	S	100	100.00
5	p-Methylbenzaldehyde	−10	6	S	92	67.00
6	o-Methylbenzaldehyde	−10	16	S	100	95.69
7	3-Phenylpropionaldehyde	50	6	S	60	67.98
8	3-Phenylpropionaldehyde	20	6	S	59	67.98
9	3-Phenylpropionaldehyde	−10	40	S	29	41.70
10	Cinnamaldehyde	−10	18	S	91	100.00
11	Phenylpropargyl aldehyde	−10	18	S	24	24.32
12	Heptanal	20	16	S	67	66.75
13	Heptanal	−10	40	S	40	30.78
14	Cyclohexanecarboxaldehyde	−10	16	S	90	95.53
15	2-Furaldehyde	−10	8	S	33	32.98
Correlation coefficient (r)						0.9334
Squared cross-validated correlation coefficient (q^2)						0.8570

[a] The reactions were carried out in diethyl ether using 6 mol% of catalyst and 1 M hexane solution of diethylzinc.

Scheme 5. Amines catalyst used by Falorni and co-workers.

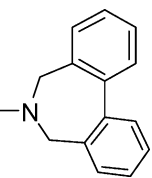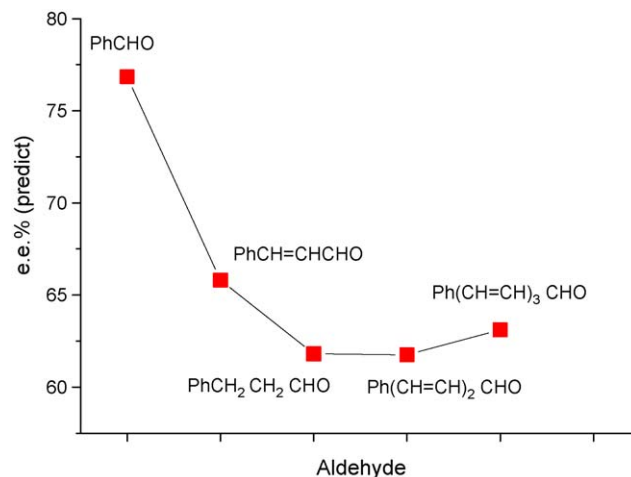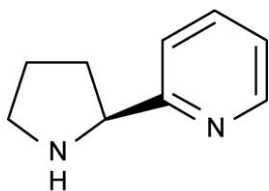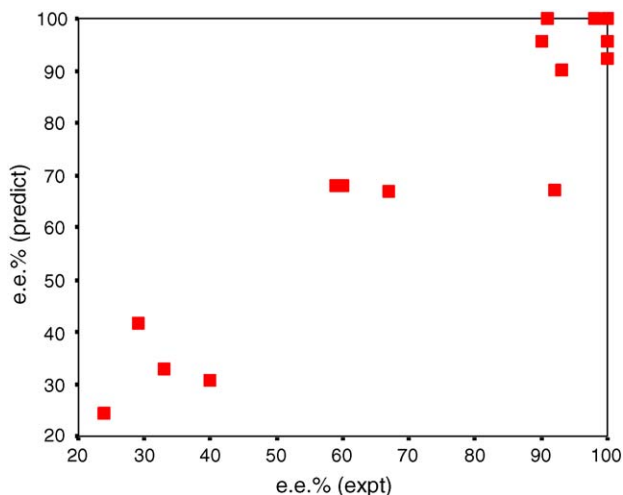Fig. 4. Plot of experimental vs. ANN predicted e.e.% of diethylzinc additions to aldehyde catalyzed by amines.

Fig. 5. Relation of reaction time vs. ANN predicted e.e.% of diethylzinc additions to aldehyde catalyzed by amine at 20 °C.

validation on five groups consisted of three entries, each of which was performed by a BP network with two hidden neurons. Results are shown in Table 3 and Fig. 4.

Reactions with 3-phenylpropionaldehyde were further investigated. A test set of reaction time varied from 5 to 60 h at 20 °C were input into the trained network. Predicted values of e.e.% are shown in Fig. 5.

An interesting result was gained. The left half of the curve tallies with our normal knowledge as the enantiomeric excess increases with prolonged reaction time. But after 40 h, the predicted e.e.% drops dramatically. The phenomenon may be expla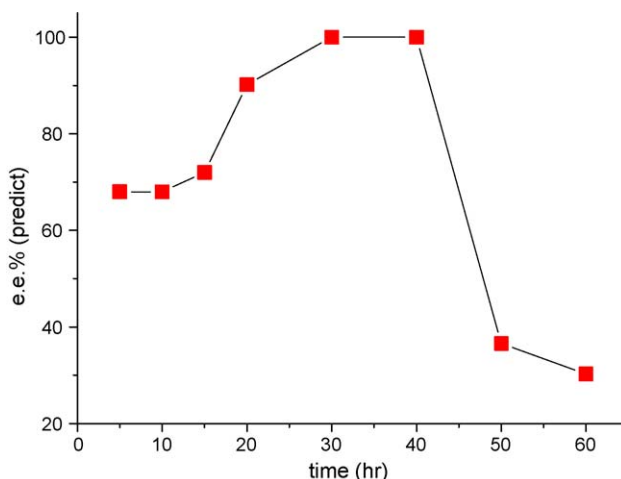ined as follows. The experimental data with long reaction time shown in Table 3 are all obtained at low temperature, such as entries 9 and 13. The reaction speed will not be as fast as the ones at 20 or 50 °C and the system may still not reach its balance after 40–50 h. So that an illusion may be observed that the longer reaction time leads to lower e.e.%. If the reactions at −10 °C last long enough, such as 80 or 100 h, better e.e.% may be achieved. The neural network may be misled by this set of specific data to a deficient prediction result after 40 h. However, there is still another probability implied by the network that the disgusting racemism may occur after long hours. Experimental workers should check it further. Anyway, it can be clearly observed from the prediction of the model that e.e.% will not increase after 40 h at 20 °C, and the longer time will not actually be beneficial to the enantiomeric excess. The preferred reaction time is about 30–40 h.
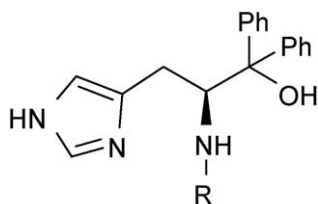
3.4. Diethylzinc additions to aldehydes catalyzed by amino alcohol containing multiple haptos

Lastly, a set of experimental data by our own laboratory were calculated to help design new catalysts. The framework of this

Table 4

Diethylzinc additions to aldehyde catalyzed by amino alcohol with multiple haptos[a]

Entry	R in catalyst	Aldehyde	Configuration	e.e.% (experimental)	e.e.% (predicted)
1	H	Benzaldehyde	R	17	18
2	*i*-Pr	Benzaldehyde	R	42	44
3	Benzyl	Benzaldehyde	R	38	57
4	*c*-Hexyl	Benzaldehyde	R	52	57
5	*c*-Hexyl	4-Tolualdehyde	R	56	58
6	*c*-Hexyl	4-Methoxybenzaldehyde	R	57	60
7	*c*-Hexyl	2-Chlorobenzaldehyde	R	49	57
8	*c*-Hexyl	4-Chlorobenzaldehyde	R	63	56
9	*c*-Hexyl	2,4-Dichlorobenzaldehyde	R	58	57
10	*c*-Hexyl	2-Thiophenecarboxadehyde	R	56	55
11	*c*-Hexyl	1-Naphthaldehyde	R	66	72
Correlation coefficient of prediction (r)					0.8833
Squared cross-validated correlation coefficient (q^2)					0.5451

[a] The reactions were carried out in toluene at room temperature for 48 h.

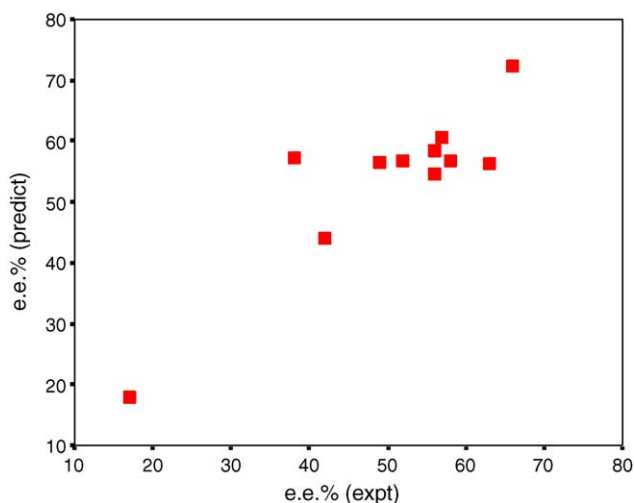Scheme 6. Catalysts derived from L-histidine with multiple haptos.

Fig. 6. Plot of experimental vs. ANN predicted e.e.% of diethylzinc additions to aldehyde catalyzed by amino alcohol with multiple haptos.

set of catalysts derived from L-histidine containing multiple haptos is shown in Scheme 6. All reactions were carried out in toluene at room temperature for 48 h, enantiomer excesses of products were determined by gas chromatography, and configurations were determined by the retention times of the major peaks.
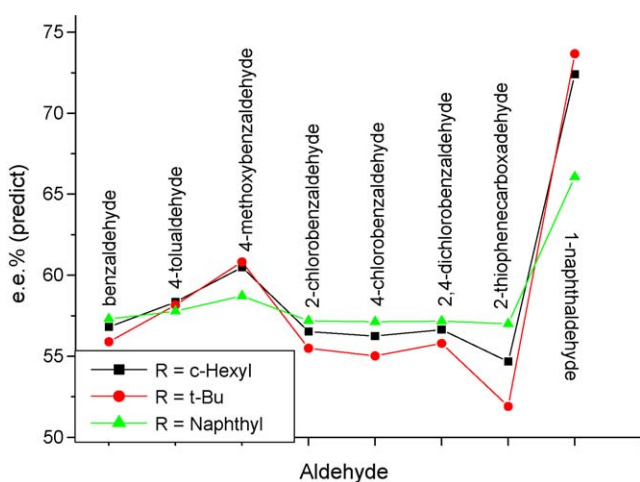

Fig. 7. ANN predicted e.e.% of diethylzinc additions to aldehyde catalyzed by amino alcohol with different substituents.

With e.e.% as the usual output variable, topological indices Randic order 2 for aldehydes and R in catalysts and the partial charge on the O atom of the aldehyde were chosen as the input variables. Calculation was performed by a BP network with two hidden neurons. Results are shown in Table 4 and Fig. 6.

We intended to predict new substituents in ligand leading to better e.e.%. Disappointedly, no common groups were found to improve the enantiomer-selectivity significantly by the network compared with cyclohexyl. However, it was predicted that naphthyl and *t*-Bu were acceptable substituents resulting in small enhancement of e.e.% in some instances as shown in Fig. 7. Further experimental work is still under way.

4. Conclusion

Different models were developed for varieties of asymmetric catalytic diethylzinc addition reactions by using molecule topological indices and feed-forward neural network calculation. Then the models were further used to validate reaction mechanisms, determine the proper reaction environment, and predict the activities of new catalysts. This method can be extended to other asymmetric reactions and would help to design novel asymmetric catalysts.

Acknowledgement

References

[1] N. Oguni, T. Omi, Tetrahedron Lett. 25 (1984) 2823.
[2] M. Kitamura, S. Suga, K. Kawai, R. Noyori, J. Am. Chem. Soc. 108 (1986) 6071.
[3] P. Lin, Y. Hongbin, Chem. Rev. 101 (2001) 757.
[4] M. Randic, J. Am. Chem. Soc. 97 (1975) 6609.
[5] L.B. Kier, L.H. Hall, Molecular Connectivity in Chemistry and Drug Research, Academic Press, New York, 1976.
[6] L.B. Kier, Computational Chemical Graph Theory, Nova Science Publishers, New York, 1990, pp. 151–174.
[7] M.J. Dewar, E.G. Zoebisch, F.E. Healy, J. Am. Chem. Soc. 107 (1985) 3902.
[8] (a) J. Zupan, J. Gasteiger, Neural Networks for Chemists: An Introduction, VCH, Weinheim, Germany, 1993, pp. 119–148;
(b) D. Svozil, V. Kvasnicka, J. Pospichal, Chemom. Intell. Lab. Syst. 39 (1997) 43.
[9] K.S. Reddy, L. Solà, A. Moyano, M.A. Pericàs, Synthesis 1 (2000) 165.
[10] (a) J. Kang, J.W. Kim, J.W. Lee, D.S. Kim, J.I. Kim, Bull. Korean Chem. Soc. 17 (1996) 1135;
(b) J. Kang, D.S. Kim, J.I. Kim, Synlett 10 (1994) 842.
[11] M. Yamakawa, R. Noyori, Organometallics 18 (1999) 128.
[12] M. Yamakawa, R. Noyori, J. Am. Chem. Soc. 117 (1995) 6327.
[13] (a) G. Chelucci, S. Conti, M. Falorni, G. Giacomelli, Tetrahedron 47 (1991) 8251;
(b) S. Conti, M. Falorni, G. Giacomelli, F. Soccolini, Tetrahedron 48 (1992) 8993.