Structural Analysis of Transition Metal β-X Substituent Interactions. Toward the Use of Soft Computing Methods for Catalyst Modeling

Thomas R. Cundari,*,# Jun Deng,# Horia F. Pop,† and Costel Sârbu†

Department of Chemistry, Computational Research on Materials Institute (CROMIUM), The University of Memphis, Memphis, Tennessee 38152-6060, and Departments of Chemistry and Computer Science, Babes-Bolyai University, Cluj-Napoca, Romania

Received January 11, 2000

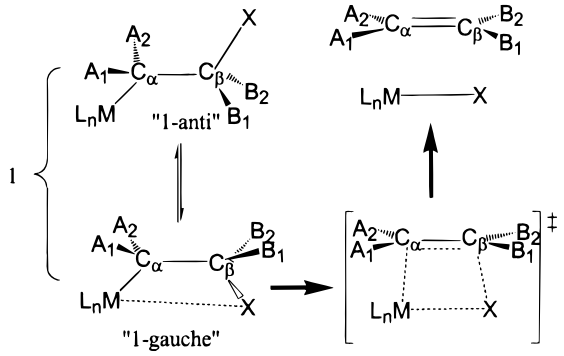
Fuzzy logic and neural network techniques are used to classify intramolecular interactions between transition metals (M) and β-X substituents in the following structural motif ($L_nMC_\alpha(A_1)(A_2)$-$C_\beta(B_1)(B_2)X$). These interactions are relevant to the direct polymerization of functionalized olefins by Ziegler−Natta (ZN) catalysis. The efficiency and effectiveness of different soft computing techniques are compared. These methods give not only encouraging results with respect to general data mining issues but also insight into the factors that effect interactions between transition metals and β-X substituents.

INTRODUCTION

Catalysis is of tremendous importance in industrial and biological applications. From in vivo catalytic oxidation of toxic substances by the enzyme cytochrome P-450[1] to the industrial polymerization of ethylene and propylene by Ziegler−Natta (ZN) catalysis,[2,3] transition metals play a pivotal role in catalysis. Global production of polyolefins in 1997 was 75 million tons, two-thirds of which is polyethylene.[2] A major recent focus of catalysis research has been to obtain a molecular-level understanding of how chemical properties affect catalyst activity and selectivity.[4] Despite a massive research effort,[5] experimental and computational, major challenges exist in ZN catalysis, primarily the inability of current catalysts to permit direct polymerization of functionalized olefins.[3] The identification of novel chemical motifs that hinder β-X elimination, Scheme 1, would constitute a major step forward in ZN catalysis.

Important hurdles exist in the routine application of computational modeling for the rational design of transition metal catalysts. For example, the relationship between catalyst properties ($Prop_{cat}$) and catalyst performance ($Perf_{cat}$) is often unclear and decidedly nonlinear. Changing one component of a catalyst system may, for example, increase activity but diminish selectivity. Many catalysts are complex, with a variety of components, chemical (transition metal, ligands, solvent, etc.) and physical (temperature, pressure, catalyst loading, etc.), that effect the outcome. Identification of the optimum component mixture to maximize catalyst activity and selectivity is thus a multidimensional optimization, which can thus result in numerous minima and complicated hypersurfaces. Finally, in terms of maximizing the utility of computation, it is desirable to isolate predictions that can be stated as succinctly as possible using natural linguistic rules (e.g., if component A is high then activity is high).

Scheme 1



In this context, soft computing (SC)—neural networks (NNs[6]), genetic algorithms (GAs[7]), and fuzzy set theory (FST[8])—seems an ideal option for catalyst modeling, an option that has been little explored.[9] SC seeks, albeit very simply, to mimic natural processes with software or hardware. SC research comprises a massive effort with growing interest for chemical applications. The most prevalent chemical applications of SC are in analytical (e.g., NNs for spectral peak identification; FST for classification of spectroscopic data) and biological (e.g., GAs that predict protein structure from amino acid sequence) chemistry.[6−8] SC techniques have been widely used in industrial chemistry for process control.[10]

Neural Networks. A neural network is software or hardware influenced by the operation of neurons in the brain. NNs have been intensively studied for their pattern recognition ability.[6] In a catalyst design scenario one might seek to train an NN to associate $Prop_{cat}$ with $Perf_{cat}$. Additionally, NNs could also be used to automatically classify $Prop_{cat}$ (and/or $Perf_{cat}$) to yield families or clusters of similar materials. Once suitably trained, NNs can then be used to assess properties that maximize, for example, activity and selectivity.

Genetic Algorithms. Genetic (or evolutionary) algorithms[7] are a programming paradigm inspired by Darwin's theory

* Corresponding author phone: (901)678-2629; fax: (901)678-3447; e-mail: tcundari@cc.memphis.edu.
# The University of Memphis.
† Babes-Bolyai University.

10.1021/ci0000023 CCC: $19.00
Published on Web 04/25/2000

of natural selection. Solutions to a problem (e.g., what structural properties minimize β-X elimination?) are treated as individuals in a population. Discrete $Prop_{cat}$ represent a "gene," the aggregate of which yields a chromosome. The GA then seeks through simple mathematical functions that mimic biological processes—crossover or mating, selection, mutation—to isolate genes ($Prop_{cat}$) that yield the fittest individual (maximize $Perf_{cat}$). Much of the utility of GAs lies in their simplicity, and as such they have been found to be powerful tools for optimization when dealing with complicated, problem hypersurfaces.

Fuzzy Set Theory. FST[8] seeks to go beyond classical binary (yes, no; true, false; 1, 0) logic. $Perf_{cat}$ would be assigned values from 0 (worst) to 1 (best) using any number of criteria—activity, selectivity, cost, environmental impact, etc. In terms of using FST to classify catalysts, one would seek to identify the strongest candidates, but perhaps more importantly those properties that serve to distinguish good (e.g., $Perf_{cat} > 0.9$) from bad (e.g., $Perf_{cat} < 0.1$) catalysts. One would then seek to synthesize new catalysts that maximize desirable $Prop_{cat}$ and minimize undesirable $Prop_{cat}$. A distinct advantage of FST is the ability to cast such analyses in natural linguistic terms such as "If the steric bulk of X is HIGH then β-X elimination is LOW".

COMPUTATIONAL METHODS

Database Searching. A search of the Cambridge Structural Database[11] is performed to identify chemical motifs such as that in Scheme 1 ($L_nMC_\alpha(A_1)(A_2)\text{-}C_\beta(B_1)(B_2)X$), **1**, in which there is a "close contact" between the transition metal M and the β substituent X ($X \neq H$). A close contact is defined when the M...X distance satisfies the relation

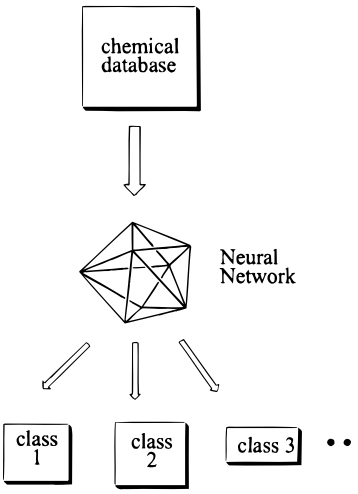
$$d(\text{M...X}) < r_{vdw}(\text{M}) + r_{vdw}(\text{X}) + \tau$$

where $r_{vdw}(M)$ and $r_{vdw}(X)$ are the van der Waals radii of the metal and β substituent atom, respectively; τ is a threshold (generally 0.5−1.0 Å) to ensure that a range of M...X interactions from weak to strong are obtained for study and analysis. For each "hit", seven metric parameters (MC, CC, CX, and MX distances; MCC and CCX angles; the absolute value of the MCCX dihedral angle) are tabulated.

Search refinements were employed to ensure the search focused on the most reliable experimental crystal structures: INSIST-ON-COORDS (to search only systems in which fractional atomic coordinates are deposited, ≈88% of the database), INSIST-NO-DISORDER (to search systems with no crystallographic disorder, ≈85% of the database), INSIST-PERFECT-MATCH (to search entries with completely matched chemical and crystallographic connectivities, ≈75% of master database), INSIST-ERROR-FREE (to include only entries whose published bond lengths agree with the recalculated values to within 0.05 Å, ≈98% of the database), and INSIST-NO-POLYMERS (to exclude entries with polymeric bonds in the crystal connectivity, ≈98% of the master database).

Statistical outliers, i.e. those structures with metrics lying more than ±3 standard deviations from the sample mean for each of the first six metrics, were removed from analysis, as were duplicate structures. This step eliminates structures with unresolved disorder and other errors. These database search yields a seven-element (or seven-characteristic) vector for each data point or object (476 in total).

Scheme 2



Neural Networks. Two different neural network types will be evaluated—competitive learning (CL) neural networks and self-organizing maps (SOMs).[6] The seven metric characteristics for each motif **1** in which there is a "close contact" between the transition metal and the β substituent X are inputs to the NNs. The outputs are the classes to which each object belongs. Classes result from mathematical operations occurring in the hidden layer neurons of the CL or SOM and define chemically similar groupings, for example, complexes with similar bond lengths and bond angles, Scheme 2. An advantage of NNs (and SC methods in general) is their ability to cope with large databases, large both in terms of the number of examples and the dimensionality of the data. Analysis of the different classes obtained from the NN simulations will be used to assess whether certain chemical motifs (ligands, metals, substituents) inhibit or promote the strength of the interaction between M and X.

For CL and SOM networks, a variety of operational parameters (e.g., learning rate, definition of neighborhood, etc.) are possible and were studied. These tests indicated that the results of the SOM and CL simulations are not inordinately sensitive to these computational details. Analysis of the effect of the number of neurons and hidden layer topology is discussed in the text. Neural network simulations employ the MATLAB program.[12]

Fuzzy Hierarchical Cross-Classification. The fuzzy hierarchical classification (FHC) or cross-classification (FHCC) algorithms developed by Pop, Sârbu, Dumitrescu, and their collaborators are applied to the classification of **1**.[13c] The FHC and FHCC algorithms are conceptually similar to that outlined above for CL and SOM neural networks, i.e., objects **1** with similar metric characteristics are assigned to different classes. However, for the FHC and FHCC techniques partitioning is accomplished using fuzzy set theory. In this way, the membership degree of each individual object **1** for a particular class can be calculated. The object then belongs to that class in which it has the highest membership degree. For typical NN classification approaches an object is assigned to a single class. Hence, the FHC and FHCC algorithms provide more information than typical CL or SOM simulations as to the relationship between various chemical classes.

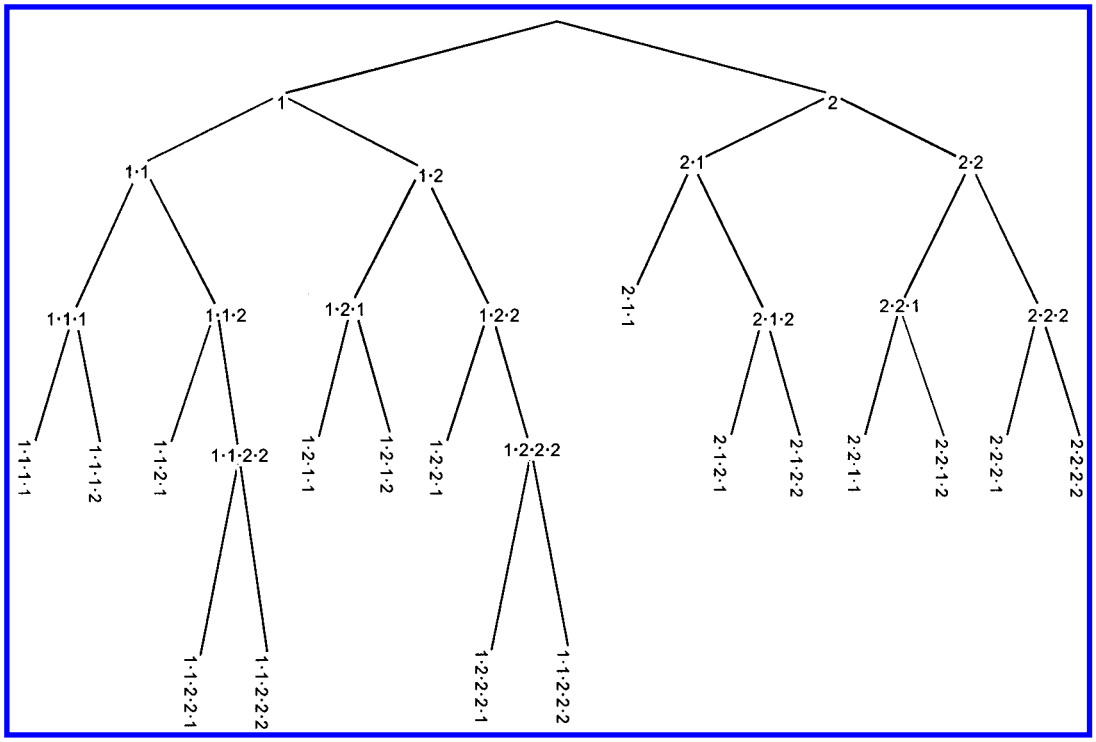


Figure 1. The binary tree obtained from fuzzy hierarchical clustering of objects. Each terminus represents an object cluster.

This approach and subsequent refinements have been validated for a range of chemical and nonchemical applications.[13] A major strength of the FHC(C) approach in relation to other SC methods is that it provides not only a partitioning of the individual objects but also the characteristics (i.e., the bond lengths and angles in **1**) used to describe them. Additionally, the FHCC algorithm can be used to simultaneously classify objects *and* characteristics. To quote,[13a] "In this way it is possible to identify which properties (variables) are responsible for the observed similarities or dissimilarities between different elements".

Several features of the FHC and FHCC algorithms are especially notable and are expected to make them powerful analysis tools. First, by decomposing the data into a "tree" format (see Figure 1 for an example) the relationships between clusters is apparent. The closer the branching point between two clusters the more similar they are expected to. Second, analysis of the simulations can yield information as to what characteristic or property is responsible for the branching of clusters. Third, the level at which a branching point is observed can yield insight into the importance of a structural characteristic in distinguishing one cluster from another. Returning to our tree, Figure 1, the property that discriminates cluster "1" from "2" is expected to be more crucial for classification than that which causes the division between the "1.1.1.1" and "1.1.1.2" clusters.

The FHC and FHCC algorithms have been implemented into a program[13] and are used for classification of the CSD database subsets studied in this paper.

RESULTS AND DISCUSSION

1. Effect of Normalization. The first issue addressed is that of data normalization. This need arises from the fact that bond lengths are on the order of 10^0 Å,while bond and dihedral angles are typically in the vicinity of 10^2 degrees. Two schemes were investigated—linear scaling (on a range from zero to one, eq 1) and Z-scoring (eq 2), where X is the unnormalized metric, X_{min} and X_{max}

$$X_{scaled} = (X - X_{min})/(X_{max} - X_{min}) \quad (1)$$

$$X_{scored} = (X - X_{mean})/\sigma \quad (2)$$

are the minimum and maximum values for a particular metric, X_{mean} is the average value for a metric, and σ is the sample standard deviation for a metric.

As found by Pop et al. in their study of the chemical and physical properties of the elements,[13] the normalization of characteristics results in the isolation of far fewer (one-third to one-quarter) different classes whether the technique employed is FHC, FHCC, or a CL neural network. Hence, to maximize the chemical information content resulting from the database mining, subsequent simulations described below are based on the use of unnormalized bond lengths and angles as originally obtained from the CSD.

2. Fuzzy Classification. i. Fuzzy Hierarchical Clustering of Objects. Use of the FHC algorithm results in a partitioning of the 476 objects (unnormalized) of the database subset into 17 distinct classes for motif **1**. The binary tree is shown in Figure 1.

The calculated centroids (weighted by the membership function of the objects belonging to a particular cluster) are organized in Table 1. The overall average for each characteristic is also listed in Table 1 as is the number of members for each cluster.

Inspection of Table 1 yields one point of immediate interest. It is evident that the first branch in the binary tree, i.e., the one that distinguishes the "1" from the "2" clusters, is mainly effected by differences in M- - -X distances. The "1" clusters have MX centroids ranging from 3.35 Å (cluster 1.2.2.2.2) to 3.80 Å (cluster 1.1.2.1). The "2" clusters have MX centroids that are considerably higher, 4.43 Å (2.2.2.1

Table 1. Centroids for FHCC Clusters[a]

cluster	MC	CC	CX	MCC	CCX	|MCCX|	MX	#
1.1.1.1	2.10	1.52	1.49	126	111	66	3.70	30
1.1.1.2	2.09	1.54	1.46	120	111	65	3.56	48
1.1.2.1	2.12	1.53	1.51	126	111	72	3.80	20
1.1.2.2.1	2.10	1.54	1.49	119	111	73	3.66	22
1.1.2.2.2	2.07	1.53	1.51	121	112	79	3.75	20
1.2.1.1	2.08	1.50	1.45	126	113	58	3.60	35
1.2.1.2	2.11	1.53	1.47	120	112	59	3.51	49
1.2.2.1	2.13	1.53	1.48	126	112	50	3.61	23
1.2.2.2.1	2.06	1.55	1.43	120	111	51	3.39	23
1.2.2.2.2	2.07	1.53	1.45	120	111	46	3.35	19
2.1.1	2.09	1.53	1.52	121	112	171	4.48	35
2.1.2.1	2.13	1.53	1.53	123	110	163	4.51	14
2.1.2.2	2.13	1.53	1.53	118	111	164	4.48	17
2.2.1.1	2.10	1.51	1.50	123	110	176	4.48	27
2.2.1.2	2.11	1.53	1.52	127	109	176	4.53	26
2.2.2.1	2.07	1.50	1.49	118	115	176	4.43	32
2.2.2.2	2.10	1.53	1.49	118	109	177	4.43	36
all	2.10	1.53	1.49	122	111	105	3.93	476

[a] Bond lengths (MC, CC, CX, and MX) are reported in Angstrom units. Bond angles (MCC and CCX) and dihedral angles (|MCCX|) are reported in deg.

and 2.2.2.2) to 4.53 Å (2.2.1.2). The overall average M- - -X distance for the entire set of 476 data points is 3.93 Å, and hence this average value is intermediate for the "1" and "2" clusters obtained from the FHC algorithm. Hence, the first branch point separates the database for motif **1** into classes ("1" clusters) ranging from very short to slightly shorter than the average M- - -X distance, while "2" clusters have M- - -X distances significantly larger than average. These short M- - -X distances (and thus small MCCX dihedrals) are suggestive of some chemical interaction between the β substituent and the metal for objects that belong to the "1" clusters.

Closer inspection of the cluster centroids in Table 1 also suggests that differentiation at the fourth-level branching points (x.y.z → x.y.z.1 and x.y.z.2) is a result of differences in bond angles, primarily MCC, but also CCX for 2.2.2 → 2.2.2.1 and 2.2.2.2.

Differentiation at the final level (i.e., between 1.1.2.2.1 and 1.1.2.2.2 as well as 1.2.2.2.1 and 1.2.2.2.2) is mainly effected by the absolute value of the MCCX dihedral. The 1.1.2.2.1. and 1.2.2.2.1 clusters are slightly closer to archetypal gauche dihedrals (|MCCX| = 60°) than 1.1.2.2.2. and 1.2.2.2.2, respectively, $(1.1.2.2.1)_{cent} = 73°$, $(1.1.2.2.2)_{cent} = 79°$, $(1.2.2.2.1)_{cent} = 51°$, and $(1.2.2.2.2)_{cent} = 46°$. As is evident from the |MCCX| centroid values in Table 1, the 1.1.2.2.x clusters are distinguished from the 1.2.2.2.x clusters by the former having |MCCX| greater than gauche, while the latter, on average, have |MCCX| below a standard gauche angle of 60°. This results in the 1.2.2.2.x clusters having appreciably shorter M- - -X separations than 1.1.2.2.x clusters: $(1.1.2.2.x)_{cent} \approx 3.7$ Å, $(1.2.2.2.x)_{cent} \approx 3.4$ Å. This is expected from simple geometric considerations given the similarity in MC, CC, MCC, and CCX among motifs **1**. Indeed, a reasonable correlation exists between M- - -X distance and |MCCX|: Spearman rank correlation coefficient,[14] ρ_s, for the cluster centroids = 0.86. Calculation of ρ_s for the other cluster centroids yields only one other significant correlation ($\rho_s = 0.91$), that between $|MCCX|_{cent}$ and $(CX)_{cent}$. This interesting correlation will be discussed below.

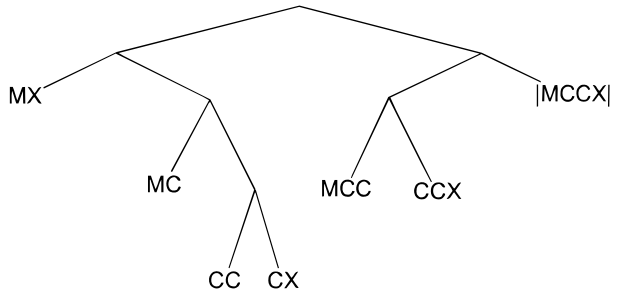


Figure 2. The binary tree obtained from fuzzy hierarchical clustering of characteristics. Each terminus represents a characteristic cluster.

Thus, considerable detail can be obtained from fuzzy hierarchical clustering of objects alone. We now turn our attention to fuzzy clustering of the characteristic metrics and subsequently cross-clustering of the objects and characteristics simultaneously.

ii. Fuzzy Hierarchical Clustering of Characteristics. Although as seen above the calculation of cluster centroids gives insight into the chemical nature of the partitioning as well as the properties primarily responsible for these partitioning. However, it is desirable to address this issue in a more systematic fashion. Thus, the FHC algorithm is applied to classification of the characteristics. Conceptually, this process can be thought of applying the same FHC methodology discussed in the previous section to the transpose of database **1**. This process results in the fuzzy binary tree shown in Figure 2.

It is observed from Figure 2 that the bond lengths occupy branches on the left side of the binary tree, while the bond angles occupy those to the right. This is likely due to the difference in magnitudes of bond lengths (measured in Angstrom units) and bond angles (measured in degrees). At the highest level in the binary tree are both |MCCX| and M- - -X, which is not surprising in light of the close correlation between these two metrics as discussed above. The next terminal clusters are the MC bond length and the MCC and CCX bond angles. The latter two are derived from the same branching point. Finally, at the lowest level are the CC and CX bond lengths. The branching profile in Figure 2 thus suggests that MX (and hence |MCCX|) is the primary discriminant between the fuzzy clusters, while CC and CX are least important in differentiating one cluster from another. The MCC and MCX angles as well as the MC distance are of roughly equal importance as differentiating characteristics. Hence, the fuzzy hierarchical characteristic clustering supports the conclusions drawn above from an analysis of the cluster centroids. The main advantage of application of the FHC algorithm to the characteristics is that it is more transparent than analyses based on centroids obtained from object clustering.

iii. Fuzzy Hierarchical Cross-clustering of Objects and Characteristics. We now seek to effect a simultaneous fuzzy partitioning of the objects and characteristics, hence the classification is referred to as fuzzy hierarchical cross-clustering (FHCC). At each node in the hierarchy, two fuzzy partitions are calculated—a fuzzy partition P for the objects and a fuzzy partition Q for the characteristics. At each node, the process is iterative and can be loosely described as follows.

Table 2. Intersection Matrix for Versions "a" and "c" of the FHCC Algorithms[a]

		1.1-a M- - -X	1.2.1-a MC	1.2.2.1-a CC	1.2.2.2-a CX	2.1.1-a MCC	2.1.2-a CCX	2.2-a |MCCX|
M- - -X	1.1-c	121	13	0	0	0	0	0
MC	1.2.1-c	19	65	12	4	0	0	0
CC	1.2.2.1-c	0	6	8	15	0	0	0
CX	1.2.2.2-c	0	0	3	23	0	0	0
MCC	2.1.1-c	0	0	0	0	7	12	37
CCX	2.1.2-c	0	0	0	0	10	14	25
|MMCX|	2.2-c	0	0	0	0	18	5	59

[a] The three versions of the fuzzy hierarchical cross-clustering (FHCC) algorithms are described in the text. Version "a" and "b" yield identical classifications.

Step 1: Based on the original data, produce a fuzzy partition P_1.

Step 2: Compute modified objects that take into account the information from Step 1.

Step 3: Based on the new objects, produce a fuzzy partition Q_1.

Step 4: Compute modified characteristics utilizing the information from Step 3.

Step 5. Based on the new characteristics, produce a new partition P_2.

Step 6: Repeat Steps 2−5 until P_i and P_{i+1} are close enough.

In version "a", the partitioning is as described above: at each node of the classification hierarchy, at Step 1, the partition P_1 is produced based on the original data. In version "b", at each node of the classification hierarchy, except for the top node, at Step 1, the partition P_1 is produced based upon the modified objects developed at Step 3 at the classification of the parent node. In version "c", at each node of the classification hierarchy, at Step 5, the partition P_2 is produced based upon the original data (the same as at Step 1) and not upon the modified data developed at Step 4. Applying the three different versions of the FHCC algorithm to the database subset **1** results in identical results for versions "a" and "b" with "c" being different from the first two cross-clustering partitionings. The latter is different with respect to the partitioning of the objects although it too yields a characteristic clustering identical to "a" and "b". For all the FHCC results there are seven terminal clusters with the same structure as that shown in Figure 2, which was obtained from clustering of the characteristics only.

It is interesting that approaches "a" and "c" of the FHCC algorithm yield identical characteristic clustering but different object clustering. To assess the similarity of the object clusters, the intersection matrix was calculated. This matrix shows the number of points that are common to pairs of clusters obtained from versions "a" (and thus "b") and "c" of the FHCC algorithm. The intersection matrix is given in Table 2. One point that is evident from Table 2 is that regardless of the specifics of the FHCC clustering there is no "migration" of data points between "1" and "2" clusters. This is likely due to the large differences in the magnitudes of the bond lengths ("1" clusters) and bond angles ("2" clusters). There is, however, more extensive mixing within the respective "1" and "2" clusters. Approximately 75% (218 of 289) of the "1" clusters belong to the same class whether FHCC-a/b or FHCC-c is utilized. Migration is more considerable within the "2" (angle) clusters; 43% (80/187) are partitioned to the same cluster by FHCC-a/b or FHCC-c. Thus, the majority switch clusters depending on the specifics of the cross-clustering algorithm, suggesting that the partitioning is loose within the "2" clusters.

Table 3. Cluster Averages Obtained from Fuzzy Hierarchical Cross-Clustering

	MC	CC	CX	MCC	CCX	MCCX	MX	cluster	#
					FHCC-a				
"MX"	2.11	1.53	1.50	124	111	70	3.73	1.1	134
"MC"	2.09	1.53	1.45	122	112	58	3.53	1.2.1	100
"CC"	2.08	1.53	1.46	122	111	50	3.45	1.2.2.1	29
"CX"	2.08	1.53	1.41	120	111	44	3.34	1.2.2.2	26
"MCC"	2.13	1.54	1.53	120	107	173	4.49	2.1.1	56
"CCX"	2.10	1.51	1.49	115	113	170	4.40	2.1.2	49
"|MCCX|"	2.07	1.52	1.51	125	113	173	4.50	2.2	82
					FHCC-c				
"MX"	2.10	1.53	1.48	122	111	70	3.68	1.1	140
"MC"	2.09	1.52	1.46	123	112	58	3.57	1.2.1	84
"CC"	2.14	1.54	1.49	128	112	48	3.66	1.2.2.1	23
"CX"	2.08	1.54	1.45	119	111	48	3.36	1.2.2.2	42
"MCC"	2.10	1.53	1.51	121	112	169	4.49	2.1.1	35
"CCX"	2.13	1.53	1.53	119	110	160	4.47	2.1.2	31
"|MCCX|"	2.09	1.52	1.50	122	111	177	4.47	2.2	121

The number of classes resulting from FHCC (less or equal to 7) is limited by the number of characteristics. Average values of each metric obtained from versions "a" and "c" of the FHCC algorithm are collected in Table 3. In version c, significant correlations only exist between M- - -X and |MCCX| ($\rho_s = 0.90$), CX and |MCCX| ($\rho_s = 0.87$), which are similar to the FHC result. In version "a", correlations are found between M- - -X and |MCCX| ($\rho_s = 0.99$), CX and |MCCX| ($\rho_s = 0.92$), CC and CCX ($\rho_s = -0.93$), and CX and M- - -X ($\rho_s = 0.89$). The observation of high correlations may result from fewer number of classes. The correlation of CX and |MCCX| is not unreasonable since both of these metrics show a strong correlation with M- - -X. The correlation between CC and CCX is possible based solely on geometric considerations—the longer the CC bond length, the smaller the CCX bond angle.

3. Neural Network Classification. i. Competitive Learning (CL) Networks. For CL networks, classification of the data is dependent on the number of hidden layer neurons specified for the network (which in turn determines the maximum number of classes that can be obtained from a simulation) and the number of epochs the simulations are run. Generally, as the number of epochs increased the data became more "spread out" among the classes. Put another way, the median class size decreases. To determine an optimal number of neurons/classes, a series of simulations were begun with a very large number (100) of CL hidden neurons. For such large networks the neurons are "too far

Table 4. Centroids Obtained from Competitive Learning Neural Network Classification[a]

class	MC av	CC av	CX av	MCC av	CCX av	|MCCX| av	MX av	#
A	2.10	1.53	1.46	121	111	59	3.49	53
B	2.11	1.53	1.51	127	111	70	3.78	45
C	2.10	1.55	1.47	117	112	71	3.61	33
D	2.11	1.54	1.48	123	112	43	3.45	29
E	2.13	1.54	1.52	118	107	176	4.46	28
F	2.11	1.53	1.52	127	108	177	4.52	27
G	2.09	1.48	1.47	121	113	177	4.44	24
H	2.08	1.54	1.43	118	111	52	3.37	24
I	2.11	1.52	1.50	115	112	167	4.42	21
J	2.06	1.48	1.39	123	115	61	3.56	20
K	2.13	1.53	1.52	132	110	170	4.58	20
L	2.10	1.53	1.49	136	112	60	3.85	19
M	2.10	1.54	1.53	120	110	170	4.49	18
N	2.13	1.53	1.50	127	113	53	3.69	18
O	2.08	1.51	1.47	114	114	177	4.38	17
P	2.10	1.55	1.47	116	112	62	3.49	17
Q	2.00	1.52	1.53	120	120	174	4.49	17
R	2.07	1.51	1.51	121	111	86	3.81	17
S	2.15	1.52	1.54	119	111	154	4.48	15
T	2.06	1.54	1.43	121	110	65	3.54	14

[a] This CL Simulation was run for 10K epochs with 20 hidden neurons. Bond lengths (MC, CC, CX, MX) and bond angles (MCC, CCX, |MCCX|) are given in Å and deg, respectively. The classes are labeled from A (most members) to S (least members).

apart", and few classes are obtained even for very long (>10K epochs) simulations. The number of neurons is then incrementally reduced by 10 and a CL simulation run (each with different random initial weights) for 10K epochs. The results of this process suggest that the CL seeks to classify the data into ca. 25 distinct classes with about four or five of them being sparse (i.e., less than five members). For this reason and to maximize comparability with the FHC and FHCC results, it was settled upon 20 neurons to define the maximum number of classes for the present CL simulations as well as the SOM networks discussed in the next section.

CL networks divided the dataset into 20 object classes. The centroid (average value) of each class is listed in Table 4. The classes are labeled from A (most members) to T (least members). The median population size for a class from the CL network is 20 members. By this measure, roughly half of the classes are "major", i.e., populations greater than the median value.

Correlations among the seven metric variables were analyzed using the Spearman rank correlation coefficient (ρ_s) given the large differences in the magnitudes of bond angles and bond lengths. The largest ρ_s [14] values in magnitude are found between MCCX and MX ($\rho_s = 0.83$) and between MCCX and CX ($\rho_s = 0.80$). The first correlation suggests that the larger the dihedral angle, the larger the distance between M and X. This is expected on the basis of geometric considerations alone. Of greater interest is the second correlation, that of CX with MX. The positive sign of ρ_s suggests that larger CX yield larger MX. In terms of soft computing issues, it is satisfying that these are the same correlations obtained from analysis of the FHC and FHCC cluster centroids. Additionally, as with the fuzzy classifications the main distinction between the object clusters lies in their separation between those with large M- - -X distance (and hence large |MCCX| dihedrals) and small M- - -X distances (i.e., small |MCCX| distances), Table 4.

Table 5. Comparison of Different SOM Network Structures[a]

hidden layer	2 × 2 × 5	4 × 5	10 × 2	20 × 1
number of object classes	14	15	18	18
median class size	31	29	24.5	24
major classes[b]	7	6	9	9
ρ_s (MX-vs-CX)	0.88	0.91	0.81	0.83
ρ_s (MX-vs-MCCX)	0.84	0.85	0.85	0.83
ρ_s-major (MX-vs-CX)	0.87	0.88	0.85	0.95
ρ_s-major (MX-vs-MCCX)	0.86	0.77	0.85	0.80

[a] For each SOM, a hexagonal topology was utilized. [b] Classes with more members than the median class size.

Table 6. Centroids Obtained from Self-Organized Map Neural Network Classification[a]

class	MC av	CC av	CX av	MCC av	CCX av	|MCCX| av	MX av	#
A	2.10	1.54	1.46	118	111	62	3.51	51
B	2.08	1.54	1.44	122	111	52	3.45	35
C	2.07	1.50	1.44	124	113	60	3.58	32
D	2.13	1.54	1.52	121	108	176	4.50	29
E	2.13	1.53	1.50	124	112	43	3.50	29
F	2.06	1.50	1.49	118	117	177	4.44	28
G	2.09	1.53	1.50	134	112	61	3.83	27
H	2.10	1.52	1.50	127	109	176	4.49	26
I	2.10	1.53	1.45	118	112	56	3.43	26
J	2.10	1.52	1.49	116	109	177	4.40	25
K	2.12	1.53	1.51	128	111	70	3.80	23
L	2.11	1.54	1.48	117	112	72	3.64	22
M	2.09	1.53	1.46	121	111	69	3.62	21
N	2.07	1.51	1.50	118	116	171	4.45	20
O	2.12	1.53	1.52	133	110	172	4.58	19
P	2.10	1.53	1.53	118	111	167	4.48	19
Q	2.08	1.52	1.51	121	111	80	3.76	18
R	2.14	1.53	1.54	120	109	161	4.49	17
S	2.08	1.52	1.51	123	111	97	3.98	5

[a] The SOM network with the hidden layer of 20 × 1 topology was trained for 100K epochs. Bond lengths (MC, CC, CX, MX) and bond angles (MCC, CCX, |MCCX|) are given in Å and deg, respectively. The classes are labeled from A (most members) to S (least members).

ii. Self-Organizing Map (SOM) Networks. As with CL networks, different computational parameters were tested for SOM networks. The tests show that the number of hidden layer neurons and the topology of the hidden layer most significantly influence the number of classes obtained. Carrying out different SOM simulations in an analogous fashion to that outlined above for the CL networks indicates that 20 neurons is also near optimal.

For a SOM with 20 neurons in the hidden layer, one can build the following topologies: 20 × 1, 10 × 2, 4 × 5, or 2 × 2 × 5. It is seen that lower dimensionality in the hidden layer topology tends to result in more object classes. Simulations of 10K epochs for each structure results in 18 (9 major), 18 (9 major), 15 (6 major), and 14 (7 major) classes for 20 × 1, 10 × 2, 4 × 5, and 2 × 2 × 5 network, respectively. However, despite the different hidden layer topologies, the largest centroid ρ_s are all found between MCCX and MX and between CX and MX, as shown in Table 5. The centroids for the 19 object classes resulting from the SOM network with the hidden layer of 20 × 1 topology (trained for 100K epochs) are listed in Table 6. The classes are labeled from A (most members) to S (least members).

iii. Neural Networks in Clustering of Characteristics. Inspired by the research of Pop et al.,[13] neural networks (CL,

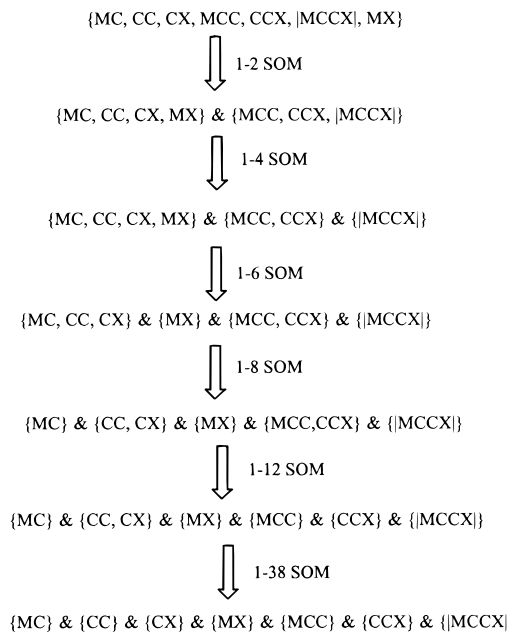


Figure 3. The characteristic clustering results obtained from a self-organizing map (1-X hidden layer topologies are utilized, and X is the number of the hidden neurons). Each parenthesis represents one cluster.

SOM) are employed to classify characteristics in addition to the object clustering just discussed. The procedure is the same as that used in clustering objects, except the data matrix (476 objects by seven characteristics) is transposed to interchange characteristics and objects.

With CL networks, only two clusters are obtained from characteristic clustering no matter how many neurons are used in the hidden layer. One cluster is bond lengths (MC, CC, CX, and MX) and the other is the cluster of angles (MCC, CCX, and |MCCX|). Since the data are not normalized, the large difference in the magnitude between the bond lengths ($\approx 10^0$ Å) and angles ($\approx 10^2$ deg) make it difficult for the CL to simultaneously differentiate bond lengths and bond angles in terms of characteristic classification. To solve this problem, it is possible to classify bond lengths and bond angles individually. If this is done, the same binary tree as for the FHCC (Figure 2) technique is obtained.

In SOM networks, the characteristic clustering results are dependent on the number of hidden neurons as shown in Figure 3. As the number of hidden neurons increases ($1 \times$ X hidden layer topologies are utilized), the characteristics are further delineated into different classes. The first level (1×2 SOM) has two classes, one is comprised of bond lengths and the other of angles. Upon increasing the number of SOM hidden neurons to 4, three classes resulted; the angle class is subdivided into {MCC, CCX} and {|MCCX|}, which suggests the importance of the latter in discriminating between clusters. The third level (1×6 SOM) has four classes—{MC, CC, CX}, {MX}, {MCC,CCX}, and {|MCCX|}—thus the MX distance is separated from the other bond lengths. This result also reminiscent of that obtained by the FHC(C) technique. Further increasing the number of hidden layer SOM neurons further discriminates the characteristics as shown in Figure 3. Upon sequentially increasing the SOM hidden layer by two neurons one finally obtains seven distinct classes, one for each characteristic, when 38 neurons are used, see Figure 3.

For characteristic classification, particularly in the present case in which there is a large difference in their magnitude, SOM networks seem to be better than CL networks for finding distinctions among characteristics. One could attempt to normalize the characteristics, but this was not done in the present case. Finally, it is found that incrementally increasing the size of the SOM hidden layer provides a technique for assessing the evolution of the characteristics as distinguishing factors. For instance, CC and CX can be differentiated only using many hidden neurons (38), which indicates they are very similar to each other. It must be stated that in terms of characteristic clustering, that the FHCC provides a more straightforward discrimination in terms of being able to perform this task with a single simulation.

SUMMARY AND CONCLUSIONS

i. Comparison of Classification Results. Several trends have emerged with respect to the SC classification schemes. The methods are compared and contrasted in this section.

Normalization tends to reduce the "structure" of the data, making it more difficult to isolate distinct classes, particularly for the fuzzy hierarchical clustering and competitive learning techniques. Relatively speaking, self-organizing maps are more capable of treating data sets with large differences in magnitude among the characteristics as, for example, in the present case between bond lengths and bond angles.

With respect to object classification, the FHC, SOM, and CL methods all indicate that the 476 objects **1** can be reduced to approximately 15−20 major object classes. These classes are distinguished from each other by the differences in their metric characteristics. Hence, while the specifics of the techniques may be different, it is encouraging that the methods are flexible and robust enough to discern the same multidimensional trends.

Classification of characteristics using FHC, SOM, and CL indicates similar ordering of characteristics in terms of their importance in discriminating the object classes **1**: MX $\approx$ |MCCX| > MC > MCC $\approx$ CCX > CC $\approx$ CX. Simultaneous clustering of objects and characteristics by the FHCC technique yields a similar ranking of the metrics. Competitive learning networks are least capable of classifying characteristics because of the unnormalized data. The algorithms employing fuzzy set theory are able to perform this task within a single simulation due to their binary tree format. For the SOMs, several simulations, each with a different number of hidden neurons (in order to control the number of resulting classes), are needed to evaluate the importance of the discriminating characteristics.

For each of the soft computing techniques utilized in this research, correlations derived from the centroids of the resulting classes are found between MX and |MCCX| (expected from simple geometric considerations) but more interestingly between MX and CX. Shorter (longer) CX bond lengths correlate with shorter (longer) MX distances. Even for FHCC where the number of classes produced is limited by the number of characteristics, correlations between MX and |MCCX| and MX and CX are found. Different SC classification schemes may thus have different operational parameters, but the "chemical" results do not seem inordinately sensitive to computational details.

No matter which method is used, |MCCX| (or M- - -X distance) is the major differentiating characteristics among

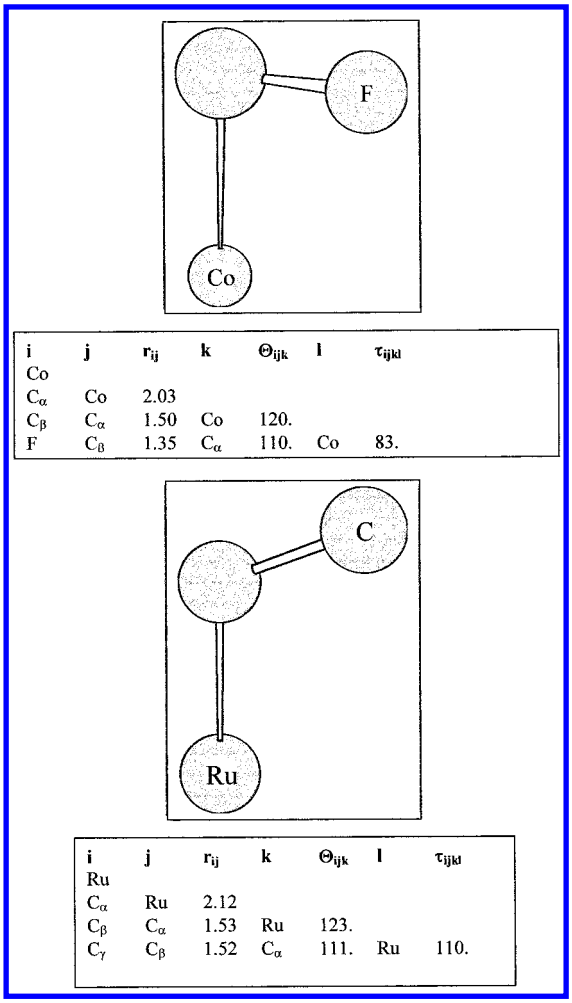
the objects **1**. To find new interesting relationships among characteristics, the 476 objects are first manually divided into two subsets based on the value of |MCCX|/M- - -X of each class center. In subset **A**, |MCCX| is below 100° (M- - -X < 4.0 Å), and in subset **B**, |MCCX| is above 100° (M- - -X > 4.0 Å). Only the object classification results from FHC, CL, and SOM are considered since they have enough object classes (around 20). Analysis of correlations among characteristics for each subset is still based on ρ_s.[14] It is found that the best correlations in subset **A** still exist between CX and MX (e.g. $\rho_s = 0.91$ in FHC method) or MCCX and M- - -X ($\rho_s = 0.78$), while in subset **B**, the correlation coefficient between MCC and M- - -X is higher than the others ($\rho_s = 0.86$). The latter suggests that enlarging the angle of MCC increases the M- - -X distance in the anti-conformation (Scheme 1). While the relationship is likely partially geometric in origin, it also identifies the inherently reasonable suggestion that a profitable way to weaken β-X interactions is to enlarge the MCC angle. This could be accomplished by (a) increasing the coordination number of the metal or (b) using bulky ligands.

Finally, it is worth noting that a limitation of this database mining study, and indeed all similar efforts, is that the objects (and hence their characteristics) are limited to available experimental data, which will exclude, for example, compounds that undergo β-X elimination. Additionally, the complexes studied may not fully span the total chemical space and may be concentrated on certain metal/substituent combinations. As such, it is of interest to augment experimental databases with data obtained from computational chemistry to fill in "gaps" and provide a more diverse database for mining.

ii. Chemical Aspects. The SC-based methods discussed here are encouraging with respect to general data mining research issues. The techniques are not only able to confirm/quantify an expected correlation (i.e., M- - -X and |MCCX|) but also to identify one that is not expected, i.e., the correlation that exists between M- - -X distance and CX bond length.

In our dataset, X is either C, N, O, or F. Fluorine (19% of the 476 objects) and carbon (79% of the 476 objects) substituents comprise the majority of data set **1**, so our analysis will focus on these. For data objects X = C and F the average M- - -X distances are 4.03 and 3.57 Å, respectively. The difference of ≈0.5 Å may be partially a result of geometric factors due to differences in the size and electronic properties of X. The C−X bond length will depend primarily on the covalent radius (r_x) and electronegativity (χ_X) of X. Since r_C (0.77 Å) > r_F (0.72 Å) and χF (3.98) > xC (2.55), one expects C−C to be longer than C−F. The average C−X bond lengths in data set **1** are 1.52 Å (X = C) and 1.35 Å (X = F). Since the C−X bond is not collinear with respect to the MX vector, this C−X bond length difference of 0.17 Å translates appropriately into an ≈0.1 Å difference on M- - -X.

Another metric factor that will influence the M- - -X distance is the M−C bond length; the average CC bond length and MCC and CCX bond angles are nearly identical for X = C and X = F, see Figure 4. However, the average MC distance for carbon-substituted systems **1** is 2.12 Å, substantially shorter than the average MC of 2.03 Å for the fluorine-substituted cases. As before this translates into a roughly 0.1 Å difference in the M- - -X distance assuming all other metric parameters remaining unchanged. Analysis of the carbon- and fluorine-substituted objects indicates the latter are more heavily dominated by lighter, later transition metals than the former, thus yielding the shorter average MC bond length for the MCCF objects.



i	j	r_{ij}	k	Θ_{ijk}	l	τ_{ijkl}
Co						
C_α	Co	2.03				
C_β	C_α	1.50	Co	120.		
F	C_β	1.35	C_α	110.	Co	83.

i	j	r_{ij}	k	Θ_{ijk}	l	τ_{ijkl}
Ru						
C_α	Ru	2.12				
C_β	C_α	1.53	Ru	123.		
C_γ	C_β	1.52	C_α	111.	Ru	110.

Figure 4. Newman projections of prototypical MCCX objects for X = C and F. M is Co (or Ru) and X is F (or C) in the upper (or lower) figure. These are viewed along the $C_\alpha-C_\beta$ bond axis. The average metrics for the MCCX and MCCF clusters are also given: r_{ij} is the distance (Å) between atom i and atom j, Θ_{ijk} is the angle (deg) formed by i−j-k, and τ_{ijkl} is the dihedral angle (deg) between i−j and k−l.

The major geometric characteristic distinguishing MCCC from MCCF objects is the dihedral angle formed by the MC and CX bonds. Using the average values for each metric, the prototypes in Figure 4 are constructed. The average |MMCX| are 110° (X = C) and 83° (X = F). Using average bond lengths and angles for the MCCC and MCCF objects, this change in dihedral corresponds to approximately 0.3 Å (or 60%) of the difference in the M- - -X distance for fluorine- and carbon-substituted systems. Hence, the major origin of the shorter M- - -F than M- - -C distances lies not in inherent size differences of the substituents but rather their preferences for different dihedral angles. While the MCCC objects are closer to gauche, the MCCF objects are closer to perpendicular. Newman projections are shown in Figure 4.

At the separation distances being considered (≈ 3.6−4.0 Å) two electronic effects seem plausible as the cause of an

inherently stronger M- - -F than M- - -C interaction: electrostatic or van der Waals effects.[15] We prefer the former as the primary determinant for the systems, since we would expect the van der Waals interaction to track with the polarizability of the substituent atoms involved. The atomic polarizability of carbon is 1.76 Å^3, while that for fluorine is one-third this value (0.56 Å^3).[16] Other research in our group on atomic contributions to polarizability also support the lesser contribution made by a fluorine substituent in comparison to carbon-based substituents.[17] Additionally, the MCCF objects are more heavily weighted toward lighter, later metals. Cobalt is the prototypical metal using the average series and triad values for the MCCF dataset; its polarizability is 7.5 Å^3. For the MCCC objects, the prototypical metal is Ru (polarizability = 9.6 Å^3). Hence, the prototypes CoCCF and RuCCC immediately suggests electrostatics as the origin for the observed structural differences as the former will involve "hard" (ionic/electrostatic) interactions between the fluorine and a light TM and the latter would be characterized as a "soft" (covalent) interaction between carbon and a heavy TM. The data mining effort suggests the former to be substantially stronger. Hence, efforts to reduce electrostatic interactions between transition metals and β-X substituents is likely to be an effective criterion for the design of improved Ziegler−Natta catalysts for direct polymerization of substituted olefins.

ACKNOWLEDGMENT

The authors acknowledge the support of the National Science Foundation through grant CHE-9614346**.** Calculations at The University of Memphis employed the Computational Chemistry Resource at The University of Memphis, funded by a grant (CHE-9708517) from the Chemical Research and Instrumentation Facilities program of the NSF. The authors further acknowledge the NSF for their support through the Academic Research Infrastructure program (CHE-9602656).

Supporting Information Available: Database of metal β-substituent "close contacts" (Table S-1). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) *Cytochrome P-450: Structure, Mechanism, and Biochemistry*; Ortiz de Montellano, P. A., Ed.; Plenum: New York, 1995.

(2) Bennett, A. M. A. B. Novel, Highly Active Iron and Cobalt Catalysts for Olefin Polymerization. *CHEMTECH* **1999**, *29*, 24−28.

(3) Various approaches (polymerization of protected olefins followed by deprotection, functionalization of the poly(olefin) after polymerization, etc.) have been pursued to employ ZN catalysis for the synthesis of functionalized poly(olefins). For a discussion of ZN catalysis of functionalized poly(olefins), see: Chung, T. C. Functional Polyolefins via Borane Monomers. *CHEMTECH* **1991**, *21*, 496−499. See, also: Kesti, R. M.; Coates, G. W.; Waymouth, R. M. Homogeneous Ziegler−Natta Polymerization of Functionalized Monomers Catalyzed by Cationic Group IV Metallocenes. *J. Am. Chem. Soc.* **1992**, *114*, 9679−9680. Yang, X.; Jia, L.; Marks, T. J. C−C activation at Electrophilic d^0/f^n Centers. Facile, Regioselective β-alkyl Shift-based Ring-opening Polymerization Reactions of Methylenecyclobutane. *J. Am. Chem. Soc.* **1993**, *115*, 3392−3393.

(4) Cundari, T. R.; Deng, J.; Fu, W.; Klinckman, T. R.; Yoshikawa, A. Molecular Modeling of Catalysts and Catalytic Reactions. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 941−948, and reference therein.

(5) (a) An informative overview, with an interesting historical perspective, of Ziegler−Natta chemistry can be found in the following: Yamamoto, A. *Organotransition Metal Chemistry*; Wiley: New York, 1986; pp 305−322. (b) Reviews of constrained-geometry Ziegler−Natta catalysts: Brintzinger, H. H.; Fischer, D.; Mulhaupt, R.; Rieger, B.; Waymouth, R. M. Stereospecific Olefin Polymerization with Chiral Metallocene Catalysts. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 1143−1170. McKnight, A. L.; Waymouth, R. M. Group 4 Ansa-cyclopentadienyl-amido Catalysts for Olefin Polymerization. *Chem. Rev.* **1998**, *98*, 2587−2598. See, also: Schaverian, C. J.; Schut, P.; Skiff, W. M.; Resconi, L.; Barbassa, E.; Balboni, D.; Dubitsky, Y. A.; Orpen, A. G.; Mercandelli, P.; Moret, M.; Sironi, A. A New Class of Chiral Bridged Metallocene: Synthesis, Structure, and Olefin (Co)polymerization Behavior of rac- and meso-1,2-CH_2CH_2 4-(7-Me-indenyl)$_2$-$ZrCl_2$. *J. Am. Chem. Soc.* **1998**, *120*, 9945−9946. Lanza, G.; Fragala, I.; Marks, T. J. Highly Electrophilic Olefin Polymerization Catalysts. Counteranion and Solvent Effects on Constrained Geometry Catalyst Ion Pair Structure and Reactivity. *J. Am. Chem. Soc.* **1998**, *120*, 8257−8258. Deck, P. A.; Beswick, C. L.; Marks, T. J. Highly Electrophilic Olefin Polymerization Catalysts. Quantitative Reaction Coordinates for Fluoroarylborane/alumoxane Methide Abstraction and Ion-pair Reorganization in Group 4 Metallocene and "Constrained Geometry" Catalysts. *J. Am. Chem. Soc.* **1998**, *120*, 1772−1784. Sun, Y.; Spence, R. E. H.; Piers, W. E.; Parvez, M.; Yap, G. P. A. Intramolecular Ion-ion Interactions in Zwitterionic Metallocene Olefin Polymerization Catalysts Derived from "Tucked-in" Catalyst Precursors and the Highly Electrophilic Boranes $XB(C_6F_5)_2$ (X = H, C_6F_5). *J. Am. Chem. Soc.* **1997**, *119*, 5132−5143. Leclerc, M. K.; Brintzinger, H. H. Zr-alkyl Isomerization in Ansa-zirconocene-catalyzed Olefin Polymerizations. Contributions to Stereoerror Formation and Chain Termination. *J. Am. Chem. Soc.* **1996**, *118*, 9024−9032. Fu, P. F.; Marks, T. J. Silanes as Chain Transfer Agents in Metallocene-mediated Olefin Polymerization. Facile in situ Catalytic Synthesis of Silyl-terminated Polyolefins. *J. Am. Chem. Soc.* **1995**, *117*, 10747−10748. Yang, X.; Stern, C. L.; Marks, T. J. Cationic Zirconocene Olefin Polymerization Catalysts based on the Organo-Lewis Acid Tris(pentafluorophenyl)borane. A Synthetic, Structural, Solution Dynamic, and Polymerization Catalytic Study. *J. Am. Chem. Soc.* **1994**, *116*, 10015−10031. Deck, P. A.; Marks, T. J. Cationic Metallocene Olefin Polymerization Catalysts. Thermodynamic and Kinetic Parameters for Ion Pair Formation, Dissociation, and Reorganization. *J. Am. Chem. Soc.* **1994**, *116*, 6128−6129. Sishta, C.; Hathorn, R. M.; Marks, T. J. Group 4 Metallocene-alumoxane Olefin Polymerization Catalysts. CPMAS NMR Spectroscopic Observation of "Cation-like" Zirconocene Alkyls. *J. Am. Chem. Soc.* **1992**, *114*, 1112−1114. Resconi, L.; Piemontesi, F.; Francesconi, G.; Abis, L.; Fiorani, T. Olefin Polymerization at Bis(pentamethylcyclopentadienyl)zirconium and -hafnium Centers: Chain-transfer Mechanisms. *J. Am. Chem. Soc.* **1992**, *114*, 1025−1032. Yang, X.; Stern, C. L.; Marks, T. J. "Cation-like" Homogeneous Olefin Polymerization Catalysts based upon Zirconocene Alkyls and Tris(pentafluorophenyl)borane. *J. Am. Chem. Soc.* **1991**, *113*, 3623−3625. (c) Just a few of the numerous computational studies are given here and the references therein. Boero, M.; Parrinello, M.; Terakura, K. First Principles Molecular Dynamics Study of Ziegler−Natta Heterogeneous Catalysis. *J. Am. Chem. Soc.* **1998**, *120*, 2746−2752. Bernardi, F.; Bottoni, A.; Miscione, G. P. A Theoretical Study of Homogeneous Ziegler−Natta Catalysis. *Organometallics* **1998**, *17*, 16−24. Puhakka, E.; Pakkanen, T. T.; Pakkanen, T. A.; Iiskole, E. *J. Organomet. Chem.* **1996**, *511*, 19. Cavallo, L.; Guerra, G. A Density Functional and Molecular Mechanics Study Of β-Hydrogen Transfer in Homogeneous Ziegler−Natta Catalysis. *Macromolecules* **1996**, *29*, 2729−2737. Yoshida, T.; Koga, N.; Morokuma, K. Ab Initio Theoretical Study on Ethylene Polymerization with Homogeneous Silylene-Bridged Group 4 Metallocene Catalysts. Ethylene Insertion and -Elimination. *Organometallics* **1995**, *14*, 746−758. Yoshida, T.; Koga, N.; Morokuma, K. A Combined ab initio MO-MM Study on Isotacticity Control in Propylene Polymerization with Silylene-Bridged Group 4 Metallocenes. C_2 Symmetrical and Asymmetrical Catalysts. *Organometallics* **1996**, *15*, 766−777. Castonguay, L. A.; Rappe, A. K. Ziegler−Natta Catalysis. A Theoretical Study of the Isotactic Polymerization of Propylene. *J. Am. Chem. Soc.* **1992**, *114*, 5832−5842. Hart, J. R.; Rappe, A. K. Predicted Structure Selectivity Trends: Propylene Polymerization with Substituted rac-(1,2-ethylenebis(η^5-Indenyl))-Zirconium(IV) Catalysts. *J. Am. Chem. Soc.* **1993**, *115*, 6159−6164. Woo, T. K.; Fan, L.; Ziegler, T. A Density Functional Study of Chain Growing and Chain Terminating Steps in Olefin Polymerization by Metallocene and Constrained Geometry Catalysts. *Organometallics* **1994**, *13*, 2252−2261. Fan, L.; Harrison, D.; Woo, T. K.; Ziegler, T. A Density Functional Study of Ethylene Insertion into the M-CH_3) Bond of the Constrained Geometry Catalysts $((SiH_2-C_5)H_4-NH)MCH_3)^+$ (M = Ti, Zr, Hf). *Organometallics* **1995**, *14*, 2018−2026. Kawamura-Kuribayashi, H.; Koga, N.; Morokuma, K. An ab initio MO and MM Study of Homogeneous Olefin Polymerization with Silylene-bridged Zirconocene Catalyst and its Regio- and Stereoselectivity. *J. Am. Chem. Soc.* **1992**, *112*, 2359, 8687−8694.

(6) (a) Zupan, J.; Gasteiger, J. *Neural Networks for Chemists*; V. C. H.: Weinheim, 1993. (b) Sumpter, B. G.; Getino, C.; Noid, D. W. Theory and Applications of Neural Computing in Chemical Science. *Annu. Rev. Phys. Chem.* **1994**, *45*, 439−481.

(7) Judson, R. Genetic Algorithms and Their Use in Chemistry. In *Reviews in Computational Chemistry*; Boyd, D. B., Lipkowitz, K. B., Eds.; 1997; Vol. 10, pp 1−73.

(8) *Fuzzy Logic in Chemistry*; Rouvray, D. H., Ed.; Academic Press: New York, 1997.

(9) (a) Kito, S.; Hattori, T.; Murakami, Y. Estimation of Performance by Neural Network − Product Distribution in Oxidative Dehydrogenation of Ethylbenzene. *Appl. Catal. A: General* **1994**, *114*, L173−L178. (b) Hattori, T.; Kito, S. Neural Network as a Tool for Catalyst Development. *Catal. Today* **1995**, *23*, 347−355.

(10) (a) Tendulkar, S. B.; Tambe, S. S.; Chandra, I.; Rao, P. V.; Naik, R. V.; Kulkarni, D. B. Hydroxylation of Phenol to Dihydroxybenzenes: Development of Artificial Neural-Network-Based Process Identification and Model Predictive Control Strategies for a Pilot Plant Scale Reactor. *Ind. Eng. Chem. Res.* **1998**, *37*, 2081−2085. (b) Whaley, A. K.; Bode, C. A.; Ghosh, J.; Eldridge, R. B. HETP and Pressure Drop Prediction for Structured Packing Distillation Columns Using a Neural Network Model. *Ind. Eng. Chem. Res.* **1999**, *38*, 1736−1739. (c) Tayal, M. C.; Fu, Y.; Diwekar, U. M. Optimal Design of Heat Exchangers: A Genetic Algorithm Framework. *Ind. Eng. Chem. Res.* **1999**, *38*, 456−467. (d) Wang, C.; Quan, H.; Xu, X. Optimal Design of Multiproduct Batch Chemical Process Using Genetic Algorithms. *Ind. Eng. Chem. Res.* **1996**, *35*, 3560−3566. (e) Adroer, M.; Alsina, A.; Aumatell, J.; Poch, M. Wastewater Neutralization Control Based on Fuzzy Logic: Experimental Results. *Ind. Eng. Chem. Res.* **1999**, *38*, 2709−2719. (f) Wang, F. S.; Jing, C. H.; Tsao, G. T. Fuzzy-Decision-Making Problems of Fuel Ethanol Production Using a Genetically Engineered Yeast. *Ind. Eng. Chem. Res.* **1998**, *37*, 3434−3443. (g) Garrido, R.; M Adroer, R.; Poch, M. Wastewater Neutralization Control Based in Fuzzy Logic: Simulation Results. *Ind. Eng. Chem. Res.* **1997**, *36*, 1665−1674.

(11) Allen, F. H.; Kennard, O. 3D Search and Research using the Cambridge Structural Database. *Chem. Design Autom. News*, **1993**, *8*, 31−37.

(12) Neural Network Toolbox, version 3, 5th printing; MATLAB, The MathWorks, Inc.: Natick, MA, 1998.

(13) (a) Pop, H. F.; Sârbu, C.; Horowitz, O.; Dumitrescu, D. A Fuzzy Classification of the Elements. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 465−482. (b) Sârbu, C.; Horowitz, O.; Pop, H. F. A Fuzzy Cross-Classification of the Chemical Elements Based on Their Physical, Chemical and Structural Features. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1098−1108. (c) Pop, H. F.; Sârbu, C. The Fuzzy Hierarchical Cross-Clustering Algorithm. Improvements and Comparative Study. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 510−516. (d) Pop, H. F.; Sârbu, C. A New Fuzzy Regression Algorithm. *Anal. Chem.* **1996**, *68*, 771−778. (e) Pop, H. F.; Dumitrescu, D.; Sârbu, C. A Study of Roman Pottery (terra sigillata) Using Hierarchical Fuzzy Clustering. *Anal. Chim. Acta* **1995**, *310*, 269−279. (f) Pop, H. F.; Dumitrescu, D.; Sârbu, C. Fuzzy Hierarchical Cross-Classification of Greek Muds. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 851−857. (g) Dumitrescu, D.; Sârbu, C.; Pop, H. F. A Fuzzy Divisive Hierarchical Clustering Algorithm for the Optimal Choice of Sets of Solvent Systems. *Anal. Lett.* **1994**, *27*, 1031−1054. (h) Sârbu, C.; Pop. H. F. Fuzzy Clustering Analysis of the First 10 MEIC Chemicals. *Chemosphere* in press. (i) Dumitrescu, D.; Pop, H. F. Degenerate and Nondegenerate Convex Decomposition of Finite Fuzzy Partitions. II. *Fuzzy Sets Systems* **1998**, *96*, 111−118. (j) Dumitrescu, D.; Pop, H. F. Degenerate and Nondegenerate Convex Decomposition of Finite Fuzzy Partitions. I. *Fuzzy Sets Systems* **1995**, *73*, 365−376.

(14) Siegel, S. *Nonparametric Statistics for the Behavioral Sciences*; McGraw−Hill: New York, 1956.

(15) Reed, A. E.; Curtiss, L. A.; Weinhold, F. Intermolecular Interactions from a Natural Bond Orbital, Donor−Acceptor Viewpoint. *Chem. Rev.* **1988**, *88*, 899−926. In particular, see ref 110 and the attendant discussion.

(16) Clardy, J.; Sinofsky, S. MacMendeleev, version 2.2; Cornell University, 1988.

(17) Dykstra, C. E.; Zhou, T. Additivity and transferability of Atomic Contributions to the Molecular Second Dipole Polarizabilities. *J. Phys. Chem.* in press.

CI0000023