การทำนายแอกทิวิตีและสมบัติของตัวเร่งปฏิกิริยาเซอร์โคโนซีนสำหรับเอทิลีนและ 1-เฮกซีน โคพอลิเมอไรเซชันโดยการสร้างแบบจำลองคิวเอสเออาร์/คิวเอสพีอาร์

นายปริญญา เจริญไข

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาปิโตรเคมีและวิทยาศาสตร์พอลิเมอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2556

PREDICTION OF ACTIVITY AND PROPERTIES OF ZIRCONOCENE CATALYST FOR ETHYLENE AND 1-HEXENE COPOLYMERIZATION BY QSAR/QSPR MODELLING

Mr. Prinya Charoenkhai

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Petrochemistry and Polymer Science
Faculty of Science
Chulalongkorn University
Academic Year 2013

Thesis Title	PREDICTION OF ACTIVITY AND PROPERTIES OF ZIRCONOCENE CATALYST FOR ETHYLENE AND 1-HEXENE COPOLYMERIZATION BY QSAR/QSPR MODELLING
By	Mr. Prinya Charoenkhai
Field of Study	Petrochemistry and Polymer Science
Thesis Advisor	Associate Professor Vudhichai Parasuk, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

.....Dean of the Faculty of Science

(Professor Supot Hannongbua, Dr.rer.nat.)

THESIS COMMITTEE

.....Chairman

(Assistant Professor Warinthorn Chavasiri, Ph.D.)

.....Thesis Advisor

(Associate Professor Vudhichai Parasuk, Ph.D.)

.....Examiner

(Associate Professor Wimonrat Trakarnpruk, Ph.D.)

.....Examiner

(Assistant Professor Somsak Pianwanit, Ph.D.)

.....External Examiner

(Tossapol Khamnaen, Ph.D.)

ปริญญา เจริญไข : การทำนายแอกทิวิตีและสมบัติของตัวเร่งปฏิกิริยาเซอร์โคโนซีนสำหรับเอทิลีนและ 1-เฮกซีนโคพอลิเมอไรเซชันโดยการสร้างแบบจำลองคิวเอสเออาร์/คิวเอสพีอาร์. (PREDICTION OF ACTIVITY AND PROPERTIES OF ZIRCONOCENE CATALYST FOR ETHYLENE AND 1-HEXENE COPOLYMERIZATION BY QSAR/QSPR MODELLING) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: รศ. ดร. วุฒิชัย พาราสุข, 84 หน้า.

เทคนิคการสร้างความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างหรือสมบัติต่อแอกทิวิตีของสารถูกนำมาประยุกต์ใช้ในการอธิบายความสัมพันธ์ของข้อมูลจากการทดลองระหว่างโครงสร้างกับสมบัติของตัวเร่งปฏิกิริยาเซอร์โคโนซีน ความสัมพันธ์ของโคพอลิเมอไรเซชันระหว่างเอทิลีนและ 1-เฮกซีนโคพอลิเมอร์กับโคพอลิเมอไรเซชันแอกทิวิตีจากการทดลองของชุดตัวเร่งปฏิกิริยาเซอร์โคโนซีนที่มีหมู่แทนที่บนลิแกนด์ไซโคลเพนทาไดอีนิลและอินดีนิลและสมบัติเชิงกายภาพแตกต่างกันถูกนำมาใช้ ในงานวิจัยนี้ศึกษาตัวเร่งปฏิกิริยาเซอร์โคโนซีนจำนวน 19 สารประกอบ โดยแบ่งกลุ่มออกเป็น 3 กลุ่มในการสร้างแบบจำลองความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างหรือสมบัติต่อแอกทิวิตีของสาร ตัวเร่งปฏิกิริยาเซอร์โคโนซีนทุกสารประกอบถูกปรับโครงสร้างสารประกอบด้วยทฤษฎีฟังชันนัลความหนาแน่น GGA-PW91 ร่วมกับเบสิกเซ็ต DNP 3.5 ในโมดูล Dmol3 ของโปรแกรมแม็ททีเรียลสตูดิโอ 5.5 สมบัติของตัวเร่งปฏิกิริยาเซอร์โคโนซีนที่คำนวณได้ทั้งหมด 33 ชนิด นำมาสร้างแบบจำลองความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างหรือสมบัติต่อแอกทิวิตีของสารด้วยการวิเคราะห์ถดถอยพหุคูณเชิงเส้นตรง แบบจำลองที่มีความสามารถในการทำนายที่ดีจะนำมาใช้ในการออกแบบตัวเร่งปฏิกิริยาเซอร์โคโนซีนสำหรับเอทิลีนและ 1-เฮกซีนโคพอลิเมอไรเซชันให้ดียิ่งขึ้น

สาขาวิชา ปิโตรเคมีและวิทยาศาสตร์พอลิเมอร์ ลายมือชื่อนิสิต .....

ปีการศึกษา 2556 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก .....

5372534323 : MAJOR PETROCHEMISTRY AND POLYMER SCIENCE
KEYWORDS: QSAR / QSPR / ZIRCONOCENE / ETHYLENE/1-HEXENE COPOLYMERIZATION

PRINYA CHAROENKHAI: PREDICTION OF ACTIVITY AND PROPERTIES OF ZIRCONOCENE CATALYST FOR ETHYLENE AND 1-HEXENE COPOLYMERIZATION BY QSAR/QSPR MODELLING. ADVISOR: ASSOC. PROF. VUDHICHAI PARASUK, Ph.D., 84 pp.

Quantitative structure-activity/property relationships (QSA/PR) methodology has been applied to explain the experimental data correlation of the structure related properties of zirconocene catalysts. The relationship between experimental ethylene and 1-hexene copolymerization activity of the series of zirconocene catalysts with different substitution patterns in the cyclopentadienyl and indenyl rings and physical properties were obtained. In this work, 3 sets of zirconocene catalyst consisting of 19 zirconocene complexes in total were used to build QSA/PR model. All zirconocene structures were optimized using the GGA-PW91 density functional theoretical method with DNP 3.5 basic set within DMol3 module of Materials Studio 5.5 program. Totally 33 properties related to the catalyst activity were calculated and fitted to QSA/PR model using the multiple linear regression (MLR) procedure. The good predictive ability of the models allows us to design better zirconocene catalyst for ethylene and 1-hexene copolymerization.

Field of Study: Petrochemistry and Polymer Science
Academic Year: 2013

Student's Signature .....

Advisor's Signature .....

ACKNOWLEDGEMENTS

My utmost gratitude goes to my thesis advisor, Assoc. Prof. Vudhichai Parasuk, for his expertise, kindness, support, and most of all, for his patience during the course of research including completing this thesis. His advise is always worthwhile and without him this work could not be possible.

I am sincerely grateful to a chairman and the member of the thesis committee, Prof. Tharapong Vitidsant, Assoc. Prof. Wimonrat Trakarnpruk, Assist. Prof. Somsak Pianwanit, and Dr. Tossapol Khamnaen for their valuable comments and suggestions.

Sincere thanks are given to SCG Chemicals Co.,Ltd.

I gratefully acknowledge the members of the ATC research groups, Mahamakut building for their companionship and friendship. Finally, I would like to take this opportunity to express my sincere appreciation and thanks to my parents for their encouragement, understanding and support throughout the entire study.

CONTENTS

Page

Page

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

atm	: Atmosphere
Cp	: Cyclopentadienyl
°C	: Degree Celsius
DFT	: Density Functional Theory
DNP	: Double Numerical Basis Set with Polarization
ESP	: Electrostatic Potential
GGA	: Generalized Gradient Approximation
h	: Hour
HOMO	: Highest Occupied Molecular Orbital
Ind	: Indenyl
LUMO	: Lowest Unoccupied Molecular Orbital
LLDPE	: Linear Low Density Polyethylene
LOO	: Leave One Out
M	: Molar
MAO	: Methyl Aluminoxane
mg	: Milligram
min	: Minute
MLR	: Multiple Linear Regression
$\bar{M}_n$	: Number average molecular weight

$\bar{M}_w$	: Weight average molecular weight
PC	: Principal Component
PCA	: Principal Component Analysis
PEH	: Polyethylene hexane copolymer
PLS	: Partial Least Squares
PRESS	: Predictive Residual Sum of Squares
q^2	: Cross-validated Correlation Coefficient
QSAR	: Quantitative Structure-Activity Relationship
QSPR	: Quantitative Structure-Property Relationship
r^2	: Correlation Coefficient
Rac	: Racemic
SCBD	: Short Chain Branching Distribution
SD	: Standard Deviation
VLDPE	: Very Low Density Polyethylene

CHAPTER I

INTRODUCTION

Plastics have an influence on our daily life in nearly every level of society today. Their resistance to corrosion and tremendous technological flexibility has enabled them to replace conventional materials such as metals and woods. Thus the demand for the plastics increases in huge quantities every year. Polyethylene is one of the most common and the world largest volume bulk plastic [1].

The first polymerization of ethylene to form polyethylene can be performed by a free radical initiation of high pressure process, then low pressure process was invented using metal alkyls as catalysts [2]. Ziegler-Natta catalyst, discovered by Karl Ziegler and Giulio Natta in 1953, was the first generation of the catalyst [3]. This catalyst is a complex formed by reaction of a transition metal compound of group IV-VII transition metal as the catalyst with a metal alkyl or alkyl halide of group I-III base metals as the cocatalyst. $MgCl_2$ and/or donor supported Ziegler-Natta transition catalyst system was the second generation, which was at least 100 times more active.

The new technology that may change the polyolefins industry was the introduction of metallocene catalyst. Such the catalyst consists of a transition metal group IV (Zr, Ti and Hf) which was sandwiched between parallel planar of two organic molecules such as cyclopentadienyl group (Cp; C_5H_5), substituted Cp group, the indenes, fluorenes and a co-catalyst such as methylaluminoxane (MAO) and borane compounds [1].

Copolymer of ethylene with 1-alkenes such as propene, 1-butene, 1-hexene, and 1-octene are very important commercial products classified as linear low density polyethylene (LLDPE) and/or very low density polyethylene (VLDPE). Rheological and mechanical properties of polymers do not depend only on their average molecular weight, but also on their molecular weight distribution (MWD) and short chain branching distribution (SCBD). Even though the control of MWD and SCBD is very

important, conventional heterogeneous Ziegler-Natta catalysts have limitations in controlling MWD and SCBD because polymers produced by these catalysts show broad MWD and SCBD due to the presence of multiple types of active sites on the catalyst. The metallocene catalyst permits the synthesis of polymers with a narrow and a well controlling MWD and SCBD at high polymerization rates.

The petrochemical industry is very competitive. Obtaining new catalysts, which give better yield and selectivity will create an advantage. Thus, the industry continues to search for new catalysts. A promising technique for designing new catalyst is the quantitative structure-activity relationship (QSAR) which is mainly used for the drug design [4]. Systematic experiments performing with different catalysts under fixed polymerisation conditions are thus highly suitable for QSAR study.

Very few authors have applied the QSAR concept to the polymerization activity of catalysts. One of the reasons is the difficulty in obtaining accurate experimental polymerization activity data. Moring and Coville [5] reported a quantitative relationship between the catalytic activity of $(CpR)_2ZrCl_2$ (where R is a cyclopentadienylmonoalkyl ring substituent) and steric and electronic descriptors. These authors noted the increase in activity of the catalytic system with the increase in the size of the R ligands as well as the increase in the electron-donating capacity of the substituents. Yao *et al*. [6] performed a QSAR analysis on the influence of the catalyst/co-catalyst ion pair structure on polymerization activity. They used molecular mechanics and dynamics to describe the molecular geometries of the ion pair. Activity differences in their model could be explained by the space between the catalyst and co-catalyst systems. Linnolahti*et al*. [7] undertook an extensive qualitative study to evaluate the effect of the ligand structure of zirconocene catalytic systems on the accessibility and relative stability of the active reaction centres. By comparing experimental and theoretical results, these authors demonstrated a qualitative correlation between active reaction centre accessibilities and polymerisation activity.

In 2004, V. L. Cruz *et al.* used 3D-QSAR method to analyse a set of metallocene catalysts which increases Cp-Zr-Cp angle to determine ethylene polymerization activity and polymer molecular weight. It was found that the

calculated model predicts that an increase in the Cp-Zr-Cp angle and/or incorporation of bulky ligands will enhance catalytic activity. The effect of electronic interaction was confirmed by correlations found between activity and the LUMO molecular orbital and between activity and local softness. The model revealed that the arrangement of the aromatic ligands around the metal center as well as the chemical nature of the ligand significantly contribute to explaining the variance shown by the experimental data. The structure of the bridge is not directly implicated in these two fields (LUMO and local softness) except that it forces the Cp-Zr-Cp angle and thus to a greater or lesser extent promotes electronic interaction between the metal centre and the atoms of the ligands [4]. In 2005, they applied QSAR to study a group of metallocene catalyst in ethylene polymerization. It was found that the steric and electrostatic fields can be interpreted in terms of catalyst/cocatalyst ion-pair interactions. Steric hindrance in specific positions and charge distribution around the aromatic ligands are correlated with an increase in activity. This could be explained by a weakening of the catalyst/cocatalyst interaction, which results in more room for the ethylene insertion reaction [8].

In 2012, Sonia Martínez*et al*. used 3D-QSAR to predict a series of zirconocene single-site catalysts with different substitution patterns in the cyclopentadienyl rings. The catalysts have an alkyl substituent or an alkenyl substituent with a C=C double bond or a silyl fragment. It was found that the best model was obtained with a training set of 37 catalysts, is composed of an 87% steric contribution and a remaining 13% electrostatic components. The steric hindrance should not be too large since the cocatalyst needs sufficient room to extract the methyl group for catalyst activation. The electrostatic component of the model points also in the same direction: negative charge around the active site will weaken the catalyst–cocatalyst intermolecular interaction by repulsive electrostatic effect between the two components, allowing the monomer to easily approach the active site. The addition of positive charge in the terminal groups of the Cp substituents will improve activity, most probably due to an attraction effect with the negatively charged cocatalyst that moves it away from the active center [9].

In this study, the influence of substituent in zirconocene catalyst towards activities of ethylene and 1-hexene copolymerization and some properties of the ethylene and 1-hexene copolymer were investigated by QSAR/QSPR technique. The results could be used for the design of a more potent zirconocene catalyst in ethylene and 1-hexene copolymerization.

1.1 Objective of the Thesis

Build QSAR/QSPR model to predict activity and properties of zirconocene catalyst for ethylene and 1-hexene copolymerization and to design zirconocene catalyst for high productivity and high activity with desired properties from the QSAR/QSPR model.

1.2 Scope of the Thesis

1. Investigate the correlation between catalyst structure and polymerization activity/polymerization properties and build the QSAR/QSPR models.
2. Design zirconocene catalysts for high polymerization activity from QSAR model.
3. Design zirconocene catalysts with desired copolymer properties such as number average molecular weight, weight average molecular weight, and 1-hexene incorporation.

CHAPTER II

THEORETICAL BACKGROUND

2.1 Background on Metallocene Catalyst

Polyolefins can be produced using Phillips type with free radical initiators, Ziegler-Natta, and metallocene catalysts. Ziegler-Natta catalysts have been widely used because of their broad range of applications. However, the Ziegler-Natta catalyst provides polymers with the broad molecular weight distribution (MWD) and the composition distribution due to the existence of multiple active sites [10].

Metallocene catalysts have been used to polymerize ethylene and **α**-olefins commercially. It can control composition distribution, incorporation of various comonomers, MWD, and stereoregularity [11].

2.1.1 Catalyst Structure

Metallocene is a class of compounds in which cyclopentadienyl or substituted cyclopentadienyl ligands are **π**-bonded to the metal atom. The structure of biscyclopentadienyl (or substituted cyclopentadienyl)-metal bis (unibidentate ligand) complexes can be most simply described as distorted tetrahedral, with each η^5-L group (L = ligand) occupying a single co-ordination position, as in **Figure 2.1** [12].

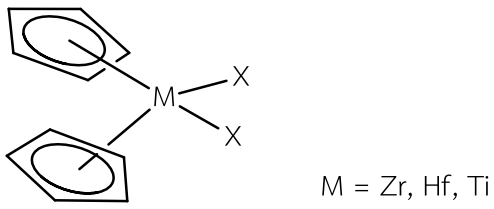


Figure 2.1 Molecular structure of metallocene.

Metallocene catalysts can be classified into 5 types. Representative examples of each category of metallocenes and some of zirconocene catalysts are shown in **Table 2.1** and **Figure 2.2**, respectively.

Table 2.1 Representative Examples of Metallocenes [12].

Category	Catalysts
[A] Nonstereorigid metallocenes	1) Cp_2MCl_2 (M = Ti, Zr, Hf)
	2) Cp_2ZrR_2 (M = Me, Ph, CH_2Ph, CH_2SiMe_3)
	3) $(Ind)_2ZrMe_2$
[B] Nonstereorigid ring-substituted metallocenes	1) $(Me_5C_5)_2MCl_2$ (M = Ti, Zr, Hf)
	2) $(Me_3SiCp)_2ZrCl_2$
[C] Stereorigid metallocenes	1) $Et(Ind)_2ZrCl_2$
	2) $Et(Ind)_2ZrMe_2$
	3) $Et(IndH_4)_2ZrCl_2$
[D] Cationic metallocenes	1) $Cp_2MR(L)^+[BPh_4]^-$ (M = Ti, Zr)
	2) $[Et(Ind)_2ZrMe]^+[B(C_6F_5)_4]^-$
	3) $[Cp_2ZrMe]^+[(C_2B_9H_{11})_2M]^-$ (M = Co)
[E] Supported metallocenes	1) Al_2O_3-$Et(IndH_4)_2ZrCl_2$
	2) $MgCl_2.Cp_2ZrCl_2$
	3) $SiO_2.Et(Ind)_2ZrCl_2$

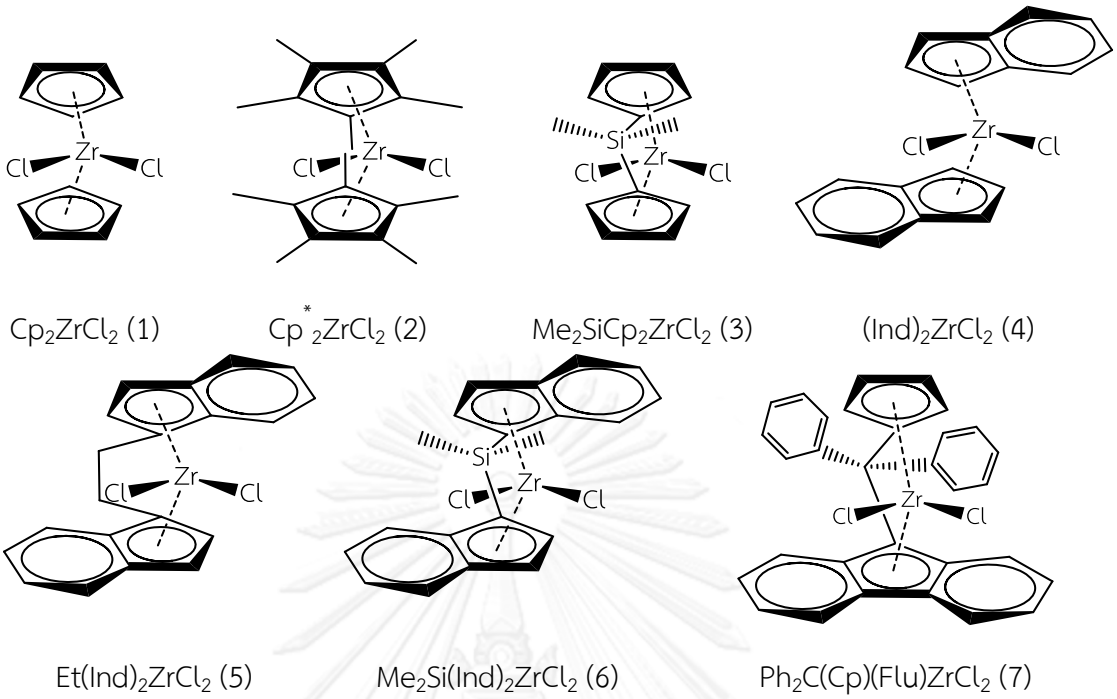


Figure 2.2 Some of zirconocene catalyst structures [13].

There are many types of metallocene, when the two cyclopentadienyl (Cp) rings on either side of the transition metal are unbridged, the metallocene is nonstereorigid and it is characterized by C_{2v} symmetry. The Cp_2M (M = metal) fragment is bent back with the centroid-metal-centroid angle **θ** about 140^{o} due to an interaction with the other two **σ** bonding ligands [14]. When the Cp rings are bridged (two Cp rings arranged in a chiral array and connected together with chemical bonds by a bridging group), the metallocene is stereorigid, so-called ansa-metallocene, and it could be characterized by either a C_1, C_2, or C_s symmetry depending upon the substitutions on two Cp rings and the structure of the bridging unit as schematically illustrated in **Figure 2.3** [12].

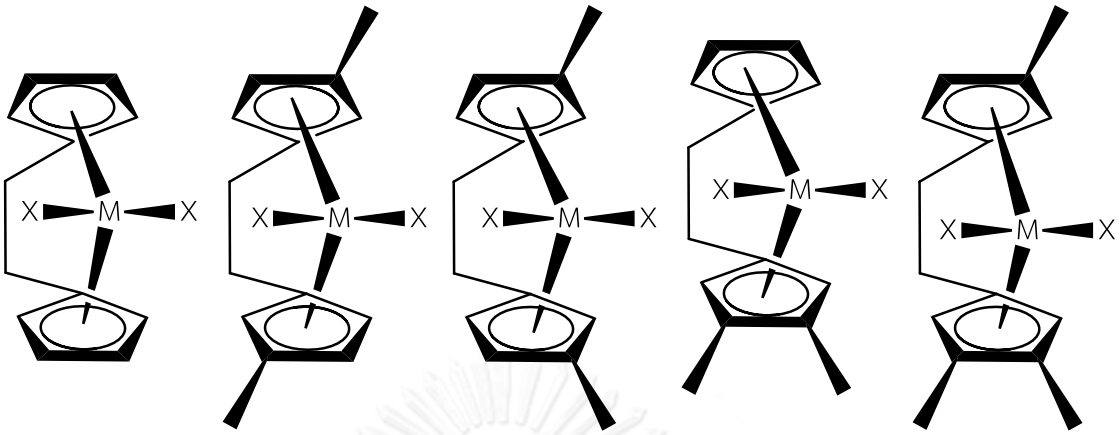


Figure 2.3 Scheme of the different metallocene complex structures [12]. Type 1 is C_{2v} symmetric; Type 2 is C_2-symmetric; Type 3 is C_s-symmetric; Type 4 is C_s-symmetric; Type 5 is C_1-symmetric.

2.1.2 Polymerization Mechanism

The mechanism of catalyst activation is not clearly understood. However, alkylation and reduction of the metal site by a cocatalyst (generally alkyl aluminum or alkyl aluminoxane) is believed to generate the cationic active catalyst species.

The polymerization mechanism is invalues 3 step, the initiation, the propagation, and the chain termination. The initiation step starts with the formation of the cationic species catalyst as shown below.

$$Cp_2ZrCl_2 + Al(CH_3)_3 \rightarrow Cp_2ZrClMe + Al(CH_3)_2Cl \qquad (2.1)$$

The propagation proceeds by the coordination and the insertion of new monomer unit into the metal carbon bond. Cossee mechanism [15] is still one of the most generally accepted polymerization mechanism **(Figure 2.4)**. First, monomer forms a complex with the vacant coordination site at the active catalyst center. Then, through a four-centered transition state, bond between monomer and metal center and between monomer and polymer chain are formed. This follows immediately by the insertion of monomer to the polymer chain increasing the length of the polymer chain by one monomer unit and generating another vacant site.

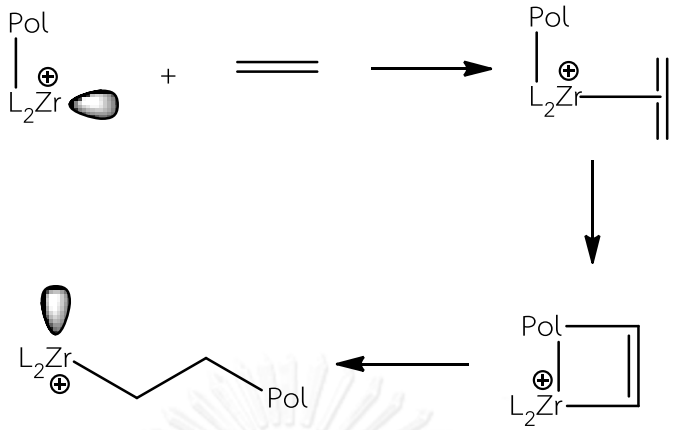


Figure 2.4 Cossee mechanism for Ziegler-Natta olefin polymerization [15].

Alternatively, the trigger mechanism has been proposed for the polymerization of **α**-olefin with Ziegler-Natta catalysts [16]. In this mechanism, two monomers interact with one active catalytic center in the transition state. A second monomer is required to form a new complex with the existing catalyst-monomer complex, thus trigger a chain propagation step. No vacant site is involved in this model. The trigger mechanism has been used to explain the rate enhancement effect observed when ethylene is copolymerized with **α**-olefins.

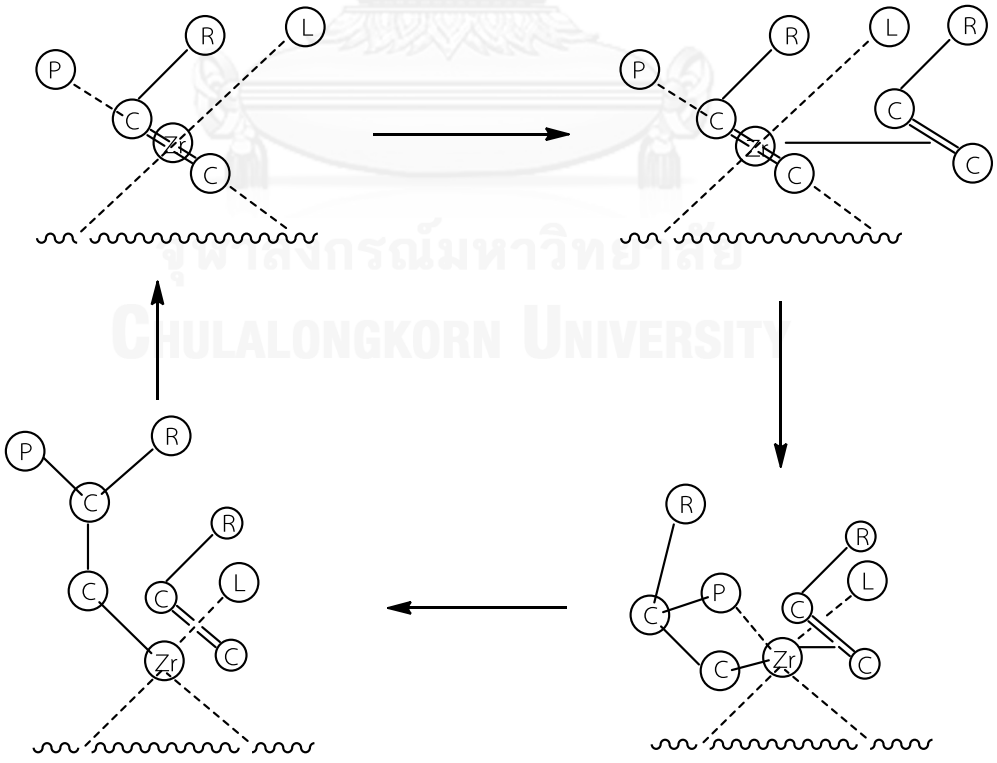


Figure 2.5 The propagation step according to the trigger mechanism [16].

The propagation mechanism in polymerization is schematically shown in **Figure 2.6**.

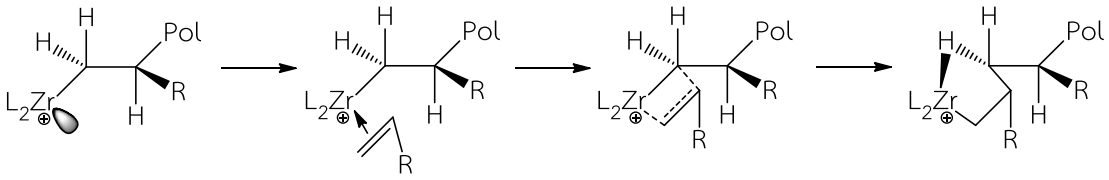


Figure 2.6 Propagation mechanism in polymerization.

Finally, the termination of polymer chains can be formed by 1) chain transfer via **β**-H elimination, 2) chain transfer via **β**-Me elimination, 3) chain transfer to aluminum, 4) chain transfer to monomer, and 5) chain transfer to hydrogen as displayed in **Figure 2.7-2.11** [12]. The first two transfer reactions form the polymer chains with terminal double bonds.

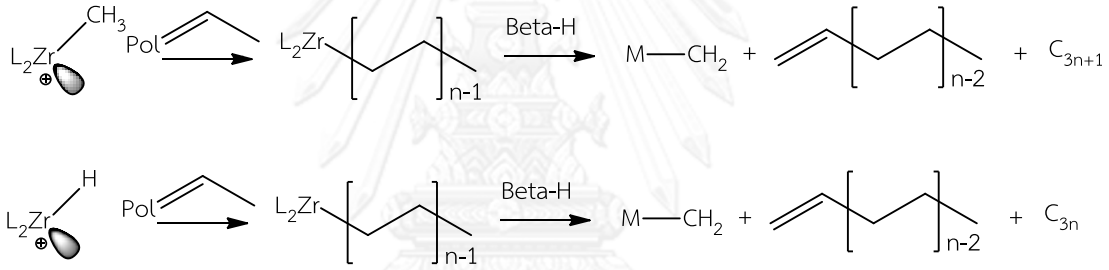


Figure 2.7 Chain transfer via **β**-H elimination [12].

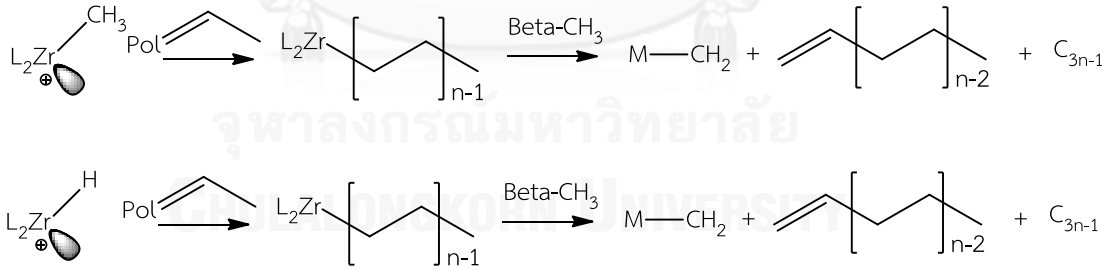


Figure 2.8 Chain transfer via **β**-CH_3 elimination [12].

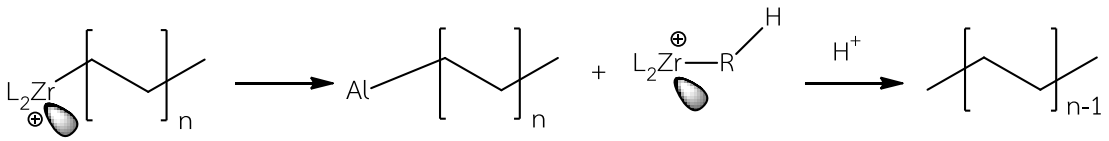


Figure 2.9 Chain transfer to aluminum [12].

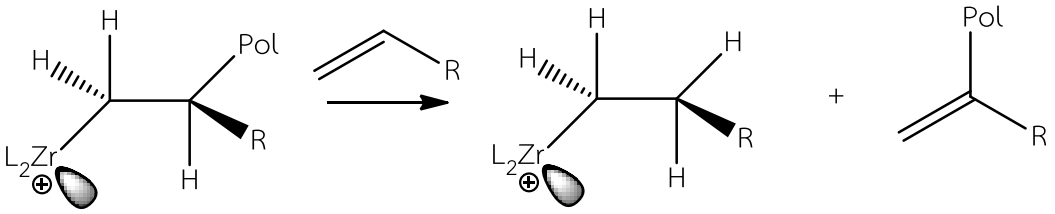


Figure 2.10 Chain transfer to monomer [12].

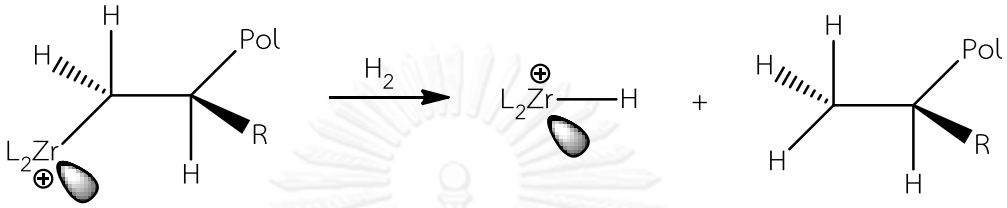


Figure 2.11 Chain transfer to hydrogen [12].

2.1.3 Cocatalysts

Metallocene catalysts have to be activated by a cocatalyst. The most common types of cocatalysts are alkyl aluminums including methyl aluminoxane (MAO), trimethylaluminum (TMA), triethylaluminum (TEA), triisobutylaluminum (TIBA) and cation forming agents such as $(C_6H_5)_3C^+(C_6F_5)_4B^-$ and $B(C_6F_5)_3$ [17].

Among these, MAO is a very effective cocatalyst for metallocene. However, due to the difficulties and costs involved in the synthesis of MAO, there has been considerable effort done to reduce or elimination the use of MAO. Due to difficulties in separation, most commercially available MAO contains a significant fraction of TMA (about 10-30%) [18]. This TMA in MAO could be substantially eliminated by toluene evaporation at 25°C.

Indeed, the difficulty in understanding the important factors for an efficient activation are mainly due to the poor knowledge of MAO compositions and structures. Several types of macromolecular arrangements, involving linear chains, monocycles and/or various three-dimensional structures have been successively postulated. These are shown in **Figure 2.12**. In recent work, a more detailed image of MAO was proposed as a cage molecule, with a general formula $Me_{6m}Al_{4m}O_{3m}$ (m equal to 3 or 4) [19].

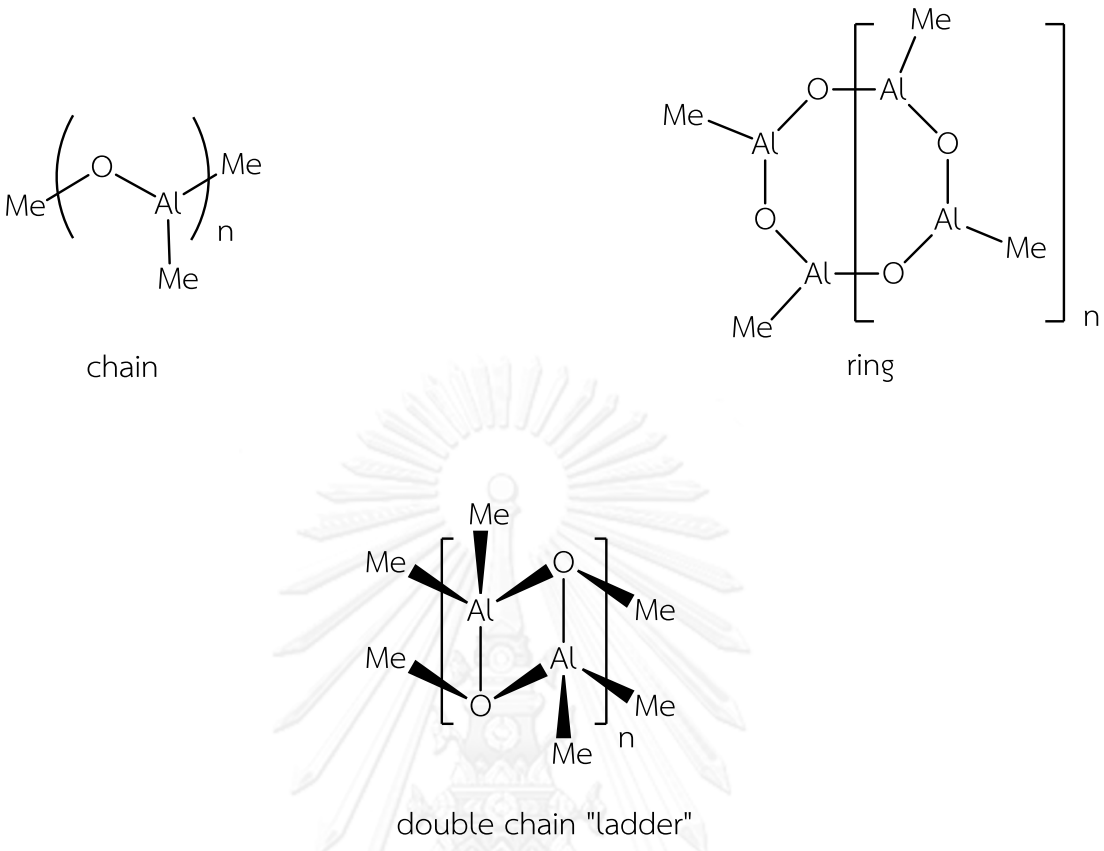


Figure 2.12 Early structure models for MAO [19].

In the case of *rac*-Et(Ind)$_2$ZrMe$_2$ as precursor, the extracted methyl ligands do not yield any modification in the structure and reactivity of the MAO counter-anion, thus allowing zirconium coordination site available for olefin that presented in **Figure 2.13** [20].

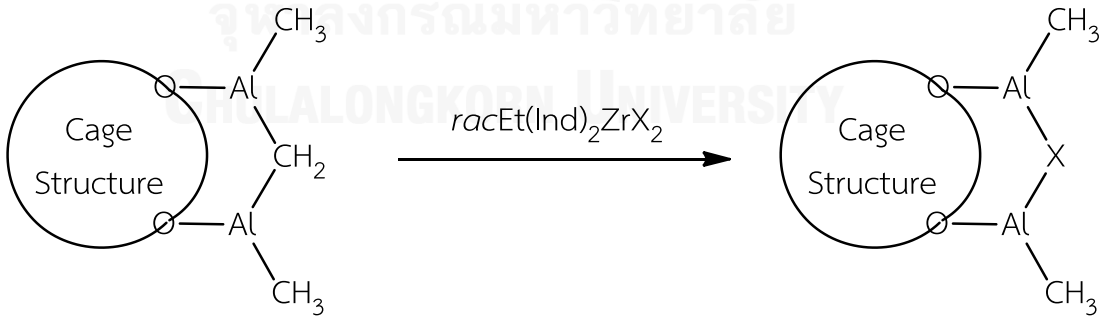


Figure 2.13 Representation of MAO showing the substitution of one bridging methyl group by X ligand extracted from *rac*Et(Ind)$_2$ZrCl$_2$ (X = Cl, NMe$_2$, CH$_2$Ph) [20].

2.2 Theoretical Background

2.2.1 Quantitative Structure-Activity/Property Relationship (QSA/PR)

The classical QSA/PR method is based on the statistical correlation of activities (Biological activity or properties for drugs and catalytic activity for catalysts) with physicochemical properties or scalar descriptors that encode certain structural features of the ligands [21]. Once a correlation model between molecular structure and activity/property is found, any number of compounds, including those not yet synthesized, could be screened on a computer. This allows one to select the most promising precursors having a set of desired properties. Finally, these compounds could be synthesized and tested in the laboratory. The typical QSA/PR model can be described as

$$\textbf{Activity/Property} = f\,\textbf{(Structural/Physicochemical properties)} \qquad (2.2)$$

In general, the regression equation can be accepted in QSA/PR studies if the following five criteria are met [22-24].

1. $r^2 \geq 0.64$ for *in vivo* data or $r^2 \geq 0.81$ for *in vitro* data.
2. The standard deviation (s) is not much larger than the standard deviation of the experimental data.
3. The overall significance level is better than 95 % as indicated by the F value.
4. The descriptors should not be intercorrelated, i.e., interdescriptor correlation coefficients should be less than 0.7.
5. A model with the q^2 value of greater than 0.50 is accepted as a good model.

The r^2 statistics is a ratio of the variance explained by the regression model to the total variance. Hence, r^2 gives information on how many percentage of the variation in the activity (Y variable) can be explained by the properties (X variables) displayed in the equation.

$$r^2 = \Sigma(y_{calculate} - y_{mean})^2 / \Sigma(y_{observe} - y_{mean})^2 \quad (2.3)$$

or

$$r^2 = 1 - \Sigma(y_{observe} - y_{calculate})^2 / \Sigma(y_{observe} - y_{mean})^2 \quad (2.4)$$

If r^2 =1, the equation can explain all 100% of the variation in the activity.

However, the r^2 gives information only on the reproducibility. Thus, the predictability should be an ability to predict the activity of a new compound outside the model. The most widely used method is the cross-validation (r^2_{cv} or q^2). The q^2 is estimated by repeatedly leaving out one (or more) compound(s) at a time until each compound is excluded exactly once.

$$q^2 = 1 - PRESS/SST \quad (2.5)$$

$$PRESS = \Sigma(y_{observed} - y_{predicted})^2 \quad (2.6)$$

$$SST = \Sigma(y_{observe} - y_{mean})^2 \quad (2.7)$$

The predictive residual sum of squares (PRESS) is the sum of the squared prediction errors. A larger q^2 indicates the model's good predictability.

2.2.2 Regression Techniques

2.2.2.1 Multiple Linear Regression (MLR)

The MLR is a method used to model the linear relationship between a dependent variable and one or more independent variables. The dependent variable is sometimes also called the predictand, and the independent variables the predictors. The MLR is based on the least squares method, The model is fited in such a way that the sum-of-squares of differences of observed and predicted values is minimized. In the process of fitting, or estimating the model, statistics are computed that summarize the *accuracy* of the regression model for the calibration period. The performance of the model on data not used to fit the model is usually checked in some way by a process called *validation*.

The predictors in any regression problem might be intercorrelated. Intercorrelation of predictors does not invalidate the use of regression, but can make it difficult or impossible to assess the relative importance of individual predictors from the estimated coefficients of the regression equation. Extremely high intercorrelation of predictors, or *multicolinearity*, exacerbates any difficulty of interpreting the regression coefficients, and may call for combination of subsets of predictors into a new set of less-intercorrelated predictors.

Regression models are generally not intended to be applied to predictor data outside the range encountered in the calibration period.

The model expresses the value of a predictand variable as a linear function of one or more predictor variables and an error term:

$$y_i = b_0 + b_1 x_{i,1} + b_2 x_{i,2} + \ldots + b_k x_{i,k} + e_i \qquad (2.8)$$

$x_{i,k}$ = value of k^{th} predictor in year i

b_0 = regression constsnt

b_k = coefficient of the k^{th} predictor

K = total number of predictors

y_i = predictant in year i

e_i = error term

The model (2.8) is estimated by least squares, which yields parameter estimates such that the sum of squares of errors is minimized. The resulting prediction equation is

$$\hat{y}_i = \hat{b}_0 + \hat{b}_1 x_{i,1} + \hat{b}_2 x_{i,2} + \ldots + \hat{b}_k x_{i,k} + e_i \qquad (2.9)$$

Where the variables are defined as in (2.8) except that "^" denotes estimated values

The error term in equation (2.8) is unknown because the true model is unknown. Once the model has been estimated, the regression residuals are defined as

$$\hat{e}_i = y_i - \hat{y}_i \tag{2.5}$$

$y_i y_i$ = observed value of predictant in year i

$\hat{y}_i \hat{y}_i$ = predicted value of predictant in year i

The residuals measure the closeness of fit of the predicted values and actual predictand in the calibration period. The algorithm for estimating the regression equation (solution of the normal equations) guarantees that the residuals have a mean of zero for the calibration period. The variance of the residuals measures the "size" of the error, and is small if the model fits the data well.

2.2.2.2 Partial Least Squares (PLS)

The PLS analyses were performed for different combinations of field descriptors. PLS calculations with the combined field were performed using the so-called autoscaling, where each field is scaled to have unit variance. The software calculates the standard deviation (SD) of each field and divides each value by the corresponding SD. The effect is to give each variable the same prior importance in the analysis. Leave one out (LOO) [4] cross-validated PLS analysis was initially performed to determine both the robustness of the statistical models and the optimal number of components or LVs. This can be achieved by examining the predictive residual sum of squares (PRESS) and the cross-validated regression coefficient (q^2) as guidelines. PLS analysis is a regression using principal component-like quantities derived from the explanatory variables, called *latent variables*. It is particularly useful for data sets with a high level of redundancy due to collinearity or multicollinearity. The PLS algorithm implemented here is as described by Stahle and Wold (1988) [25, 26].

Because MLR and PLS analysis are similar techniques, the decision whether to use MLR or PLS on a particular QSAR problem. MLR is most useful for large numbers of samples with a few descriptors. PLS is most useful for large numbers of descriptors. One way to reduce risk of overfitting the data normally use principle components analysis before performing the main statistical analysis.

2.2.3 Principal Component Analysis (PCA)

PCA is known a Principle Component Analysis. It is a method that can be used to reduce a large number of variables (descriptors) to a smaller number without losing information.

PCA is performed on a data set. The original variables are transformed into a new orthogonal set of linear combinations. This new orthogonal set can be preserve the variance of the original data set. The main feature of PCA is to reduce the size of large data matrix to a few features by capturing the variance in terms of Principal Components (PCs).

For example,

$$PC_1 = c_{(1,1)}v_1 + c_{(1,2)}v_1 + \ldots + c_{(1,n)}v_n \quad (2.8)$$

$$PC_2 = c_{(2,1)}v_1 + c_{(2,2)}v_1 + \ldots + c_{(2,n)}v_n \quad (2.9)$$

$$PC_3 = c_{(3,1)}v_1 + c_{(3,2)}v_1 + \ldots + c_{(3,n)}v_n \quad (2.10)$$

Each principal component (PCs) is a combination of the *original variables (v)*, defined using *loading coefficients (c).* Additionally, each principal component has an associated eigenvalue, which shows how much of the variance of the original data set is explained by that component. The first component or PC_1 always explains the greatest variance, the second the next greatest, and so on.

Generally, if a data matrix contains **S samples** and **D descriptors**, the maximum number of components is equal to S when D > S or D, in case of S > D. In PCA a data matrix (X) is decomposed into a principal components consisting of a

scores matrix (S_m), a loadings matrix (L_m) and a residual matrix (R_m), shown in equation as [$X = S_m L_m + R_m$] which represented graphically in **Figure 2.14**.

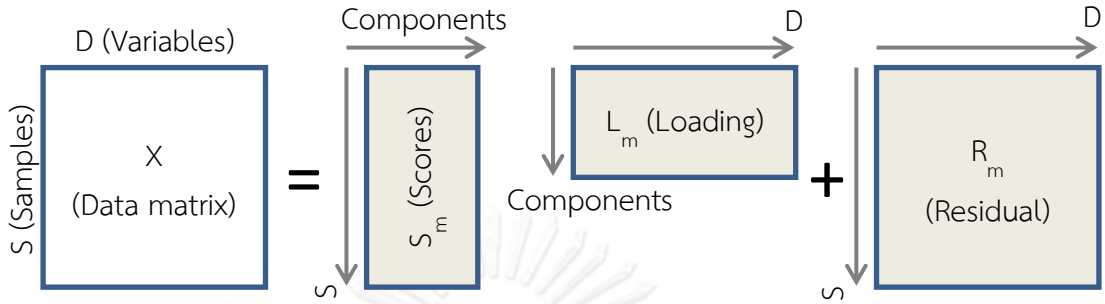


Figure 2.14 The decomposition of a data matrix (X) into a Scores matrix (S_m) and a Loadings matrix (L_m) with principal components (PCs).

The PCA results can be interpreted by the variance plot, the loading plot and the scores plot.

2.2.3.1 Variance Plot

The variance plot shows how much variance in the dataset is explained by which PC (as bar) and how much variance is explained by the first n PCs (as line) as shown in **Figure 2.16**.

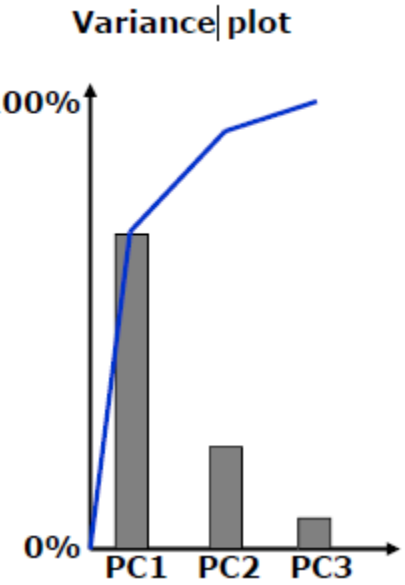


Figure 2.15 Variance plot in 3 PC.

2.2.3.2 Loading Plot

The Loading Plot is a plot of the relationship between original variables and subspace dimensions. It is used for interpreting relationships among variables (descriptors/properties).

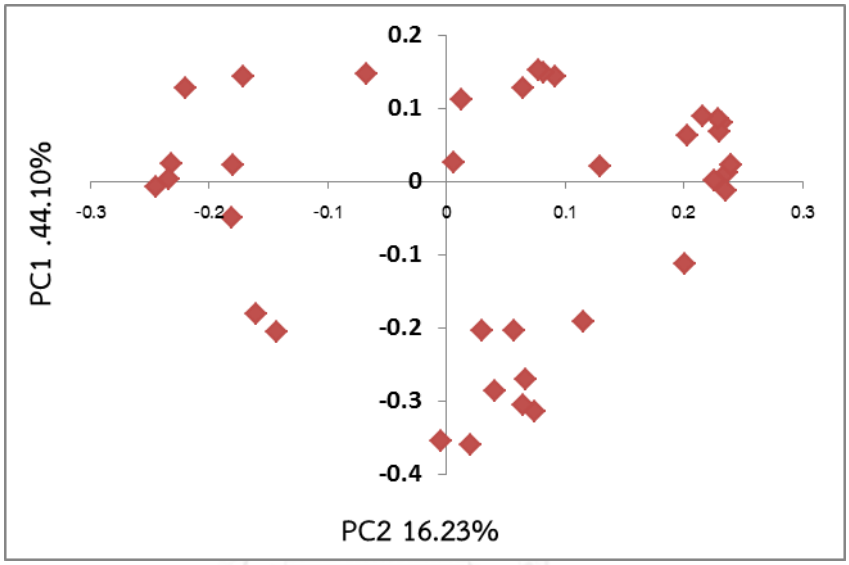


Figure 2.16 The plot of the loading of the variables (37 descriptors).

2.2.3.3 Scores Plot

The score plot is a projection of data onto subspace. It is used for interpreting relations among observations (or samples).

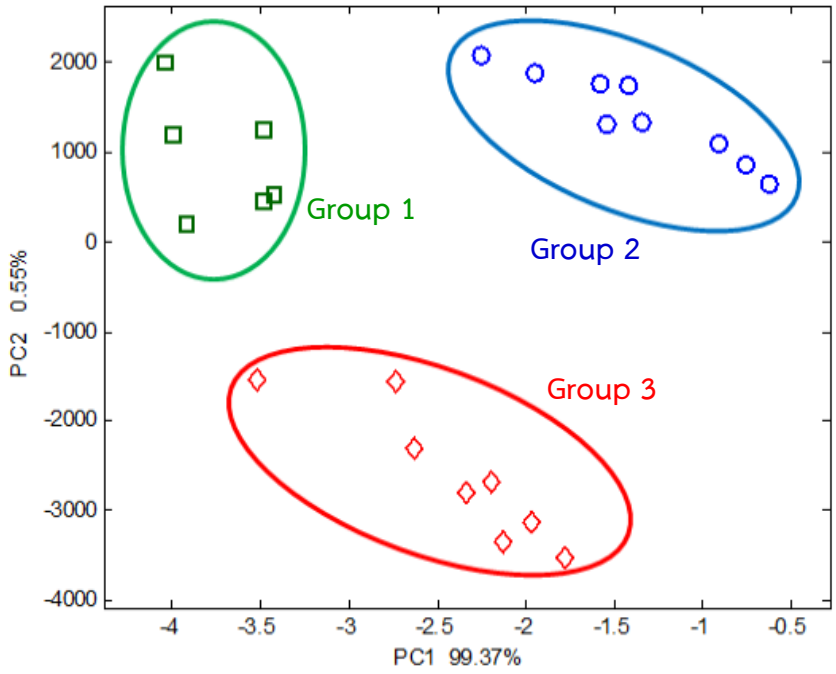


Figure 2.17 The plot of the score of the samples (3 groups).

CHAPTER III

COMPUTATIONAL DETAILS

3.1 Source of Zirconocene Catalyst and Polymerization Data

We used 3 groups of zirconocene catalyst consisting 19 zirconocene compounds to build QSA/PR model. The experimental catalytic activities and properties such as % 1-hexene incorporation, $\bar{M}_W$ (Weight average molecular weight), and $\bar{M}_n$ (Number average molecular weight) of these compounds were given in **Table 3.1** Structures of compounds from the 3 groups were illustrated in **Scheme 3.1**.

For group 1, the position and the degree of methyl substitution on cyclopentadienyl ring of zirconocene were varied. The copolymerization of ethylene and 1-hexene was carried out at 80°C, 1 hour, olefin pressure of 2 bar, and molar ratio MAO/zirconocene of 3000 [27].

For group 2, the position and the degree of aryl-substitution (Phenyl and pentafluorophenyl (C_6F_5)) on cyclopentadienyl ring of zirconocene were varied. The copolymerization of ethylene and 1-hexene was carried out at 1 atm of C_2H_4, 0.4 M 1-hexene, 50°C, 5 min and molar ratio MAO/zirconocene of 2000 [27].

For group 3, the position and the degree of ethyl and propyl substitution on indenyl ring with *meso* and *rac* formed of zirconocene were varied. The copolymerization of ethylene and 1-hexene was carried out at 85°C, an olefin pressure of 4 bar, 10 ml of 1-hexene, 80 mg of catalyst and used time of 30 min [28].

3.2 Geometry Optimization

All zirconocene structures as displayed in scheme 3.1 were optimized using the GGA-PW91 [29] density functional theoretical method (DFT) [30] and DNP 3.5 basic set within $Dmol^3$ [31] module of Materials Studio 5.5 program [32].

3.3 Molecular Descriptors

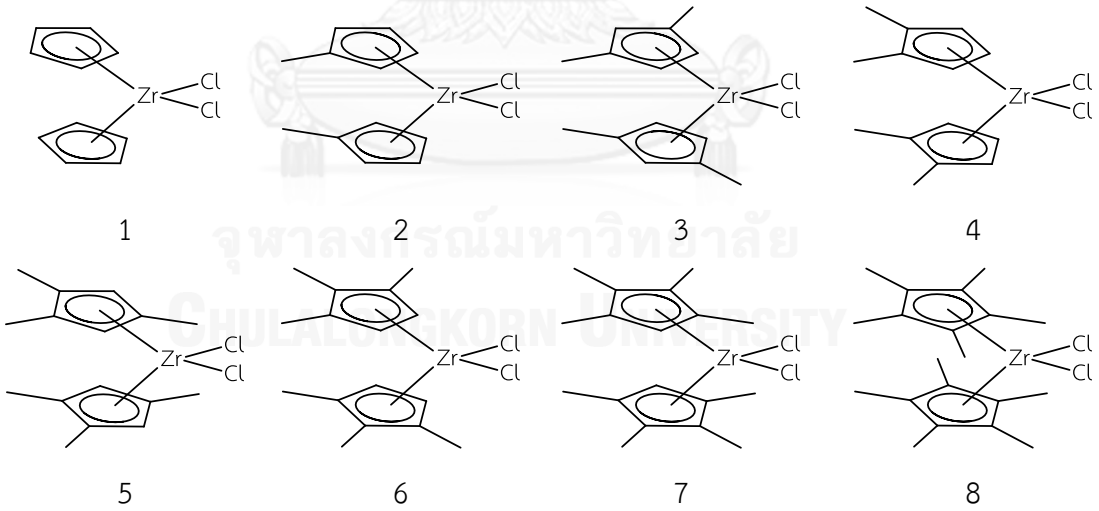
Molecular descriptors such as geometrical properties, crystal similarity, atomistic, spatial, electronic, and energies were considered and calculated within the QSAR module [31] of Materials Studio 5.5 program [32].

3.4 PCA analysis

The PCA analyses with mean/SD standardize data of data reduction module in Materials Studio 5.5 program.

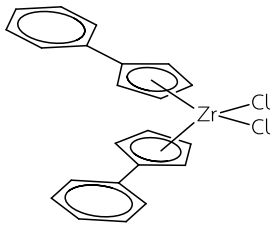
3.5 PLS analysis

For MLR calculations, the partial least squares analyses of QSA/PR model building module with mean/SD standardize data and omit 3 groups of rows cross-validation in Materials Studio 5.5 were chosen.

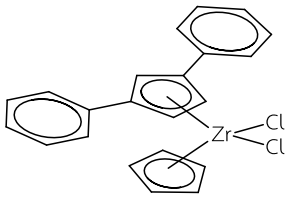


Group 1

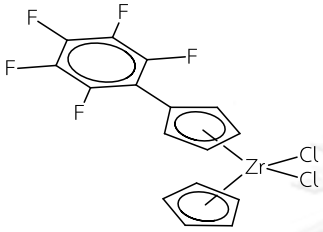
Figure 3.1 Scheme of zirconocene structures for group 1, 2 and 3.



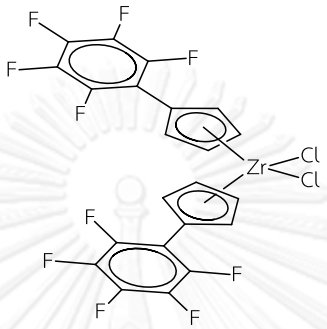
9



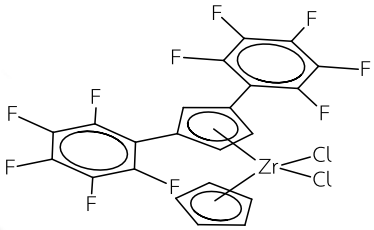
10



11

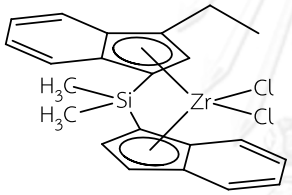


12

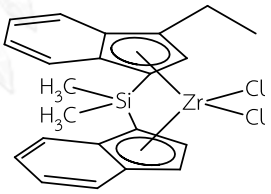


13

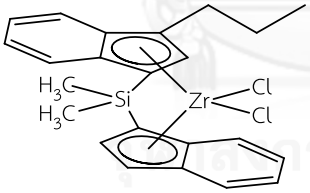
Group 2



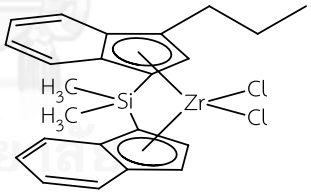
14



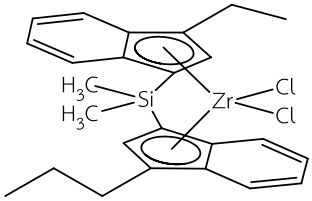
15



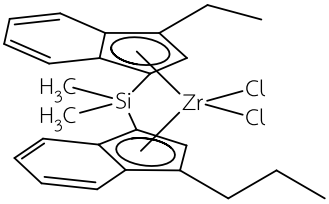
16



17



18



19

Group 3

Figure 3.1 Scheme of zirconocene structures for group 1, 2 and 3 (Continue).

Table 3.1 Ethylene and 1-hexene copolymerization data of zirconocene catalysts [27, 28].

No.	Catalyst in group 1	Activity[a]	1-hx incorp[c]	$\overline{M}_n$	
1	Cp_2ZrCl_2	22	2.7	11000	
2	$(MeCp)_2ZrCl_2$	88	2.5	25500	
3	$[(1,3\text{-}Me)_2Cp]_2ZrCl_2$	29	4.0	34100	
4	$[(1,2\text{-}Me)_2Cp]_2ZrCl_2$	314	1.8	57200	
5	$[(1,2,4\text{-}Me)_3Cp]_2ZrCl_2$	361	5.5	51600	
6	$[(1,2,3\text{-}Me)_3Cp]_2ZrCl_2$	334	1.7	65200	
7	$[(1,2,3,4\text{-}Me)_4Cp]_2ZrCl_2$	99	1.4	73000	
8	$[(1,2,3,4,5\text{-}Me)_5Cp]_2ZrCl_2$	31	0.3	47300	
No.	**Catalyst in group 2**	**Activity[b]**	**1-hx incorp[c]**	**M_w**	
9	$(PhCp)_2ZrCl_2$	26	5.9	230000	
10	$[(1,3\text{-}Ph)_2Cp]CpZrCl_2$	26	5.4	265000	
11	$[(C_6F_5)Cp]CpZrCl_2$	12	6.0	114000	
12	$[(C_6F_5)Cp]_2ZrCl_2$	10	11.3	37500	
13	$[1,3\text{-}(C_6F_5)_2Cp]_2ZrCl_2$	7	19.0	17000	
No.	**Catalyst in group 3**	**Activity[b]**		**M_w**	**MWD**
14	*rac*-$[Me_2Si(EtInd)Ind]ZrCl_2$	740		122500	2.71
15	*meso*-$[Me_2Si(EtInd)Ind]ZrCl_2$	190		120100	1.99
16	*rac*-$[Me_2Si(PrInd)Ind]ZrCl_2$	1100		70300	2.86
17	*meso*-$[Me_2Si(PrInd)Ind]ZrCl_2$	350		63100	1.87
18	*rac*-$[Me_2Si(EtInd)(PrInd)]ZrCl_2$	598		108200	2.82
19	*meso*-$[Me_2Si(EtInd)(PrInd)]ZrCl_2$	450		98900	1.92

[a] Catalyst activity in $kg_{PEH}/mol_{zr}.h.bar$

[b] Catalyst activity in $kg_{PEH}/mol_{zr}.h$

[c] mol%

CHAPTER IV

RESULTS AND DISCUSSION

4.1 Group 1

4.1.1 QSAR Model

The degree of methylation on cyclopentadienyl ring has an effect the activity of the process. However, there seems to be no regular tendency towards the degree of methylation and activity. The zirconocene catalyst with tri-substituted cyclopentadienyl ring gives the highest activity.

4.1.1.1 Correlation Analysis

Totally 33 descriptors were used for QSAR fitting. Values of 33 descriptors were given in the appendix (**Table 1A**). Only 10 descriptors with high correlation to activity were listed in **Table 4.1** These descriptors were further fitted to yield QSAR models.

Table 4.1 Descriptors with high correlation to activity.

Descriptor	r
Total dipole	0.73
Dipole x	-0.65
Cp ring Up Charge (Mulliken)	-0.50
Rank in cluster	-0.35
Cp ring Up Charge (ESP)	0.31
Cluster number	0.28
Dipole y	0.25
Zr Charge (Hirshfeld)	0.24
Cp-Zr-Cp Angle	0.23
Cp ring Down Charge (Mulliken)	0.22

4.1.1.2 PCA Analysis

The dot in the loading plot represents the projection of the variables (33 descriptors) on the PC1–PC2 plane as shown in **Figure 4.1**. Total variances of PC axis are 80.66% and 7.79% which corresponding to PC1 and PC2 respectively. The two variances are closed to 100% which refers to suitable information visualized from the data set.

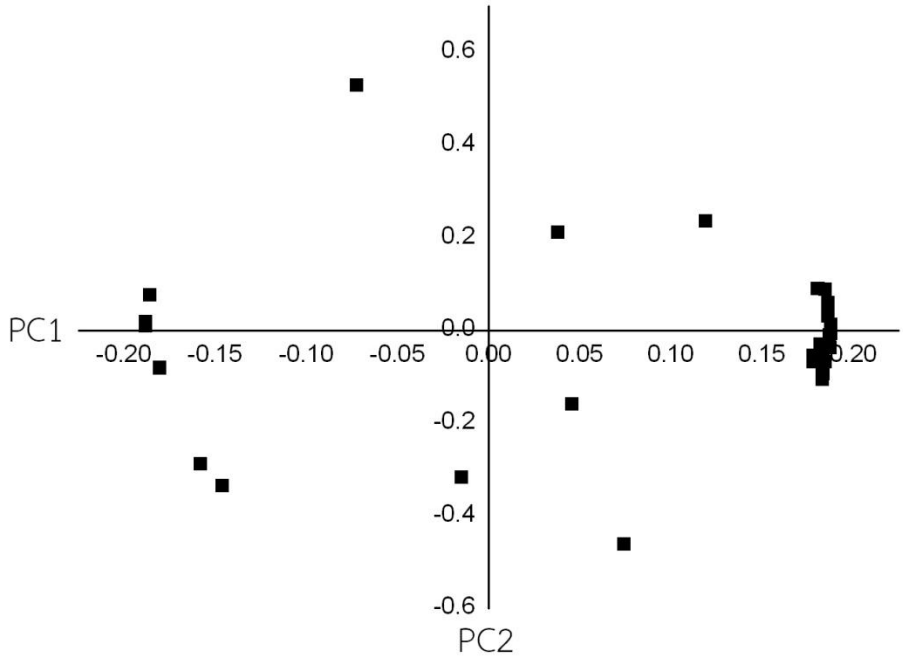


Figure 4.1 PCA loadings of QSAR for zirconocene catalyst 1-8 in group 1.

From the loading plot, it can be seen that there is apparent clusters of descriptors. Thus, the descriptors in the same clusters (these descriptors are highly inter-correlated) should be omitted for further QSAR analysis. The total energy and binding energy descriptors are obvious the same clusters of descriptors thus two descriptors should not be omitted or selected only one descriptor for fitting the QSAR model.

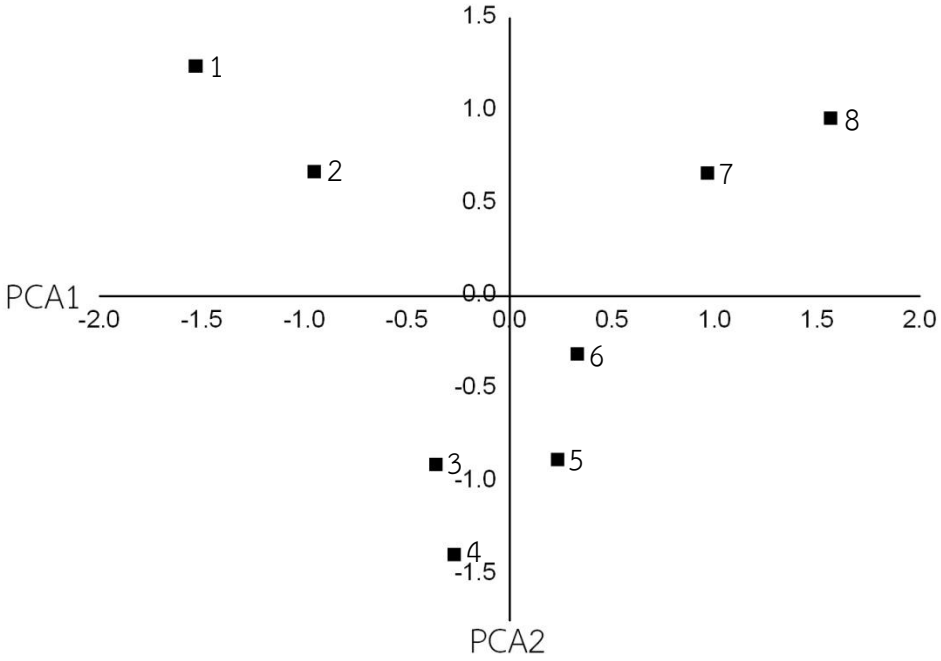


Figure 4.2 PCA scores of QSAR for zirconocene catalyst 1-8 in group 1.

The dot in the scores plot represents each zirconocene catalysts of group 1 is presented in **Figure 4.2**. From the score plot, it can be seen that there is obvious cluster of zirconocene catalyst on the score plots. This suggests the structural information of zirconocene catalyst 1, 2, 7, and 8 are different from the catalyst 3-5. The results from the score plot is obviously unrelated due to the difference structures of zirconocene catalyst, hence the further QSAR analysis will be performed.

4.1.1.3 Model

The descriptors after PCA analysis was fitted to the QSAR models by MLR procedure. The best model obtained from MLR fit were given in **Table 4.2**.

Table 4.2 QSAR models from MLR analysis.

No.	Model	r^2	q^2
1	Y = 516.56 * Total dipole - 850.69	0.54	0.33
2	Y = -352.95 * Dipole x - 415.07	0.42	0.11
3	Y = 99.70 * Cluster number + 6954.96 * Zr Charge (Hirshfeld) - 3074.94	0.79	0.56
4	Y = 512.35 * Total dipole - 279.47 * Cp ring Up Charge (Mulliken) - 999.30	0.78	0.54
5	Y = -284.45 * Cp ring Up Charge (Mulliken) + 0.10	0.25	-0.60
6	Y = -147.14 * Rank in cluster + 325.29	0.12	-1.14
7	Y = 159.34 * Cp ring Up Charge (ESP) + 202.20	0.09	-0.66
8	Y = 18.64 * Cluster number + 78.20	0.08	-0.56
9	Y = 329.29 * Dipole y + 144.08	0.06	-1.02
10	Y = 1135.80 * Zr Charge (Hirshfeld) - 297.27	0.06	-0.44

From **Table 4.2**, model (3) which contains 2 descriptors, cluster number and Zr charge (Hirshfeld) was chosen since it has very high correlation coefficient r^2 of 0.79 and q^2 of 0.56. One–descriptor models have r^2 less than 0.6. The model (4) though having high r^2 and q^2 values but contains only electronic descriptors while model (3) contains both steric and electronic descriptors.

Values of cluster number and Zr charge (Hirshfeld) of compounds from group 1 together with activities were listed in **Table 4.3**.

Table 4.3 Calculated properties and activity of QSAR model.

No.	Cluster number	Zr Charge (Hirshfeld)	Activity[a]
1.	1	0.425	22
2.	2	0.421	88
3.	3	0.423	29
4.	4	0.429	314
5.	5	0.42	361
6.	6	0.394	334
7.	7	0.351	99
8.	7	0.356	31

[a] Catalyst activity in kgPEH/mol_{zr}.h.bar

From the model, the cluster number representing steric effect and zirconium charge (Hirshfeld) representing electronic effect. The coefficients in front of these two descriptors are positive suggestion that the activities will increase as values of descriptors increases. The increment in the degree of methyl substitutions raises both cluster number and Zr charge (Hirshfeld) hence the activity. The compound 5 has large value for cluster number and Zr charge (Hirshfeld) and it is resulted in high activity. The Zr charge could be enhanced by having electron donating group as substituents.

4.1.1.4 Catalyst Design

From the information in **4.1.1.3**, we have change substituent group from methyl to ethyl, normal-butyl, tertiary-butyl and $SiMe_3$ to see the electronic effect of substituents. Activities were predicted activity according to model (3) and there values were given in **Table 4.4**.

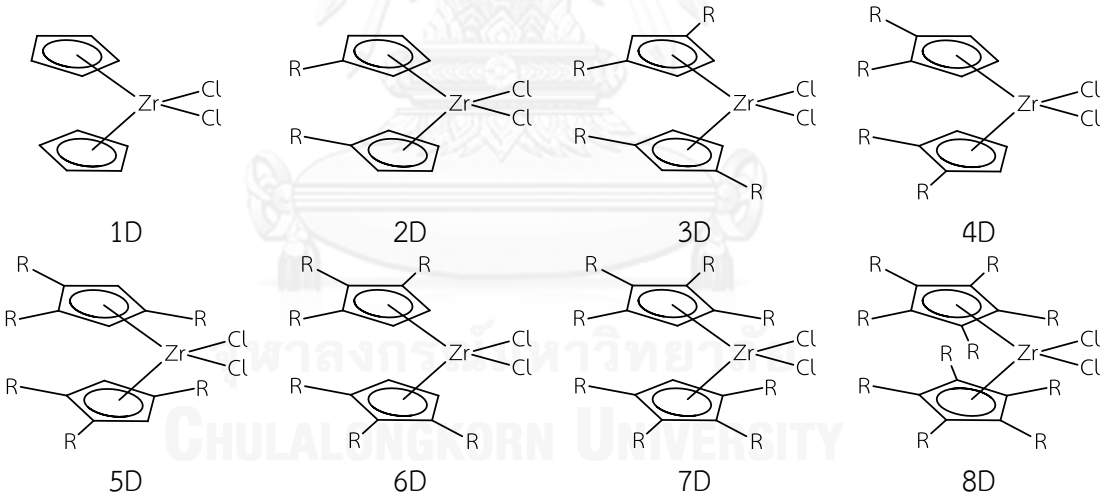


Figure 4.3 Scheme of designed catalysts in group 1.

From **Table 4.4**, ligands with tert-butyl substitution on cyclopentadienyl have better zirconium charge and activity. The order is tert-butyl>$SiMe_3$>normal-Butyl>ethyl>methyl.

From the prediction, our suggestion (tert-butyl) has the activity of 1.8 times better than compound 5.

Table 4.4 Calculated properties and prediction activity of design zirconocene catalyst in group 1.

No.	Cluster Number					Zr Charge (Hirshfeld)					Predictive Activity				
	Me	Et	^{n}Bu	^{t}Bu	$SiMe_3$	Me	Et	^{n}Bu	^{t}Bu	$SiMe_3$	Me	Et	^{n}Bu	^{t}Bu	$SiMe_3$
1D	1	1	1	1	1	0.425	0.425	0.425	0.425	0.425	-19.3867	-19.3867	-19.3867	-19.3867	-19.3867
2D	2	2	2	2	2	0.421	0.423	0.429	0.435	0.434	52.49501	66.40493	108.1347	149.8644	142.9094
3D	3	3	3	3	3	0.423	0.431	0.434	0.453	0.437	166.1064	221.7461	242.6109	374.7551	263.4758
4D	4	4	4	4	4	0.429	0.431	0.435	0.48	0.459	307.5376	321.4476	349.2674	662.2404	516.1863
5D	5	5	5	5	5	0.42	0.433	0.437	0.465	0.468	344.6445	435.059	462.8788	657.6176	678.4824
6D	6	6	6	6	6	0.394	0.435	0.438	0.316	0.461	263.5172	548.6704	569.5352	-278.969	729.4992
7D	7	7	7	7	7	0.351	0.436	0.424	0.343	0.472	64.15555	655.3268	571.8673	8.515896	905.7052
8D	7	8	8	8	8	0.356	0.43	0.441	0.504	0.493	98.93033	713.2986	789.8031	1227.965	1151.461

4.1.2 QSPR Model

In an ethylene/1-hexene copolymerization study, the electron-donating methyl group initially increases copolymer molecular weight. However, when eight or ten methyl groups are present (7 and 8), steric effects overcome the electronic effects resulting in lowered copolymer molecular weight. The hexene incorporation data suggest that the more heavily substituted catalysts exhibit a steric effect which lowers the amount of hexene incorporation. A low degree of methyl substitution seems to enhance comonomer incorporation, however, the electronic effect is not straightforward.

4.1.2.1 Correlation Analysis

Totally 33 descriptors were used for QSPR fitting. Values of 33 descriptors were given in the appendix (**Table 1A**). Only 10 descriptors with high correlation to $\bar{M}_n$ and %1-hexene incorporation were listed in **Table 4.5-4.6** These descriptors were further fitted to yield QSPR models.

Table 4.5 Descriptors with high correlation to $\bar{M}_n$.

Descriptor	r
Cluster number	0.87
Total dipole	0.82
Molecular density	-0.80
Cp ring Up Charge (ESP)	0.79
Cp ring Down Charge (ESP)	0.77
Principal moment of inertia X	0.75
Binding energy	-0.75
Total energy	-0.75
Element count	0.75
Total molecular mass	0.75

Table 4.6 Descriptors with high correlation to 1-hexene incorporation.

Descriptor	r
Cp ring Up Charge (Mulliken)	-0.85
Cp-Zr-Cp Angle	0.72
Zr Charge (Hirshfeld)	0.65
Dipole y	-0.60
LUMO-HOMO energy	0.56
Rank in cluster	-0.54
HOMO energy	-0.53
LUMO energy	-0.49
Cluster number	-0.41
Principal moment of inertia X	-0.41

4.1.2.2 PCA Analysis

The dot in the loading plot represents the projection of the variables (33 descriptors) on the PC1–PC2 plane as shown in **Figure 4.4**. Total variances of PC axis are 81.38% and 7.47% which corresponding to PC1 and PC2 respectively. The two variances are closed to 100% which refers to suitable information visualized from the data set.

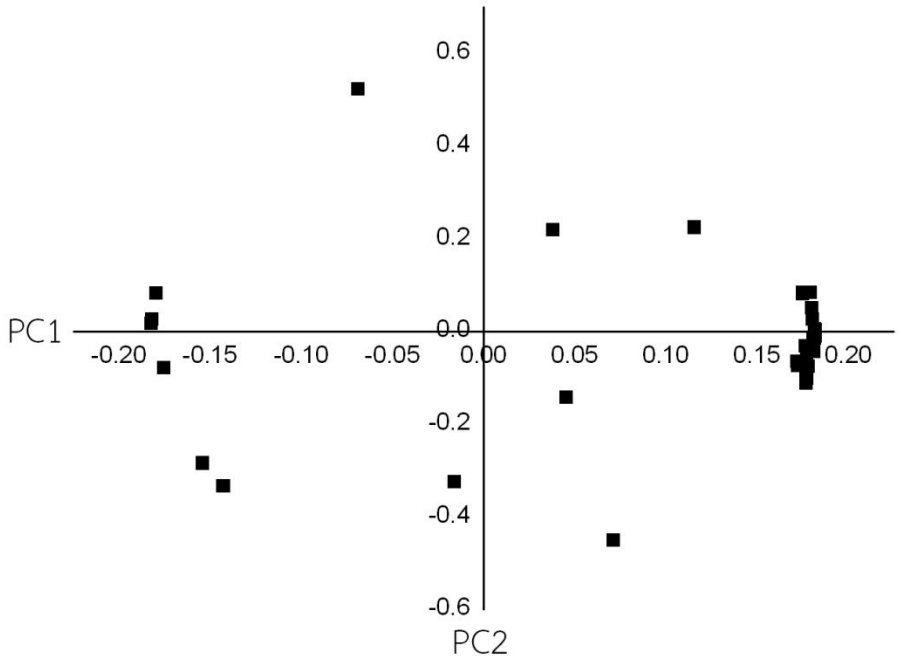


Figure 4.4 PCA loadings of QSPR for zirconocene catalyst 1-8 in group 1.

From the loading plot, it can be seen that there is apparent clusters of descriptors. Thus, the descriptors in the same clusters (these descriptors are highly inter-correlated) should be omitted for further QSAR analysis. The total energy and binding energy descriptors are obvious the same clusters of descriptors thus two descriptors should not be omitted or selected only one descriptor for fitting the QSAR model.

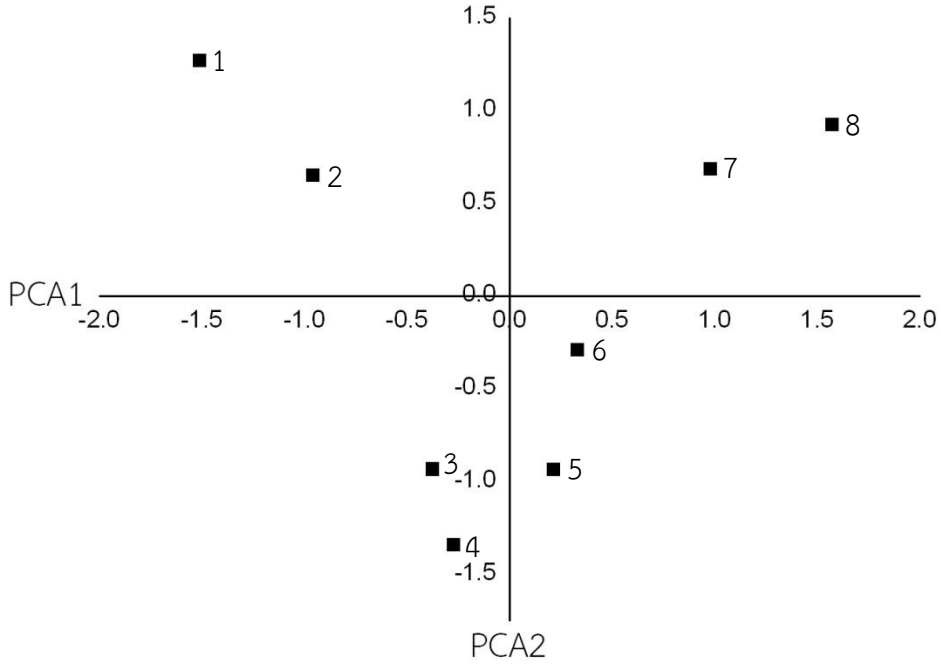


Figure 4.5 PCA scores of QSPR for zirconocene catalyst 1-8 in group 1.

The dot in the scores plot represents each zirconocene catalysts of group 1 is presented in **Figure 4.5**. From the score plot, it can be seen that there is obvious cluster of zirconocene catalyst on the score plots. This suggests the structural information of zirconocene catalyst 1, 2, 7, and 8 are different from the catalyst 3-5. The results from the score plot is obviously unrelated due to the difference structures of zirconocene catalyst, hence the further QSAR analysis will be performed.

4.1.2.3 Model

The descriptors after PCA analysis was fitted to the QSPR models by MLR procedure. The best models obtained from MLR fit of one and two descriptors from **Table 4.5** and **Table 4.6** for $\overline{M}_n$ and %1-hexene incorporation were given in **Table 4.7** and **Table 4.9** respectively.

Table 4.7 QSPR models in $\overline{M}_n$ from MLR analysis.

No.	Model	r^2	q^2
1	Y = 80257.55 * Total dipole - 111378.73	0.67	0.63
2	Y = 57480.69 * Cp ring Up Charge (ESP) + 60923.92	0.62	0.51
3	Y = 57276.55 * Total dipole + 26.78 * Principal moment of inertia X - 103196.28	0.88	0.86
4	Y = 19.02 * Cp ring Up Charge (ESP) + 29140.05 * Principal moment of inertia X + 27254.43	0.60	0.44
5	Y = -2524.58 * Cp-Zr-Cp Angle + 78126.28 * Total dipole + 213429.03	0.91	0.87

From **Table 4.7**, two descriptors models show higher correlation with $\overline{M}_n$. The model (5) gives correlation coefficient r^2 of 0.91 and q^2 of 0.87. And it is selected to represent the QSPR model.

Values of Cp-Zr-Cp angle and total dipole moment together with $\overline{M}_n$ were list in **Table 4.8**.

Table 4.8 Calculated properties and activity of QSPR model in $\overline{M}_n$.

No.	Cp-Zr-Cp Angle	Total dipole	$\overline{M}_n$
1.	128.948	1.665316	11000
2.	129.146	1.688157	25500
3.	130.726	1.99795	34100
4.	130.564	2.199438	57200
5.	129.961	2.022165	51600
6.	124.816	2.218295	65200
7.	120.839	2.036801	73000
8.	121.053	1.820622	47300

From the model, Cp-Zr-Cp angle representing steric effect and total dipole representing electronic effect. Contribute to the property, the large values of total dipole with smaller values of Cp-Zr-Cp angle constitute to larger $\overline{M}_n$. The tri/tetra methyl substitutions have the largest Cp-Zr-Cp angle and total dipole. The optimum condition is for compound 7 (Tetra-substitution).

Table 4.9 QSPR models in 1-hexene incorporation from MLR analysis.

No.	Model	r^2	q^2
1	Y = 0.19 * Cp-Zr-Cp Angle - 3.49 * Cp ring Up Charge (Mulliken) - 23.16	0.83	0.72
2	Y = 0.14 * Cp-Zr-Cp Angle + 16.58 * Zr Charge (Hirshfeld) - 22.09	0.47	0.47
3	Y = 0.24 * Cp-Zr-Cp Angle - 7.22 * Dipole y - 27.24	0.74	0.71
4	Y = 0.15 * Cp-Zr-Cp Angle + 70.88 * LUMO-HOMO energy - 24.53	0.45	0.41
5	Y = 0.15 * Cp-Zr-Cp Angle - 31.23 * HOMO energy - 23.29	0.43	0.40
6	Y = 0.15 * Cp-Zr-Cp Angle - 52.47 * LUMO energy - 22.48	0.41	0.39

From **Table 4.8**, model (1) show the best fit for %1-hexene incorporation with correlation coefficient r^2 of 0.83 and q^2 of 0.72. And it is selected to represent the QSPR model.

Values of Cp-Zr-Cp angle and Cp ring up charge (Mulliken) together with %1-hexene incorporation were listed in **Table 4.10**.

Table 4.10 Calculated properties and activity of QSPR model in 1-hexene incorporation.

No.	Cp-Zr-Cp Angle	Cp ring Up Charge (Mulliken)	1-hx incorp[a]
1.	128.948	-0.6	2.7
2.	129.146	-0.534	2.5
3.	130.726	-0.48	4.0
4.	130.564	-0.507	1.8
5.	129.961	-1.185	5.5
6.	124.816	-0.43	1.7
7.	120.839	-0.416	1.4
8.	121.053	-0.338	0.3

[a] mol%

From the model, Cp-Zr-Cp angle representing steric effect and cyclopentadienyl upper ring charge (Mulliken) representing electronic effect. The large values of Cp-Zr-Cp angle and the high negative values of Cp ring up charge (Mulliken) constitutes the higher value of %1-hexene incorporation. The compound 5 has large Cp-Zr-Cp angle and high negative Cp ring up charge (Mulliken) and hence demonstrating the highest %1-hexene incorporation. The Cp-Zr-Cp depends on the

degree of methyl substitution where tri/tetra substitution gives the largest Cp-Zr-Cp angle. The Cp ring up charge (Mulliken) also depends on the degree of substitution. However, the results are varied.

4.2 Group 2

4.2.1 QSAR Model

In an ethylene/1-hexene copolymerization study, the catalysts with Phenyl substituent group is the most active catalysts (9 and 10), the activities decrease with increasing C_6F_5 functionalization (11, 12 and 13).

4.2.1.1 Correlation Analysis

Totally 34 descriptors were used for QSAR fitting. Values of 34 descriptors were given in the appendix (**Table 2A**). Only 10 descriptors with high correlation to activity were listed in **Table 4.11** These descriptors were further fitted to yield QSAR models.

Table 4.11 Descriptors with high correlation to activity.

Descriptor	r
Molecular density	-0.99
LUMO energy	0.99
HOMO energy	0.97
Zr Charge (Hirshfeld)	0.95
Cp ring Up Charge (ESP)	-0.91
Total energy	0.89
LUMO-HOMO energy	0.89
Cluster number	-0.86
Dipole x	-0.83
Total molecular mass	-0.81

4.2.1.2 PCA Analysis

The dot in the loading plot represents the projection of the variables (34 descriptors) on the PC1–PC2 plane as shown in **Figure 4.6**. Total variances of PC axis are 48.73% and 26.77% which corresponding to PC1 and PC2 respectively. The two variances are closed to 100% which refers to suitable information visualized from the data set.

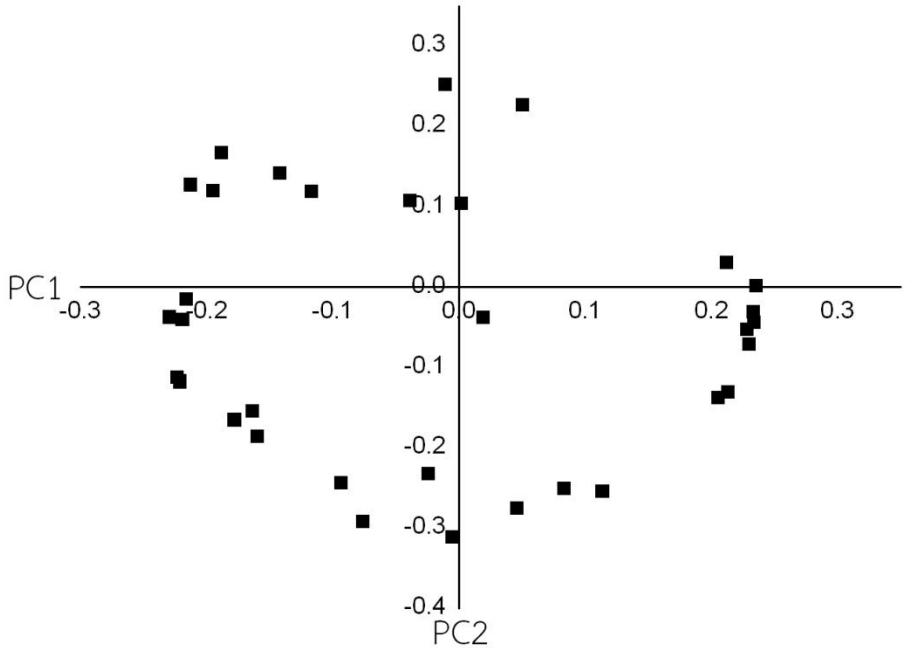


Figure 4.6 PCA loadings of QSAR for zirconocene catalyst 9-13 in group 2.

From the loading plot, it can be seen that there is apparent clusters of descriptors. Thus, the descriptors in the same clusters (these descriptors are highly inter-correlated) should be omitted for further QSAR analysis. The total molecular mass, principal moment of inertia X, and Cp ring up charge (ESP) descriptors are obvious the same clusters of descriptors thus two descriptors should not be omitted or selected only one descriptor for fitting the QSAR model.

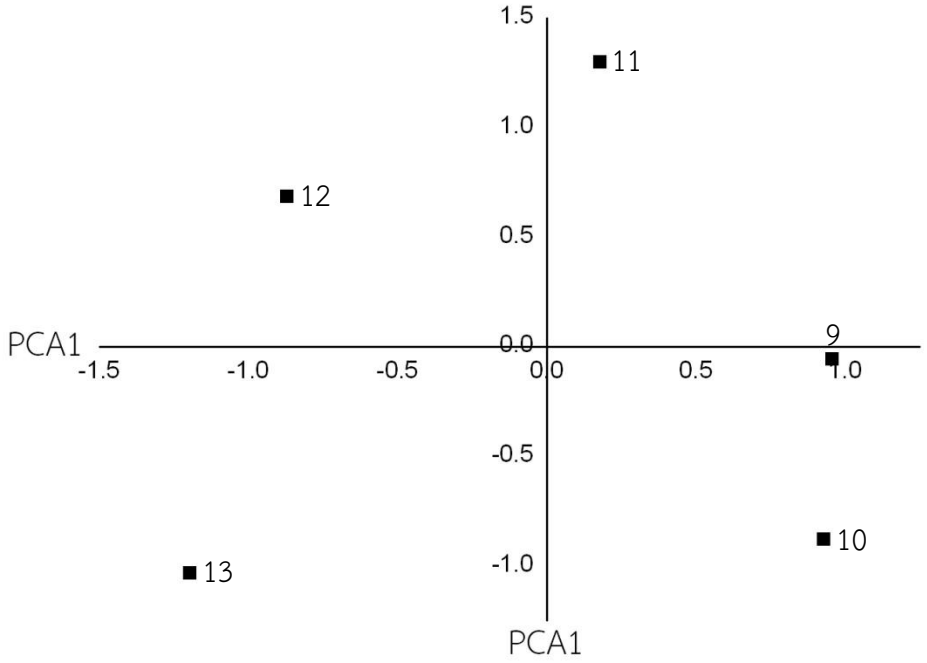


Figure 4.7 PCA scores of QSAR for zirconocene catalyst 9-13 in group 2.

The dot in the scores plot represents each zirconocene catalysts of group 2 is presented in **Figure 4.7**. From the score plot, it can be seen that there is obvious cluster of zirconocene catalyst on the score plots. This suggests the structural information of zirconocene catalyst 1 is different from other catalyst (2-5). The results from the score plot is obviously unrelated due to the difference structures of zirconocene catalyst, hence the further QSAR analysis will be performed.

4.2.1.3 Model

The descriptors after PCA analysis was fitted to the QSAR models by MLR procedure. The best model obtained from MLR fit were given in **Table 4.12**.

Table 4.12 QSAR models from MLR analysis.

No.	Model	r^2	q^2
1	Y = 540.61 * LUMO energy + 0.00071 * Principal moments of inertia + 91.02	0.99	0.96
2	Y = 531.74 * LUMO energy + 0.0010 * Principal moment of inertia Y + 90.18	0.99	0.95
3	Y = 531.77 * LUMO energy + 0.00089 * Principal moment of inertia Z + 90.27	0.99	0.95
4	Y = 433.70 * LUMO energy + 80.89	0.97	0.97
5	Y = -0.0020 * Principal moments of inertia + 32.84	0.64	0.56
6	Y = -0.0031 * Principal moment of inertia Y + 32.50	0.62	0.55
7	Y = -0.0028 * Principal moment of inertia Z + 32.40	0.63	0.56

The model (1), (2), and (3) are similar models all with correlation coefficient r^2 of 0.99 and $q^2 > 0.95$. The model (1) was selected prediction of the activities.

Values of LUMO energy and principal moments of inertia (PMI) together with activity data were given in **Table 4.13**.

Table 4.13 Calculated properties and activity of QSAR model.

No.	LUMO energy	Principal moments of inertia	Activity[a]
9.	-0.12722	5127.16	26
10.	-0.12822	5950.549	26
11.	-0.15288	6148.823	12
12.	-0.16512	9832.758	10
13.	-0.17231	13600.1	7

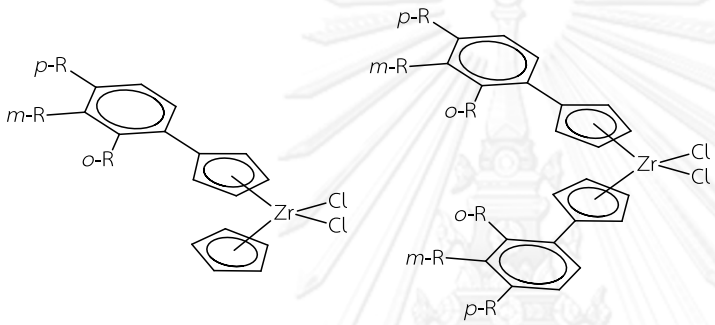
[a] Catalyst activity in $kgPEH/mol_{Zr}.h$

From the model, the principal moments of inertia representing the steric effect and LUMO energy representing the electronic effect. Larger values of LUMO energy (less negative) and principal moments of inertia cause higher activity. With phenyl substitution on cyclopentadienyl ring, the catalysts yields the higher LUMO energy but the lower principal moments of inertia. However, with pentafluorophenyl substitution the catalysts yields the lower LUMO energy (more negative) and the

higher principal moments of inertia. With bulky electron donating group on phenyl ring one can raise the LUMO energy and enhance principal moments of inertia.

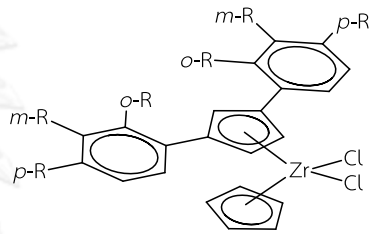
4.2.1.4 Catalyst Design

From the information in **4.2.1.3**, we took 3 structures from the experimental data (11-13) and modified their structures with electron donating groups on the phenyl or pentafluorophenyl ring. The modified structures were named according to their parents i.e. 11D, 12D, and 13D.

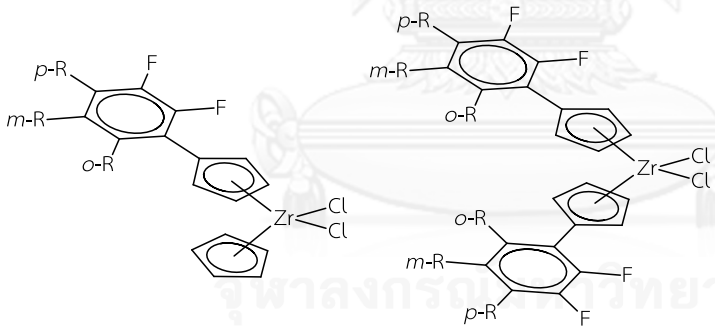


11D on phenyl

12D on phenyl

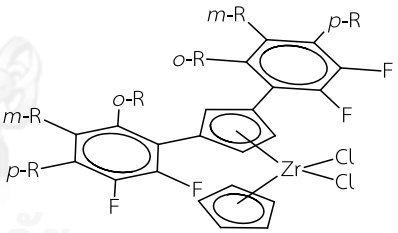


13D on phenyl



11D on C_6F_5

12D on C_6F_5



13D on C_6F_5

Figure 4.8 Scheme of mono-substitution in group 2.

4.2.1.4.1 Mono-substitution

Ortho, meta and para substitutions on phenyl and pentafluorophenyl ring with methyl, iso-propyl and tertiary-butyl group were generated. LUMO energy and principal moments of inertia of the compounds were computed. Calculated properties and activities were predicted with model (1) of design zirconocene catalysts for mono-substitution are shown in **Table 4.14**.

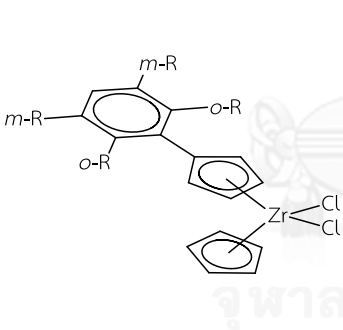
Table 4.14 Calculated properties and prediction activity of design zirconocene catalyst for mono-substituent in group 2.

No.	LUMO energy				Principal moments of inertia				Predictive Activity			
	$Me(C_6F_5)$	Me(Ph)	^{i}Pr(Ph)	^{t}Bu(Ph)	$Me(C_6F_5)$	Me(Ph)	^{i}Pr(Ph)	^{t}Bu(Ph)	$Me(C_6F_5)$	Me(Ph)	^{i}Pr(Ph)	^{t}Bu(Ph)
11D-o	-0.120	-0.118			6300.282	3893.257			30.604	29.882		
11D-m	-0.120	-0.117			6200.104	4146.963			30.477	31.004		
11D-p	-0.120	-0.116	-0.114	-0.112	6164.916	4285.813	5980.933	7029.152	30.428	31.339	33.250	35.289
12D-o	-0.130	-0.115			11277.41	5913.77			28.270	32.900		
12D-m	-0.125	-0.113			10564.65	6365.113			30.622	34.406		
12D-p	-0.125	-0.112	-0.112	-0.111	9814.849	6618.061	10409.97	11892.86	30.346	34.934	37.659	39.343
13D-o	-0.130	-0.119			12809.51	6480.01			29.721	30.858		
13D-m	-0.124	-0.117			12479.2	7110.353			33.079	32.536		
13D-p	-0.124	-0.116	-0.116	-0.115	12363.94	7724.365	12370.51	14529.82	32.721	33.787	37.104	38.939

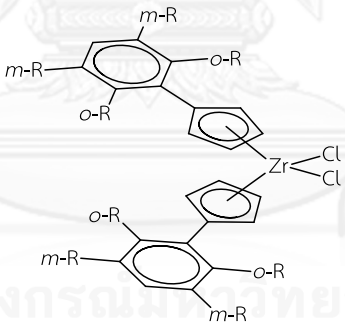
From the group 2 QSAR model, LUMO energy showed more influence to the catalyst activity than the principal moments of inertia. So we tried to increases LUMO energy by substituent with more electron donating groups on phenyl and pentafluorophenyl ring. See **Table 4.14** when substitution at the meta position with tert-butyl results to LUMO energy and principal moments of inertia increased. So the predicted activity when substitution with tert-butyl on phenyl ring gives the best activity (t-Bu>i-Pr>Me). For ortho and para substitution, the similar trend in catalyst activity were observed but the increase in value of properties due to the meta substitution is more obvious.

4.2.1.4.2 Di-substitution

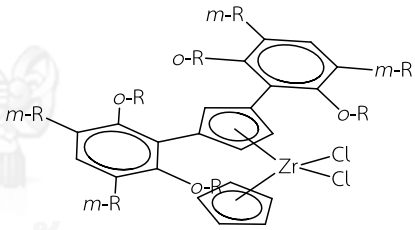
Ortho and meta substitutions on phenyl and pentafluorophenyl ring with methyl group were generated. LUMO energy and principal moments of inertia of the compounds were computed. Calculated properties and predicted activities was predicted with model (1) of design zirconocene catalysts were shown in **Table 4.15**.



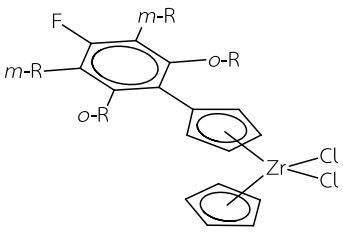
11D on phenyl



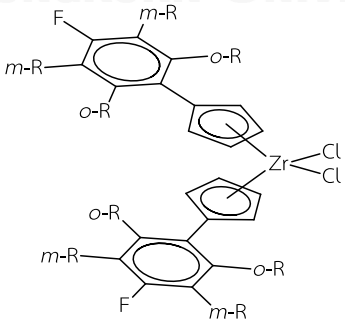
12D on phenyl



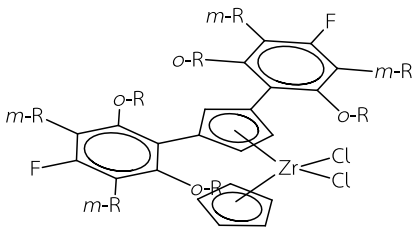
13D on phenyl



11D on C_6F_5



12D on C_6F_5



13D on C_6F_5

Figure 4.9 Scheme of di-substitution in group 2.

Table 4.15 Calculated properties and prediction activity of design zirconocene catalyst for di-substituent in group 2.

No.	LUMO energy		Principal moments of inertia		Predictive Activity	
	C_6F_5	Ph	C_6F_5	Ph	C_6F_5	Ph
11D-o,o	-0.12305	-0.1138	6534.482	4323.575	29.16678	32.59189
11D-m,m	-0.11475	-0.11475	6102.44	4596.202	33.34507	32.27343
12D-o,o	-0.13065	-0.11519	11009.68	7263.846	28.25623	33.93904
12D-m,m	-0.11613	-0.11196	9758.385	7265.562	35.21237	35.6881
13D-o,o	-0.13086	-0.11459	12672.6	7129.608	29.33208	34.16781
13D-m,m	-0.11488	-0.11534	12173.35	8573.231	37.61284	34.79566

All compounds have better activities than those in **Table 4.14**. From **Table 4.15** when methyl substitution on phenyl and pentafluorophenyl ring, results to LUMO energy increased. So predicted activity were better when methyl substitution on phenyl ring more than methyl substitution on pentafluorophenyl ring. Both LUMO energy and the principal moments of inertia values are improved (Become larger). The compound 12D-m,m give the highest activity. The activity when di substitution was better than mono substitution cause di substitution have more electron donating group than mono substitution.

4.2.1.4.3 Tri-substitution

Substitution at 2,4,6 and 3,4,5 position on phenyl ring were generated. LUMO energy and principal moments of inertia of the compounds were computed. Calculated properties and predicted activities according to model (1) of design zirconocene catalysts were shown in **Table 4.16**.

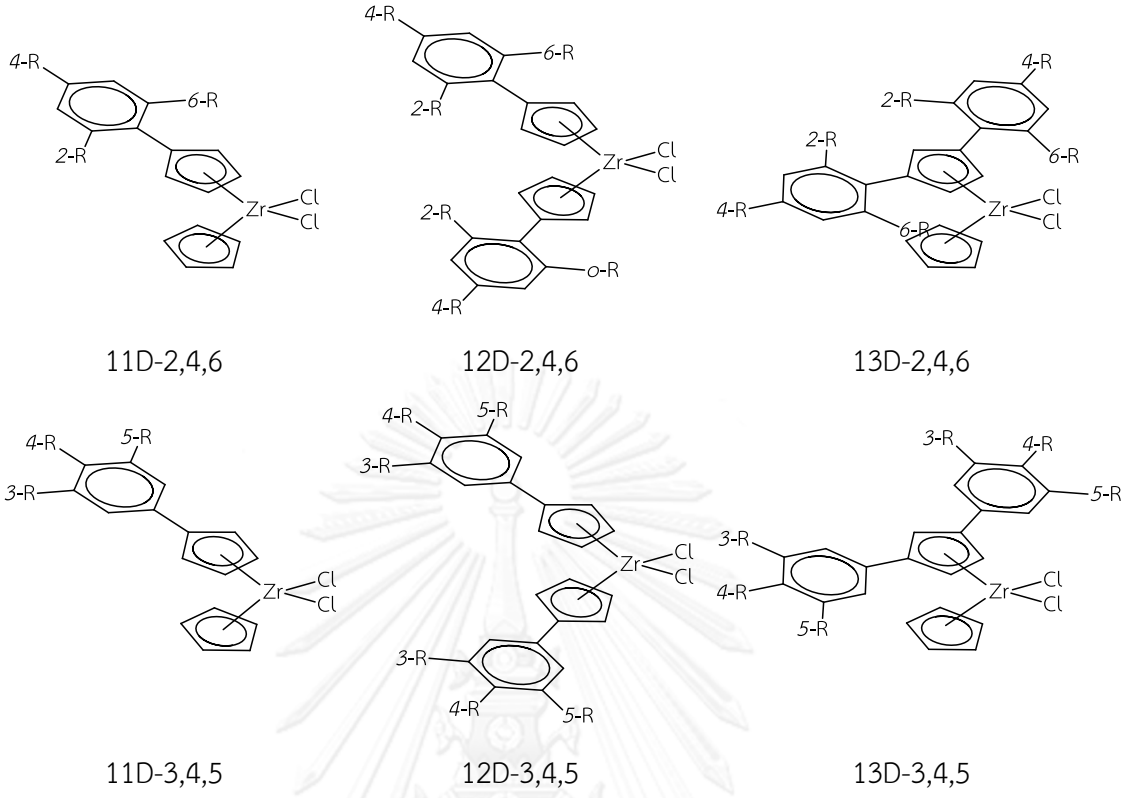


Figure 4.10 Scheme of tri-substitution in group 2.

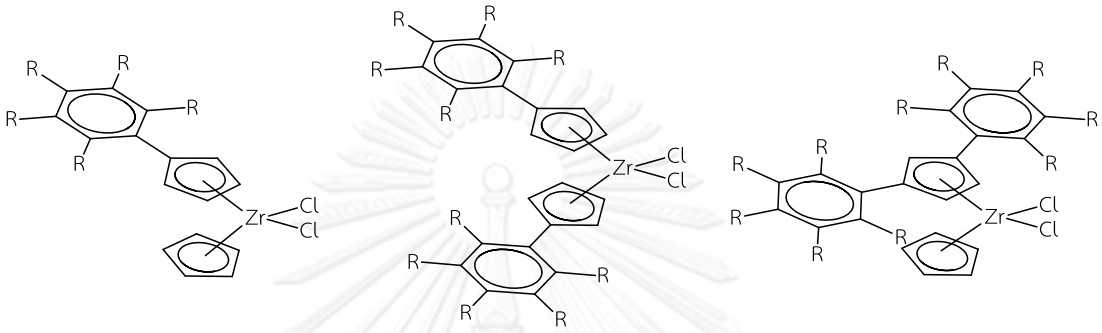
Table 4.16 Calculated properties and prediction activity of design zirconocene catalyst for tri-substituent in group 2.

No.	LUMO energy	Principal moments of inertia	Predictive Activity
11D-2,4,6	-0.11406	4989.791	32.92522
11D-3,4,5	-0.10961	5474.962	35.67731
12D-2,4,6	-0.11385	9131.181	35.99904
12D-3,4,5	-0.10749	8937.003	39.29753
13D-2,4,6	-0.11202	9051.616	36.92783
13D-3,4,5	-0.11196	10458.71	37.96801

From **Table 4.16** the best activity is from catalyst 12D. Substitution at 3,4,5-position gives better activity than 2,4,6-substitution. LUMO energy increases more for 3-position substitution than the 2-position.

4.2.1.4.4 Full Ring-substitution

Full substitution on the phenyl ring with methyl group were created. LUMO energy and principal moments of inertia of the compounds were computed. Calculated properties and predicted activity according to model (1) of design zirconocene catalysts were shown in **Table 4.17**.



11D 12D 13D

Figure 4.11 Scheme of full ring substitution in group 2.

Table 4.17 Calculated properties and prediction activity of design zirconocene catalyst for full ring in group 2.

No.	LUMO energy	Principal moments of inertia	Predictive Activity
11D	-0.10489100	6233.38	38.77067
12D	-0.09855900	9429.85	44.47793
13D	-0.10379800	11835.97	43.36505

With full methyl substitution, LUMO energy and principal moments of inertia are higher. The compound 12D has the highest activity among all group 2 compounds studied.

4.2.2 QSPR Model

The copolymer molecular weights decrease with increasing C_6F_5 functionalization (11, 12 and 13). The phenyl substituted catalysts (9 and 10) produce higher molecular weight copolymers than the other catalysts. The degree of 1-hexene incorporation for catalyst 9, 10 and 11 reveal no significant substituent effects.

4.2.2.1 Correlation Analysis

Totally 34 descriptors were used for QSPR fitting. Values of 34 descriptors were given in the appendix (**Table 2A**). Only 10 descriptors with high correlation to $\bar{M}_w$ and %1-hexene incorporation were listed in **Table 4.18-4.19**. These descriptors were further fitted to yield QSPR models.

Table 4.18 Descriptors with high correlation in $\bar{M}_w$ to activity.

Descriptor	r
LUMO energy	0.99
Zr Charge (Hirshfeld)	0.99
HOMO energy	0.97
Molecular density	-0.96
Total energy	0.95
Phenyl	0.94
Cp ring Up Charge (ESP)	-0.90
LUMO-HOMO energy	0.89
Total molecular mass	-0.89
Principal moments of inertia	-0.84

Table 4.19 Descriptors with high correlation in 1-hexene incorporation to activity.

Descriptor	r
Principal moment of inertia Z	0.99
Principal moment of inertia Y	0.99
Principal moments of inertia	0.99
LUMO-HOMO energy	-0.91
Cp ring Up Charge (ESP)	0.91
Rank in cluster	0.91
Similarity to cluster reference	0.91
Total molecular mass	0.88
Total energy	-0.87
Zr Charge (Hirshfeld)	-0.83

4.2.2.2 PCA Analysis

The dot in the loading plot represents the projection of the variables (34 descriptors) on the PC1–PC2 plane as shown in **Figure 4.12**. Total variances of PC axis are 50.8% and 25.46% which corresponding to PC1 and PC2 respectively. The two variances are closed to 100% which refers to suitable information visualized from the data set.

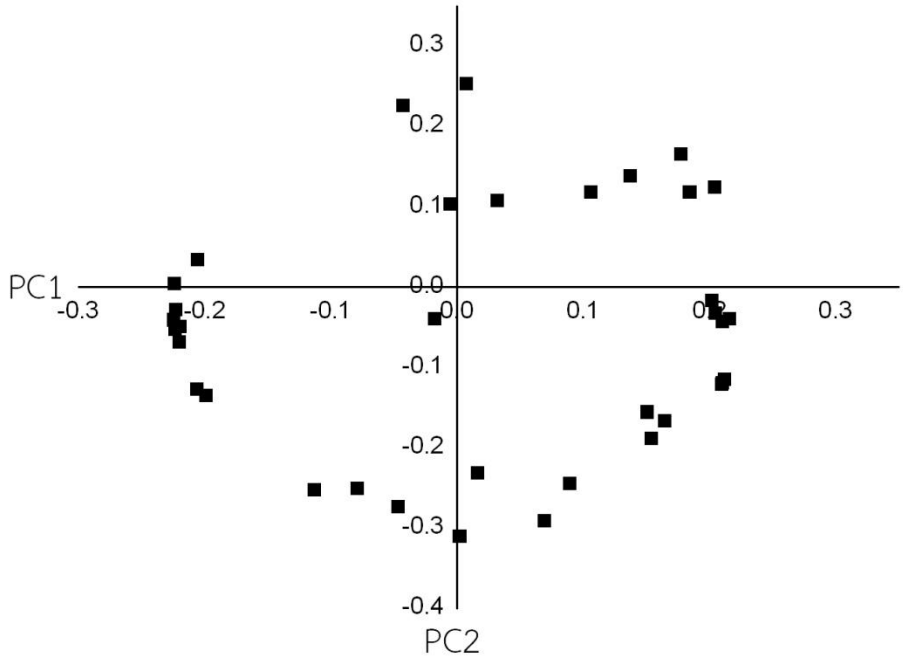


Figure 4.12 PCA loadings of QSPR for zirconocene catalyst 9-13 in group 2.

From the loading plot, it can be seen that there is apparent clusters of descriptors. Thus, the descriptors in the same clusters (these descriptors are highly inter-correlated) should be omitted for further QSAR analysis. The total molecular mass and Zr charge (Hirshfeld) descriptors are obvious the same clusters of descriptors thus two descriptors should not be omitted or selected only one descriptor for fitting the QSAR model.

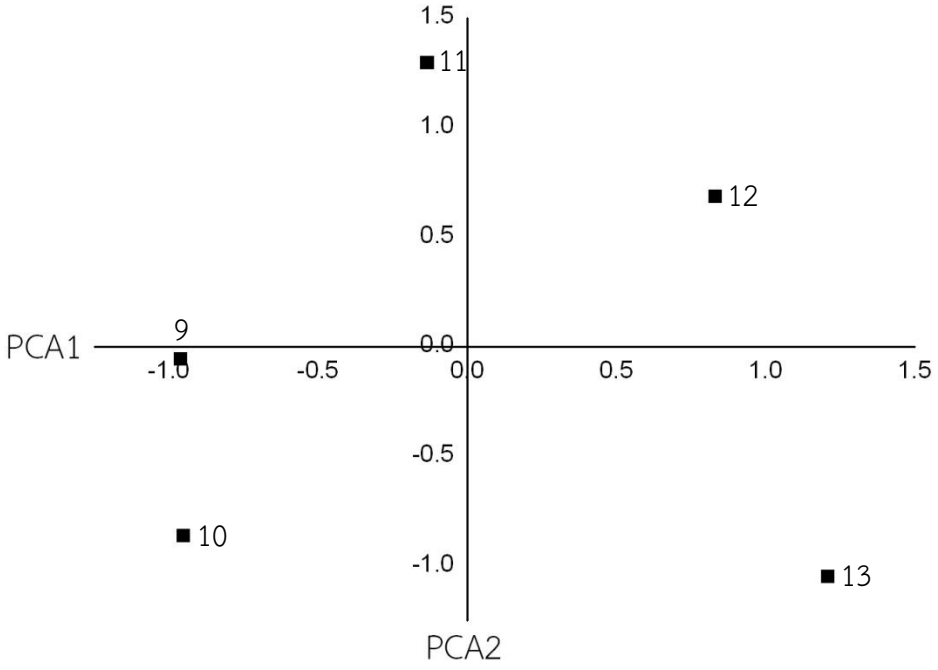


Figure 4.13 PCA scores of QSPR for zirconocene catalyst 9-13 in group 2.

The dot in the scores plot represents each zirconocene catalysts of group 2 is presented in **Figure 4.13**. From the score plot, it can be seen that there is obvious cluster of zirconocene catalyst on the score plots. This suggests the structural information of zirconocene catalyst 1 and 2 are different from other catalysts. The results from the score plot is obviously unrelated due to the difference structures of zirconocene catalyst, hence the further QSAR analysis will be performed.

4.2.2.3 Model

The descriptors after PCA analysis was fitted to the QSPR models for $\bar{M}_w$ and %1-hexene incorporation by MLR procedure. The best models obtained from the MLR fit for $\bar{M}_w$ and %1-hexene incorporation are given in **Table 4.20** and **Table 4.22** respectively.

Table 4.20 QSPR models in $\overline{M}_w$ from MLR analysis.

No.	Model	r^2	q^2
1	Y = 2740927.81 * LUMO energy + 49289.44 * Phenyl + 502080.53	0.96	0.94
2	Y = 5319911.23 * LUMO energy + 926169.98	0.98	0.93
3	Y = 38454268.29 * Zr Charge (Hirshfeld) - 16102692.07	0.97	0.96
4	Y = 50576.75 * Phenyl + 20329880.86 * Zr Charge (Hirshfeld) - 8491037.10	0.98	0.98
5	Y = 3624434.44 * HOMO energy + 49098.28 * Phenyl + 879362.25	0.94	0.92
6	Y = 52320.81 * Phenyl - 325466.08 * Cp ring Up Charge (ESP) - 70522.73	0.92	0.85
7	Y = 2855483.97 * LUMO energy - 14.22 * Principal moments of inertia + 674253.47	0.91	0.72
8	Y = 3902773.41 * HOMO energy - 14.64 * Principal moments of inertia + 1098074.55	0.92	0.80
9	Y = 8493221.88 * LUMO-HOMO energy - 13.99 * Principal moments of inertia - 328463.84	0.80	0.48
10	Y = - 0.72 * Principal moments of inertia + 37694018 * Zr Charge (Hirshfeld) - 15775843.77	0.97	0.93

The models (1) - (4) yielded the best statistic with correlation coefficient r^2 and $q^2 > 0.9$. The model (3) was selected for the prediction of $\overline{M}_w$ since it is a one descriptor model and Zr charge can easily be modified.

Values of Zr charge (Hirshfeld) for compounds 9-13 were computed and listed together with experimental $\overline{M}_w$ in **Table 4.21**.

Table 4.21 Calculated properties and activity of QSPR model in $\overline{M}_w$.

No.	Zr Charge (Hirshfeld)	$\overline{M}_w$
9.	0.424	230000
10.	0.426	265000
11.	0.422	114000
12.	0.420	37500
13.	0.419	17000

Only the electronic effect is important in this case. Perhaps because all compounds contain steric groups. There is no obvious difference in steric between compounds. The $\bar{M}_w$ will increase as the Zr charge (Hirshfeld) becomes more positive. Thus, without electron withdrawing group on the phenyl ring the Zr charge (Hirshfeld) become more positive and larger $\bar{M}_w$ is obtained.

Table 4.22 QSPR models in 1-hexene incorporation from MLR analysis.

No.	Model	r^2	q^2
1	Y = 5.68 * Dipole y + 0.0016 * Principal moments of inertia - 3.40	0.99	0.67
2	Y =5.68 * Dipole y + 0.0025 * Principal moment of inertia Y - 3.28	0.99	0.66
3	Y = 5.67 * Dipole y + 0.0022 * Principal moment of inertia Z - 3.10	0.99	0.66
4	Y = -450.69 * LUMO-HOMO energy + 0.0012 * Principal moment of inertia Z + 33.25	0.95	0.95
5	Y = 0.0024 * Principal moment of inertia Z - 3.26 * Cp ring Up Charge (ESP) - 6.14	0.98	0.96
6	Y = -0.0053 * Total energy + 0.0012 * Principal moment of inertia Z - 27.74	0.92	0.83
7	Y = -452.35 * LUMO-HOMO energy + 0.0013 * Principal moment of inertia Y + 33.26	0.96	0.95
8	Y = 0.0026 * Principal moment of inertia Y - 1.69 * Cp ring Up Charge (ESP) - 4.96	0.98	0.95
9	Y = -0.0053 * Total energy + 0.0013 * Principal moment of inertia Y - 27.81	0.92	0.83
10	Y = -452.96 * LUMO-HOMO energy + 0.00085 * Principal moments of inertia + 33.23	0.96	0.95
11	Y = 0.00084 * Principal moments of inertia + 16.43 * Cp ring Up Charge (ESP) + 10.82	0.94	0.92
12	Y = -0.0053 * Total energy + 0.00085 * Principal moments of inertia - 27.58	0.91	0.81

There are 6 models with correlation coefficient r^2 and $q^2 > 0.9$. The model (5), (8), and (11) contain similar descriptors. The model (11) were selected since principal moments of inertia (PMI) is more easily modified.

Values of principal moments of inertia and Cp ring up charge (ESP) were calculated for compound 9-13 and their value together with %1-hexene incorporation showed in **Table 4.23**.

Table 4.23 Calculated properties and activity of QSPR model in 1-hexene incorporation.

No.	Principal moments of inertia	Cp ring Up Charge (ESP)	1-hx incorp[a]
9.	5127.16	-0.665	5.9
10.	5950.549	-0.632	5.4
11.	6148.823	-0.476	6.0
12.	9832.758	-0.464	11.3
13.	13600.1	-0.242	19.0

[a] mol%

The principal moments of inertia representing steric effect and cyclopentadienyl upper ring charge (ESP) representing electronic effect. From the model (11) %1-hexene incorporation becomes larger and/or cyclopentadienyl upper ring charge (ESP) less negative. The principal moments of inertia are smaller for catalysts with phenyl substitution while the larger principal moments of inertia was observed for catalysts with pentafluorophenyl substitution. Similarly for the cyclopentadienyl upper ring charge (ESP), its value is less negative as a result of the electron withdrawing group substituted on the phenyl ring (H vs. F). The compound 13 yielded the largest %1-hexene incorporation.

4.3 Group 3

4.3.1 QSAR Model

In all cases, *rac* compounds present higher activities in copolymerization than their *meso* analogs. *Rac* mono-substituted metallocene with a propyl chain (16) proved to be the most active catalyst. For *rac* and *meso* compounds, lower activity was obtained when different chains (ethyl and propyl) were used to substitute the two indenyl rings. Thus, both the size and distribution of the alkyl substituents influence the resulting activity.

4.3.1.1 Correlation Analysis

Totally 35 descriptors were used for QSAR fitting. Values of 35 descriptors were given in the appendix (**Table 3A**). Only 10 descriptors with high correlation to activity were listed in **Table 4.24**. These descriptors were further fitted to yield QSAR models.

Table 4.24 Descriptors with high correlation to activity.

Descriptor	r
LUMO-HOMO energy	0.84
Zr Charge (Hirshfeld)	0.78
Dipole z	0.72
Dipole y	0.71
Rank in cluster	0.67
Dipole x	-0.66
Principal moment of inertia Z	0.53
Cp-Zr-Cp Angle	-0.48
LUMO energy	0.47
Cp ring Up Charge (Hirshfeld)	0.44

4.3.1.2 PCA Analysis

The dot in the loading plot represents the projection of the variables (35 descriptors) on the PC1–PC2 plane as shown in **Figure 4.14**. Total variances of PC axis are 65.67% and 19.21% which corresponding to PC1 and PC2 respectively. The two variances are closed to 100% which refers to suitable information visualized from the data set.

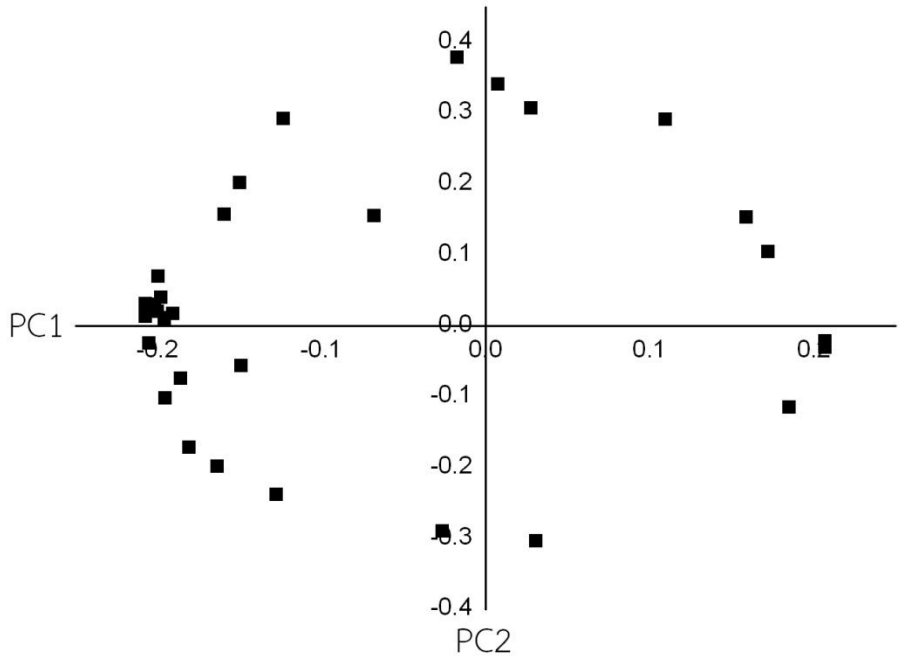


Figure 4.14 PCA loadings of QSAR for zirconocene catalyst 14-19 in group 3.

From the loading plot, it can be seen that there is apparent clusters of descriptors. Thus, the descriptors in the same clusters (these descriptors are highly inter-correlated) should be omitted for further QSAR analysis. The total energy and molecular density descriptors are obvious the same clusters of descriptors thus two descriptors should not be omitted or selected only one descriptor for fitting the QSAR model.

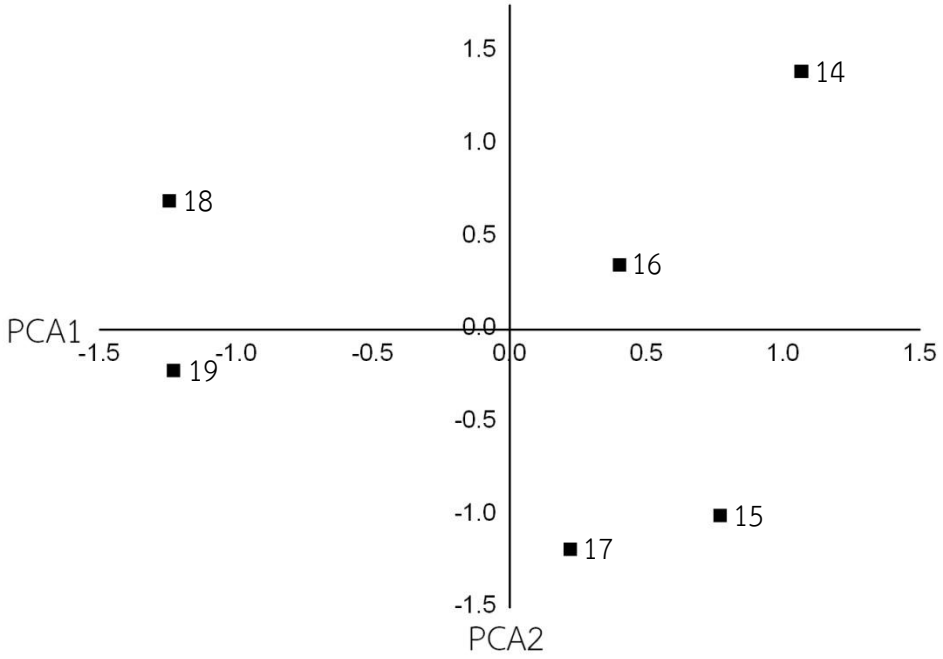


Figure 4.15 PCA scores of QSAR for zirconocene catalyst 14-19 in group 3.

The dot in the scores plot represents each zirconocene catalysts of group 3 is presented in **Figure 4.15**. From the score plot, it can be seen that there is obvious cluster of zirconocene catalyst on the score plots. This suggests the structural information of zirconocene catalyst 1, 3, and 5 of group 3 are different from the catalyst 2, 4, and 6. The results from the score plot is obviously unrelated due to the difference structures of zirconocene catalyst, hence the further QSAR analysis will be performed.

4.3.1.3 Model

The descriptors after PCA analysis was fitted to the QSAR models by MLR procedure. The best models obtained from MLR fit were given in **Table 4.25**.

Table 4.25 QSAR models from MLR analysis.

No.	Model	r^2	q^2
1	Y = 5636.92 * Dipole y + 13.13 * Molecular volume - 4394.74	0.97	0.90
2	Y = - 419.73 * Binding energy + 5623.79 * Dipole y - 3758.15	0.97	0.90
3	Y = 3619.11 * Dipole y + 0.55 * Principal moment of inertia Z - 1392.02	0.97	0.57
4	Y = 6866.60 * Cp ring Down Charge (Mulliken) + 7295.08 * Dipole y + 7562.85	0.97	0.94
5	Y = 5611.97 * Dipole y + 15.78 * Total molecular mass - 7087.70	0.97	0.90
6	Y = 221.41 * Element count + 5611.97 * Dipole y - 4435.37	0.97	0.90
7	Y = 73.80 * Atom count + 5611.97 * Dipole y - 3106.92	0.97	0.90
8	Y = 5769.84 * Dipole y + 10.95 * Molecular area - 4271.56	0.97	0.90
9	Y = 16821.05 * Cp ring Down Charge (ESP) + 7081.808426297 * Dipole y + 8050.50	0.97	0.95
10	Y = 4718.41 * Dipole y + 0.47 * Principal moments of inertia - 1942.09	0.93	0.85

All models yielded correlation coefficient $r^2 > 0.9$ and $q^2 > 0.8$, excepted the model (3). However, the model (10) was picked for predicting the activity since it is an only model which contains both steric and electronic descriptors with good statistics.

Dipole y and principal moments of inertia (PMI) of compounds 14-19 were computed and listed in **Table 4.26** together with the activity data.

Table 4.26 Calculated properties and activity of QSAR model.

No.	Dipole y	Principal moments of inertia	Activity[a]
14.	0.04442	5146.535	740
15.	-0.04358	4996.981	190
16.	0.079726	5695.82	1100
17.	-0.04105	5665.576	350
18.	-0.08225	6304.836	598
19.	-0.12778	6230.864	450

[a] Catalyst activity in $kgPEH/mol_{zr}.h$

The dipole y represents electronic effect and principal moments of inertia represents steric effect. From **Table 4.26**, the catalyst with the high and positive dipole y and high principal moments of inertia gives high activity. The *rac* form has positive dipole y excepted the compound 18 while the *meso* form has negative dipole y. The principal moments of inertia does not depend on the stereochemistry of the catalyst but rather the alkyl substitution on indenyl ring. Hence, the *rac* form provides a more active catalyst. On the alkyl substitution, with large alkyl group the compound has less (Positive) value.

4.3.1.4 Catalyst Design

From the information in **4.3.1.3**, we tried to design the catalyst structure which more electron donating group to see the effect of electronic in higher catalyst activity. There are 2 position (5 and 6) for substitution pattern on indenyl ring. The effect of the substitution position should also be investigated.

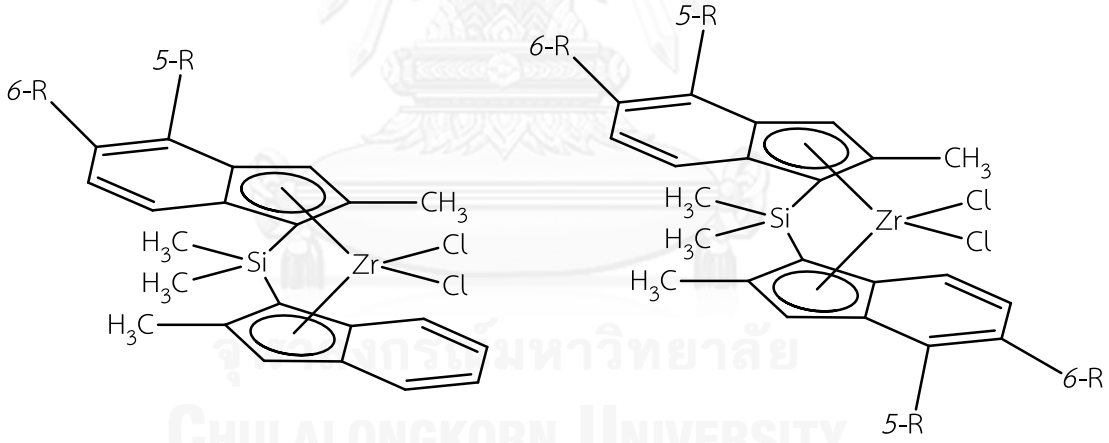


Figure 4.16 Scheme of designed catalysts in group 3.

4.3.1.4.1 5-position

Compound with methyl, ethyl, normal-propyl, iso-propyl, normal-butyl, tertiary-butyl, phenyl, C_6H_4OMe, C_6H_4Me, C_6H_4F substitution at 5 position on indenyl ring were created. Dipole y and principal moments of inertia of these compound

were computed and their activities were predicted according to model (10) are shown in **Table 4.27**.

From the group 3 QSAR model, principal moments of inertia has more influence to the catalyst activities than dipole y. The principal moments of inertia and dipole y can be increased by substituted on indenyl ring with bigger and more electron donating groups. Substitution with C_6H_4OMe on both ring gave the largest principal moments of inertia and hence the highest activities.

4.3.1.4.2 6-position

Compounds with phenyl, C_6H_4OMe, C_6H_4Me, C_6H_4F substitution at position 6 on indenyl ring were created. Dipole y and principal moments of inertia of these compound were computed and their activities were predicted using to model (10) are shown in **Table 4.28**.

From **Table 4.28,** similar to the 5-position, the compound with C_6H_4OMe substitution on both ring yielded the highest principal moments of inertia and hence activity. But it is not clear that 6-position was better predictive activity than 5-position cause the dipole y and principal moments of inertia values are not specifically different.

Table 4.27 Calculated properties and prediction activity of design zirconocene catalyst for 5-position in group 3.

No.	Dipole y									
	Me	Et	*n*Pr	*i*Pr	*n*Bu	*t*Bu	Ph	C_6H_4OMe	C_6H_4Me	C_6H_4F
20D	0.031894	-0.11193	-0.33087	0.040848	0.059077	-0.21984	-0.01271	0.611784	0.086275	-0.506
21D	-0.13599	-0.04001	-0.35621	-0.02001	-0.05828	-0.38156	-0.19267	0.818349	-0.13057	-0.25547

No.	Principal moments of inertia									
	Me	Et	^{n}Pr	^{i}Pr	^{n}Bu	^{t}Bu	Ph	C_6H_4OMe	C_6H_4Me	C_6H_4F
20D	5312.24	5950.68	6309.703	6557.678	7521.732	6958.699	8271.402	10882.75	9428.77	9612.631
21D	5734.594	6747.41	8011.508	8205.54	9494.474	9007.403	11722.76	16276.49	13143	14216.46

No.	Predictive Activity									
	Me	Et	^{n}Pr	^{i}Pr	^{n}Bu	^{t}Bu	Ph	C_6H_4OMe	C_6H_4Me	C_6H_4F
20D	687.3659	306.6942	-558.834	1310.799	1846.688	267.9216	1857.781	6023.012	2864.944	156.1503
21D	92.29828	1017.802	115.7668	1792.641	2213.529	460.8687	2619.258	9514.675	3575.043	3486.663

Table 4.28 Calculated properties and prediction activity of design zirconocene catalyst for 6-position in group 3.

No.	Dipole y				Principal moments of inertia				Predictive Activity			
	Ph	C_6H_4OMe	C_6H_4Me	C_6H_4F	Ph	C_6H_4OMe	C_6H_4Me	C_6H_4F	Ph	C_6H_4OMe	C_6H_4Me	C_6H_4F
20D	-0.12571	0.558696	0.060681	-0.60749	9832	13359	11434	11746	2053.26	6928.39	3679.93	672.82
21D	-0.23918	0.036245	0.083154	-0.26412	14978	22159	18246	19901	3919.13	8569.47	6965.09	6098.55

CHAPTER V

CONCLUSION

In this study, the QSAR and QSPR analyses were successfully applied to predict the ethylene and 1-hexene copolymerization activity and centain properties such as $\bar{M}_n$, $\bar{M}_w$, and %1-hexene incorporation of 3 groups of zirconocene catalyst consisting 19 zirconocene. The QSAR and QSPR models produced satisfactory predictivity results in terms of r^2 and q^2 values. The model provided the most signigificant correlation of steric and electronic effects with copolymerization activities.

5.1 Quantitative Structure-Activity Relationship (QSAR)

The statistical coefficients (r^2 and q^2) of 3 QSAR models obtained from 3 groups of zirconocene catalysts show high predictivity for the activity of zirconocene catalysts in ethylene and 1-hexene copolymerization. The catalyst activity could be well described by the electronic and the steric effect. Moreover, it was found that the electronic effect has more contributions to the catalyst activity (64.43%) than the steric effect (35.57%).

5.2 Quantitative Structure-Property Relationship (QSPR)

The statistical coefficients (r^2 and q^2) of 3 QSPR models obtained from 2 groups of zirconocene catalysts show high predictivity for the certain polymer properties such as $\bar{M}_n$, $\bar{M}_w$, and %1-hexene incorporation of zirconocene catalysts in ethylene and 1-hexene copolymerization. Those properties could be well described by the electronic and the steric effect. Moreover, it was found that the steric effect has more contributions to the properties (68.61%) than the electronic effect (31.38%).

5.3 Catalyst Design

The information obtained by the QSAR and QSPR study could be used to design the more potent zirconocene catalysts in ethylene and 1-hexene copolymerization. For group 1, the compound with the highest activity contains tertiary-butyl on the cyclopentadienyl ligand. It has 1.8 times activity of the best experimental catalyst (5). Thus, the substitution of the bulky electron donating group such as the branched large alkyl group on both cyclopentadienyl ligands will increase the activity. For group 2, the compound with the highest activity contains tertiary-butyl on benzene ring at *para* and 3,4,5-position. It has 5.4 times activity of the best experimental catalyst (12). The compound with the highest activity contains full-ring substitution by methyl group. It has 6.3 times activity of the best experimental catalyst (13). Thus, the substitution of the bulky electron donating group such as branched large alkyl group on both cyclopentadienyl ligands will increase activity. The electron withdrawing group such as F should be avoided. And for group 3, the compound with the highest activity contains C_6H_4OMe substitution on both indenyl ligands of the racemic form. It has 8.6 times activity of the best experimental catalyst (16). Thus, the substitution of the bulky electron donating group such as large aryl group on both indenyl ligands will increase activity.

REFERENCES

1. Robert D.J. Froese, D.G.M., Keiji Morokuma, Theoretical studies of the Cp ZrRq-catalyzed propylene 2 polymerization reactions and a comparison with ethylene polymerization. Journal of Molecular Structure, 1999. **Theochem**(461-462): p. 121-135.
2. Wei Wang, Z.F., Yinbang Zhu, Yonghui Zhang, Linxian Feng, Effects of cocatalyst on structure distribution of propylene polymers catalyzed by rac-$Me_2Si(Ind)_2ZrCl_2$/aluminoxane. European Polymer Journal, 2002. **38**: p. 1551–1558.
3. Wang, W., Z.-q. Fan, and L.-x. Feng, Ethylene polymerization and ethylene/1-hexene copolymerization using homogeneous and heterogeneous unbridged bisindenyl zirconocene catalysts. European Polymer Journal, 2005. **41**(10): p. 2380-2387.
4. Cruz, V., et al., 3D-QSAR analysis of metallocene-based catalysts used in ethylene polymerisation. Polymer, 2004. **45**(6): p. 2061-2072.
5. Möhring, P.C. and N.J. Coville, Group 4 metallocene polymerisation catalysts: quantification of ring substituent steric effects. Coordination Chemistry Reviews, 2006. **250**(1): p. 18-35.
6. Yao, S., et al., Consideration of an activity of the metallocene catalyst by using molecular mechanics, molecular dynamics and QSAR. Computational and Theoretical Polymer Science, 1999. **9**(1): p. 41-46.
7. Linnolahti, M. and T.A. Pakkanen, Theoretical study on the factors controlling the accessibility of cationic metal centers in zirconocene polymerization catalysts. Macromolecules, 2000. **33**(25): p. 9205-9214.
8. Cruz, V.L., et al., Structure-activity relationship study of the metallocene catalyst activity in ethylene polymerization. Organometallics, 2005. **24**(21): p. 5095-5102.
9. Martínez, S., et al., Polymerization Activity Prediction of Zirconocene Single-Site Catalysts Using 3D Quantitative Structure–Activity Relationship Modeling. Organometallics, 2012. **31**(5): p. 1673-1679.

10. Kashiwa, N. and J.-i. Imuta, Recent progress on olefin polymerization catalysts. Catalysis Surveys from Asia, 1997. **1**(1): p. 125-142.
11. Horton, A.D., et al., Cationic Alkylzirconium Complexes Based on a Tridentate Diamide Ligand: New Alkene Polymerization Catalysts. Organometallics, 1996. **15**(12): p. 2672-2674.
12. Gupta, V.K., et al., Titanium Complexes of Tartarate for Ethylene Polymerization. Synthesis and Reactivity in Inorganic and Metal-Organic Chemistry, 2005. **35**(7): p. 579-582.
13. Naga, N. and Y. Imanishi, Copolymerization of ethylene and cyclopentene with zirconocene catalysts: Effect of ligand structure of zirconocenes. Macromolecular Chemistry and Physics, 2002. **203**(1): p. 159-165.
14. Lauher, J.W. and R. Hoffmann, Structure and chemistry of bis (cyclopentadienyl)-MLn complexes. Journal of the American Chemical Society, 1976. **98**(7): p. 1729-1742.
15. Castonguay, L. and A. Rappe, Ziegler-Natta catalysis. A theoretical study of the isotactic polymerization of propylene. Journal of the American Chemical Society, 1992. **114**(14): p. 5832-5842.
16. Huang, J. and G. Rempel, Ziegler-Natta catalysts for olefin polymerization: mechanistic insights from metallocene systems. Progress in Polymer Science, 1995. **20**(3): p. 459-526.
17. Yang, X., C.L. Stern, and T.J. Marks, Cation-like homogeneous olefin polymerization catalysts based upon zirconocene alkyls and tris (pentafluorophenyl) borane. Journal of the American Chemical Society, 1991. **113**(9): p. 3623-3625.
18. Shen, S. and J.M. Torkelson, Miscibility and phase separation in poly (methyl methacrylate)/poly (vinyl chloride) blends: study of thermodynamics by thermal analysis. Macromolecules, 1992. **25**(2): p. 721-728.
19. Lian, B., et al., Titanium (IV) complexes containing mono-cyclopentadienyl and bulky trityloxy mixed ligands: synthesis and polymerization activity. Applied Organometallic Chemistry, 2005. **19**(5): p. 621-626.

20. Punniyamurthy, T., S. Velusamy, and J. Iqbal, Recent advances in transition metal catalyzed oxidation of organic substrates with molecular oxygen. Chemical Reviews, 2005. **105**(6): p. 2329-2364.
21. Kubinyi, H., Der Schlüssel zum Schloß II. Hansch-Analyse, 3D-QSAR und De novo-Design. Pharmazie in unserer Zeit, 1994. **23**(5): p. 281-290.
22. Kubinyi, H., QSAR, QSAR: Hansch Analysis and Related Approaches. Vol. 1. 2008: John Wiley & Sons.
23. Timmerman, H., et al., Chemometric methods in molecular design. 2008: John Wiley & Sons.
24. van de Waterbeemd, H., S. Rose, and S. Butler, Quantitative approaches to structure-activity relationships. 1996: Academic Press: London.
25. Ståhle, L., S. Segersvärd, and U. Ungerstedt, A comparison between three methods for estimation of extracellular concentrations of exogenous and endogenous compounds by microdialysis. Journal of Pharmacological Methods, 1991. **25**(1): p. 41-52.
26. Gigli, M., et al., Quantitative study of doxorubicin in living cell nuclei by microspectrofluorometry. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression, 1988. **950**(1): p. 13-20.
27. Thornberry, M.P., Synthesis, Properties, and Reactivity of Pentafluorophenyl Substituted Cyclopentadienes and Their Transition Metal Complexes, in Chemistry. 2001, Virginia State University.
28. Alonso-Moreno, C., et al., Synthesis, characterization and compared reactivity of asymmetrical ansa-metallocenes. Inorganic Chemistry Communications, 2009. **12**(2): p. 184-186.
29. Perdew, J.P., et al., Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. Physical Review B, 1992. **46**(11): p. 6671-6687.
30. Jeloaica, L. and V. Sidis, DFT investigation of the adsorption of atomic hydrogen on a cluster-model graphite surface. Chemical Physics Letters, 1999. **300**(1): p. 157-162.

31. Dnol[3]. Datasheet: Materials Studio 2011 [cited 2012 October]; Available from: http://accelrys.com/products/datasheets/dmol3.pdf.

32. Materials Studio. Datasheet: Materials Studio Overview 2011 [cited 2012 October]; Available from: http://accelrys.com/products/datasheets/materials-studio-overview.pdf.

APPENDIX

Table 1A Calculated properties of QSAR zirconocene catalyst in group 1.

No.	Cp-Zr-Cp Angle	Cluster number	Rank in cluster	Total energy	Binding energy	HOMO energy	LUMO energy	LUMO-HOMO energy	Element count	Atom count	Methyl	Total dipole
1	128.948	1	1	-4838.3	-4.76056	-0.22795	-0.11513	0.112818	10	23	0	1.665316
2	129.146	2	1	-4916.21	-5.82834	-0.223	-0.10972	0.113278	12	29	2	1.688157
3	130.726	3	1	-4994.13	-6.89126	-0.21747	-0.10783	0.109633	14	35	4	1.99795
4	130.564	4	1	-4994.11	-6.87738	-0.21455	-0.10903	0.105514	14	35	4	2.199438
5	129.961	5	1	-5072.03	-7.94462	-0.20935	-0.10358	0.105769	16	41	6	2.022165
6	124.816	6	1	-5072.01	-7.92618	-0.20202	-0.09707	0.104948	16	41	6	2.218295
7	120.839	7	1	-5149.83	-8.89295	-0.19128	-0.09436	0.096917	18	47	8	2.036801
8	121.053	7	2	-5227.75	-9.95762	-0.18526	-0.09168	0.093581	20	53	10	1.820622

Table 1A Calculated properties of QSAR zirconocene catalyst in group 1 (Continue).

No.	Dipole x	Dipole y	Dipole z	Radius of gyration	Principal moments of inertia	Principal moment of inertia X	Principal moment of inertia Y	Principal moment of inertia Z	Molecular density	Molecular volume
1	-1.16787	0.002077	-1.18716	2.712553	1525.547	751.4701	928.806	948.6344	1.515118	192.9328
2	-1.34265	-0.04299	-1.02241	3.031304	1988.339	941.8103	1157.079	1314.403	1.413505	226.6493
3	-1.8299	-0.03163	-0.80142	3.157468	2402.55	1244.29	1367.954	1533.849	1.339726	260.071
4	-1.95061	0.287813	-0.97457	3.102096	2461.553	1267.805	1336.253	1632.893	1.341605	259.7069
5	-1.71821	-0.06912	-1.06402	3.396388	2959.113	1472.593	1667.465	1951.251	1.285204	292.9325
6	-1.88704	0.05318	-1.16493	3.157572	2767.566	1502.785	1577.579	1706.547	1.288432	292.1986
7	-1.61691	0.1014	-1.23446	3.422325	3315.804	1776.85	1955.596	2003.248	1.217367	332.3007
8	-1.51592	0.079854	-1.00513	3.595242	3825.008	2028.548	2286.938	2299.043	1.177509	367.3739

Table 1A Calculated properties of QSAR zirconocene catalyst in group 1 (Continue).

No.	Molecular area	Total molecular mass	Zr Charge (ESP)	Zr Charge (Mulliken)	Zr Charge (Hirshfeld)	Cp ring Up Charge (ESP)	Cp ring Up Charge (Mulliken)	Cp ring Up Charge (Hirshfeld)	Cp ring Down Charge (ESP)	Cp ring Down Charge (Mulliken)	Cp ring Down Charge (Hirshfeld)
1	236.5423	292.316	0.625	0.82	0.425	-0.747	-0.6	-0.292	-0.751	-0.604	-0.293
2	277.1073	320.37	0.614	0.858	0.421	-0.558	-0.534	-0.257	-0.542	-0.541	-0.257
3	325.5249	348.424	0.729	0.905	0.423	-0.372	-0.48	-0.219	-0.382	-0.485	-0.221
4	315.153	348.424	0.726	0.897	0.429	-0.29	-0.507	-0.217	-0.324	-0.475	-0.223
5	355.2269	376.478	0.743	0.944	0.42	-0.07	-1.185	-0.178	-0.15	-0.426	-0.181
6	335.6916	376.478	0.755	0.909	0.394	-0.163	-0.43	-0.185	-0.151	-0.446	-0.187
7	370.594	404.532	0.766	0.96	0.351	-0.057	-0.416	-0.141	-0.016	-0.423	-0.141
8	413.5367	432.586	0.868	1.001	0.356	0.126	-0.338	-0.105	0.192	-0.343	-0.104

Table 2A Calculated properties of QSAR zirconocene catalyst in group 2.

No.	Cp-Zr-Cp Angle	Cluster number	Rank in cluster	Similarity to cluster reference	Total energy	Binding energy	HOMO energy	LUMO energy	LUMO-HOMO energy	Element count	Atom count	Phenyl
9	117.819	1	1	0	-5295.96	-8.97748	-0.19892	-0.12722	0.071694	22	43	2
10	116.671	2	1	0	-5295.97	-8.98836	-0.20276	-0.12822	0.074532	22	43	2
11	117.837	3	1	0	-5560.33	-6.90867	-0.21997	-0.15288	0.067091	16	33	0
12	117.887	4	1	0	-6282.52	-9.21135	-0.23208	-0.16512	0.066955	22	43	0
13	117.952	3	2	0.118053	-6282.51	-9.20322	-0.23051	-0.17231	0.058199	22	43	0

Table 2A Calculated properties of QSAR zirconocene catalyst in group 2 (Continue).

No.	Total dipole	Dipole x	Dipole y	Dipole z	Radius of gyration	Principal moments of inertia	Principal moment of inertia X	Principal moment of inertia Y	Principal moment of inertia Z	Total molecular mass
9	1.306816	-1.24585	0.184014	-0.34895	3.622549	5127.16	1546.973	3252.133	3649.421	444.512
10	1.272413	-1.03279	-0.19599	-0.71691	4.027453	5950.549	1557.302	3858.66	4253.774	444.512
11	1.85517	-0.12325	-0.10678	-1.84799	3.458391	6148.823	1544.485	3943.249	4457.956	458.364
12	2.668415	0.670317	-0.1101	-2.5805	3.742463	9832.758	3328.327	6161.198	6902.536	624.412
13	0.49296	-0.29261	0.018441	-0.3963	4.119134	13600.1	2870.287	8807.368	9957.63	624.412

Table 2A Calculated properties of QSAR zirconocene catalyst in group 2 (Continue).

No.	Molecular density	Molecular volume	Molecular area	Zr Charge (ESP)	Zr Charge (Mulliken)	Zr Charge (Hirshfeld)	Cp ring Up Charge (ESP)	Cp ring Up Charge (Mulliken)	Cp ring Up Charge (Hirshfeld)	Cp ring Down Charge (ESP)
9	1.262414	352.1128	378.997	0.628	0.852	0.424	-0.665	-0.532	-0.252	-0.692
10	1.254356	354.3747	394.9555	0.682	0.854	0.426	-0.632	-0.456	-0.209	-0.751
11	1.513831	302.784	343.8925	0.598	0.835	0.422	-0.476	-0.509	-0.261	-0.738
12	1.533037	407.3039	441.5908	0.617	0.842	0.42	-0.464	-0.517	-0.256	-0.467
13	1.53728	406.1797	451.8618	0.676	0.846	0.419	-0.242	-0.418	-0.231	-0.717

Table 2A Calculated properties of QSAR zirconocene catalyst in group 2 (Continue).

No.	Cp ring Down Charge (Mulliken)	Cp ring Down Charge (Hirshfeld)
9	-0.53	-0.253
10	-0.608	-0.294
11	-0.597	-0.286
12	-0.511	-0.259
13	-0.591	-0.286

Table 3A Calculated properties of QSAR zirconocene catalyst in group 3.

No.	Cp-Zr-Cp Angle	Cluster number	Rank in cluster	Similarity to cluster reference	Total energy	Binding energy	HOMO energy	LUMO energy	LUMO-HOMO energy	Ethyl	Methyl	Propyl
14	126.782	1	1	0	-5587.1	-9.99092	-0.19512	-0.11378	0.081332	1	3	0
15	127.614	2	1	0	-5587.11	-9.99273	-0.19404	-0.11563	0.078408	1	3	0
16	127.321	2	2	0.115407	-5626.06	-10.5171	-0.19506	-0.11391	0.081155	1	3	1
17	128.022	3	1	0	-5626.06	-10.5192	-0.19439	-0.11593	0.078463	1	3	1
18	127.399	3	2	0.095228	-5703.97	-11.5762	-0.19084	-0.1116	0.079243	2	4	1
19	128.819	4	1	0	-5703.97	-11.5786	-0.19166	-0.11384	0.077817	2	4	1

Table 3A Calculated properties of QSAR zirconocene catalyst in group 3 (Continue).

No.	Atom count	Element count	Total dipole	Dipole x	Dipole y	Dipole z	Principal moments of inertia	Principal moment of inertia X	Principal moment of inertia Y	Principal moment of inertia Z	Radius of gyration
14	48	22	1.347438	-1.1278	0.04442	-0.73599	5146.535	2187.433	2989.771	3572.565	3.881906
15	48	22	1.38932	-0.94097	-0.04358	-1.02121	4996.981	2427.976	3054.236	3121.921	3.845437
16	51	23	1.373849	-1.09746	0.079726	-0.82261	5695.82	2520.625	3199.014	3981.85	4.055076
17	51	23	1.389655	-0.88977	-0.04105	-1.06666	5665.576	2463.768	3581.422	3633.459	4.056975
18	57	25	1.24282	-0.83983	-0.08225	-0.91243	6304.836	2969.312	3749.286	4108.161	4.183597
19	57	25	1.200916	-0.84397	-0.12778	-0.84474	6230.864	3029.154	3604.267	4081.318	4.169749

Table 3A Calculated properties of QSAR zirconocene catalyst in group 3 (Continue).

No.	Total molecular mass	Molecular area	Molecular volume	Molecular density	Zr Charge (ESP)	Zr Charge (Mulliken)	Zr Charge (Hirshfeld)	Cp ring Up Charge (ESP)	Cp ring Up Charge (Mulliken)	Cp ring Up Charge (Hirshfeld)
14	476.629	428.0781	367.7096	1.29621	0.786	0.921	0.433	-1.351	-0.928	-0.401
15	476.629	430.8052	368.2622	1.294265	0.681	0.918	0.429	-1.197	-0.92	-0.406
16	490.656	450.2957	384.7851	1.275143	0.701	0.921	0.432	-1.176	-0.917	-0.403
17	490.656	451.6533	384.6289	1.275661	0.708	0.914	0.429	-1.202	-0.914	-0.405
18	518.71	489.8432	418.7633	1.238671	0.725	0.94	0.432	-1.263	-0.926	-0.41
19	518.71	493.8771	418.7726	1.238643	0.72	0.925	0.431	-1.226	-0.922	-0.407

Table 3A Calculated properties of QSAR zirconocene catalyst in group 3 (Continue).

No.	Cp ring Down Charge (ESP)	Cp ring Down Charge (Mulliken)	Cp ring Down Charge (Hirshfeld)
14	-1.564	-1.033	-0.45
15	-1.372	-1.019	-0.446
16	-1.407	-1.032	-0.45
17	-1.387	-1.018	-0.444
18	-1.203	-0.923	-0.405
19	-1.196	-0.903	-0.401

Table 4A PCA loading value for group 1 QSAR.

Descriptor	PC1	PC2
Cp-Zr-Cp Angle	-0.14572800	-0.33644600
Cluster number	0.18229700	-0.03085320
Rank in cluster	0.11931000	0.23391000
Total energy	-0.18803300	0.00872684
Binding energy	-0.18790700	0.01758420
HOMO energy	0.18478600	0.08736520
LUMO energy	0.18102200	0.08821190
LUMO-HOMO energy	-0.18013300	-0.08205110
Element count	0.18803400	-0.00860745
Atom count	0.18803400	-0.00860745
Methyl	0.18803400	-0.00860745
Total dipole	0.07435440	-0.46263100
Dipole x	-0.07223240	0.52759300
Dipole y	0.04605080	-0.16002200
Dipole z	-0.01448960	-0.31858800
Radius of gyration	0.17835300	-0.05562850
Principal moments of inertia	0.18724500	-0.01083790
Principal moment of inertia X	0.18804500	-0.00583152
Principal moment of inertia Y	0.18624900	0.05813040
Principal moment of inertia Z	0.18337000	-0.08030090
Molecular density	-0.18566000	0.07448430
Molecular volume	0.18801200	0.00927652
Molecular area	0.18473100	-0.06872000
Total molecular mass	0.18803400	-0.00860745
Zr Charge (ESP)	0.17814300	-0.06914920
Zr Charge (Mulliken)	0.18375800	-0.05008230
Zr Charge (Hirshfeld)	-0.15783700	-0.28890300
Cp ring Up Charge (ESP)	0.18337500	-0.10660900
Cp ring Up Charge (Mulliken)	0.03826620	0.20976600
Cp ring Up Charge (Hirshfeld)	0.18787200	-0.00747301
Cp ring Down Charge (ESP)	0.18739400	-0.03857690

Table 4A PCA loading value for group 1 QSAR (Continue).

Descriptor	PC1	PC2
Cp ring Down Charge (Mulliken)	0.18367500	-0.09443720
Cp ring Down Charge (Hirshfeld)	0.18779000	0.01135330

Table 5A PCA score value for group 1 QSAR.

PCA Vector 1	PCA Vector 2
-1.52785575	1.23625486
-0.95067578	0.66700396
-0.35738865	-0.91483900
-0.26883339	-1.39766986
0.23689255	-0.88729863
0.33279959	-0.31816806
0.96609051	0.65872563
1.56897094	0.95599110

Table 6A PCA loading value for group 1 QSPR.

Descriptor	PC1	PC2
Cp-Zr-Cp Angle	-0.14248500	-0.33462900
Cluster number	0.17663800	-0.03332240
Rank in cluster	0.11591000	0.22340800
Total energy	-0.18196300	0.01564110
Binding energy	-0.18179800	0.02459970
HOMO energy	0.17935200	0.08306220
LUMO energy	0.17553200	0.08274890
LUMO-HOMO energy	-0.17503300	-0.07933580
Element count	0.18196400	-0.01552030
Atom count	0.18196400	-0.01552030
Methyl	0.18196400	-0.01552030
Total dipole	0.07125570	-0.45247700
Dipole x	-0.06862960	0.52141100

Table 6A PCA loading value for group 1 QSPR (Continue).

Descriptor	PC1	PC2
Dipole y	0.04550560	-0.14356100
Dipole z	-0.01578970	-0.32559500
Radius of gyration	0.17210300	-0.06648160
Principal moments of inertia	0.18116000	-0.01844370
Principal moment of inertia X	0.18207200	-0.01180460
Principal moment of inertia Y	0.18038700	0.04978930
Principal moment of inertia Z	0.17711600	-0.08836430
Molecular density	-0.17943800	0.08084230
Molecular volume	0.18200800	0.00229737
Cp ring Down Charge (ESP)	0.18125800	-0.04512460
Cp ring Down Charge (Mulliken)	0.17735600	-0.10228900
Cp ring Down Charge (Hirshfeld)	0.18177700	0.00400857

Table 7A PCA score value for group 1 QSPR.

PCA Vector 1	PCA Vector 2
-1.50798775	1.26777735
-0.95547662	0.64695830
-0.37392047	-0.93755631
-0.26918709	-1.34738410
0.21771933	-0.93799498
0.33243255	-0.29499128
0.98223340	0.68377231
1.57418665	0.91941872

Table 8A PCA loading value for group 2 QSAR.

Descriptor	PC1	PC2
Cp-Zr-Cp Angle	-0.14156100	0.14066500
Cluster number	-0.19446100	0.11908700
Rank in cluster	-0.15958300	-0.18667900
Similarity to cluster reference	-0.15958300	-0.18667900

Table 8A PCA loading value for group 2 QSAR (Continue).

Descriptor	PC1	PC2
Total energy	0.23587100	9.592010e-004
Binding energy	0.05060160	0.22608400
HOMO energy	0.22983200	-0.07217600
Element count	-0.02393760	-0.23321900
Atom count	-0.02393760	-0.23321900
Phenyl	0.20560100	-0.13850000
Total dipole	-0.01079270	0.25156000
Dipole x	-0.18755500	0.16683000
Dipole y	0.01969940	-0.03873270
Dipole z	0.08362270	-0.25142300
Radius of gyration	-0.07596610	-0.29237400
Principal moments of inertia	-0.22302500	-0.11262600
Principal moment of inertia X	-0.21574800	-0.01552130
Principal moment of inertia Y	-0.22050400	-0.11954200
Principal moment of inertia Z	-0.22084100	-0.11716700
Total molecular mass	-0.22913500	-0.03817510
Molecular density	-0.21238000	0.12665500
Molecular volume	-0.16317100	-0.15522300
Molecular area	-0.17751100	-0.16609100
Zr Charge (ESP)	-0.00487807	-0.31224600
Zr Charge (Mulliken)	0.11393400	-0.25533700
Zr Charge (Hirshfeld)	0.22843400	-0.05354450
Cp ring Up Charge (ESP)	-0.21881400	-0.04093470
Cp ring Up Charge (Mulliken)	-0.09324060	-0.24464900
Cp ring Up Charge (Hirshfeld)	0.04616260	-0.27609100
Cp ring Down Charge (ESP)	-0.11652300	0.11818000
Cp ring Down Charge (Mulliken)	-0.03830290	0.10710600
Cp ring Down Charge (Hirshfeld)	0.00218103	0.10339600

Table 9A PCA score value for group 2 QSAR.

PCA Vector 1	PCA Vector 2
0.95766950	-0.06052276
0.92912388	-0.88147288
0.17936346	1.29513712
-0.87040597	0.68354784
-1.19575087	-1.03668932

Table 10A PCA loading value for group 2 QSPR.

Descriptor	PC1	PC2
Cp-Zr-Cp Angle	0.13755000	0.13727300
Cluster number	0.18470700	0.11742100
Rank in cluster	0.15407000	-0.18922000
Similarity to cluster reference	0.15407000	-0.18922000
Total energy	-0.22365400	0.00360047
Binding energy	-0.04236460	0.22557800
HOMO energy	-0.21947000	-0.06945700
LUMO energy	-0.22408400	-0.04159530
LUMO-HOMO energy	-0.20524700	0.03321320
Element count	0.01692650	-0.23239200
Atom count	0.01692650	-0.23239200
Phenyl	-0.19840100	-0.13554400
Total dipole	0.00755895	0.25220300
Dipole x	0.17793000	0.16499500
Dipole y	-0.01746890	-0.04024100
Dipole z	-0.07836980	-0.25109300
Radius of gyration	0.06968550	-0.29211700
Principal moments of inertia	0.21225900	-0.11525500
Principal moment of inertia X	0.20247000	-0.01747670
Principal moment of inertia Y	0.21000800	-0.12216100
Principal moment of inertia Z	0.21043400	-0.11983000
Total molecular mass	0.21634500	-0.04056600
Molecular density	0.20461600	0.12363000

Table 10A PCA loading value for group 2 QSPR (Continue).

Descriptor	PC1	PC2
Molecular volume	0.15082200	-0.15627900
Molecular area	0.16507400	-0.16729100
Zr Charge (ESP)	0.00252552	-0.31126900
Zr Charge (Mulliken)	-0.11258400	-0.25310900
Zr Charge (Hirshfeld)	-0.21891200	-0.05012950
Cp ring Up Charge (ESP)	0.21088900	-0.04406080
Cp ring Up Charge (Mulliken)	0.08961170	-0.24526400
Cp ring Up Charge (Hirshfeld)	-0.04626720	-0.27413000
Cp ring Down Charge (ESP)	0.10627600	0.11761900
Cp ring Down Charge (Mulliken)	0.03247910	0.10661800
Cp ring Down Charge (Hirshfeld)	-0.00496363	0.10289600

Table 11A PCA score value for group 2 QSPR.

PCA Vector 1	PCA Vector 2
-0.97033275	-0.04029972
-0.96227011	-0.85266924
-0.09001152	1.28306877
0.81268184	0.68500029
1.20993254	-1.07510010

Table 12A PCA loading value for group 3 QSAR.

Descriptor	PC1	PC2
Cp-Zr-Cp Angle	-0.12734300	-0.23901400
Cluster number	-0.18062200	-0.17202200
Rank in cluster	-0.06761370	0.15416300
Total energy	0.20697600	-0.03080830
Binding energy	0.20701900	-0.03048220
HOMO energy	-0.19546000	0.01013240
LUMO energy	-0.12307500	0.29130100
LUMO-HOMO energy	0.10971400	0.29087400
Ethyl	-0.19946800	0.06933140

Table 12A PCA loading value for group 3 QSAR (Continue).

Descriptor	PC1	PC2
Methyl	-0.19946800	0.06933140
Propyl	-0.14867000	-0.05713980
Atom count	-0.20697500	0.03081280
Dipole x	-0.16344900	-0.19864900
Dipole y	0.15892400	0.15271600
Dipole z	0.00775451	0.34059800
Principal moments of inertia	-0.19756000	0.03954550
Principal moment of inertia X	-0.20485500	-0.02499380
Principal moment of inertia Y	-0.18562800	-0.07481990
Principal moment of inertia Z	-0.15903400	0.15649100
Radius of gyration	-0.19035600	0.01684540
Total molecular mass	-0.20697500	0.03081280
Molecular area	-0.20720700	0.01246850
Molecular volume	-0.20714300	0.02954670
Molecular density	0.20703600	-0.02259100
Zr Charge (ESP)	0.02777600	0.30616300
Zr Charge (Mulliken)	-0.14978000	0.20128500
Zr Charge (Hirshfeld)	-0.01723940	0.37801800
Cp ring Up Charge (ESP)	-0.02592040	-0.28987900
Cp ring Up Charge (Mulliken)	0.03077770	-0.30378900
Cp ring Up Charge (Hirshfeld)	0.17232700	0.10437800
Cp ring Down Charge (ESP)	-0.19517500	-0.10232200
Cp ring Down Charge (Mulliken)	-0.19986200	0.01964950
Cp ring Down Charge (Hirshfeld)	-0.20124800	0.02942870

Table 13A PCA score value for group 3 QSAR.

PCA Vector 1	PCA Vector 2
1.06853640	1.38874273
0.77181657	-1.00614839
0.40482279	0.34304386
0.22398390	-1.18986469
-1.24179616	0.68919577
-1.22736350	-0.22496927

Table 14A PCA loading value for group 3 QSPR.

Descriptor	PC1	PC2
Cp-Zr-Cp Angle	0.12437400	0.23110500
Cluster number	0.17644000	0.16314100
Rank in cluster	0.06477980	-0.16453300
Total energy	-0.20205200	0.03165910
Binding energy	-0.20210000	0.03130880
HOMO energy	0.19247000	-0.00711454
LUMO energy	0.12028700	-0.28012800
LUMO-HOMO energy	-0.10897700	-0.28291700
Ethyl	0.19572100	-0.06273880
Methyl	0.19572100	-0.06273880
Propyl	0.14313500	0.04170250
Atom count	0.20205100	-0.03166400
Element count	0.20205100	-0.03166400
Total dipole	-0.18129800	0.10390700
Dipole x	0.16129400	0.19129000
Dipole y	-0.15727500	-0.15409500
Dipole z	-0.00889274	-0.32024200
Principal moments of inertia	0.19212700	-0.04333430
Principal moment of inertia X	0.20045300	0.02151900
Principal moment of inertia Y	0.18121200	0.06707260
Principal moment of inertia Z	0.15343000	-0.15652100
Radius of gyration	0.18476600	-0.02331330

Table 14A PCA loading value for group 3 QSPR (Continue).

Descriptor	PC1	PC2
Total molecular mass	0.20205100	-0.03166400
Molecular area	0.20226600	-0.01401520
Molecular volume	0.20223800	-0.03044200
Molecular density	-0.20208300	0.02434130
Zr Charge (ESP)	-0.02742590	-0.28020500
Zr Charge (Mulliken)	0.14729900	-0.19399400
Zr Charge (Hirshfeld)	0.01564310	-0.35994100
Cp ring Up Charge (ESP)	0.02439990	0.26219800
Cp ring Up Charge (Mulliken)	-0.03172920	0.27736000
Cp ring Up Charge (Hirshfeld)	-0.17017000	-0.09713100
Cp ring Down Charge (ESP)	0.19150400	0.09469950

Table 15A PCA score value for group 3 QSPR.

PCA Vector 1	PCA Vector 2
-1.06633362	-1.32581613
-0.74474475	1.03147795
-0.43292988	-0.44095876
-0.23244710	1.17810952
1.25258656	-0.72032346
1.22386880	0.27751087

VITA

Mister Prinya Charoenkhai was born on April 12, 1988 in Nakhon Ratchasima, Thailand. I received a Bachelor's degree of science from Department of Chemistry, Faculty of Science, Rajamangkala University of Technology, Thailand in 2010. I was admitted to a Master's Degree Program in Petrochemistry and Polymer science, Faculty of Science, Chulalongkorn University and completed the program in 2014. I have presented a poster presentation in the title of QSAR Study of Zirconocene Catalysts in Ethylene and 1-hexene Copolymerization at PACCON2014 conference. My address is 579 Moo 8, Muenwai Sub-district, Mueng District, Nakhon Ratchasima 30000