Available online at www.sciencedirect.com
ScienceDirect
Polymer 48 (2007) 4663–4674
polymer
www.elsevier.com/locate/polymer

3D-QSAR study of *ansa*-metallocene catalytic behavior in ethylene polymerization

V.L. Cruz [a,*], S. Martinez [a], J. Martinez-Salazar [a], D. Polo-Cerón [b], S. Gómez-Ruiz [b], M. Fajardo [b], S. Prashar [b]

[a] *Instituto de Estructura de la Materia, CSIC, Serrano, 113bis, E-28006 Madrid, Spain*
[b] *Departamento de Química Inorgánica y Analítica, E.S.C.E.T., Universidad Rey Juan Carlos, 28933 Móstoles (Madrid), Spain*

Received 5 February 2007; received in revised form 30 May 2007; accepted 31 May 2007
Available online 9 June 2007

Abstract

ansa-Zirconocene complexes have been studied by Three-Dimensional Quantitative Structure–Activity Relationship (3D-QSAR). Using the comparative molecular field analysis (CoMFA) method, the experimental results obtained for catalytic activity and polymer molecular weight have been successfully correlated with the steric and electrostatic 3D structural descriptors, which were calculated by density functional theory (DFT) methods.

Keywords: *ansa*-Zirconocene; 3D-QSAR; Polymerization

1. Introduction

The discovery by Sinn and Kaminsky [1] of the group 4 metallocene/MAO catalytic system for the polymerization of α-olefins has brought about a swift and unabating development in this field [2]. The main efforts in this research area are focused on the design of metallocene complexes [3] with high catalytic activities and capable of producing materials with ''user defined'' physical properties [2]. Small changes in the structural make-up of the metallocene catalysts are well known to have notable effects on the polymerization. However, to date, much of the chemistry carried out has been of a ''hit or miss'' nature and therefore in order to reduce time and cost, it is essential the capacity to predict for hypothetical metallocene systems their catalytic behavior or propose possible modifications to existing catalysts to improve activity and polymer molecular weight.

Quantitative Structure–Activity Relationship (QSAR) techniques can be applied to search for correlations between the catalyst structure and its behavior in the polymerization process. An adequate exploration of structure–activity relationship can be expected considering our previous experience in the application of Three-Dimensional QSAR (3D-QSAR) methodology in the field of metallocene-based polymerization catalysis.

In our initial study [4a], we used a set of metallocene catalysts for determining polymerization activity and polymer molecular weight under the same experimental reaction conditions. Descriptors obtained from DFT (Density Functional Theory) calculations, such as LUMO (Lowest Unoccupied Molecular Orbital), electrostatic, steric and local softness fields, were used in order to describe the structure of the catalysts. The electronic interaction was confirmed by correlations found between activity and LUMO as well as between activity and local softness. The model revealed that the experimental variance in catalytic activity is well explained in terms of the arrangement of the ligands around the metal center or the type of aromatic ligands (i.e., cyclopentadienyl or indenyl). Furthermore, it was found that in *ansa*-metallocene catalysts

* Corresponding author. Tel.: +34 915616800.
E-mail address: victor.cruz@iem.cfmac.csic.es (V.L. Cruz).

0032-3861/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.polymer.2007.05.081

the bridges were not directly involved in LUMO and local softness fields, but rather in the cyclopentadienyl centroid–zirconium–cyclopentadienyl centroid (Cent–Zr–Cent) angle promoting electronic interaction between the metal center and the atoms of the ligands to a greater or lesser extent. Polymer molecular weight was found to correlate also with those fields where the Cent–Zr–Cent angle is the key geometric variable. However, steric fields were not able to explain the variance in the molecular weight data.

In our second work [4b] we performed a 3D-QSAR analysis to a larger and more varied type of metallocene catalyst systems extracted from an experimental study conducted by the Kaminsky group [5]. The results obtained in this 3D-QSAR analysis were in agreement with the correlations found in our initial study [4b], however, with an additional implication of the electrostatic field, which was not previously observed.

We include in this paper a section named ''Problem Statement'' where we explain the global problem of using chemometric tools in the field of homogeneous catalysis and its formal analogy with the drug design field.

As a continuation of our computational chemistry studies [4,6] applied to single site polymerization catalysts and our work related to zirconocene complexes [7], we present in this paper the study of a set of *ansa*-bis(cyclopentadienyl) zirconium catalysts where the structural variability, mainly focused on the nature of the substituents attached to the silicon bridge atom, has been considered. The structure–activity correlations found in this work show an interesting parallelism with the behavior generally observed in supported catalysis [8].

1.1. Problem statement

QSAR is an essential tool in the designing of new drugs in the medicinal chemistry field since 1960s [9]. The fundamental hypothesis supporting the application of these tools is that drug activity is modulated by its chemical structure. In fact, it is generally observed that small changes in the structure yield very different responses in bioactivity. This is confirmed by X-ray crystallography which shows unequivocally how some drug's structural features are connected with the activation or inhibition of some protein or enzyme. This prompted researchers in drug design to use chemometric tools to rationalize the study of these structure–activity relationships. The application of QSAR methodologies is today an essential step in the process of designing a new drug [10]. Among the different tools, Three-Dimensional Quantitative Structure–Activity Relationship (3D-QSAR) [11] is one of the most useful methodologies due to its capability to locate in the three-dimensional space the main structural features associated to the particular biological response.

One of the main criticisms posed towards the usage of 3D-QSAR concerns to the inclusion of thousands of variables used to correlate with a few dependent variables. However, it should be taken into account that these variables are no longer independent and cannot be used as in Multivariate Linear Correlation techniques. The advent of statistical techniques such as Partial Least Squares (PLS) solves this situation by deriving true independent variables which are then used for the fitting procedure.

The successful application of 3D-QSAR methodologies in drug design has stimulated us to apply them to the study of single site polymerization catalysis [4]. In this case the structural variability resides in the catalyst itself which would correspond with the receptor in the drug design arena, where it is common for each case. The monomer would be equivalent to the ligand, but in this case this component will be the same for all the catalyst species considered. As can be deduced, the participants in the ligand–receptor interactions are equivalent both in the drug design and in the homogeneous catalysis arena, but the target of the modeling process is the drug/ligand in the former case and the receptor/catalyst in the latter one.

Other criticism commonly made by researchers in the homogeneous catalysis field to the usage of QSAR tools concerns the complexity of the real system. Of course, this is the fundamental drawback inherent to any simulation or theoretical methodology. There is an additional point of debate with the application of QSAR, i.e., the calculated descriptors have to be directly correlated with experimental values. In general, many factors other than molecular structure influence the action of any drug or catalyst. However, careful and systematic experiments can be done to isolate the structural effects as much as possible, for example, by performing experiments in the same conditions. The validity of this approach has been demonstrated by the numerous successful cases in the drug design field. The area of homogeneous catalysis should not present, in principle, further complications to get similar results. The complexity of the real biological system can be yet superior to the artificial polymerization system, even if simpler ''in vitro'' experiments are contemplated.

The above arguments have motivated us to apply the 3D-QSAR paradigm to the study of homogeneous polymerization catalysis with the aim of obtaining useful information to gain further knowledge about these processes and to design better catalysts. Our previous works [4] show how useful can be such QSAR models, both to interpret the polymerization mechanisms and to predict the properties of catalytic systems not present in the training set, i.e., the set of compounds used to derive the models. These models are easy to interpret and are also in agreement with the experimental and theoretical observations reported in the literature.

2. Methods

2.1. Source of catalysts and polymerization data

A training set of 22 metallocene complexes were taken from previous work carried out by us [7g,i]. The compounds considered here (see Fig. 1) along with the corresponding activities and polymer molecular weights are shown in Table 1. As can be seen, this training set covers silylene bridged zirconocene complexes with different substituents at the silicon atom. Permethyl substitution at the aromatic cyclopentadienyl ligand has been also considered as a structural variable.

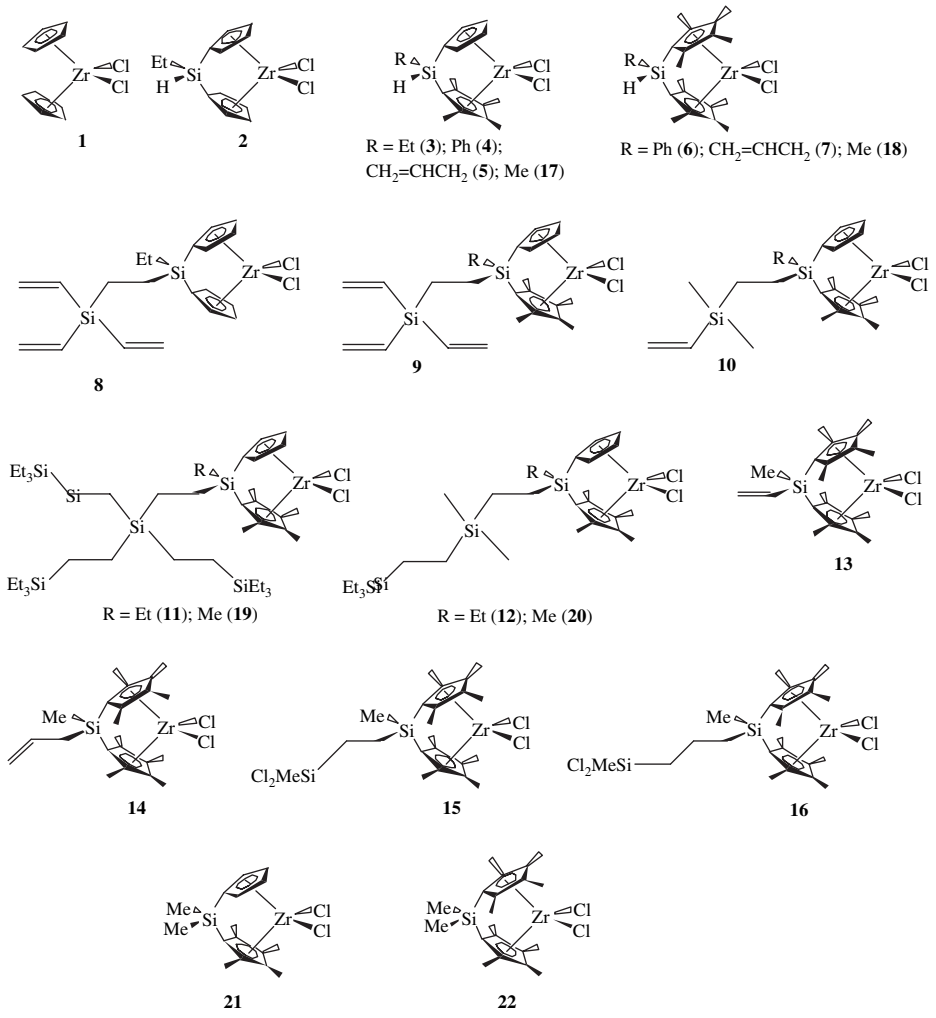

Fig. 1. Metallocene catalysts **1**–**22**.

Table 1
Ethylene polymerization results

No.	Catalyst	Activity[a]	Log activity	M_w	Log M_w
1	$(\eta^5\text{-}C_5H_5)_2ZrCl_2$	23 750	4.37	1 014 670	6.01
2	$[Et(H)Si(\eta^5\text{-}C_5H_4)_2]ZrCl_2$	11 100	4.05	210 200	5.32
3	$[Et(H)Si(\eta^5\text{-}C_5Me_4)(\eta^5\text{-}C_5H_4)]ZrCl_2$	8760	3.94	186 200	5.27
4	$[Ph(H)Si(\eta^5\text{-}C_5Me_4)(\eta^5\text{-}C_5H_4)]ZrCl_2$	7053	3.85	185 240	5.27
5	$[CH_2{=}CHCH_2(H)Si(\eta^5\text{-}C_5Me_4)(\eta^5\text{-}C_5H_4)]ZrCl_2$	3120	3.49	133 500	5.13
6	$[Ph(H)Si(\eta^5\text{-}C_5Me_4)_2]ZrCl_2$	15 967	4.20	194 650	5.29
7	$[CH_2{=}CHCH_2(H)Si(\eta^5\text{-}C_5Me_4)_2]ZrCl_2$	6373	3.80	156 800	5.20
8	$[Et\{(CH_2{=}CH)_3SiCH_2CH_2\}Si(\eta^5\text{-}C_5H_4)_2]ZrCl_2$	3307	3.52	139 460	5.14
9	$[Et\{(CH_2{=}CH)_3SiCH_2CH_2\}Si(\eta^5\text{-}C_5Me_4)(\eta^5\text{-}C_5H_4)]ZrCl_2$	5513	3.74	420 000	5.62
10	$[Et\{(CH_2{=}CH)Me_2SiCH_2CH_2\}Si(\eta^5\text{-}C_5Me_4)(\eta^5\text{-}C_5H_4)]ZrCl_2$	7213	3.86	141 300	5.15
11	$[Et\{(Et_3SiCH_2CH_2)_3SiCH_2CH_2\}Si(\eta^5\text{-}C_5Me_4)(\eta^5\text{-}C_5H_4)]ZrCl_2$	3367	3.53	717 000	5.86
12	$[Et\{(Et_3SiCH_2CH_2)Me_2SiCH_2CH_2\}Si(\eta^5\text{-}C_5Me_4)(\eta^5\text{-}C_5H_4)]ZrCl_2$	10 147	4.01	669 000	5.83
13	$[Me(CH_2{=}CH)Si(\eta^5\text{-}C_5Me_4)_2]ZrCl_2$	15 510	4.19	236 000	5.37
14	$[Me(CH_2{=}CHCH_2)Si(\eta^5\text{-}C_5Me_4)_2]ZrCl_2$	16 472	4.22	136 790	5.14
15	$[Me(Cl_2MeSiCH_2CH_2)Si(\eta^5\text{-}C_5Me_4)_2]ZrCl_2$	26 898	4.43	483 000	5.68
16	$[Me(Cl_2MeSiCH_2CH_2CH_2)Si(\eta^5\text{-}C_5Me_4)_2]ZrCl_2$	17 707	4.25	558 000	5.75
17	$[Me(H)Si(\eta^5\text{-}C_5Me_4)(\eta^5\text{-}C_5H_4)]ZrCl_2$	9633	3.98	162 900	5.21
18	$[Me(H)Si(\eta^5\text{-}C_5Me_4)_2]ZrCl_2$	26 407	4.42	151 300	5.18
19	$[Me\{(Et_3SiCH_2CH_2)_3SiCH_2CH_2\}Si(\eta^5\text{-}C_5Me_4)(\eta^5\text{-}C_5H_4)]ZrCl_2$	3907	3.59	743 300	5.87
20	$[Me\{(Et_3SiCH_2CH_2)Me_2SiCH_2CH_2\}Si(\eta^5\text{-}C_5Me_4)(\eta^5\text{-}C_5H_4)]ZrCl_2$	11 100	4.05	341 200	5.53
21	$[Me_2Si(\eta^5\text{-}C_5Me_4)(\eta^5\text{-}C_5H_4)]ZrCl_2$	16 073	4.21	175 500	5.24
22	$[Me_2Si(\eta^5\text{-}C_5Me_4)_2]ZrCl_2$	11 500	4.06	332 800	5.52

[a] In kg pol (mol Zr h)$^{-1}$.

All polymerization reactions were carried out under the same experimental conditions: a reactor temperature of 20 °C, an ethylene pressure of 2 bar, 3×10^{-5} mol L^{-1} of metallocene concentration and a molar ratio MAO (methyl aluminoxane)/metallocene of 3000. Polymer molecular weights were determined by GPC (Waters 150C Plus) in 1,2,4-trichlorobenzene at 145 °C.

2.2. Molecular modeling

The active species of the catalyst is a cationic organometallic complex with a vacant coordination site where polymerization takes place and can interact with the co-catalyst. Thus, the DFT calculations are based on the cationic species rather than on the metallocene precursor itself. The reactivity descriptors considered in this work contemplate some characteristics of the isolated reactants. This implies that the former are relevant only with respect to the initial interaction between the catalyst and the other species in the reaction medium. This type of information can be related to the reactivity only if the reaction has an early transition state, as it is the case with the ethylene insertion into the metal alkyl bond of the metallocene catalyst. Furthermore, the possible co-catalyst effect on the activity and molecular weight is assumed to be the same in all cases although the co-catalyst is not taken into account explicitly in the present work. This assumption can be valid as far as the same co-catalyst and the same Al/Zr ratio are used for all polymerization reactions. For each organometallic cationic species, geometry optimization at B3LYP [12]/LANL2DZ [13] was performed using the Gaussian 03 package [14]. The following 3D fields were evaluated in Cartesian grids: electron densities, electrostatic potential, HOMO and LUMO molecular orbitals. We also calculated electrostatic charges by fitting the electrostatic potential to nuclear positions according to the CHELPG scheme [15]. Steric and electrostatic 3D fields are calculated in the CoMFA method through the interaction between each catalyst and a probe atom. The probe atom should have specific charge and steric properties to evaluate the interaction energy at each particular point in the grid. The probe atom selected was an sp^3 C atom with a −1 point charge. This atom corresponds to atom C(3) in the Tripos Force Field [16], which was used to calculate van der Waals (steric) interactions. The value for the probe atom charge was selected to represent the effect of the electrostatic nature of either the ethylene or the anionic co-catalyst.

2.3. CoMFA details

All the reported 3D-QSAR analyses were carried out with the CoMFA module [17] implemented in the SYBYL 7.2 package [18].

2.3.1. Alignment rule

It is essential in 3D-QSAR to align all structures in a common framework in order to make possible the comparison between all cationic active species of the tested catalysts [19]. In this study the catalyst molecules were aligned in such a way that the active site presented similar orientations. Zirconium atom, cyclopentadienyl centroids, Cp(1) and Cp(2) and the alkyl carbon atom C(3) attached to the metal were used for the molecular alignment as represented in Fig. 2a. The resulting alignment is shown in Fig. 2b along with the cubic region used to calculate the molecular fields. Several grid spacings were used, but the best results are obtained with lattices of 1.0 Å and 2.0 Å spacing. These values represent a compromise in the sense that a higher precision in the evaluation of the 3D field represented by a finer grid increases the so-called "brown noise" due to the sensitivity of the statistical technique applied to generate the models [20].

2.3.2. PLS analysis

Partial Least Squares (PLS) [21] 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 (stdev) of each field and divides each value by the corresponding stdev. The effect is to give each variable the same prior importance in the analysis. Leave one out (LOO) [22] cross-validated PLS analysis was initially performed to determine both the robustness of the statistical models and the optimal number of components. This can be achieved by examining the Predictive Residual Sum of Squares (PRESS) and the cross-validated regression coefficient (q^2) as guidelines. The q^2 statistics are defined as:

$$\mathrm{PRESS} = \sum_{i=1}^{N} \left(Y_{\mathrm{obs},i} - Y_{\mathrm{pred},i}\right)^2 \tag{1}$$

$$q^2 = 1 - \mathrm{PRESS}/\mathrm{SSD} \tag{2}$$

where $Y_{\mathrm{obs},i}$ and $Y_{\mathrm{pred},i}$ are, respectively, the actual and predicted dependent variables and SSD is the sum of the squared deviations of each dependent variable from the mean of all dependent variables. It has been estimated by some authors [23] that a q^2 value greater than 0.3 has a 95% confidence limit. The usual practice in drug design is to consider valid a model with a q^2 greater than 0.5, i.e., half way between perfect predictions ($q^2 = 1.0$) and no model at all ($q^2 = 0.0$). The optimum number of components was determined by minimizing PRESS while maximizing q^2 values. Whenever the increase in q^2 with

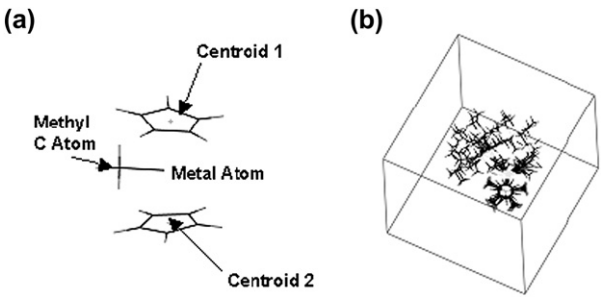

Fig. 2. (a) Points used for the alignment rule. (b) Alignment of the training set molecules and cubic region used to calculate the 3D fields.

an additional component was less than 5%, the model with fewer components was selected. Addition of more components improves the fitting statistics but has two disadvantages: on one hand, complicating the model and on the other loosing its predictive ability. Finally, subsequent non-cross-validated PLS analysis was carried out for the optimum number of components to obtain a final model.

The CoMFA results are graphically represented as 3D maps and for the sake of clarity only one catalyst (number **11**) of the set has been depicted in the corresponding figures.

3. Results

For clarity, this section has been divided into two parts according to the different experimental data considered in this study, i.e., polymerization activity and polymer molecular weight. A plot of both experimental variables is shown in Fig. 3. The unbridged zirconocene complex, used as a reference catalyst in the polymerization study, appears to be an outlier, clearly visible in the molecular weight data. This behavior has been confirmed in the analysis described in the following paragraphs. The experimental data have been transformed to the logarithmic scale (see Table 1) to give a uniform distribution as is recommended for the statistical analysis.

3.1. Polymerization activity

One of our previous studies [4a] showed a positive linear correlation between the activity data and the Cent–Zr–Cent angle, the wider the angle the better the activity. A somewhat similar observation can be made in the present work. Fig. 4 shows a graphical representation of the logarithm of the activity versus the Cent–Zr–Cent angle. If the outlier value is not considered, then the angle values can be grouped into two sets.

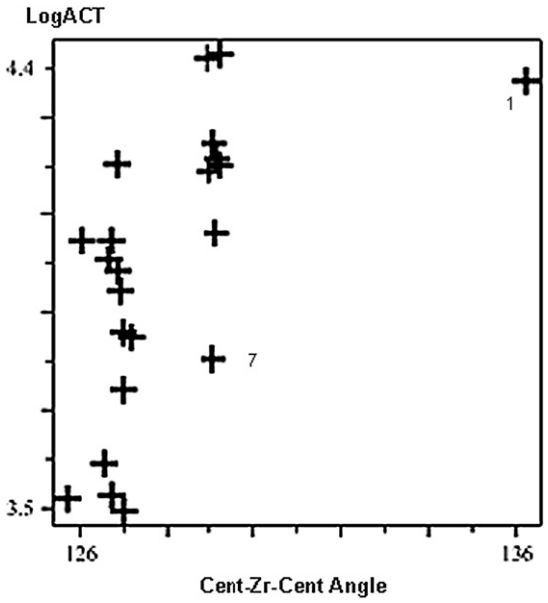

Fig. 4. Plot representing the logarithm of experimental activity versus the Cent–Zr–Cent angle (in degrees).

It can be observed that compounds containing two permethylated cyclopentadienyl rings have angle values around 129° and correspond to larger activity values. The least active compound in this group corresponds to catalyst number **7**, which contains a vinyl group as one of the substituents at the silicon atom *ansa*-bridge.

The presence of vinyl groups seems to have an influence on polymerization activity as will be shown later in this section. Conformational analysis performed on catalysts containing vinyl groups has confirmed that it is possible to find low energy conformers where the double bond can interact with the metal center via intra- or intermolecular interactions. This would probably give rise to deactivated catalyst molecules. For example, a conformer showing zirconium–vinyl

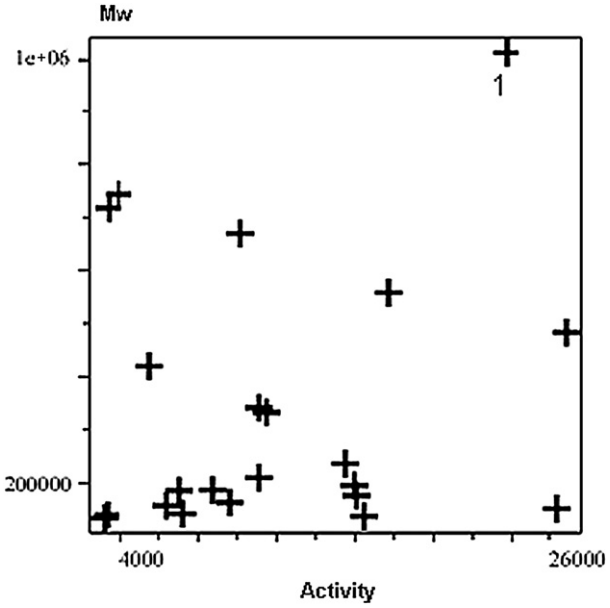

Fig. 3. Experimental activity versus polymer molecular weight plot.

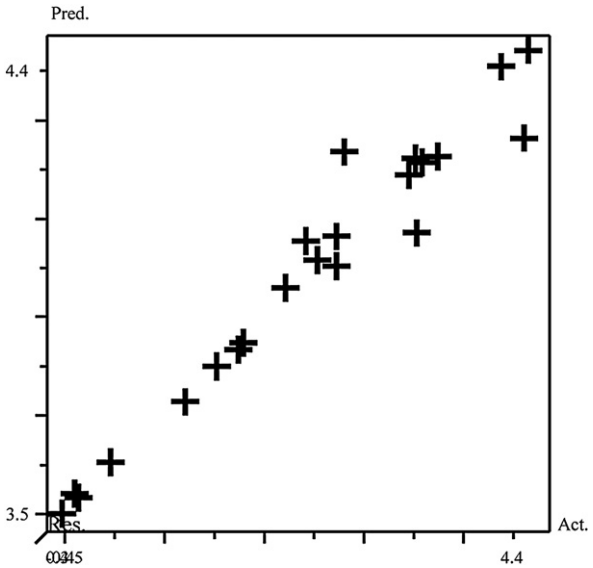

Fig. 5. Experimental versus predicted logarithm of activity plot.

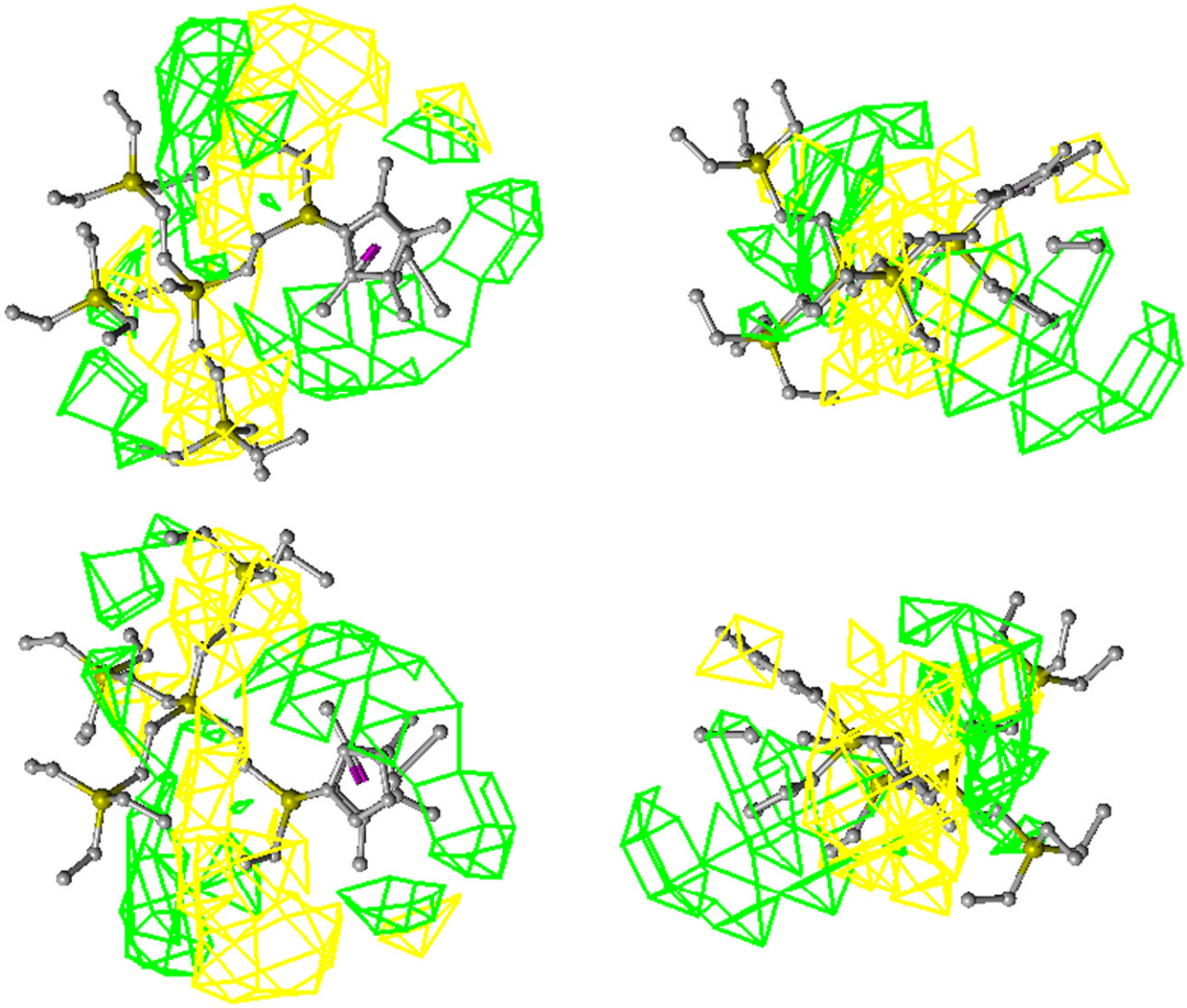

Fig. 6. Stdev × coeff. field isosurface for steric contribution mapped onto catalyst number **11**. Each image corresponds to different viewpoints. H atoms have been omitted for clarity.

group interaction has been found for catalyst number **8**. This conformer is 10 kcal/mol lower in energy than the extended conformation. The latter conformation was, however, selected to perform the QSAR study in order to be consistent in maintaining a similar shape for all molecules. The contribution of the conformational variability of the vinyl containing catalysts to the QSAR models was, however, small due to the low activity showed by these complexes. On the other hand, the high strength of this interaction is due to the absence of methyl groups in the cyclopentadienyl ligands in this catalyst. In this manner, it has been observed that zirconium–vinyl group interaction in catalyst number **10**, which contains one permethylated cyclopentadienyl ligand, is much more sterically hindered.

Different 3D fields such as steric, electrostatic potential, frontier molecular orbitals and local softness have been calculated and used to search correlations with the polymerization activity. CoMFA models including only steric and electrostatic fields have shown to give good correlations with the experimental activity. Those fields related with electronic features give poor models due to the large similarity found for these amongst the set of catalysts considered.

The best model found corresponds to a combination of steric and electrostatic fields evaluated in a 2 Å grid spacing region. The resulting cross-validated statistics were: cross-validated correlation coefficient, $q^2 = 0.501$ for 3 components; standard error of prediction, SEP = 0.246 log units. The final model with three "latent variables" including all catalysts gives the following statistics: correlation coefficient, $q^2 = 0.952$; standard error of prediction, SEP = 0.075 log units. The contribution of each field to the global model was 91% and 9% for the steric and electrostatic fields, respectively. The corresponding plot of experimental versus predicted polymerization activity is shown in Fig. 5. As can be seen a good correlation is obtained between both sets.

The PLS analysis gives information about the QSAR equation and its characteristics in terms of 3D fields. In this respect, it is generally accepted that the most informative field is that

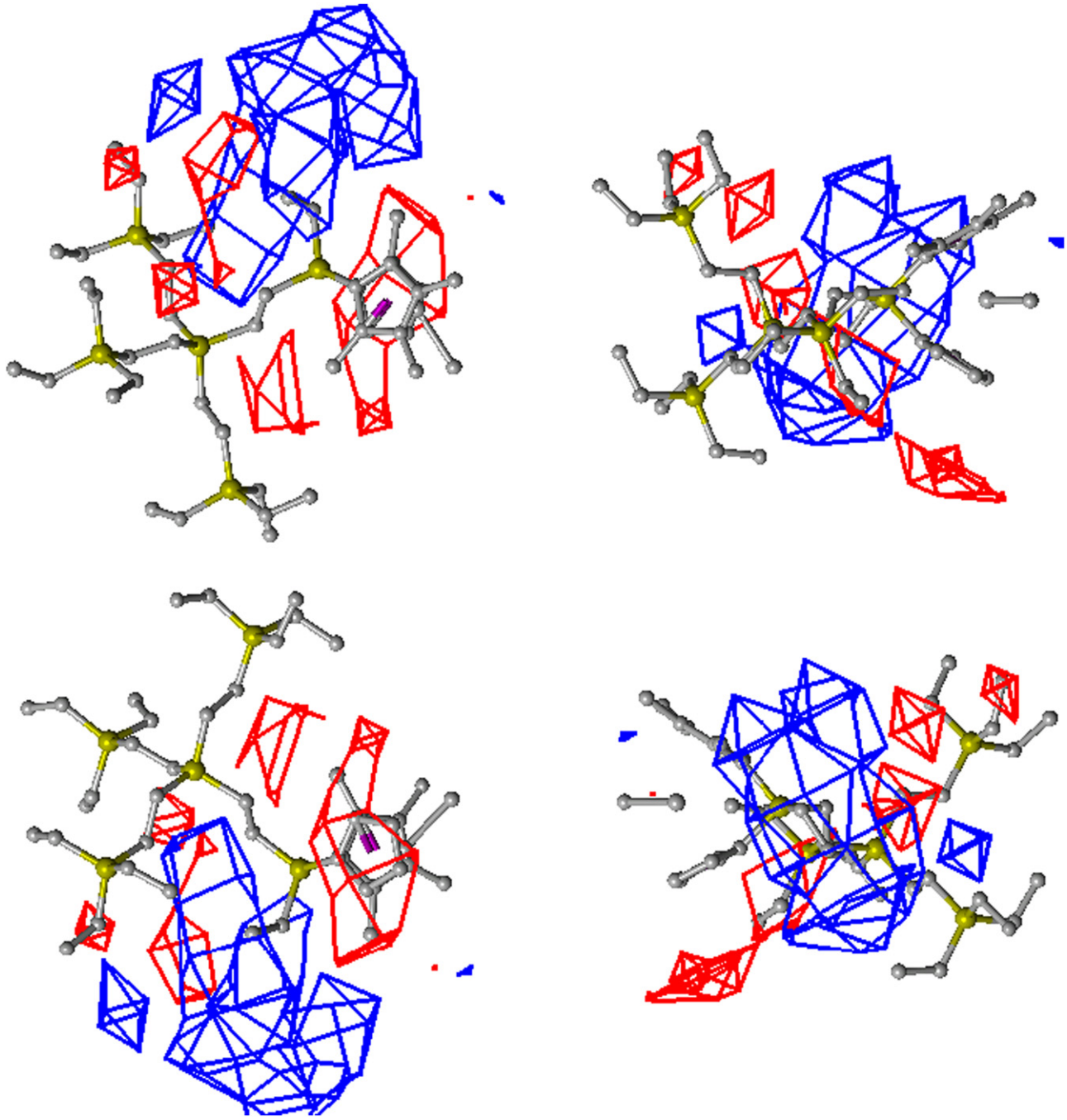

Fig. 7. Stdev × coeff. field isosurface for electrostatic contribution mapped onto catalyst number **11**. Each image corresponds to different viewpoints. H atoms have been omitted for clarity.

resulting from the product of the standard deviation at each grid point by the coefficient obtained for the same point in the QSAR equation.

The standard deviation (stdev) times QSAR coefficients (β) field gives a rough location where structure–activity relationship statements can be inferred discriminating areas where the local descriptor is important from those that have no significance [24]. Fig. 6 shows isosurfaces for the stdev × β field contoured at 0.4 (green, positive values) and −0.4 (yellow, negative values) onto catalyst number **11**. The positive values indicate where an increase of steric field corresponds to an increase in activity. Two main areas of positive steric influence can be observed, one around substituent positions in the aromatic cyclopentadienyl ligands, indicating that addition of bulky substituents should enhance activity. This could be interpreted in terms of the catalyst–co-catalyst interaction. A possible explanation could be that bulky substituents in these positions would prevent the approximation of the anionic co-catalyst species to the metal, so that complexation and insertion of the ethylene monomer are facilitated. This observation is in agreement with previous results obtained in our group on different catalyst series [4]. Another area of positive

contribution of the steric field is located on the opposite side of the active site (*ansa*-bridge region), although at some distance from the silicon bridge atom. This contribution is, however, more diffused than the permethyl contribution explained above, however, a plausible interpretation for this cannot be given at the moment.

On the other hand, negative values are associated with locations where it is necessary to release steric impediment to increase activity. The yellow isosurfaces in Fig. 6 correspond to those areas with a negative influence of the steric field on polymerization activity. These regions are located in the proximity of the silicon bridge atom indicating that steric hindrance around the bridge will be detrimental to the activity. Monomer insertion can be difficult if steric bulk is near to the active site.

The electrostatic field exerted by a molecule outside its van der Waals radius is supposed to be a principle descriptor of intermolecular interaction in 3D-QSAR [25]. The electrostatic nature of the catalyst active species can play an important role in the case of ethylene polymerization with metallocene catalysts where several molecules (monomer, solvent or cocatalyst) can interact through electrostatic forces with the metallocene compound.

The stdev $\times \beta$ field for the electrostatic component of the CoMFA model is shown in Fig. 7 contoured at 0.05 (blue, positive values) and -0.05 (red, negative values). The red mesh represents areas where an increase in negative charge will also contribute to improved polymerization activity. The negative areas are positioned in the vicinity of the aromatic ring substituents. These areas correspond to the most external surface of the catalyst above the active site. These observations suggest that negatively charged substituents in the aromatic rings would be able to facilitate the ion-pair separation, making more room for ethylene coordination to the metal center. This might result in an improvement of the catalyst activity as was mentioned in the previous section on steric field. This result is in agreement with models for the electrostatic effect on the catalyst activity obtained in a previous study [4b]. The areas corresponding to positive values of the stdev $\times \beta$ field can be associated with those molecular regions where vinyl groups are located. The model is thus indicating that release of these groups, which contain negative charge concentration, would be beneficial for the polymerization activity.

3.2. Polymer molecular weight

Reliable 3D-QSAR models have been obtained for the correlation of the experimental polymer molecular weight and a combination of steric and electrostatic molecular fields. Other fields related with electronic features such as Frontier Molecular Orbitals or Local Softness give models with very low correlation coefficients.

The activity versus molecular weight plot of experimental values in Fig. 3 showed catalyst **1** to be an outlier more pronounced in the molecular weight data set. This has been subsequently confirmed during the derivation of PLS models. The uncommon nature of the unbridged zirconocene complex with respect to the other compounds in the set is most likely responsible for the differential behavior of this catalytic species. Therefore, models to explain the molecular weight variability have been derived without the outlier catalyst.

The best model found corresponds to a combination of steric and electrostatic fields evaluated in a 1 Å grid spacing region. The resulting cross-validated statistics were: cross-validated correlation coefficient, $q^2 = 0.558$ for two components; standard error of prediction, SEP = 0.19 log units. The final model with two latent variables including all catalysts gives the following statistics: correlation coefficient, $q^2 = 0.776$; standard error of prediction, SEP = 0.132 log units. The contribution of each field to the global model was 94% and 6% for the steric and electrostatic fields, respectively. The model for the molecular weight gives better predictive capability, measured by the q^2 value, than the model obtained for the activity although the final statistics are worse. The corresponding plot of experimental versus predicted polymer molecular weight is shown in Fig. 8. It can be seen a dispersion of points at lower molecular weight values. The farthest points from the diagonal line correspond to some of the catalysts containing a vinyl group as one of the silicon bridge atom substituents (numbers **7**, **9**, **10** and **14**). It seems that molecular weight data are more sensible to the presence of such groups.

The stdev $\times \beta$ field for the steric component of the CoMFA model is shown in Fig. 9 as an isosurface contoured at 0.3 (green, positive values) and -0.3 (yellow, negative values). Two areas of positive steric contribution can be observed. On one hand, steric bulk around the cyclopentadienyl ring substituents seems to be beneficial in increasing the polymer molecular weight. This contribution to the model is, however, smaller than the corresponding analogue in the activity case described in the previous section. On the other hand, a large area of positive steric effect is located in the region occupied by the silicon bridge atom substituents. The meaning of this contribution is, however, not clear. Some parallelism with the supported metallocene catalysis can be observed. The effect of the support on

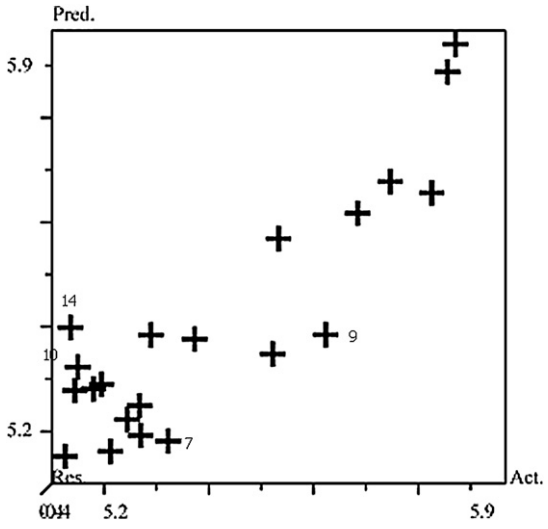

Fig. 8. Experimental versus predicted logarithm of M_w plot.

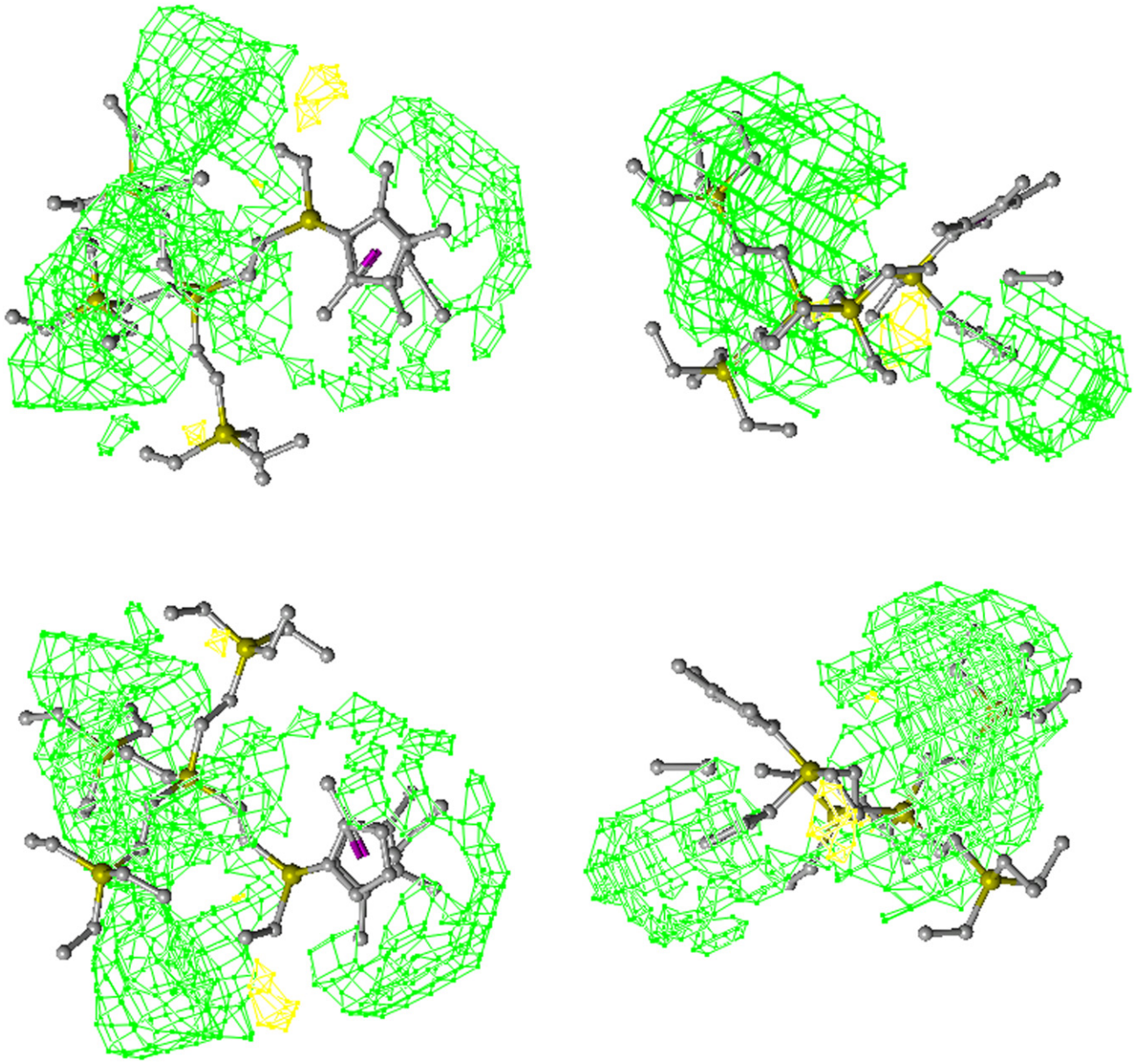

Fig. 9. Stdev × coeff. field isosurface for steric contribution mapped onto catalyst number **11**. Each image corresponds to different viewpoints. H atoms have been omitted for clarity.

olefin polymerization is in general to increase the polymer molecular weight with respect to the unsupported case. It can be supposed that the steric hindrance built up by the support disfavors the polymerization termination reactions [8]. The effect of the catalyst volume on the polymer molecular weight can also be observed by plotting the experimental values against the total molecular volume as in Fig. 10. A positive correlation between both variables can be observed which is in agreement with the 3D-QSAR model. This correlation can be improved after removing those catalysts containing vinyl group substituents. The negative steric contribution is very small and does not deserve further comment.

The stdev × β field for the electrostatic component of the CoMFA model is shown in Fig. 11 contoured at 0.05 (blue, positive values) and −0.05 (red, negative values). The contribution

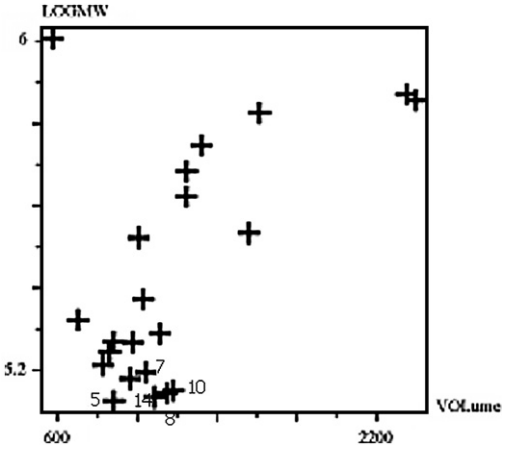

Fig. 10. Plot representing the logarithm of M_w versus the catalyst molecular volume (in angstrom).

of positive electrostatic potential can be observed in those areas occupied by vinyl group substituents, as was observed in the case of catalytic activity. The same interpretation can also be given here. The electrostatic negative areas are of minor relevance and are spread around the bridge atom substituent area.

4. Conclusions

Some interesting 3D-QSAR models have been found to explain the variance in certain experimental data such as polymerization activity and polymer molecular weight obtained for a set of silylene bridged bis(cyclopentadienyl) zirconium complexes. The CoMFA models derived in the present work are in agreement with those models obtained in previous 3D-QSAR studies applied to metallocene-based homogeneous catalysis [4].

The positive correlation between polymerization activity and steric hindrance in the cyclopentadienyl substituents has been confirmed. This result is in agreement with the dependence observed between Cent–Zr–Cent angle and activity. In particular, the presence of two permethylated cyclopentadienyl rings gives rise to larger Cent–Zr–Cent angles. This can be associated to a weakening of the cation–anion interaction, which allows a more efficient olefin insertion. The steric contribution

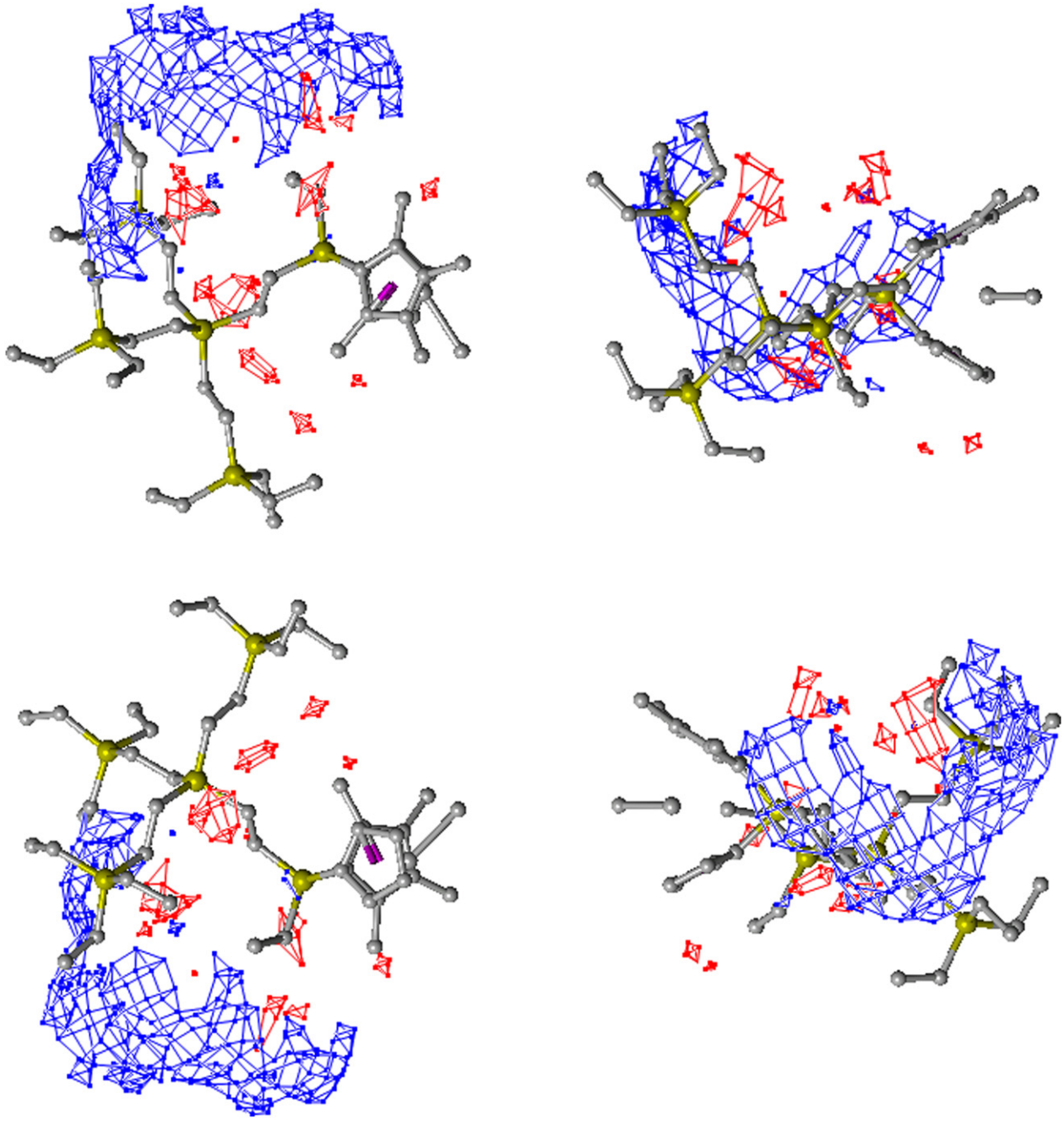

Fig. 11. Stdev × coeff. field isosurface for electrostatic contribution mapped onto catalyst number **11**. Each image corresponds to different viewpoints. H atoms have been omitted for clarity.

constitutes more than 90% of the CoMFA model for the activity as well as for the molecular weight dependent variables.

There are, however, some special peculiarities of the QSAR models described in this work not encountered in previous studies. It has been found that an increase in steric bulk in the region defined by the silicon bridge atom substituents is beneficial for larger polymer molecular weights. This result is completely new, as previous 3D-QSAR studies did not take into account this molecular area. The positive correlation between molecular weight and catalyst volume can be also directly observed by a bivariate plot (Fig. 10). No clear interpretation is given for this correlation but the analogy with the observed polymer molecular weight increase in supported metallocene-based catalysis has been remarked upon.

On the other hand, correlation between dependent variables and electronic fields associated to nucleophilic reactivity like frontier molecular orbitals or local softness has not been found. This result can be explained by the high similarity of these fields for the catalyst set considered in this work. In fact, all the compounds contained very similar aromatic ligands coordinated to the zirconium atom. It should be stressed again that the main source of structural variability is focused on the substituents attached to the Si bridge atom.

The good predictive ability of the models allows one to consider the application of the 3D-QSAR methodology as a valuable tool in the design of better catalysts for olefin polymerization.

The fundamental methodologies used in 3D-QSAR are well established and can provide robust models from a statistical point of view. However, some details should be improved in order to get better results.

One concern is about the large number of descriptors that are used for the correlations. Possible solutions have been pointed out by Cruciani's group regarding the application of variable selection methods in addition to the PLS methodology.

Other improvements in the 3D-QSAR technique try to overcome the molecular alignment problem. The requirement is that all molecules in the set should be aligned so that each point in the grid had the same meaning for each compound. This introduces some arbitrarities which greatly affect the outcome of the final model. The implementation of Auto-correlation and Cross-correlation techniques simplifies this problem to the requirement of only one arbitrary point to be the common origin for all compounds. This point can be easily identified in the homogeneous catalysis case to be the metallic center, for example.

In summary, we believe that chemometric tools can play an important role to help in the rational design of polymerization catalysts. In particular, 3D-QSAR can yield useful and precise information about the influence of molecular structure details on the catalytic properties of the organometallic compounds, much as it is done in the drug design field.

Acknowledgments

We gratefully acknowledge financial support from the Ministerio de Educación y Ciencia, Spain (Grants nos. CTQ2005-07918-C02-02/BQU and MAT2006-0400), the Comunidad de Madrid (S-0505/PPQ-0328) and the Universidad Rey Juan Carlos (graduate fellowship for D.P-C). The authors also acknowledge Centro Técnico de Informática (CTI-CSIC, Madrid, Spain), Centro de Supercomputación de Galicia (CESGA, Santiago de Compostela, Spain) and Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT; Madrid, Spain) for the use of their computational resources.

Appendix. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.polymer.2007.05.081.

References

[1] Andresen A, Cordes H-G, Herwig J, Kaminsky W, Merck A, Mottweiler R, et al. Angew Chem Int Ed Engl 1976;15:630.
[2] See for example:
(a) Alt HG, Köppl A. Chem Rev 2000;100:1205;
(b) Coates GW. Chem Rev 2000;100:1223;
(c) Resconi L, Cavallo L, Fait A, Piemontesi F. Chem Rev 2000;100:1253;
(d) Brintzinger HH, Fischer D, Mülhaupt R, Rieger B, Waymouth RM. Angew Chem Int Ed Engl 1995;36:1143.
[3] Prashar S, Antiñolo A, Otero A. Coord Chem Rev 2006;250:133.
[4] (a) Cruz V, Ramos J, Muñoz-Escalona A, Lafuente P, Peña B, Martínez-Salazar J. Polymer 2004;45:2061;
(b) Cruz V, Ramos J, Martinez S, Muñoz-Escalona A, Martínez-Salazar J. Organometallics 2005;24:5095.
[5] Kaminsky W. Macromol Chem Phys 1996;197:3907.
[6] (a) Muñoz-Escalona A, Ramos J, Cruz V, Martínez-Salazar J. J Polym Sci Part A Polym Chem 2000;38:571;
(b) Ramos J, Muñoz-Escalona A, Cruz V, Martínez-Salazar J. Polymer 2000;41:6161;
(c) Ramos J, Cruz V, Muñoz-Escalona A, Martínez-Salazar J. Polymer 2001;42:7275;
(d) Ramos J, Muñoz-Escalona A, Cruz V, Martínez-Salazar J. Polymer 2001;42:8019;
(e) Ramos J, Cruz V, Muñoz-Escalona A, Martínez-Salazar J. Polymer 2002;43:3635;
(f) Ramos J, Muñoz-Escalona A, Cruz V, Martínez-Salazar J. Polymer 2003;44:2169;
(g) Ramos J, Muñoz-Escalona A, Cruz V, Martínez-Salazar J. Polymer 2003;44:2177;
(h) Exposito MT, Martinez S, Ramos J, Cruz V, Lopez M, Muñoz-Escalona A, et al. Polymer 2004;45:9029;
(i) Ramos J, Muñoz-Escalona A, Martínez S, Martínez-Salazar J, Cruz V. J Chem Phys 2005;122:74901;
(j) Martínez S, Exposito MT, Ramos J, Cruz V, Martínez MC, Muñoz-Escalona A, et al. J Polym Sci Part A Polym Chem 2005;43:711;
(k) Martínez S, Ramos J, Cruz V, Martínez-Salazar J. J Polym Sci Part A Polym Chem 2006;44:4752;
(l) Martínez S, Ramos J, Cruz VL, Martínez-Salazar J. Polymer 2006;47:883.
[7] (a) Antiñolo A, López-Solera I, Orive I, Otero A, Prashar S, Rodríguez AM, et al. Organometallics 2001;20:71;
(b) Antiñolo A, López-Solera I, Otero A, Prashar S, Rodríguez AM, Villaseñor E. Organometallics 2002;21:2460;
(c) Alonso-Moreno C, Antiñolo A, López-Solera I, Otero A, Prashar S, Rodríguez AM, et al. J Organomet Chem 2002;656:129;
(d) Antiñolo A, Fernández-Galán R, Orive I, Otero A, Prashar S. Eur J

Inorg Chem 2002;2470;
(e) Antiñolo A, Fernández-Galán R, Gallego B, Otero A, Prashar S, Rodríguez AM. Eur J Inorg Chem 2003;2626;
(f) Antiñolo A, Fajardo M, Gómez-Ruiz S, López-Solera I, Otero A, Prashar S, et al. J Organomet Chem 2003;683:11;
(g) Antiñolo A, Fajardo M, Gómez-Ruiz S, López-Solera I, Otero A, Prashar S. Organometallics 2004;23:4062;
(h) Antiñolo A, Fernández-Galán R, Otero A, Prashar S, Rivilla I, Rodríguez AM, et al. Organometallics 2004;23:5108;
(i) Gómez-Ruiz S, Prashar S, Fajardo M, Antiñolo A, Otero A, Maestro MA, et al. Polyhedron 2005;24:1298;
(j) Alonso-Moreno C, Antiñolo A, Carrillo-Hermosilla F, Carrión P, López-Solera I, Otero A, et al. Eur J Inorg Chem 2005;2924;
(k) Antiñolo A, Fernández-Galán R, Otero A, Prashar S, Rivilla I, Rodríguez AM. J Organomet Chem 2006;691:2924.

[8] (a) Hlatky GG. Chem Rev 2000;100:1347;
(b) Bialek M, Czaja K, Reszka A. J Polym Sci Part A Polym Chem 2005;43:5562.

[9] Muir RM, Hansch C. Nature 1962;194:178.

[10] Shen M, Beguin C, Golbraikh A, Stables JP, Kohn H, Tropsha A. J Med Chem 2004;47:2356.

[11] Kubinyi H, editor. 3D QSAR in drug design: theory methods and applications. Dordrecht: Kluwer–ESCOM; 2000.

[12] Becke AD. Phys Rev A 1998;38:3098.

[13] Hay PJ, Wadt WR. J Chem Phys 1985;82:299.

[14] Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, et al. Gaussian 03 revision C.02. Wallingford, CT: Gaussian Inc.; 2004.

[15] Breneman CM, Wiberg KB. J Comput Chem 1990;11:361.

[16] Clark M, Cramer III RD, Van Opdenbosch N. J Comput Chem 1989;10:982.

[17] Cramer III RD, Patterson DE, Bunce JD. J Am Chem Soc 1988; 110:5959.

[18] SYBYL Molecular Modeling System, Tripos Inc., 1699 S Hanley Rd, St Louis, MO 63144.

[19] Martin YC, Kim KH, Liu CT. In: Charton M, editor. Advances in quantitative structure–property relationships, vol. 1. Greenwich, CT: JAI Press; 1996. p. 1–52.

[20] Rännar S, Lindgren F, Geladi P, Wold S. J Chemom 1994;8:111.

[21] Wold S, Ruhe A, Wold H, Dunn WJ. SIAM J Sci Stat Comput 1984;5:735.

[22] Cramer III RD, Bunce JD, Patterson DE, Frank IE. Quant Struct-Act Relat 1988;7:18.

[23] Clark M, Cramer III RD. Quant Struct-Act Relat 1993;12:137.

[24] Cramer III RD, DePriest SA, Patterson DE, Hecht P. In: Kubinyi H, editor. 3D QSAR in drug design: theory methods and applications. Dordrecht: Kluwer–ESCOM; 2000. p. 443–85.

[25] Wade RC. In: Kubinyi H, editor. 3D QSAR in drug design: theory methods and applications. Dordrecht: Kluwer/ESCOM; 2000. p. 486–505.