Received: 10 May 2008,

Revised: 30 September 2008,

Published online in Wiley InterScience: 3 December 2008

(www.interscience.wiley.com) DOI 10.1002/poc.1491

Quantitative structure-reactivity relationship studies on the catalyzed Michael addition reactions

Bahram Hemmateenejad^{a,b*}, Mahmood Sanchooli^c and Ahmadreza Mehdipour^b

Quantitative structure-reactivity relationship (QSRR) can be considered as a variant of quantitative structureproperty relationship (QSPR) studies, where the chemical reactivity of reactants or catalysts in a specified chemical reaction is related to chemical structure. In this manner, the Michael addition of some different substrates using different catalysts (SDS, silica gel, and ZrOCl₂) was subjected to structure-reactivity relationship, quantitatively. Multiple linear regression (MLR) and partial least square (PLS) were used to perform the QSRR analysis. The resulted models for different catalyzed reactions showed that the catalysts probably act in different mechanisms since the models obtained for the catalysts included different parameters from substrate and enones. Overall, it was found that the reactivity in Michael addition reactions is controlled by coulombic (dipole and charge) interactions as well as the orbital energetic parameters. In the presence of different catalysts, the relative importance of these parameters is changed and hence the catalytic activity is changed. Copyright © 2008 John Wiley & Sons, Ltd.

Keywords: quantitative structure-reactivity relationship; catalysis; Michael addition; quantum chemical; PLS

INTRODUCTION

The acid or base-induced conjugate addition of nucleophiles to α , β -unsaturated carbonyl compounds, Michael addition, is among the most useful carbon–carbon and carbon–heteroatom bond forming reactions. Michael addition reactions generally require the activation of the starting materials under acidic or basic conditions.^[1] This led to development of catalytic methods especially using a large number of Lewis acid catalysts.^[2–4]

Design of catalysts with high catalytic efficiency is an important research in chemistry and chemical catalysis.^[5,6] Catalyst designers use rational knowledge based on previous catalytic activities reported from old catalysts. In this way, a large number of chemical catalysts are prepared; among which some of them represent desirable catalytic activity. Therefore resources and time are wasted during discovery of new catalysts. On the other hand, the use of computational methods for designing of molecules with desired reactivity or property has been growing area in chemistry. Thus, computational methods are now becoming the expedite source of introducing new catalysts.^[7,8]

Quantitative structure property relationships (QSPR), as emerging computational techniques in chemistry, make a mathematical connection between chemical property of interest for a class of compounds (such as boiling and melting point, acid–base behavior, chromatographic retention indices, partitioning phenomena, reaction kinetics and equilibriums and so on) and encoded molecular structural parameters named molecular descriptors.^[9,10] By the use of different chemometrics methods, a reasonable relationship between chemical property and structural parameters is discovered, by which chemists can obtain a deeper knowledge about the chemical system under study in one hand and predicting the chemical property of interest for new or even non-synthesized molecules on the other hand. This leads to the design of molecules with optimized property such as catalytic activity.

Quantitative structure-reactivity relationship (QSRR) can be considered as a variant of QSPR studies, where the chemical reactivity of reactants or catalysts in a specified chemical reaction is related to chemical structure.^[11–14] The history of the structure-reactivity relationship modeling goes back to the end of 1970 decade when Carpenter and coworkers qualitatively studied the effects of substituents and of benzannelation on the rates of pericyclic reactions. Then, they used semi-empirical quantum chemical methods and found quantitative models for a variety of unsaturated hydrocarbons.^[15,16]

^e Correspondence to: B. Hemmateenejad, Department of Chemistry, Shiraz University, Shiraz, Iran. F-mail: hemmath@sums.ac.ir

a B. Hemmateenejad

- Department of Chemistry, Shiraz University, Shiraz, Iran
- b B. Hemmateenejad, A. Mehdipour Medicinal & Natural Products Chemistry Research Center, Shiraz University of Medical Sciences, Shiraz, Iran
- c M. Sanchooli Department of Chemistry, Zabol University, Zabol, Iran

Very recently, Firouzabadi and coworkers designed many Lewis acid catalysts for different carbon–carbon and carbon–heteroatom bond formation by Michael addition reactions.^[17–22] They used some different catalysts such as $ZrOCI_2 \cdot 8H_2O$, silica gel and sodium dodecyl sulfate (SDS) and the results showed high yields under solvent-free condition.^[20–22] They found that catalytic activity not only depended on chemical structure of substrates but also depended on the type of the catalyst.

In this article we conducted a QSRR study on the catalyzed Michael addition reactions, performed by Firouzabadi and coworkers, in order to find the quantitative effects of reactants structures on the efficiency of these reactions. In the case of each catalyst, separate QSRR models were developed using subset of descriptors calculated by quantum chemical calculations. The models were used to describe the effects of molecular structures as well as catalyst on the reaction yields.

COMPUTATIONAL METHODS

Data set

The data sets used were the yield of Michael addition of different substrate to three enones in the presence of ZrOCl₂·8H₂O, silica gel and SDS as catalyst. The data were refined from the papers of Firouzabadi and coworkers.^[17–22] The chemical structures of the substrates and enones are represented in Fig. 1 and the reactivity data are listed in Table 1. The reactivity index (RI) was taken as the logarithm of yield over reaction time, that is, reactions having higher yield in lower time considered as highly efficient.

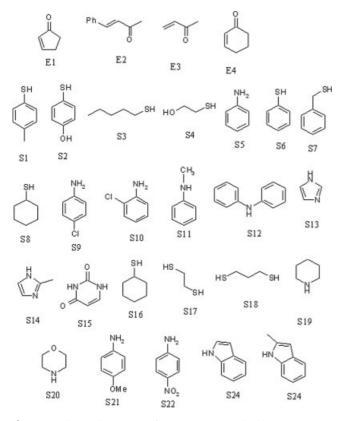


Figure 1. Chemical structures of enones (E1–E4) and Substrates (S1–S24) used in this study

Quantum chemical computations

We used the electronic properties of both substrate and enones to derive QSRR models. Hyperchem software (Hypercube, Inc., version 7) was used to build the molecular structures into computer, which then optimized utilizing GUASSIAN 98.[23] The structures were optimized using 6-31 G basis set for all atoms. Considering no molecular symmetry constraint, all bond lengths and angles optimization were carried out at level of UHF/6-31 G. The calculated electronic descriptors for each molecule are briefly described in Table 2. Hardness (HD), softness (SOF), electronegativity (EN), and electrophilicity (EPH) are the important electronic futures used to describe stability, reactivity, chemical potential and other related properties of molecules. Hardness has been used to understand chemical reactivity and stability of molecules.^[24] Electronegativity was introduced by Pauling as a power of an atom in a molecule to attract electron to itself. Softness is a property of molecule that measures the extent of chemical reactivity. Electrophilicity was proposed by Parr et al.[24] as a measure of energy lowering due to maximal electron flow between donor and acceptor. Taking into account ionization potential (I) $\approx -E_{HOMO}$ and electron affinity (A) $\approx -E_{LUMO}$, the quantum chemical indices of HD, SOF, EN and EPH can be calculated as $HD = (I - A)/2 = (E_{LUMO} - E_{HOMO})/2$, $EN = (I + A)/2 = -(E_{LUMO} + A)/2$ E_{HOMO})/2, SOF = 1/HD, and EPH = -EN²/2HD.^[26-28]

The calculated descriptors can be classified into three different electronic categories including local charges, dipoles, and orbital energies. For each molecule (both substrates and enones) 12 electronic descriptors (Table 2) were calculated. In addition, two parameters including the difference between HOMO and LUMO levels of substrate and enone were used. Therefore, the descriptor data matrix of each reaction (i.e., each entry of Table 1) composed of 26 columns (each 12 columns for electronic descriptors of substrate and enone and two additional parameters of differences between LUMO and HOMO energies).

Data processing

In the case of each catalyst, two type of data were available; a descriptor data matrix of $(m \times n)$ dimension (**X**), where m and n denote the number of reaction entry and number of descriptors, respectively, and a column vector (y) of size m, whose elements were the reactivity indices (RI). The guantitative relationships between RI and the calculated descriptors were obtained by partial least squares (PLS) and multiple linear regression (MLR) combined with genetic algorithm (GA) as feature selection method.^[29,30] Because the calculated descriptors have different scales, they were scaled to zero mean and unit variances (autoscaling).^[31] The GA-PLS procedure was similar to our previous studies.^[29,30] In the PLS model development steps, cross-validated correlation coefficient (Q^2) was used to validate the models. Both leave-one-out and leave-some-out crossvalidation methods were employed. To test the true prediction ability of the resulted models, a prediction set composed of 25% of the available data was selected randomly. In addition, the robustness of resulted models was checked by input scrambling test.^[32,33] For each set of reactivity data, the RI values were randomly attributed to the reaction entries and the PLS modeling was repeated with the randomized data for 10 times. If the statistical qualities of these models are much lower than the original model, it can be considered that the model is reasonable and had not been obtained by the chance.^[34]

Table 1. Experimental and leave-one-out cross-validation predicted values of reactivity index of the Michael addition reactions in the presence of different catalysts

Substrate S1 S2 S2 S3 S3 S4 S5 S1 S1 S6 S6	Enone E1 E2 E3 E4 E3 E4 E3 E3 E3 E3 E3 E4 E3	Exp 0.968 0 1.676 1.123 0.199 -0.154 0.009 0.477 1.681	Silica gel Pred ^a 0.945 0.199 1.395 0.994 0.021 -0.379 0.120 0.741	Pred ^b 1.124 0.100 1.649 1.170 0.142 0.346	Exp	SDS Pred ^a	Pred ^b 	Exp 	ZrOCl ₂ Pred ^a	Pred ^b
S1 S1 S2 S3 S3 S4 S5 S1 S1 S6 S6	E1 E2 E3 E4 E3 E4 E3 E3 E3 E3 E4	0.968 0 1.676 1.123 0.199 -0.154 0.009 0.477 1.681	0.945 0.199 1.395 0.994 0.021 -0.379 0.120	1.124 -0.100 1.649 1.170 0.142		Pred ^a 	Pred ^b — —		Pred ^a —	_
S1 S2 S3 S3 S4 S5 S1 S1 S6 S6	E2 E3 E4 E3 E4 E3 E3 E3 E3 E4	0 1.676 1.123 0.199 -0.154 0.009 0.477 1.681	0.199 1.395 0.994 0.021 -0.379 0.120	-0.100 1.649 1.170 0.142	 			_	_	_
S2 S2 S3 S4 S5 S1 S1 S6 S6	E3 E4 E3 E4 E3 E3 E3 E4	1.676 1.123 0.199 -0.154 0.009 0.477 1.681	1.395 0.994 0.021 0.379 0.120	1.649 1.170 0.142		_	_			
S2 S3 S4 S5 S1 S1 S6 S6	E4 E3 E4 E3 E3 E3 E3 E4	1.123 0.199 0.154 0.009 0.477 1.681	0.994 0.021 0.379 0.120	1.170 0.142	_	_	_			
S3 S3 S4 S5 S1 S1 S6 S6	E3 E4 E3 E3 E3 E4	0.199 -0.154 0.009 0.477 1.681	0.021 0.379 0.120	0.142	—			_	_	_
S3 S4 S5 S1 S1 S6 S6	E4 E3 E3 E3 E4	-0.154 0.009 0.477 1.681	-0.379 0.120			—	—	—		—
S4 S5 S1 S1 S6 S6	E3 E3 E3 E4	0.009 0.477 1.681	0.120	-0.346	—	—		—	—	_
S5 S1 S1 S6 S6	E3 E3 E4	0.477 1.681			—	—	—	—	—	—
S1 S1 S6 S6	E3 E4	1.681	0.741	-0.028	—	_	—	—	—	—
S1 S6 S6	E4			0.814	0.788	0.047	-0.010	1.070	0.880	1.086
S6 S6			1.233	1.256	1.283	1.154	1.200	—	—	—
S6	F3	0.977	0.831	0.778	0.959	0.571	0.626	—	—	—
		1.274	1.236	1.190	1.265	1.667	1.715	—		
	E4	0.977	0.834	0.711	1.283	1.085	1.141	—	—	_
S7	E3	0.653	1.164	1.017	1.260	0.911	0.882	—	—	—
S7								—	—	
S8		0.124	0.050	0.254				—	—	
S9		—	—	_				—	—	
S9		—	—	_				—	—	
S10		—	—					—	—	—
S11	E3	—	—	_			0.131	—	—	_
S12	E3	_	—		0.815	0.741	0.989	—	_	—
S13	E3	_	—		0.462	0.303	0.077	—	_	—
S13	E4	—	—	_	-0.875	-0.279	-0.496	—	—	—
S14	E3	—	—	_	0.681	0.233	0.159	—	—	_
S14	E4	_	—		-1.141	-0.348	-0.414	—	_	—
S15	E3	_	—		0.098	-0.159	0.058	—	_	—
S6	E1	_	—		1.255	1.279	1.332	—	_	—
S6	E2	_	—		-1.459	-1.632	-1.536	—	_	—
S16	E4	—	—	_	0.954	0.797	0.865	—	—	—
S17	E3	_	—		0.426	0.956	0.762	—	_	—
S17	E4	_	—		0.007	0.373	0.195	—	_	—
S18	E3	_	—		0.386	0.981	0.867	—	_	_
S18	E4	_	—		-0.015	0.398	0.293	—	_	—
S5	E4	_	—	—	-0.222	-0.535	-0.583	0.748	0.880	0.652
S19	E3	—	—	—	1.181	0.780	1.014	1.278	1.019	1.171
S20	E3	—	—	—	1.190	0.833	1.109	1.074	1.118	1.298
S19	E4	_	—		—		—	1.070	1.019	0.736
S20	E4	_	—	—	—	—	—	0.893	1.118	0.864
S5	E1	_	—	—	—	—	—	0.753	0.880	0.761
S21	E3	—	—	—	—	—		1.676	1.394	1.558
S21	E4	—	—	—	—	—		1.070	1.394	1.124
S22	E3	—	—	—	—	—		0.898	0.646	1.016
S22	E4	—	—	_	_	_	_	0.672	0.646	0.770
S11	E3	—	—	_	_	_	_	0.574	0.282	0.450
S23	E3	—	—	_	_	_	_	-0.192	0.315	-0.170
S24	E3	—	—	_	_	_	_	0.500	0.382	0.737
S24	E4	_	_	_	_	_	_	0.274	0.382	0.303
	57 58 59 59 510 511 512 513 513 514 514 514 515 56 56 516 517 518 518 55 519 520 519 520 55 521 521 522 521 522 521 522 522 511 522 522	57 E4 58 E3 59 E4 50 E3 59 E4 50 E3 510 E3 511 E3 512 E3 513 E4 514 E4 515 E3 56 E1 56 E1 56 E2 516 E4 517 E3 518 E4 519 E3 520 E3 519 E4 520 E3 521 E4 522 E3 521 E4 522 E3 521 E4 522 E3 522 E4 521 E3 522 E3 523 E3 524 E3	57 E4 0.176 58 E3 0.124 59 E3 $ 59$ E4 $ 510$ E3 $ 511$ E3 $ 512$ E3 $ 512$ E3 $ 513$ E3 $ 513$ E4 $ 513$ E4 $ 514$ E3 $ 514$ E4 $ 515$ E3 $ 566$ E1 $ 576$ E4 $ 516$ E4 $ 517$ E3 $ 518$ E4 $ 518$ E4 $ 520$ E3 $ 520$ E3 $ 521$ E3 $ 521$ E3 $ 522$ E3 $ 522$ E3 $ 522$ <td>57 E4 0.176 0.762 58 E3 0.124 0.050 59 E3 - - 59 E4 - - 510 E3 - - 511 E3 - - 512 E3 - - 513 E3 - - 513 E4 - - 513 E3 - - 514 E3 - - 514 E4 - - 514 E4 - - 516 E3 - - 516 E4 - - 517 E3 - - 518 E4 - - 519 E3 - - 520 E3 - - 521 E3 - - 521 E3 - - 521 E4 - -</td> <td>57 E4 0.176 0.762 0.532 58 E3 0.124 0.050 0.254 59 E3 59 E4 510 E3 511 E3 512 E3 512 E3 513 E3 514 E3 514 E3 514 E3 514 E3 516 E4 517 E3 518 E4 519 E4 520 E4</td> <td>37 E4 0.176 0.762 0.532 0.973 38 E3 0.124 0.050 0.254 1.241 39 E3 -0.130 39 E4 -0.130 310 E3 -0.130 511 E3 -0.023 512 E3 0.023 513 E3 0.023 513 E3 0.023 514 E3 0.023 514 E3 0.023 514 E3 0.681 514 E3 0.098 56 E1 1.255 56 E2 0.0954 517 E3 0.0386 518 E4 $-$</td> <td>57E4$0.176$$0.762$$0.532$$0.973$$0.329$$58$E3$0.124$$0.050$$0.254$$1.241$$1.380$$59$E3$-0.130$$0.044$$59$E4$-0.135$$-0.538$$510$E3$-0.150$$0.018$$511$E3$0.023$$0.108$$512$E3$0.815$$0.741$$513$E3$0.875$$-0.279$$514$E3$0.681$$0.233$$514$E4$-0.875$$-0.279$$514$E3$-0.681$$0.233$$514$E3$-0.681$$0.233$$514$E3$-0.681$$0.233$$515$E3$-0.098$$-0.159$$56$E1$-0.225$$1.632$$516$E4$0.007$$0.373$$518$E3$0.386$$0.981$$518$E4$520$E4$521$E3$521$E3$-$<</td> <td>$\begin{array}{cccccccccccccccccccccccccccccccccccc$</td> <td>$\begin{array}{cccccccccccccccccccccccccccccccccccc$</td> <td>$\begin{array}{cccccccccccccccccccccccccccccccccccc$</td>	57 E4 0.176 0.762 58 E3 0.124 0.050 59 E3 - - 59 E4 - - 510 E3 - - 511 E3 - - 512 E3 - - 513 E3 - - 513 E4 - - 513 E3 - - 514 E3 - - 514 E4 - - 514 E4 - - 516 E3 - - 516 E4 - - 517 E3 - - 518 E4 - - 519 E3 - - 520 E3 - - 521 E3 - - 521 E3 - - 521 E4 - -	57 E4 0.176 0.762 0.532 58 E3 0.124 0.050 0.254 59 E3 $ 59$ E4 $ 510$ E3 $ 511$ E3 $ 512$ E3 $ 512$ E3 $ 513$ E3 $ 514$ E3 $ 514$ E3 $ 514$ E3 $ 514$ E3 $ 516$ E4 $ 517$ E3 $ 518$ E4 $ 519$ E4 $ 520$ E4	37 E4 0.176 0.762 0.532 0.973 38 E3 0.124 0.050 0.254 1.241 39 E3 $ -0.130$ 39 E4 $ -0.130$ 310 E3 $ -0.130$ 511 E3 $ -0.023$ 512 E3 $ 0.023$ 513 E3 $ 0.023$ 513 E3 $ 0.023$ 514 E3 $ 0.023$ 514 E3 $ 0.023$ 514 E3 $ 0.681$ 514 E3 $ 0.098$ 56 E1 $ 1.255$ 56 E2 $ 0.0954$ 517 E3 $ 0.0386$ 518 E4 $-$	57E4 0.176 0.762 0.532 0.973 0.329 58 E3 0.124 0.050 0.254 1.241 1.380 59 E3 $ -0.130$ 0.044 59 E4 $ -0.135$ -0.538 510 E3 $ -0.150$ 0.018 511 E3 $ 0.023$ 0.108 512 E3 $ 0.815$ 0.741 513 E3 $ 0.875$ -0.279 514 E3 $ 0.681$ 0.233 514 E4 $ -0.875$ -0.279 514 E3 $ -0.681$ 0.233 514 E3 $ -0.681$ 0.233 514 E3 $ -0.681$ 0.233 515 E3 $ -0.098$ -0.159 56 E1 $ -0.225$ 1.632 516 E4 $ 0.007$ 0.373 518 E3 $ 0.386$ 0.981 518 E4 $ 520$ E4 $ 521$ E3 $ 521$ E3 $-$ <	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$

^b Predicted by GA-PLS model.

Table 2. List of substituent electronic descriptors calculated in this work					
No.	Notation	Definition			
1	RMSC	Root mean square error of charges			
2	SSPC	Sum of positive charges			
3	SSNC	Sum of negative charges			
4	MPC	Most positive charge			
5	LNC	Least negative charge			
6	DM	Dipole moment			
7	НОМО	Energy of the highest occupied molecular orbital			
8	LUMO	Energy of the lowest unoccupied molecular orbital			
9	SOF	Softness			
10	HD	Hardness			
11	EPH	Electrophilicity			
12	EN	Electronegativity			
13	ΔΗΟΜΟ	Difference between HOMO level of substrate and enone			
14	ΔLUMO	Difference between LUMO level of substrate and enone			
15	Δ HOMO $-$ LUMO	Difference between HOMO level of substrate and LUMO level of enone			
16	Δ LUMO — HOMO	Difference between LUMO level of substrate and HOMO level of enone			

RESULTS AND DISCUSSION

Modeling of the silica gel catalyzed reaction

As it is shown in Table 1, there are available 15 Michael addition entries catalyzed by silica gel. The least RI value is related to entry #6 correspond to addition of S3 to E4 whereas the additions of S2 or S1 to E3 (reaction entry #3 and #9, respectively) represents the highest reactivity.

When stepwise variable selection-based MLR analysis was employed to obtain the structure-reactivity relationships, the following three-parametric equations was resulted:

$$\begin{aligned} \mathsf{RI} &= 0.77(\pm 0.06) - 0.50(\pm 0.06) \ \Delta \mathsf{LUMO} - 0.24(\pm 0.06) \\ \mathsf{DM}_\mathsf{E} &+ 0.16(\pm 0.07) \mathsf{DM}_\mathsf{S} \ \textit{N} = 15, \textit{R}^2 = 0.846, \mathsf{SE} = 0.26, \end{aligned} \tag{1} \\ Q^2_{\mathsf{LOO}} &= 0.795, Q^2_{\mathsf{LOO}} = 0.797, \textit{F} = 32.26 \end{aligned}$$

The values in the parenthesis represent the standard deviation of the coefficients. *N*, R^2 , SE, and *F* are number of components, correlation coefficient, standard error of regression and Fisher's *F*-ratio. The correlation coefficients of leave-one-out and leave three-out cross-validations are denoted by Q_{LOO}^2 and Q_{LTO}^2 , respectively. The R^2 value of 0.846 describes that the resultant equation can explain about 85% of variance in the RI data of the silica gel catalyzed Michael addition reaction whereas the high values of cross-validated correlation coefficients and also their closeness to each other explain the high predictivity and stability of the model.

Among the selected molecular structural parameters, the difference between LUMO levels of substrates and enones represented the highest impact on the RI. The relative importance of the variables was measured by the standardized regression coefficient; however, for the sake of simplicity they are not represented here. Instead, the parameters in this and the subsequent equations have been ranked based on their relative importance so that the most important parameter has been appeared first. The negative sign of Δ LUMO implies that for a silica gel catalyzed Michael addition reactions the small difference between LUMO energies of substrate and enone is favorable for obtaining higher reactivity. The presence of dipole moments of both substrate and enones (DM_S and DM_E , respectively) suggests the significance of dipolar coulombic interactions in the silica gel catalyzed reactions. Interestingly, the coefficients of the parameters have reverse sign so that less polar enones and high polar substrates lead to high reactivity index. The presence of coulombic interactions can be attributed to the ionic nature of the silica gel catalysis.

In the next step, GA-PLS modeling was used to derive the QSRR equation. The resulting model is summarized in Table 3. As seen, by using five electronic parameters as predictor variables, the resulted model exhibited more significant statistical parameters. The selected variables by GA-PLS (i.e., SSNC_s, DM_s, EN_s, RMSC_e, and Δ LUMO) are similar to those appeared in the MLR-based QSRR model (Eqn. 1). As it is observed from Table 3, two PLS latent variables have been used as optimum number of factors. This suggests that the RI of the studied silica gel catalyzed reaction can be affected by two main factors, and according to the selected molecular parameters these factors can be attributed to (1) orbital energy related parameters (described by Δ LUMO and EN_s) and (2) columbic interactions (described by $SSNC_s$, DM_s , and QRMS_E). It should be noted that similar results were concluded from the MLR equation, however by taking into account more variable by GA-PLS, the resulted QSRR model possessed better performances. To find the relative importance of included variables in GA-PLS model, they were subjected to the variable importance in projection (VIP)^[35] and the results are plotted in Fig. 2A. According to the VIP values $RMSC_E$ and $\Delta LUMO$ can be considered as highly influential parameters.

Modeling of the SDS catalyzed reaction

As it is shown in Table 2, there are available RI values of 28 reaction entries for the SDS catalyzed reaction, among which 8 reaction entries are in common with those of the silica gel catalyzed reaction. The RI values of the SDS catalyzed reactions are ranged between -1.46 and 1.28. A comparison between the

Tabl	Table 3. Statistical parameters for different PLS models								
No.	Data set	Selected variables	N ^a	LVs ^b	R ² ^c	$R_{\rm P}^{2{\rm d}}$	Q ^{2 e}	$Q_{\rm MC}^{2}$ f	
1 2 3	Silica gel catalyzed reaction SDS catalyzed reaction ZrOCl ₂ catalyzed reaction	SSNQ _E , DM _S , EN _S , QRMS _E , ΔLUMO SSNQ _E , LUMO _S , HD _E , EPH _S , LNC _S , HOMO _E QRMS _S , SNQ _E , SPQ _S , QRMS _E , ΔLUMO	15 28 15	2 4 4	0.880 0.823 0.934	0.760 0.834 0.853	0.832 0.731 0.845	0.530 0.302 0.498	
^b Nu ^c Squ ^d Squ ^e Cro	 ^a Number of molecules in the data set. ^b Number of PLS latent variables is represented by LV. ^c Square of correlation coefficient for calibration. ^d Square of correlation coefficient for prediction. ^e Cross-validated square of correlation coefficient. ^f Maximum of the cross-validated square of correlation coefficient for 20 times random input scrambling (chance correlation). 								

reactivity of the Michael addition reactions catalyzed by silica gel and SDS may be interesting. Among the eight RI values reported for both silica gel and SDS catalyzed reactions, SDS represented higher reactivity for five entries and only in one case silica gel showed higher efficiency. For two remaining reactions no significant difference is observed between two catalysts. Thus, SDS can be considered as more reactive comparing with silica gel. This difference between reactivity of the catalysts can be attributed to the difference in substrate–catalyst, enone–catalyst,

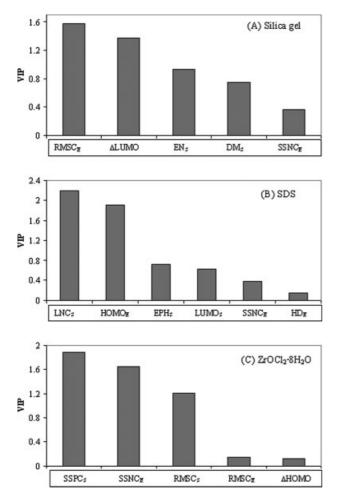


Figure 2. Plot of variable important in projection (VIP) for the descriptors selected by GA-PLS model for different catalyzed reactions

and substrate–enone interactions. QSRR will help to identify these differences.

By using stepwise-based MLR analysis, a two-parametric equation was obtained between structural parameters of enones and substrates in one hand the RI of the SDS catalyzed Michael addition reactions on the other hand:

 $\begin{aligned} \mathsf{RI} &= 0.41 (\pm 0.08) - 0.64 (\pm 0.08) \; \mathsf{HOMO_E} + 0.59 (\pm 0.08) \; \mathsf{LNC_S} \\ \textit{N} &= 28, \textit{R}^2 = 0.786, \mathsf{SE} = 0.42, \textit{Q}_{\mathsf{LOO}}^2 = 0.689, \textit{Q}_{\mathsf{LFO}}^2 = 0.670, \textit{F} = 10.8 \end{aligned} \label{eq:nonlinear}$

This two-parametric equation could explain and reproduce 78% and about 70% of variances in the reactivity index of the SDS catalyzed reactions, respectively. Although the statistical quality of this model is lower than that of previous section, it can clearly demonstrate the differences in the mechanism of the silica gel and SDS catalyzed reactions. In the QSRR equation found for silica gel catalyzed reaction the difference between LUMO energy of the enones and substrates were identified as important factors whereas in that of SDS catalyzed reactions the HOMO energy of enones has been appeared as influencing parameter. This explains that the studied catalysts can change the mechanism of electron transfer between substrates and enones. Besides, coulombic interaction, which has been identified as influential parameter in the QSRR model of the silica gel catalyzed reactions, has not been appeared in that of SDS catalyzed reactions. On the other hand, the most negative charge on the substrate atoms, which is already on the donor atom of the substrates, has been appeared in Eqn. (2). The positive sign of this parameter suggests that by increasing negative charge on the donor atom of substrate its reactivity will be increased.

GA-PLS modeling of the SDS catalyzed reactions resulted in a 6-parametric equation having better statistical quality with respect to the MLR-based QSRR model (see Table 3). As it is observed from Table 3, in addition to the two parameters selected by MLR method (i.e., LNC₅ and HOMO_E) four extra parameters (i.e., SSNC₅, LUMO₅, HD₅, EPH₅) has been selected by GA-PLS for modeling the reactivity index of the SDS catalyzed reactions. The plot of the calculated VIP values for these parameters (shown in Fig. 2) indicates the higher importance of MNC₅ and HOMO_E in modeling the reactivity index of the SDS catalyzed reactions. This confirms the significance of the variables selected by MLR analysis. By considering extra parameters in GA-PLS, a QSRR model with improved statistical quality was obtained.

Modeling of the ZrOCl₂·8H₂O catalyzed reaction

Table 2 shows that the ZrOCl₂·8H₂O catalyzed reactivity index is available for 15 reaction entries, among which 3 reaction entries are in common with those of the SDS catalyzed reaction and one reaction entry is in common between SDS and silica gel catalyzed reactions. Therefore, it is hard to compare the catalysts for their reactivity in the Michael addition reaction. Nevertheless, a deep look on the reported common RI data reveals that ZrOCl₂·8H₂O resulted in higher reactivity index (for instance see reaction entries numbers 8 and 33–35). QSRR analysis on the ZrOCl₂·8H₂O catalyzed reactions and comparing with those found for other catalysts can help us to investigate the differences in catalytic activity.

A three-parametric equation was obtained using MLR based on stepwise selection of molecular parameters:

$$\begin{split} &\mathsf{RI}=6.37(\pm0.71)-17.28(\pm1.73)\;\mathsf{MPC}_{\mathsf{S}}+2.09\;(\pm0.26)\\ &\mathsf{EPH}_{0}-3.43(\pm0.61)\;\mathsf{LNC}_{\mathsf{E}}\;\mathsf{N}=14, \mathsf{R}^{2}=0.855, \mathsf{SE}=0.11, \quad \textbf{(3)}\\ &Q_{\mathsf{LOO}}^{2}=0.823, Q_{\mathsf{LFO}}^{2}=0.819, \mathsf{F}=45.1 \end{split}$$

Among the studied reaction entries, the entry #44 represented significant deviation from the regression equation and therefore it was not included in the model. The resulted QSRR equation for the ZrOCl₂·8H₂O catalyzed reaction represent high statistical quality so that it can reproduce about 82% of variances in the reactivity indices. Interestingly, the molecular descriptors appeared in this equation are mainly related to charge parameters, indicating the significance role of coulombic interactions in the Michael addition reactions catalyzed by ZrOCl₂·8H₂O. In addition, molecular parameters of substrates (MPC and EPH) represented higher impacts on RI with respect to that of enones (LNC), which is in a reverse direction of what was found for two other catalysts, in which those of enones represented higher impacts on the RI. These observations suggest that for the SDS and silica gel catalyzed Michael addition reactions, the catalyst efficiency is mainly controlled by the molecular and electronic structure of the enones whereas in the case of ZrOCl₂·8H₂O catalyzed reactions those of substrates are more significant.

By considering more molecular descriptors in the QSRR model obtained by GA-PLS for the ZrOCl₂·8H₂O catalyzed reaction (Table 3) not only all reaction entries were lied in the regression model without significant outlier but also better predictive ability was obtained. The resulted GA-PLS model contains six descriptors, among which orbital energetic parameters are also appeared. This indicates that the HOMO and LUMO levels are significant in the ZrOCl₂·8H₂O catalyzed reaction but their importance is so low that could not be detected by MLR analysis. The VIP values of these descriptors (Fig. 2) again indicate that charge descriptors are more significant.

CONCLUSIONS

The QSRR analyses of the Michael addition of different substrates to some enones catalyzed by SDS, silica gel and ZrOCl₂·8H₂O revealed that the pronounced effect of catalyst on the reactivity can be primarily related to changes in the ways that substrates and enones are interact. The catalyst can interact with both enone and substrate so that its electronic and physicochemical properties change the substrate–enone interactions. In the case of silica gel catalyzed reactions, it was found that orbital energy parameters and permanent dipole–dipole coulombic interactions are the controlling factors whereas for the SDS catalyzed reactions the charge interactions were detected as significant parameters. On the other hand, for the ZrOCl₂·8H₂O catalyzed reactions, the charge interactions represented the main role and a diminished effect was found for orbital energy. In addition to the descriptive ability of the resulted QSRR model, they could also show high ability to predict the reactivity indices of addition of a given substrate to an enone, as measured by cross-validation method. This helps the organic chemist to predict (before synthesis) if a reaction under investigation is resulted in the satisfied yield.

Acknowledgements

We are thankful to Professor Habib Firouzabadi for his discussion, appointment, and encouragement.

REFERENCES

- [1] J. Christoffers, Eur. J. Org. Chem. 1998, 7, 1259.
- [2] J. Comelles, M. Moreno-Mañas, A. Vallribera, Arkivoc 2005, 9, 207.
- [3] G. Lu, Q. Zhang, Y.-J. Xu, Chin, J. Org. Chem. 2004, 24, 600.
- [4] N. Krause, A. Hoffmann-Roder, Synthesis 2001, 171.
- [5] G. C. Fu, J. Org. Chem. 2004, 69, 3245.
- [6] D. Yang, Y.-C. Yip, G.-S. Jiao, M.-K. Wong, J. Org. Chem. 1998, 63, 8952.
- [7] B. F. Yates, Ann. Rep. Prog. Chem. B 2005, 101, 210.
- [8] Y. Zho, D. G. Drueckhammer, J. Org. Chem. 2005, 70, 7755.
- [9] A. R. Katritzky, A. Lomaka, R. Petrukhin, R. Jain, M. Karelson, A. E. Visser, R. D. Rogers, J. Chem. Inf. Comput. Sci. 2002, 42, 71.
- [10] M. Staikova, P. Messih, Y. D. Lei, F. Wania, D. J. Donaldson, J. Chem. Eng. Data 2005, 50, 438.
- [11] A. Nigan, M. T. Klien, Ind. Eng. Chem. Res. 1993, 32, 1297.
- [12] C. F. Wilcox, B. K. Carpenter, J. Am. Chem. Soc. 1979, 101, 3897.
- [13] S. C. Korre, M. T. Klein, Catal. Today 1996, 31, 79.
- [14] E. Y. Tshuva, I. Goldberg, M. Kol, Z. Goldschmidt, Organometallics 2001, 20, 3017.
- [15] B. K. Carpenter, *Tetrahedron* **1978**, *34*, 1877.
- [16] C. F. Wilcox, B. K. Carpenter, W. R. Dolbier, Tetrahedron 1979, 35, 707.
- [17] H. Firouzabadi, N. Iranpoor, H. Hazarkhani, J. Org. Chem. 2001, 66, 7527.
- [18] H. Firouzabadi, N. Iranpoor, M. Jafarpour, A. Ghaderi, J. Mol. Catal. A: Chem. 2006, 253, 249.
- [19] H. Firouzabadi, N. Iranpoor, A. A. Jafari, S. Makarem, J. Mol. Catal. A: Chem. 2006, 250, 237.
- [20] H. Firouzabadi, N. Iranpoor, A. A. Jafari, Adv. Synth. Catal. 2005, 347, 655.
- [21] H. Firouzabadi, N. Iranpoor, M. Jafarpour, A. Ghaderi, J. Mol. Catal. A: Chem. 2006, 252, 150.
- [22] H. Firouzabadi, N. Iranpoor, M. Jafarpour, A. Ghaderi, J. Mol. Catal. A: Chem. 2006, 249, 98.
- [23] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, *et al. GAUSSIAN 98, Revision A.7*, Gaussian, Inc. Pittsburgh, PA, **1998**.
- [24] R. G. Parr, P. K. Chattaraj, J. Am. Chem. Soc. 1991, 113, 1854.
- [25] I. Pauling, *The Nature of Chemical Bond*, Cornell University Press, NY, **1960**.
- [26] R. G. Pearson, in *Theoretical Models of Chemical Bonding Part II*, (Ed.: Z. B. Maksic,). Springer, Berlin, **1990**.
- [27] B. Hemmateenejad, M. A. Safarpour, F. Taghavi, J. Mol. Struct. (Theochem) 2003, 635, 183.
- [28] P. Thanikaivelan, V. Subramanian, J. R. Rao, B. U. Nair, Chem. Phys. Lett. 2000, 32, 359.
- [29] B. Hemmateenejad, M. Sancholi, J. Chemometr. 2007, 21, 96.
- [30] A. Mohejeri, B. Hemmateenejad, A. R. Mehdipour, R. Miri, J. Mol. Graph. Model. 2008, 26, 1057–1065.
- [31] S. Wold, M. Sjostrom, L. Eriksson, Chemom. Int. Lab. Sys. 2001, 58, 109.
- [32] J. G. Topliss, R. J. Costello, J. Med. Chem. 1972, 15, 1066.
- [33] D. J. Livinstone, D. W. Salt, J. Med. Chem. 2005, 48, 661.
- [34] K. Baumann, QSAR. Comb. Sci. 2005, 24, 1033.
- [35] M. Olah, C. Bologa, T. I. Oprea, J. Comput. Aided. Mol. Des. 2004, 18, 43.