PREDICTION OF MONOMER REACTIVITY RATIOS IN RADICAL COPOLYMERIZATION OF VINYL MONOMERS

Xinliang YU[a1,*], Wenhao YU[b], Bing YI[a2] and Xueye WANG[c]

[a] *Department of Chemistry and Chemical Engineering, Hunan Institute of Engineering, Xiangtan, Hunan 411104, China; e-mail: [1] yxl@hnie.edu.cn, [2] bingyi2004@126.com.cn*

[b] *School of Resources and Environmental Science, Wuhan University, Wuhan, Hubei 430079, China; e-mail: xzm5602@sina.com.cn*

[c] *College of Chemistry, Xiangtan University, Xiangtan, Hunan 411105, China; e-mail: wxueye@xtu.edu.cn*

Received December 12, 2008
Accepted July 22, 2009
Published online September 3, 2009

Quantitative structure-property relationship (QSPR) models are developed to predict monomer reactivity ratios (log r_{12}) in radical copolymerization with monomers M_1 (styrene, methyl methacrylate and acrylonitrile) and M_2 (vinyl monomers). The quantum chemical descriptors are calculated by the density functional theory (DFT) at B3LYP level of theory with 6-31G(d) basis set. Stepwise multiple linear regression analysis and artificial neural network (ANN) were used to generate Model S (monomer 1: styrene), Model MM (monomer 1: methyl methacrylate) and Model A (monomer 1: acrylonitrile). Simulation results show that the predicted log r_{12} values are in good agreement with the experimental data, with the test sets possessing correlation coefficients of 0.972 for Model S, 0.933 for Model MM and 0.946 for Model A.

Keywords: Artificial neural network; Density functional theory; Radical copolymerization; Monomer reactivity ratios; QSPR; Quantum chemical.

When a vinyl monomer M_1 is copolymerized with a second monomer M_2, the relationship between the composition of the initially formed copolymer and the initial monomer mixture is given by[1]

$$R_p = R_m(r_{12}R_m + 1)/(r_{21} + R_m) \quad (1)$$

where R_m is equal to $[M_1]/[M_2]$ in the monomer mixture and R_p is equal to $[M_1]/[M_2]$ in the polymer formed, r_{12} and r_{21} are the monomer reactivity ratios (known as reactivity ratios).

The monomer reactivity ratios, r_{12} and r_{21}, for any monomer pair are the ratios of the rate constants of the different propagation reactions

doi:10.1135/cccc2008215

$$\sim M_1^{\bullet} + M_1 \rightarrow \sim M_2^{\bullet} \qquad k_{11}$$

$$\sim M_1^{\bullet} + M_2 \rightarrow \sim M_2^{\bullet} \qquad k_{12}$$

$$\sim M_2^{\bullet} + M_1 \rightarrow \sim M_1^{\bullet} \qquad k_{21}$$

$$\sim M_2^{\bullet} + M_2 \rightarrow \sim M_2^{\bullet} \qquad k_{22}$$

with $r_{12} = k_{11}/k_{12}$, $r_{21} = k_{22}/k_{21}$, where $\sim M^{\bullet}$ represents a polymer chain ending with a radical derived from monomer M.

One of the most important aspects of the study of copolymerization is the relationship between the composition of the monomer feed and that of the resulting copolymer. It would be extremely useful to predict the values of r_{12} and r_{21} and hence the composition of any copolymer produced from any pair of monomers at any concentration ratios[2]. The monomer reactivity ratios can be predicted using the empirical *Q*–*e* scheme[2–5] or the revised patterns scheme[2,4,5]. However, these schemes are limited as the parameter values (*Q*, *e*, *u*, *v*) are not known.

The development of reliable quantitative structure–property relationship (QSPR) models for prediction of the monomer reactivity ratios is of real interest, particularly for new monomers for which experimental investigation would be expensive. The QSPR approach can conserve resources and accelerate the process of development of new polymers[6].

In principle, quantum chemical theory can provide precise quantitative descriptors of molecular structures and their chemical properties[6]. The goal of this paper is to produce robust QSPR models which could predict the monomer reactivity ratios log r_{12} in radical copolymerization with monomers M_1 (styrene, methyl methacrylate and acrylonitrile) and M_2 ($C^1H_2{=}C^2H{\cdot}X$ and $C^1H_2{=}C^2X{\cdot}Y$). The quantum chemical descriptors calculated by the DFT method are correlated with the monomer reactivity ratios using the nonlinear artificial neural network (ANN) method.

MATERIALS AND METHODS

Table I shows three data sets for monomer reactivity ratios of radical copolymerization of vinyl monomers with structures $C^1H_2{=}C^2H{\cdot}X$ (or $C^1H_2{=}C^2X{\cdot}Y$)[7,8]. Monomer 1 is styrene, methyl methacrylate and acrylonitrile. Monomers 2 show a high degree of structural variety. For example, the functionalities present in the side chains include aldehydes, nitriles, ketones, halides, esters, aromatic rings, non-aromatic rings, etc. The loga-

rithms of monomer reactivity ratios r_{12} are used because the spread of the data sets is consistently even when log r_{12} is used instead of r_{12}. Moreover, the logarithmic form of monomer reactivity ratios provides a more convenient linear solution for the *Q–e* scheme[2–5] and the revised patterns scheme[2,4,5].

All experimental data of monomer reactivity ratios (see Table I) were randomly divided into a training set and a validation set. The training set was used to build the QSPR model which was evaluated with the validation set.

All the model structures studied in this work were fully optimized by the density functional theory (DFT) using Gaussian 03 program[9] at B3LYP level of theory with 6-31G(d) basis set[10,11], in order to obtain some quantum chemical descriptors to fit monomer reactivity ratios. Frequency calculations were carried out with the optimized geometries to assure that they were indeed stationary minima at the same level and basis set. Nineteen descriptors were calculated – the Mulliken and atomic polar tensor (APT)[12] charges of C^1, C^2 and C^3 (attached to C^2 directly) (q_{MC^1}, q_{MC^2}, q_{MC^3}, q_{AC^1}, q_{AC^2} and q_{AC^3}), mean positive APT atomic charge (q_M^+), total negative APT atomic charge (q_T^-), mean negative Mulliken atomic charge with hydrogens summed into heavy atoms (q_M^-), total dipole moment (μ), mean quadrupole moment (*Q*), mean octapole moment (ω), energies of the highest occupied molecular orbital (E_{HOMO}) and the lowest unoccupied molecular orbital (E_{LUMO}), LUMO and HOMO orbital energy difference ($\Delta E_g = E_{LUMO} - E_{HOMO}$)[13,14].

The mean quadrupole moment is calculated using Eq. (*2*)

$$Q = \frac{Q_{XX} + Q_{YY} + Q_{ZZ}}{3} \qquad (2)$$

where Q_{XX}, Q_{YY} and Q_{ZZ} are the quadrupole moment tensors in the X-, Y-, and Z-coordinates, respectively.

The mean octapole moment is defined as

$$\omega = \frac{|\omega_{XX}| + |\omega_{YY}| + |\omega_{ZZ}|}{3} . \qquad (3)$$

Similarly, ω_{XX}, ω_{YY} and ω_{ZZ} are the octapole moment tensors in the X-, Y- and Z-coordinates, respectively.

In addition, four descriptors (R_{TM}, R_{MA}, $R_{\omega C^2}$ and $R_{\mu H}$) are defined as follows.

TABLE I
Experimental and calculated monomer reactivity ratios (log r_{12})[a]

Monomer 1[a]	Monomer 2	log r_{12}(exp)	log r_{12}(calc)	Δlog r_{12}[b]
Model S (monomer1:styren)		Training set		
S	Acrylaldehyde	–0.658	–0.717	0.060
S	Methylacrylaldehyde	–0.699	–0.558	–0.141
S	*N*-(Hydroxymethyl)acrylamide	0.190	0.200	–0.010
S	Benzyl acrylate	–0.284	–0.440	0.156
S	Ethyl acrylate	–0.102	–0.218	0.116
S	Methyl acrylate	–0.125	–0.140	0.015
S	Acryloyl chloride	–1.000	–0.973	–0.027
S	Allyl acetate	1.875	1.832	0.043
S	Diallyl phthalate	1.362	1.322	0.040
S	Ethene	1.176	1.158	0.018
S	Isopropenyl methyl ketone	–0.357	–0.446	0.089
S	Benzyl methacrylate	–0.222	–0.301	0.079
S	2-Chloroethyl methacrylate	–0.456	–0.511	0.055
S	Glycidyl methacrylate	–0.347	–0.334	–0.013
S	Methyl methacrylate	–0.301	–0.297	–0.004
S	Phenyl methacrylate	–0.602	–0.533	–0.069
S	2-Vinylpyridine	–0.268	–0.278	0.011
S	4-Vinylpyridine	–0.284	–0.375	0.091
S	α-Methylstyrene	0.041	0.061	–0.020
S	4-Methylstyrene	–0.050	–0.056	0.006
S	Vinyl chloroacetate	1.477	1.508	–0.030
S	Vinyl chloride	1.301	1.233	0.068
Model S (monomer 1:styrene)		Test set		
S	Acrylamide	0.079	–0.017	0.096
S	Butyl acrylate	–0.102	–0.381	0.278
S	Acrylonitrile	–0.420	–0.518	0.098
S	Allyl chloride	1.558	1.781	–0.223
S	Isoprene	–0.337	–0.027	–0.310

TABLE I
(*Continued*)

Monomer 1[a]	Monomer 2	log r_{12}(exp)	log r_{12}(calc)	Δlog r_{12}[b]
S	Isobutyl methacrylate	–0.268	–0.332	0.065
S	2-Hydroxyethyl methacrylate	–0.276	–0.037	–0.239
S	Methacrylonitrile	–0.420	–0.325	–0.095
S	4-Acetoxy styrene	–0.051	–0.047	–0.003
S	Vinyl benzoate	1.519	1.480	0.038
S	Vinylidene chloride	0.255	0.414	–0.159
Model MM (monomer 1:methyl methacrylate)		Training set		
MM	Methylacrylaldehyde	–1.000	–0.5365	–0.464
MM	Benzyl acrylate	0.348	0.0303	0.318
MM	Methyl acrylate	0.332	0.3821	–0.050
MM	Allyl acetate	1.996	1.9219	0.074
MM	Acrylamide	0.447	0.1883	0.259
MM	Butyl acrylate	0.301	0.2159	0.085
MM	Acrylonitrile	0.121	0.1309	–0.010
MM	Allyl chloride	1.558	1.5251	0.033
MM	Acryloyl chloride	–0.347	–0.5263	0.179
MM	Benzyl methacrylate	–0.071	–0.1097	0.039
MM	Glycidyl methacrylate	–0.125	0.1879	–0.313
MM	Methacrylonitrile	–0.125	–0.2378	0.113
MM	Styrene	–0.337	–0.5798	0.243
MM	Diallyl phthalate	1.362	1.3417	0.020
MM	Isobutyl methacrylate	–0.036	–0.0506	0.015
MM	2-Vinylpyridine	–0.420	–0.4094	–0.011
MM	α-Methylstyrene	–0.284	–0.2139	–0.070
MM	4-Vinylpyridine	–0.260	–0.3302	0.070
MM	Vinylidene chloride	0.342	0.2879	0.054
MM	Vinyl chloride	0.954	0.9459	0.008
Model MM (monomer 1:methyl methacrylate)		Test set		
MM	Isoprene	–0.602	–0.0778	–0.524
MM	2-Chloroethyl methacrylate	0.063	–0.0321	0.095

TABLE I
(*Continued*)

Monomer 1[a]	Monomer 2	log r_{12}(exp)	log r_{12}(calc)	Δlog r_{12}[b]
MM	Phenyl methacrylate	–0.252	–0.4535	0.202
MM	4-Methylstyrene	–0.462	–0.301	–0.161
MM	Ethyl acrylate	0.301	0.3703	–0.069
MM	2-Hydroxyethyl methacrylate	–0.125	0.1504	–0.275
MM	Acrylaldehyde	0.079	0.2088	–0.130
MM	Vinyl benzoate	1.307	1.6657	–0.359
Model A (monomer 1:acrylonitrile)		Training set		
A	Acrylaldehyde	–0.222	–0.172	–0.050
A	Benzyl acrylate	0.176	–0.155	0.331
A	Butyl acrylate	0.079	0.070	0.009
A	Acryloyl chloride	0.079	0.023	0.056
A	Allyl acetate	0.818	0.644	0.174
A	Allyl chloride	0.431	0.630	–0.199
A	Ethene	0.845	0.760	0.085
A	2-Chloroethyl methacrylate	–0.854	–0.840	–0.014
A	Glycidyl methacrylate	–0.854	–0.801	–0.053
A	2-Hydroxyethyl methacrylate	–0.699	–0.693	–0.006
A	Methacrylonitrile	–0.367	–0.487	0.120
A	Isopropenyl methyl ketone	–0.420	–0.617	0.197
A	2-Vinylpyridine	–1.097	–1.257	0.160
A	4-Vinylpyridine	–1.000	–1.020	0.020
A	Vinylidene chloride	–0.222	–0.229	0.007
A	Isoprene	–1.523	–1.419	–0.104
A	Methylacrylaldehyde	–0.824	–0.515	–0.309
A	Acrylamide	–0.046	0.013	–0.059
A	Ethyl acrylate	–0.155	0.010	–0.165
A	Methyl acrylate	–0.071	–0.179	0.108

TABLE I
(*Continued*)

Monomer 1[a]	Monomer 2	log r_{12}(exp)	log r_{12}(calc)	Δlog r_{12}[b]
Model A (monomer 1:acrylonitrile)		Test set		
A	Benzyl methacrylate	–0.699	–0.742	0.043
A	Diallyl phthalate	0.544	0.767	–0.223
A	Methyl methacrylate	–0.854	–0.869	0.015
A	Phenyl methacrylate	–0.444	–0.725	0.281
A	4-Acetoxystyrene	–1.155	–1.476	0.321
A	4-Methylstyrene	–1.301	–1.507	0.206
A	Styrene	–1.398	–1.505	0.107
A	Vinyl chloride	0.477	0.014	0.463
A	α-Methylstyrene	–1.398	–1.481	0.083
A	*N*-(Hydroxymethyl)acrylamide	–0.097	0.418	–0.515
Model A (monomer 1:acrylonitrile)		Prediction set		
A	3-Bromostyrene	–	–1.503	–
A	3-Chlorostyrene	–	–1.174	–
A	3-Nitrostyrene	–	0.582	–
A	4-Bromostyrene	–	–1.504	–
A	4-Cyanostyrene	–	0.664	–
A	4-Methoxystyrene	–	–1.503	–
A	2,3,4,5,6-Pentachlorostyrene	–	–1.172	–
A	4-Fluoro-2-(trifluoromethyl)styrene	–	–1.490	–
A	2,6-Dichlorostyrene	–	–1.327	–
A	2-Bromo-4-(trifluoromethyl)styrene	–	–1.394	–
A	4-Isopropylstyrene	–	–1.517	–

[a] S, styrene; MM, methyl methacrylate; A, acrylonitrile. [b] $\Delta\log r_{12} = \log r_{12}(\text{exp}) - \log r_{12}(\text{calc})$.

$$R_{\mathrm{TM}} = \frac{q^{-}}{q_{\mathrm{MC}^3}} \tag{4}$$

$$R_{\mathrm{MA}} = \frac{q_{\mathrm{MC}^3}}{q_{\mathrm{M}}^{+}} \tag{5}$$

$$R_{\omega\mathrm{C}^2} = \frac{\omega}{q_{\mathrm{MC}^2}} \tag{6}$$

$$R_{\mu\mathrm{H}} = \frac{\mu}{E_{\mathrm{HOMO}}} \tag{7}$$

Stepwise multiple linear regression (MLR) analysis[15] has proved to be an extremely useful computational technique in seeking an optimum linear combination of variables from the subsets of the N variables. The technique only adds one parameter to a model at a time and always in the order from the most significant to the least significant. Thus, stepwise MLR method was used to select their respective best subsets of descriptors for Model S (monomer 1: styrene), Model MM (monomer 1: methyl methacrylate) and Model A (monomer 1: acrylonitrile) from the 19 descriptors mentioned above. The subset size was increased until addition of another descriptor did not significantly improve the root-mean-square (rms) error of the estimate. Furthermore, a variance inflation factor (VIF) was calculated to see if multicollinearities exist between the descriptors in a model. Models were not accepted if they contained descriptors with VIFs over a value of 10. The smallest subset of descriptors that did not compromise the rms error of the estimate and VIF-test was identified as optimal.

All experimental data in Table I were divided into a training set and a test set, respectively. In order to develop the respective artificial neural network (ANN) models from the training sets, three-layer, fully connected, back-propagation (BP) neural networks were adopted[16]. For each model, the number of neurons of the input layer was equal to the molecular descriptors taken from stepwise MLR analysis. The output layer contained one neuron representing the reactivity ratios (log r_{12}). The network geometry (i.e. the number of hidden layers and the number of nodes in per hidden layer) was optimized by the trial-and-error method with permission error 0.0001, momentum parameter 0.6 and sigmoid parameter 0.9. The optimal

number of hidden layers and the number of neurons in each hidden layer were determined by varying the number of hidden layers (from 1 to 2) and the number of hidden neurons (from 1 to 4) and observing rms errors. The sums of rms errors of the training set and the test set were used to assess the accuracy of a model. The rms error was defined as

$$\text{rms} = \sqrt{\frac{\sum (f_i - x_i)^2}{N}} \qquad (8)$$

where f_i is the predicted value for the i-th compound, x_i is the observed value for the i-th compound, and N is the total number of compounds in the data set. The number of hidden layers and the number of nodes in per hidden layer were increased successfully until improvement was observed for that model.

The ANN architecture is described with the code: N_{in}–[N_{h1}–N_{h2}]$_e$–N_{out}, where N_{in} and N_{out} are the element numbers of input and output nodes, respectively; N_{h1} and N_{h2} are numbers of nodes in the first and second hidden, respectively; e is the number of hidden layers.

RESULTS AND DISCUSSION

Correlating 19 descriptors and the three sets of monomer reactivity ratios using stepwise multiple linear regression (MLR) analysis[15], three MLR models were obtained – Model S (monomer 1: styrene), Model MM (monomer 1: methyl methacrylate) and Model A (monomer 1: acrylonitrile).

$$\log r_{12} = 1.064 + 0.011Q + 20.276E_{\text{LUMO}} - 1.743q_{\text{MC}^3} - 0.537q_{\text{T}}^- \qquad (9)$$

$n = 33$, $R = 0.952$, se = 0.253, $F = 67.073$ (MLR Model S)

$$\log r_{12} = 0.761 + 13.776E_{\text{LUMO}} - 1.696q_{\text{M}}^- + 0.020R_{\text{TM}} - 0.201R_{\text{MA}} \qquad (10)$$

$n = 28$, $R = 0.936$, se = 0.264, $F = 40.725$ (MLR Model MM)

$$\log r_{12} = -5.668 - 0.028\omega - 0.003R_{\omega\text{C}^2} - 0.077R_{\mu\text{H}} + 21.585\Delta E \qquad (11)$$

$n = 30$, $R = 0.925$, se = 0.291, $F = 32.270$ (MLR Model A)

where R is correlation coefficient, se is the standard error, F is the Fischer ratio, and n is the number of samples. The characteristics of descriptors

appearing in each model are shown in Table II and the definitions of descriptors are described in Table III.

All the subsets of descriptors selected by stepwise MLR method were then fed to artificial neural network (ANN) as input vectors[15]. The optimal conditions of neural networks were obtained by adjusting parameters by the trial-and-error method. The architectures of the final optimum ANNs are 4–$[4]_1$–1 for Model S (monomer 1: styrene), 4–$[4–2]_2$–1 for Model MM (monomer 1: methyl methacrylate) and 4–$[4]_1$–1 for Model A (monomer 1: acrylonitrile).

Statistical parameters for ANN Model S are
n = 22, R = 0.997, rms = 0.068 (training set),
n = 11, R = 0.972, rms = 0.176 (test set).

Statistical parameters for ANN Model MM are
n = 20, R = 0.971, rms = 0.175 (training set),
n = 8, R = 0.933, rms = 0.268 (test set).

TABLE II
The characteristics of descriptors appearing in MLR models[a]

Model[a]	Descriptor	Coefficient	Std. error	Sig.-test	*t*-test	VIF
Model S	constant	1.064	0.141	0.000	7.541	–
	Q	0.011	0.004	0.007	2.927	2.408
	E_{LUMO}	20.276	2.410	0.000	8.413	1.209
	q_{MC^3}	–1.743	0.144	0.000	–12.134	1.090
	q_T^-	–0.537	0.078	0.000	–6.859	2.630
Model MM	constant	0.761	0.133	0.000	5.829	–
	E_{LUMO}	13.7761	3.063	0.000	4.719	1.302
	q_M^-	–1.718	0.432	0.001	–4.125	1.282
	R_{TM}	0.020	0.005	0.000	4.263	1.153
	R_{MA}	–0.201	0.027	0.000	–7.627	1.171
Model A	constant	–5.668	0.513	0.000	–11.042	–
	ω	–0.028	0.010	0.011	–2.763	1.475
	$R_{\omega C^2}$	–0.003	0.001	0.000	–6.124	1.201
	$R_{\mu H}$	–0.077	0.015	0.000	–5.249	1.353
	ΔE_g	21.585	2.166	0.000	9.966	1.099

[a] S, styrene; MM, methyl methacrylate; A, acrylonitrile.

Statistical parameters for ANN Model A are
n = 20, R = 0.972, rms = 0.146 (training set),
n = 10, R = 0.946, rms = 0.278 (test set).

The results calculated by each ANN method are listed in Table I and depicted in Fig. 1, respectively. They indicate that all the experimental monomer reactivity ratios values are close to the predicted values. Thus the three ANN models are useful for predicting monomer reactivity ratios in binary copolymerization with monomers M_1 (styrene, methyl methacrylate and acrylonitrile). Furthermore, the ANN Model A is used to predict the reactivity ratios for 11 monomer pairs (monomer 1: acrylonitrile, monomer 2: 11 styrenes) whose experimental values are not available. The prediction values are also listed in Table I.

Table II shows all the Sig.-test values are lower than 0.05, demonstrating that these descriptors all are significant. In addition, all the VIF (variance inflation factor) values are lower than 10, which show that multicollinearities do not exist in the descriptors in each model. The *t*-test measures the statistical significance of the regression coefficients. The higher *t*-test values correspond to the relatively more significant regression coefficients.

TABLE III
The definitions of descriptors

Number	Descriptor	Definition	Unit
1	Q	mean molecular quadrupole moment	D Å
2	E_{LUMO}	energy of the lowest unoccupied molecular orbital	a.u.
3	q_{MC^3}	Mulliken charge on C^3	a.u.
4	q_T^-	total negative APT atomic charge	a.u.
5	q_M^-	mean negative Mulliken atomic charge with hydrogens summed into heavy atoms	a.u.
6	R_{TM}	ratio q_T^-/q_{MC^3}	–
7	R_{MA}	ratio q_{MC^3}/q_M^+ (means positive APT atomic charge in a molecule	–
8	ω	mean molecular octapole moment	D $Å^2$
9	$R_{\omega C^2}$	ratio ω/q_{MC^2}(Mulliken charge on C^2)	D $Å^2$ $a.u.^{-1}$
10	$R_{\mu H}$	ratio μ (the total dipole moment)/E_{HOMO} (the energy of the highest occupied molecular orbital)	D $a.u.^{-1}$
11	ΔE_g	LUMO and HOMO orbital energy gap	a.u.

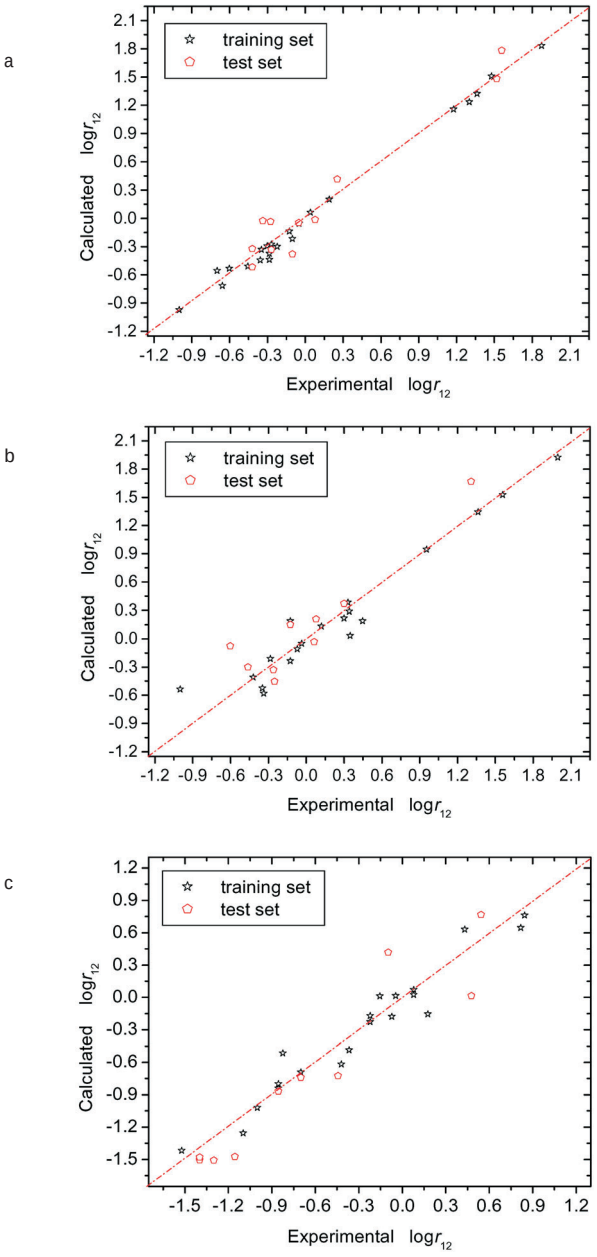

FIG. 1

Plots of the experimental versus calculated monomer reactivity ratios (log r_{12}); monomer 1: styrene (a), methyl methacrylate (b) and acrylonitrile (c)

According to the *Q–e* scheme[2–5], the rate constant for a radical monomer reaction, for example, for the reaction of a radical M_1 with a monomer M_2, can be written as

$$k_{12} = P_1 Q_2 \exp(-e_1 e_2) \quad (12)$$

and

$$\log k_{12} = \log P_1 - e_1 e_2 + \log Q_2 \quad (13)$$

where P_1 and Q_2 are measures of the resonance stabilization of M_1 and M_2, respectively; e_1 and e_2 are measures of their respective polarities.

By subtracting the corresponding expression for k_{11} from both sides of the equations, the parallel equations for the monomer reactivity ratio r_{12} are obtained

$$r_{12} = \frac{P_1}{Q_2} \exp[-e_1(e_1 - e_2)] . \quad (14)$$

Thus

$$\log r_{12} = \log P_1 - \log Q_2 + e_1 e_2 - e_1^2 . \quad (15)$$

In this paper, monomer 1 is styrene, methyl methacrylate and acrylonitrile, i.e. P_1 and e_1 are constant. Hence, the monomer reactivity ratios r_{12} depend only on the parameters Q_2 and e_2. The parameter log r_{12} increases with increasing polarity or/and decreasing general reactivity of monomer 2.

The negative Q can reflect the degree of the asymmetry of spherical charge distribution in the molecule. The decrease in Q value indicates the increase in polarity of the molecule. Thus, the descriptor Q carries a negative coefficient in the Model S (see Table II). The correlation of the molecular ω with Q is negative. It is easy to understand why the descriptor ω bears a positive coefficient in Model MM.

The total dipole moment is a measure of the asymmetry of molecular charge distribution. It is also the most widely used quantity for description of the polarity of a molecule[17,18]. Therefore, the ratio of μ and E_{HOMO}, $R_{\mu H}$, correlates with the parameter log r_{12}.

According to the frontier molecular orbital theory (FMO) of chemical reactivity, the highest occupied molecular orbital (HOMO), the lowest unoccupied molecular orbital (LUMO) and the energy gap (ΔE_g) between LUMO and HOMO play major roles in controlling many chemical reactions[13,14]. The energy of the LUMO (E_{LUMO}) is directly related to the electron affinity.

It characterizes the susceptibility of the molecule toward attack by nucleophiles. A molecule with a smaller E_{LUMO} value may be prone to formation of a radical and resulting in a larger value of parameter Q_2. Thus, there are positive terms associated with descriptors E_{LUMO} and ΔE_g in Model S, Model MM and Model A (see Table II).

In Table II, six descriptors (q_{MC^3}, q_T^-, q_M^-, R_{TM}, R_{MA} and $R_{\omega C^2}$) correlate with local electron densities or charges. By the classical chemical theory, all chemical interactions are either electrostatic (polar) or orbital (covalent). Electric charges in the molecule are obviously the driving force of electrostatic interactions. Indeed, it has been proven that local electron densities or charges are important in many chemical reactions and for physicochemical properties of compounds, and can reflect the polarity of a molecule[6]. The six descriptors therefore correlate also with the parameter log r_{12}.

Despite many different factors affecting the monomer reactivity ratios (log r_{12}), the molecular structures of monomers are the main ones. These descriptors based on the structural analysis of monomers express the important factors relating to monomer reactivity ratios. In addition, ANN is a powerful chemometrics tool to make quantitative prediction. Therefore, the ANN models based on the reasonable descriptors can accurately predict the reactivity ratios.

According to transition state theory, the reaction rate constant k for a given temperature T can be evaluated as

$$k = (\kappa k_B T/h) \exp(-\Delta^{\ddagger}G^{o}_{a}/RT) \qquad (16)$$

where $\Delta^{\ddagger}G^{o}_{a}$ is the Gibbs free energy of activation, k_B is the Boltzmann constant, h is the Planck constant, κ is the tunneling correction, and R is the gas constant. By using Eq. (*16*) for the rate constants k_{11} and k_{12}, a more detailed expression of the reactivity ratio log r_{12} is obtained (we assume that the κ and T values are the same).

$$\log r_{12} = \Delta^{\ddagger}G^{o}_{12} - \Delta^{\ddagger}G^{o}_{11} \qquad (17)$$

To use Eq. (*17*) to calculate the log r_{12} value of a monomer pair, one needs to determine the Gibbs free energy of activation for the reaction of the monomer with a radical. However, the reliable computational determination of a Gibbs free energy of activation is significantly more expensive than that of a reaction energy[19,20]. Fortunately, the computation can be simplified because there are excellent linear relationships between the activation energy and reaction energy[20–22]. Hence, if we consider monomer

pairs (monomer 1: styrene, methyl methacrylate and acrylonitrile), we may obtain an alternative expression of the reactivity ratio log r_{12}

$$\log r_{12} = A\Delta G_{a12} + C \qquad (18)$$

where ΔG_{a12} is the Gibbs free energy of reaction, and A and C are empirical constants to be determined by comparison to experimentally derived log r_{12} values.

Here we take the 20 monomer pairs in the training set of Model A (monomer 1: acrylonitrile) (see Table I) as calculation example. According to the terminal model, the Gibbs free energy of reaction ΔG_{a12} can be calculated with the propagation reaction

$$CH_3CH(CN)^{\bullet} + M_2 \rightarrow CH_3CH(CN)M_2^{\bullet}$$

where M_2 represents 20 different monomers (i.e. monomer 2) (see Table I). All optimizations and frequency calculations were carried out at the B3LYP level of theory with 6-31G(d) basis set. In the end, we obtained a relationship between the reactivity ratio log r_{12} and reaction energy ΔG_{a12} (Hartrees)

$$\log r_{12} = -0.186\Delta G_{a12} + 48.250 \qquad (19)$$

with a correlation coefficient $R = 0.639$, which is lower than that of Eq. (*11*). Equation (*19*) is only approximate. The reason may be that the reactivity of a propagating radical is determined not only by the terminal unit but also by the penultimate unit. Hence, it is necessary to develop these ANN models to predict the reactivity ratio log r_{12}.

CONCLUSIONS

The monomer reactivity ratios (log r_{12}) values of vinyl monomers in radical copolymerizations were predicted by using the QSPR models constructed by back-propagation (BP) neural networks. In the monomer pairs, monomer 1 is styrene, methyl methacrylate and acrylonitrile. The followed results indicate that the atom charges, multipole moments and frontier molecular orbital energies are the most important factors in correlation with monomer reactivity ratios. The produced artificial neural network (ANN) models are proved to be accurate. For test sets, rms errors are 0.176 for Model S (monomer 1: styrene), 0.268 for Model MM (monomer 1: methyl methacrylate) and 0.278 for Model A (monomer 1: acrylonitrile). The results encourage further applications of the proposed AAN models to other monomer pairs.

The project was supported by the Scientific Research Fund of Hunan Provincial Education Department (No. 07C205), the Scientific Research Fund of Hunan Institute of Engineering (No. 0761) and the Natural Science Foundation of Hunan Province (No. 06JJ50017).

REFERENCES

1. Mayo F. R., Lewis F. M.: *J. Am. Chem. Soc.* **1944**, *66*, 1594.
2. Jenkins A. D., Jenkins J.: *Macromol. Symp.* **2001**, *174*, 187.
3. Alfrey T., Price C. C.: *J. Polym. Sci.* **1947**, *2*, 101.
4. Jenkins A. D.: *J. Polym. Sci., Part A: Polym. Chem.* **1999**, *37*, 113.
5. Jenkins A. D., Jenkins J.: *Macromol. Symp.* **1996**, *111*, 159.
6. Karelson M., Lobanov V. S., Katritzky A. R.: *Chem. Rev.* **1996**, *96*, 1027.
7. Jenkins A. D., Hatada K., Kitayama T., Nishiura T.: *J. Polym. Sci., Part A: Polym. Chem.* **2000**, *38*, 4336.
8. Brandrup J., Immergut E. H., Grulke E. A. (Eds): *Polymer Handbook*, 4th ed. Wiley, New York 1999.
9. Frisch M. J., Trucks G. W., Schlegel H. B., Scuseria G. E., Robb M. A., Cheeseman J. R., Montgomery J. A., Jr., Vreven T., Kudin K. N., Burant J. C., Millam J. M., Iyengar S. S., Tomasi J., Barone V., Mennucci B., Cossi M., Scalmani G., Rega N., Petersson G. A., Nakatsuji H., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Klene M., Li X., Knox J. E., Hratchian H. P., Cross J. B., Bakken V., Adamo C., Jaramillo J., Gomperts R., Stratmann R. E., Yazyev O., Austin A. J., Cammi R., Pomelli C., Ochterski J. W., Ayala P. Y., Morokuma K., Voth G. A., Salvador P., Dannenberg J. J., Zakrzewski V. G., Dapprich S., Daniels A. D., Strain M. C., Farkas O., Malick D. K., Rabuck A. D., Raghavachari K., Foresman J. B., Ortiz J. V., Cui Q., Baboul A. G., Clifford S., Cioslowski J., Stefanov B. B., Liu G., Liashenko A., Piskorz P., Komaromi I., Martin R. L., Fox D. J., Keith T., Al-Laham M. A., Peng C. Y., Nanayakkara A., Challacombe M., Gill P. M. W., Johnson B., Chen W., Wong M. W., Gonzalez C., Pople J. A.: *Gaussian 03*, Revision B. 05. Gaussian Inc., Pittsburgh (PA) 2003.
10. Becke A. D.: *J. Chem. Phys.* **1993**, *98*, 5648.
11. Francl M. M., Pietro W. J., Hehre W. J., Binkley J. S., Gordon M. S., DeFrees D. J., Pople J. A.: *J. Chem. Phys.* **1982**, *77*, 3654.
12. Cioslowski J.: *J. Am. Chem. Soc.* **1989**, *111*, 8333.
13. Fukui K.: *Theory of Orientation and Stereoselection*, p. 34. Springer, New York 1975.
14. Franke R.: *Theoretical Drug Design Methods*, p. 115. Elsevier, Amsterdam 1984.
15. Tang Q. Y., Feng M. G.: *Practical Statistics and DPS Data Processing System*. Science Press, Beijing 2002.
16. Liu W. Q., Yi P. G., Tang Z. L.: *QSAR Combinat. Sci.* **2006**, *25*, 936.
17. Kikuchi O.: *Quant. Struct.–Act. Relat.* **1987**, *6*, 179.
18. Bodor N., Gabanyi Z., Wong C. K.: *J. Am. Chem. Soc.* **1989**, *111*, 3783.
19. Zhan C. G., Landry D. W., Ornstein R. L.: *J. Am. Chem. Soc.* **2000**, *122*, 1522.
20. Zhan C. G., Dixon D. A.: *J. Phys. Chem. A* **2002**, *106*, 10311.
21. Fueno T., Kamachi M.: *Macromolecules* **1988**, *21*, 908.
22. Rogers S. C., Mackrodt W. C., Davis T. P.: *Polymer* **1994**, *35*, 1258.