

Jinwei Gao · Xueye Wang · Xiaobing Li ·  
Xinliang Yu · Hanlu Wang

## Prediction of polyamide properties using quantum-chemical methods and BP artificial neural networks

Received: 23 June 2005 / Accepted: 15 November 2005 / Published online: 11 January 2006  
© Springer-Verlag 2006

**Abstract** Quantitative structure -property relationships (QSPR) for glass transition temperatures ( $T_g$ ), density ( $\rho$ ) and indices of refraction ( $n$ ) of the polyamides have been determined. All descriptors are calculated from molecular structures at the B3LYP/6-31G(d) level. These QSPR models are generated by two methods: multiple linear regression (MLR) and error back-propagation artificial neural networks (BPANN). The model obtained by MLR is used for the calculations of  $T_g$  ( $R_{\text{training}}=0.9074$ ,  $SD_{\text{training}}=22.4687$ ,  $R_{\text{test}}=0.8898$ ,  $SD_{\text{test}}=23.2417$ ),  $\rho$  ( $R_{\text{training}}=0.9474$ ,  $SD_{\text{training}}=0.0422$ ,  $R_{\text{test}}=0.8928$ ,  $SD_{\text{test}}=0.0422$ ),  $n$  ( $R_{\text{training}}=0.9298$ ,  $SD_{\text{training}}=0.0204$ ,  $R_{\text{test}}=0.9095$ ,  $SD_{\text{test}}=0.0274$ ). The model obtained by BPANN is used for the calculations of  $T_g$  ( $R_{\text{training}}=0.9273$ ,  $SD_{\text{training}}=14.8988$ ,  $R_{\text{test}}=0.8989$ ,  $SD_{\text{test}}=16.4396$ ),  $\rho$  ( $R_{\text{training}}=0.9523$ ,  $SD_{\text{training}}=0.0466$ ,  $R_{\text{test}}=0.9014$ ,  $SD_{\text{test}}=0.0512$ ),  $n$  ( $R_{\text{training}}=0.9401$ ,  $SD_{\text{training}}=0.0131$ ,  $R_{\text{test}}=0.9445$ ,  $SD_{\text{test}}=0.0179$ ). These results demonstrate that the MLR and BPANN methods can be used to predict  $T_g$ ,  $\rho$  and  $n$ . The more accurate predicted results are obtained from BPANN.

**Keywords** Polyamides · QSPR · DFT · BP Artificial neural networks

### Introduction

It is a fundamental tenet of chemistry that the structural formula of any compound contains coded within it all that compound's chemical, physical, and biological properties. Therefore, the development of alternative approaches to finding quantitative mathematical relationships between the intrinsic molecular structure and observable properties of chemical compounds will be of increasing importance in the chemistry of the 21st century. Polymers are widely-used

and important materials. The key of their molecular design is the relationship between polymer composition, structure and performance. Research into the QSPR of polymers contributes to understanding the factors that affect the properties of polymers. Knowing and grasping the regularities between them can predict the properties of polymers and design polymers with good performance. Moreover, it is of reference value to the basic theories of chemistry and physics of polymers. Therefore, such research have attracted much attention.

The glass transition is the most important transition and relaxation that occurs in amorphous polymers, and it has a significant effect on the properties and processing characteristics of this type of polymer [1].  $T_g$  is difficult to determine, because the transition happens over a comparatively wide temperature range and depends on the method, the duration and the pressure of the measuring device [2, 3]. Moreover,  $T_g$  is quite dependent on the structural, constitutional and conformational features of the polymer [4].

Numerous models for predicting the  $T_g$  of amorphous polymeric materials have been reported. Many empirical/semiempirical [5–9] and group-contribution [10, 11] methods have been used to build models for the prediction of  $T_g$ . Waegell and coworkers approach the EVM QSPR model using Energy, Volume and Mass [12, 13] for linear and branched aliphatic acrylate and methacrylate polymers. Van Krevelen uses a group additive property (GAP) theory to predict the properties of polymer. In GAP theory, the property under consideration is assumed to be the scalar sum of the corresponding properties of component chemical groups [5]. Katritzky and co-workers create a generally applicable QSPR development program, called Comprehensive Descriptors for Structural and Statistical Analysis (CODESSA). This approach has been shown in numerous studies to result in superior correlations between a wide variety of molecules and several physical and biological properties. Katritzky used the CODESSA method to predict  $T_g$  for 88 linear homopolymers using a five-descriptor equation that gave a coefficient of determination of 0.946 [14], to develop a correlation of  $R=0.970$  for a set of the refractive indices of 95 linear polymers with five de-

J. Gao · X. Wang (✉) · X. Li · X. Yu · H. Wang  
College of Chemistry, Xiangtan University,  
Xiangtan, 411105, People's Republic of China  
e-mail: wxueye@xtu.edu.cn  
Tel.: +86-0732-8292206

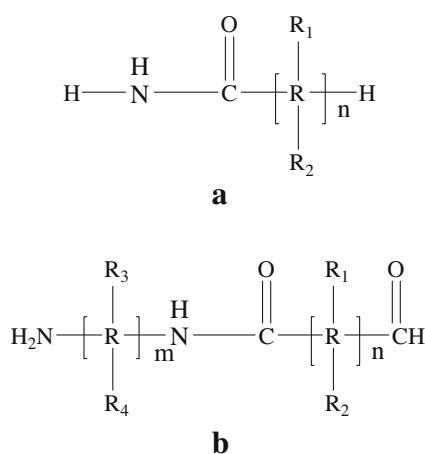
criptors [15]. Cao and Lin tested the same set of 88 linear homopolymers against a correlation involving their own chosen descriptors in an attempt to derive a more physically meaningful QSPR [16] and these descriptors are found to give a QSPR with a coefficient of determination of  $R^2=0.9056$  for  $T_g$ .

The index of refraction ( $n$ ) is a measure of the bending of a ray of light when it passes from one medium into another. The prediction of this property for polymers is valuable due to its application to the design of new optical materials. For example, the total reflection of light at the boundary between two media with different optical properties is of utmost importance in the manufacture of waveguides, films, and optical fibers. In addition,  $n$  is related to dielectric functions and other electromagnetic properties [17], such as the energy gap in semiconductors [18]. The prediction of refractive indices for polymers has been carried out using several approaches [16, 19]. Meknyan uses a modified Wiener path number, which provides finite values in the case of infinite polymer chains within homologous series (topological extrapolation method). Jenecke and coworkers developed successful group-contribution methods on the basis of the additivity principle [20].

In this paper, quantum-chemical and structural descriptors are used to correlate with three properties of  $T_g$ ,  $\rho$  and  $n$ . Three models for  $T_g$ ,  $\rho$  and  $n$  are obtained by multiple linear regression (MLR) and error back-propagation artificial neural networks (BPANN).

## Procedure

The experimental data are quoted from [21] and the web [22]. Molecular structures for all polyamides used in this paper are showed in Fig. 1. A total of 87 polyamides with extensive structural diversity was selected as a working dataset (Table 1), in which the number for  $T_g$ ,  $\rho$  and  $n$  are 78, 87 and 67, respectively. The experimental data were



**Fig. 1** Model of repeating units for polyamide ( $R$  in (a) and (b) represent the groups of  $-\text{CH}_2-$  and  $-\text{Ph}-$ ,  $R_1, R_2, R_3$  and  $R_4$  in (a) and (b) represent common groups,  $n$  and  $m$  sign the repeating number of  $R$ )

randomly divided into two groups, one is a training set and the other is a test set. Both of the data sets cover a relatively wide range and represent an extensive range of chemical groups.

It is impossible to calculate descriptors directly for an entire molecule with high molecular weight. In this paper, two kinds of models consisting of repeating units end-capped by hydrogen were chosen as study models to calculate the descriptors (Fig. 1). The single point energy and thermochemistry data were calculated by density functional theory (B3LYP/6-31G(d)) using Gaussian 03 [23].

The correlation analysis to obtain the best QSPR model was carried out by two methods. MLR generates an equation of the form  $Y = a_1 + a_2X_2 + a_3X_3 + \dots + a_nX_n$ , where  $X_2, \dots, X_n$  are the descriptors,  $a_1, a_2, \dots, a_n$  are the regression coefficient parameters. The quality of the regression model is measured using primarily four statistical parameters: correlation coefficient ( $R$ ), standard error of estimation (SD),  $T$ -value ( $T$ ) and significance level value ( $P$ ). This process was performed using the analysis software DPS 2.0 [24]. Artificial neural networks are a type of information-processing system based on modeling the neural system structures of the human brain. They have some remarkable properties such as self-learning and adaptation, a resistance to noise, a high degree of fault tolerance, which make them suitable for nonlinear problems with complex factors. They are powerful for exploiting information from a vast amount of experimental data through learning, and are especially useful for quantitative prediction. The networks consist in general of an input layer, an output layer, and any number of intermediate layers, called hidden layers. Each unit in the network is influenced by those units to which it is connected, the degree of influence being dictated by the values of the links or connections. The overall behavior of the system can be modified by adjusting the values of the connections, or weights, through the repeated application of a learning algorithm. One of the most popular algorithms is the back propagation (BP) algorithm. The number of neurons of the input layer is equal to that of the molecular descriptors taken from the best MLR analysis. The output layer contained one neuron representing the calculated value. The learning rate was 0.0001 and momentum 0.01. The number of neurons in the hidden layer was optimized by trial and error assays on calculations of the training process. The trained network was used to predict the values of the test set. The program BPANN was written in our lab.

## Results and discussion

Two structural descriptors (PMA and M) and six quantum-chemical descriptors ( $E_{\text{HOMO}}$ ,  $E_{\text{LUMO}}$ ,  $\mu$ ,  $Q_{\text{O}-}$ ,  $E_t$  and  $U$ ) were selected. The detailed descriptors are described in Table 2. For  $T_g$ , the training set and the test set were 63 and 15 polyamides. For  $\rho$ , the training set and the test set were 70 and 17 polyamides. For  $n$ , the training set and the test set were 53 and 14 polyamides, respectively. These experimental values and calculated values are listed in Table 1.

**Table 1** The experimental values and calculated results with two methods for three properties of polyamides

Polymer name	$T_g$ (K)		$\rho$ (g/cm <sup>3</sup> )		$n$	
	Exp.	MLR/ANN	Exp.	MLR/ANN	Exp.	MLR/ANN
Poly(glycine)	338	328/343	1.32	1.36/1.28	—	—/—
Poly(hydrazine-1,2 - diylethanediyl)	349	375*/347*	2.11	2.09/2.04	1.73	1.73/1.73
Poly(propano-3 - lactam)	332	322/340	1.33	1.31*/1.23*	1.70	1.67/1.68
Poly(iminomethyleneiminomalonyl )	338	300/315	1.28	1.39/1.33	1.67	1.66/1.66
Poly(2-aminoisobutyric acid)	276	301/304	1.26	1.35*/1.27*	—	—/—
Poly(3-amino-4-methylbutyric acid)	280	288/288	1.11	1.25/1.18	—	—/—
Poly(4-aminobutyric acid)	338	312*/329*	1.25	1.28/1.21	—	—/—
Poly(5-aminovaleric acid)	325	303/321	1.20	1.25/1.18	1.61	1.63/1.63
Poly(glutamic acid)	—	—/—	1.31	1.33/1.26	—	—/—
Poly[3-(glycylamino)propionic acid]	335	310/329	1.31	1.32/1.26	1.74	1.71/1.70
Poly(4-isopropyl -2,5- dioxo -1,3 - oxazolidin	259	273/277	1.18	1.25/1.18	—	—/—
Poly(hexano- 6-lactam)	323	318/318	1.14	1.23/1.15	1.59	1.62/1.62
Poly[3- (alanylalmino)propionic acid]	328	314/330	1.28	1.30/1.25	1.68	1.66/1.67
Poly(2- ethyl-2- methyl-3- aminopropionic acid)	258	282/271	1.13	1.23*/1.16*	—	—/—
Poly(imino -3- isopropyl -1- oxopropane -1,3- diyl)	265	288/288	1.14	1.25/1.18	—	—/—
Poly(iminoxalyliminotetramethylene)	332	309/323	1.37	1.38*/1.31*	1.58	1.60/1.60
Poly(iminomethyleneiminoglutaryl)	332	328/344	1.34	1.37/1.32	—	—/—
Poly/heptano -7- lactam)	335	347/352	1.10	1.20/1.12	1.57	1.59/1.59
Poly[D(-)-3- methyl- 6 - aminocaproic acid]	278	268/252	1.11	1.21*/1.14*	—	—
Poly(iminomethyleneiminohexanedioyl)	330	318*/335*	1.33	1.33/1.26	1.66	1.69/1.69
Poly(2- methyl-2- propyl-3- aminopropionic acid)	215	262/249	1.07	1.20/1.13	1.74	1.72/1.71
Poly(gamma- ethylglutamic acid)	—	—/—	1.23	1.27/1.20	—	—/—
Poly(4- methylhexano - 6 -lactam)	291	308/301	1.08	1.20/1.12	1.60	1.63/1.63
Poly(imino - 2,2,3,3- tetramethyl-1- oxopropane -1,3- diyl)	252	265/251	1.10	1.21/1.13	—	—/—
Poly(iminomalonyliminobutane -1,4- diyl)	330	326*/347*	1.32	1.35/1.29	1.63	1.63/1.63
Poly[imino(1- oxooctane -1,8- diyl)]	324	302/321	1.08	1.18/1.09	1.56	1.61/1.60
Poly(iminoethyleneliminohexanedioyl)	328	317*/333*	1.26	1.33/1.28	1.64	1.63/1.62
Poly(tetramethylene succinamide)	328	315/331	1.28	1.32*/1.26*	1.64	1.64/1.64
Poly(pentamethylene malonamide)	338	312/318	1.30	1.28/1.20	1.56	1.57*/1.59*
Poly(hexamethylene oxamide)	—	—/—	1.28	1.35/1.29	1.60	1.62/1.61
Poly(2- hydroxyphenyl D-valyl-L-alaninate)	286	292*/295*	1.27	1.34/1.28	1.76	1.74/1.72
Poly(2- butyl-2- methyl-3- aminopropionic acid)	213	253/239	1.04	1.18*/1.10*	1.71	1.71/1.70
Poly(alpha- isobutylaspartic acid)	—	—/—	1.22	1.28/1.21	1.66	1.64/1.64
Poly(gamma - propylglutamic acid)	239	273/294	1.25	1.27*/1.20	1.74	1.72*/1.71*
Poly(5-methyl-7- aminoenanthic acid)	—	—/—	1.19	1.18/1.10	1.58	1.59*/1.59*
Poly(hydrazine-1,2- diylisophthaloyl)	432	435/429	1.44	1.50/1.49	1.65	1.63/1.64
Poly(hydrazine-1,2- diylterephthaloyl)	470	452/436	1.56	1.55*/1.57	1.66	1.65/1.65
Poly(imino-3- isopropyl-2- 2- dimethyl-1- oxopropane-1,3- diyl)	242	256*/240*	1.08	1.18/1.10	—	—/—
Poly(iminomethyleneimino -1,7- dioxoheptane -1,7- diyl)	328	300/313	1.23	1.33/1.26	1.64	1.63/1.63
Poly(9- aminononanoic acid)	318	329/337	1.10	1.16/1.06	1.55	1.54/1.59
Poly[iminoxalylimino (2,2- dimethylpentane -1,5 - diyl)]	302	298/309	1.22	1.32/1.27	—	—/—
Poly[6 - (alanylalmino) caproic acid]	323	318/325	1.26	1.35/1.29	1.61	1.61/1.61
Poly(imino -1- oxo -2- phenylmethylethylene]	—	—/—	1.20	1.30/1.24	1.66	1.71*/1.70*
Poly(imino -1- oxo-3- phenylpropane-1,3- diyl)	376	373/346	1.25	1.30/1.24	1.67	1.65/1.65
Poly(iminomethyleneimino -1,8- dioxooctane -1,8- diyl)	326	313/328	1.24	1.30/1.22	1.62	1.62/1.61
Poly(iminomalonyliminohexane -1,6 - diyl)	323	310*/326	1.22	1.30/1.24	1.59	1.57/1.58
Poly(tetramethylene adipamide)	325	304/322	1.23	1.28/1.24	1.58	1.58*/1.60*
Poly(10 - aminocapric acid)	316	332/338	1.02	1.14/1.04	1.54	1.49/1.59
Poly(iminoglutaryliminopentane-1,5 - diyl	325	335*/336*	1.25	1.27/1.19	1.61	1.63/1.62
Poly(iminobutanedioyliminohexane-1,6 - diyl)	—	—/—	1.26	1.27*/1.21*	1.61	1.60/1.59
Poly(iminoethyleneliminoterephthaloyl)	422	396*/393*	1.23	1.40/1.37	1.62	1.631.62
Poly(hydrazine -1,2 - diyldecanedioyl)	325	315/315	1.18	1.36*/1.28*	1.61	1.64/1.63
Poly {imino[2-(3-pentyloxypropanoyl)-1- oxoethylene]}	233	251/282	1.11	1.24/1.16	—	—/—

**Table 1** (continued)

Polymer name		$T_g$ (K)		$\rho$ (g/cm <sup>3</sup> )		$n$
Poly(hydrazine -1,2- diylsuccinylhydrazine-1,2- diyladipoyl)	335	323/325	1.46	1.42/1.39	1.70	1.67/1.67
Poly(hydrazine -1,2- diyloxalylhydrazine -1,2- diylisophthaloyl)	440	469/433	1.68	1.53/1.61	1.67	1.66/1.67
Poly(iminopentane-1,5- diylimino adipoyl)	324	327/333	1.14	1.25/1.17	1.60	1.62/1.61
Poly[imino(1- oxoundecane-1,11- diyl)]	363	335*/336*	1.01	1.12/1.01	1.54	1.52/1.59
Poly(iminomethyleneiminodecanedioyl)	323	322/325	1.21	1.26*/1.16*	1.59	1.59*/1.59*
Poly(iminobutane-1,4- diyliminoheptanedioyl)	324	301/321	1.16	1.24/1.15	1.60	1.60/1.60
Poly(iminoglutaryl iminohexane-1,6- diyl)	324	299*/321	1.26	1.25/1.18	1.59	1.621.62
Poly(iminopropane-1,3- diyliminoterephthaloyl)	420	383/374	1.35	1.37/1.32	1.60	1.60/1.61
Poly[ethylene ( <i>p</i> - carboxyphenylene)acetamide]	355	383/374	1.35	1.36/1.30	1.64	1.67*/1.67*
Poly(azelaoyl oxalodihydrazide)	333	364/342	1.41	1.45/1.45	1.61	1.64*/1.63*
Poly( <i>p</i> -phenylene glutaramide)	386	384*/375	1.35	1.40/1.39	—	—/—
Poly(imino-1,2-phenyleneiminoglutaryl)	400	386/378	1.28	1.38/1.38	1.69	1.68/1.69
Poly(imino-1,3-phenyleneiminoglutaryl)	364	367/366	1.32	1.33*/1.32	1.70	1.69*/1.68*
Poly(imino-2,2- dimethyl-1- oxo-3-phenylpropane -1,3- diyl)	340	317/344	1.17	1.13/1.18	—	—/—
Poly(imino-2,3- dimethoxybutanedioyliminopentamethylene)	270	251*/311	1.12	1.18/1.08	1.64	1.65*/1.65
Poly(iminopropane-1,3- diyliminoisophthaloyl)	406	395/396	1.32	1.42/1.42	1.64	1.64/1.64
Poly[imino(1- oxododecane -1,12- diyl)]	299	337/332	0.99	1.11/0.98	—	—/—
Poly(trimethylene ( <i>p</i> - carboxyphenylene)acetamide)	351	380/360	1.30	1.34*/1.28*	1.61	1.64*/1.64*
Poly(iminoethyleneiminodecanedioyl)	323	318/318	1.14	1.30*/1.22*	1.59	1.621.60
Poly(iminobutane-1,4- diyliminooctanedioyl)	323	297/320	1.15	1.23/1.15	1.58	1.57/1.59
Poly(pentamethylene pimelamide)	323	305/323	1.20	1.32/1.25	1.58	1.55/1.58
Poly(iminobutane-1,4- diyliminoterephthaloyl)	419	384/361	1.30	1.36/1.30	1.60	1.61/1.61
Poly(imino-1,2-phenyleneiminoadipoyl)	—	—/—	1.25	1.29/1.28	1.68	1.65/1.65
Poly(imino-1,3-phenyleneiminoadipoyl)	361	362/354	1.28	1.30/1.29	1.68	1.66*/1.67*
Poly(imino-1,4-phenyleneiminoadipoyl)	381	361/353	1.30	1.30/1.30	1.68	1.67/1.67
Poly{imino-1-[2-(4-nitrobenzylloxycarbonyl)ethyl]-2- oxoethylene}	350	382/387	1.43	1.40/1.37	1.60	1.59/1.60
Poly(alpha- benzyl-L-glutamate)	—	—/—	1.26	1.32/1.25	—	—/—
Poly(iminohexane-1,6- diyliminoheptanedioyl)	327	317/323	1.14	1.29/1.22	1.58	1.59/1.60
Poly(iminoadipoyliminoheptane-1,7- diyl)	333	293*/320*	1.12	1.20/1.11	1.58	1.59/1.60
Poly(13- aminotridecanoic acid)	314	342/330	1.03	1.09/0.95	1.50	1.44*/1.52*
Poly(iminomesaconyliminoethylene-1,3-phenylenemethyl)	370	390/360	1.24	1.34/1.28	1.60	1.61*/1.61*
Poly(hexamethylene 3- methyladipamide)	273	271/315	1.09	1.15*/1.04*	1.57	1.63/1.61
Poly[tetramethylene ( <i>p</i> - carboxyphenylene)- acetamide]	349	381/370	1.27	1.32/1.26	—	—/—
Poly(tetramethylene azelamide)	322	295/320	1.13	1.21/1.11	—	—/—
Poly(trimethylene ( <i>p</i> - carboxyphenylene)acetamide]	—	—/—	—	—/—	1.62	1.63/1.64

\*The predicted value. The data on the left of the slash are the calculated values for the MLR method; the data on the right are the calculated values for the ANN method

The PMA value can be calculated using the following formula:  $PMA = N_{\text{methylene}} / N_{\text{acylamino}}$ , where  $N_{\text{methylene}}$  is the number of methylenes in the backbone chain, and  $N_{\text{acylamino}}$

is the number of acylamino in the backbone chain. QSPR models of three properties for training set were obtained by the MLR method as shown in Table 3.

**Table 2** Quantum-chemical and other descriptors

Number	Descriptors	Definition	Unit
1	$E_{\text{HOMO}}$	The energy of the highest occupied molecular orbital (HOMO)	au
2	$E_{\text{LUMO}}$	The energy of the lowest unoccupied molecular orbital (LUMO)	au
3	$Q_{\text{O}-}$	Atomic charge for oxygen in acylamino	au
4	$E_t$	Total molecular energy	au
5	$\mu$	Molecular dipole moment	Debye
6	$U$	The internal energy	au
7	PMA	The proportion of methylene to acylamino in the backbone chain	—
8	$M$	The relative molecular weight of the repeating unit	—

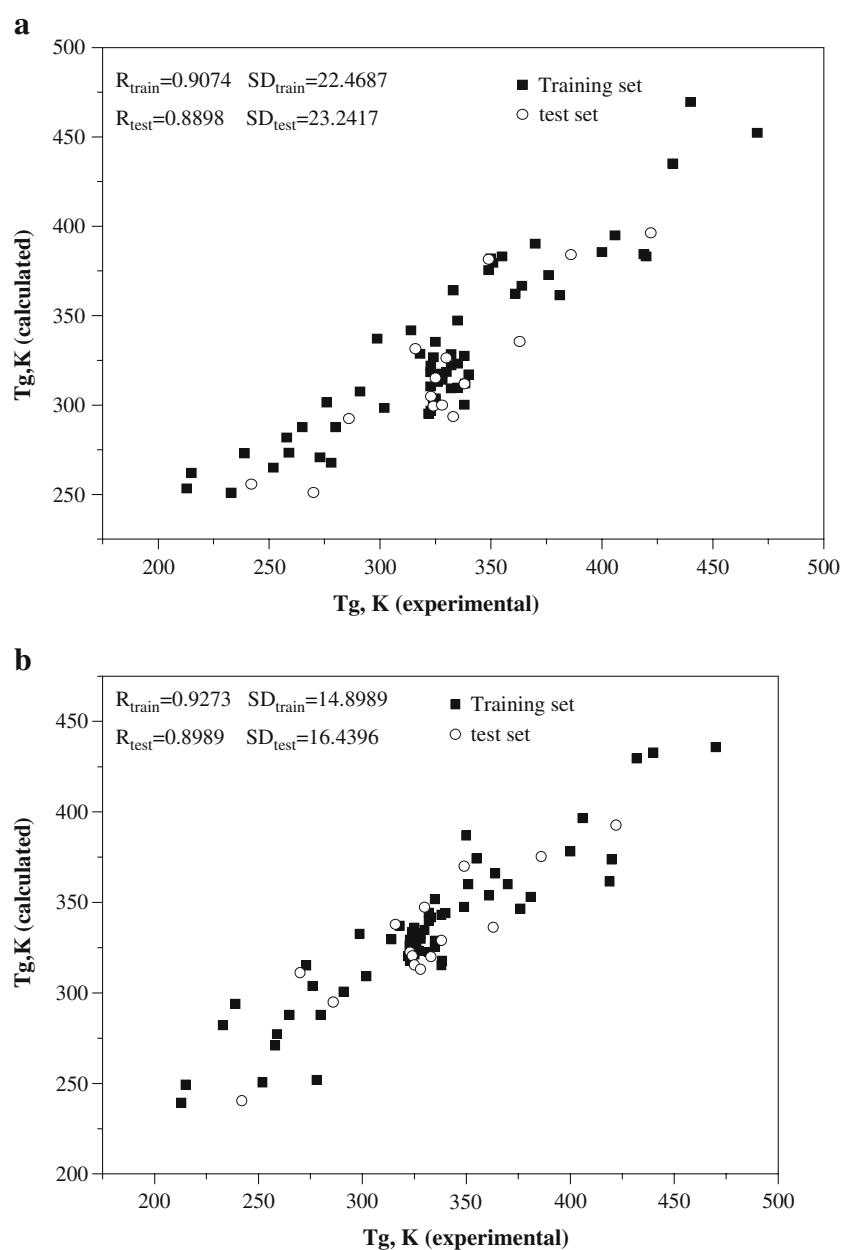
**Table 3** Three models for  $T_g$ ,  $\rho$  and  $n$  and statistics properties of the models

No	Equation and statistics property
1	$T_g = 444.7204 - 14.3153PMA - 193.1553E_{LUMO} + 178.4501Q_{O-} + 6.35783M - 2.7701U$ $R=0.9074$ ; $F$ -value=43.4878; $P$ value<0.00001; SD=22.4687; N=63
2	$\rho = 2.6854 + 3.0329E_{HOMO} + 0.8751Q_{O-} - 0.0005E_t$ $R=0.9474$ ; $F$ -value =125.0056 $P$ value<0.00001; SD=0.0422; N=70
3	$n = 1.8557 - 0.0281PMA + 0.5319E_{HOMO} + 0.9934E_{LUMO} + 0.0079\mu$ $R=0.9298$ ; $F$ -value =47.1978; $P$ value<0.00001; SD=0.0204; N=53

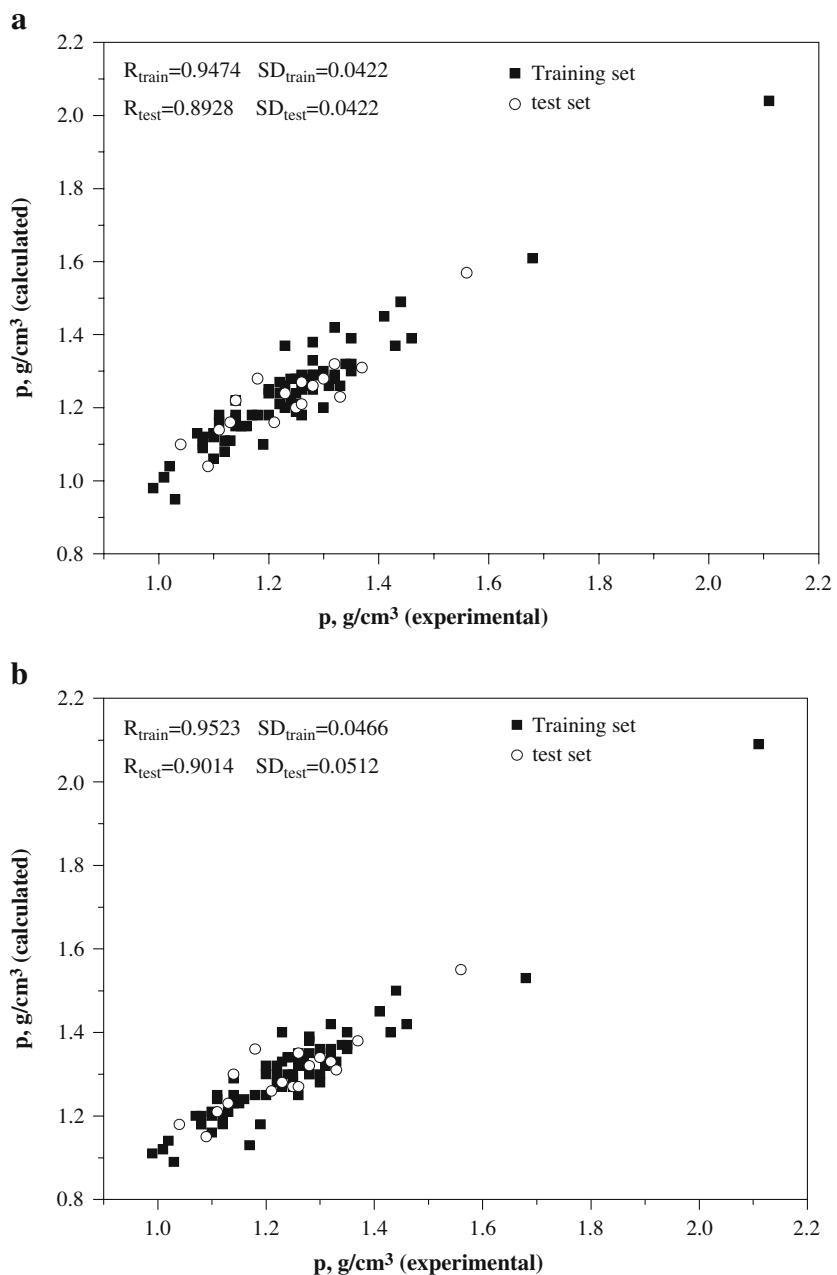
The BPANN models were generated using the same descriptors as for the MLR results.  $T_g$ ,  $\rho$  and  $n$  were calculated with five-two-one networks, three-two-one networks and four-two-one networks, respectively. The BPANN models are more accurate than the MLR models, as shown in Figs. 2, 3 and 4.

**Fig. 2** Experimental vs. calculated  $T_g$  using two methods for the training set of 53 polyamides and the test set of 13 polyamides **a** MLR  
**b** BPANN

A subset of 25 descriptors was used. After the examination by MLR, only six descriptors were selected for the best correlation models as shown in Table 4. The  $T$ -values of the selected descriptors are greater than 2, the SD does not exceed 25% of the value, and the  $P$  value is as low as 0.05, which indicate the models obtained have prediction ability.



**Fig. 3** Experimental vs. calculated  $\rho$  using two methods for training set of 70 polyamides and the test set of 17 polyamides  
**a** MLR **b** BPANN

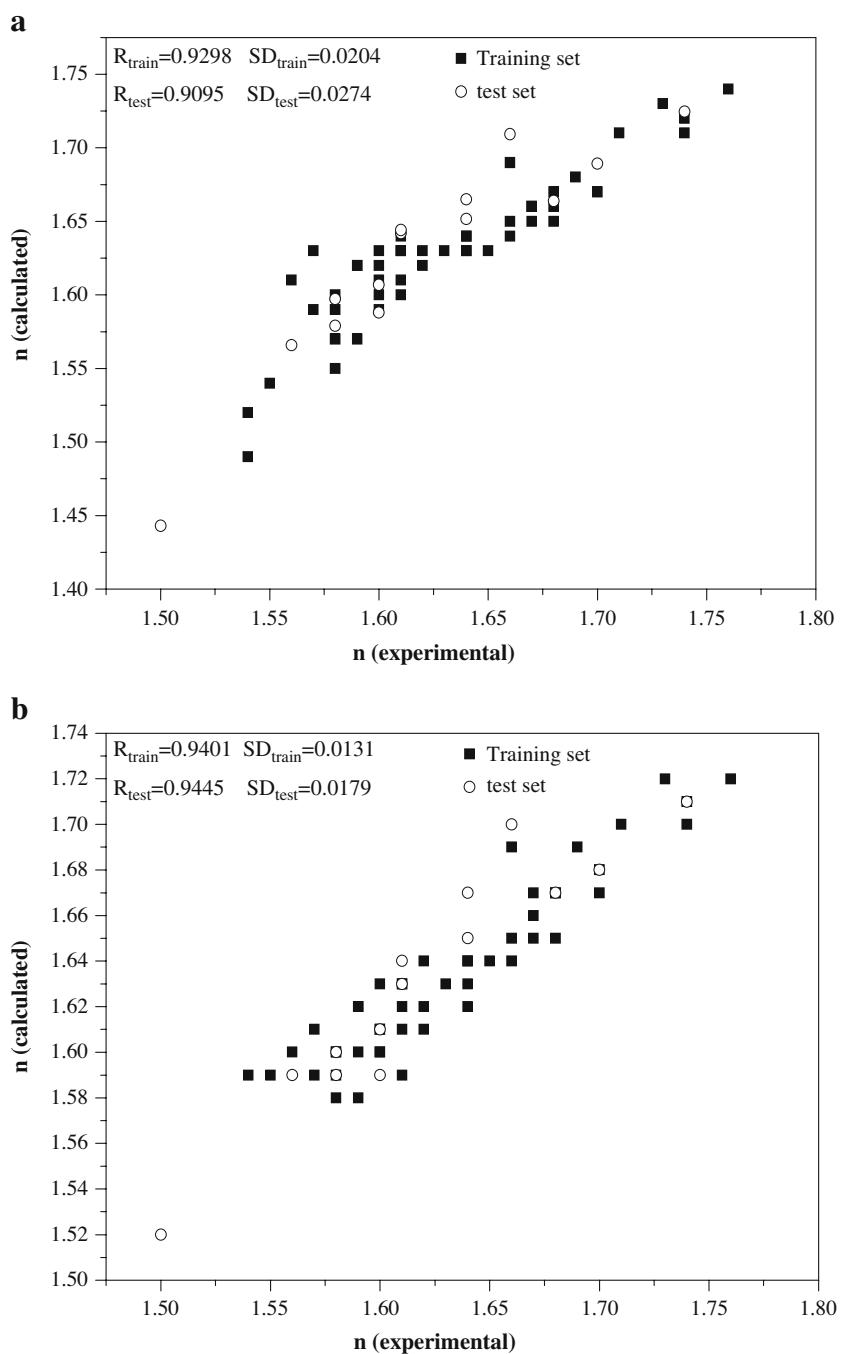


Five descriptors were chosen for the  $T_g$  model, such as PMA,  $E_{\text{LUMO}}$ ,  $Q_{\text{O}-}$ ,  $M$  and  $U$ . According to the  $T$ -test value, the most significant descriptor is  $U$ , which is the internal energy including vibration energy, rotation energy and so on. As the value of  $U$  increases, the stability of the ground state of a molecule decreases, the temperature of the molecule approaching the degree of vitrification will low,  $U$  has a negative effect on  $T_g$ . The second important descriptor is PMA, which counts the proportion of the nonpolar (methylene) to the polar groups (acylamino). In these research structures, the number of the acylamino groups is fixed. Consequently, the higher the number of methylenes, the higher the proportion of the nonpolar group (methylene) to

the polar group (acylamino). The intermolecular force between the backbone chains decreases, the free volume increases, and  $T_g$  decreases, so the PMA has the negative correlation with  $T_g$ .  $M$  is the relative molecular weight of repeating unit, which describes directly the viscosity coefficient of the polymer, the  $M$  has a positive correlation with  $T_g$ .  $Q_{\text{O}-}$  is the value of the charge of the oxygen atom of the acylamino group, correlating to the molecular polarity, which has a positive effect on  $T_g$ .

Three descriptors are employed in the QSPR model for density ( $\rho$ ), such as  $E_{\text{HOMO}}$ ,  $Q_{\text{O}-}$  and  $E_t$ . The density at room temperature  $\rho_{298 \text{ K}}$  can be obtained according to the formula  $\rho_{298 \text{ K}} = M/V_{298 \text{ K}}$ , where  $M$  is the relative molecular weight

**Fig. 4** Experimental vs. calculated  $n$  using two methods for the training set of 53 polyamides and the test set of 14 polyamides **a** MLR **b** BPANN



of repeating unit of polymer. Based on the  $T$ -value, the most important descriptor is  $E_{\text{HOMO}}$ , which has a direct effect on the ionization energy.

Four descriptors were chosen for the QSPR model for  $n$ , such as PMA,  $E_{\text{HOMO}}$ ,  $E_{\text{LUMO}}$  and  $\mu$ . As to the  $T$ -test value, the descriptor PMA is the most important for this model. As described above, the descriptor PMA measures the polarity of a molecule,  $n$  will increase as the PMA decreases. The descriptor  $\mu$ , which directly affects the refraction index, has a positive correlation with  $n$ . The descriptors  $E_{\text{HOMO}}$  and

$E_{\text{LUMO}}$ , which are the energies of the frontier molecular orbitals, have a positive correlation with  $n$ .

## Conclusion

The calculated results demonstrate that quantum-chemical and structural descriptors successfully correlate with three properties of  $T_g$ ,  $\rho$  and  $n$  of polyamides. The MLR and BPANN methods are practical methods for the prediction of  $T_g$ ,  $\rho$  and  $n$  of polyamides. The value of  $T_g$  and  $\rho$  are affected

**Table 4** Descriptors used in MLR for three properties of polyamides

	Descriptor	Coefficient	Standard error	T-statistic	P value
$T_g$	Constant	444.7204	40.5796	10.9592	0.00001
	PMA	14.3153	1.8477	7.7473	0.00001
	$E_{LUMO}$	-193.1550	57.2273	-3.3752	0.00119
	$Q_{O^-}$	178.4502	50.6704	3.5217	0.00075
	$M$	6.3578	0.9556	6.6527	0.00001
$\rho$	$U$	-2.7701	0.2091	-13.2460	0.00001
	Constant	2.6116	0.1018	25.6474	0.00001
	$E_{HOMO}$	3.0185	0.2868	10.5247	0.00001
	$Q_{O^-}$	0.7740	0.0994	7.7826	0.00001
$n$	$E_t$	-0.0006	0.0001	-9.8650	0.00001
	Constant	1.8556	0.0340	54.5329	0.00001
	PMA	-0.0259	0.0017	-15.1504	0.00001
	$E_{HOMO}$	0.5628	0.1448	3.8861	0.00025
	$E_{LUMO}$	0.9358	0.1088	8.5993	0.00001
	$\mu$	0.0077	0.0021	3.7409	0.0004

mainly by the molecular energy and molecular polarity. The value of  $n$  is affected mainly by molecular polarity.

**Acknowledgement** The authors like to thank the financial support from the Scientific Research Fund of Hunan Provincial Education Department (05A002) for the research work.

## References

- Bicerano J (2003) Encyclopedia of polymer science and technology. Wiley, New York
- Krause S, Gormley JJ, Roman N, Shetter JA, Watanade WH (1965) *J Polym Sci Part A: Polym Chem* 3:3573–3586
- Bicerano J (1996) Prediction of polymer properties, 2nd edn. Marcel Dekker, New York
- van Krevelen DW (1990) Properties of Polymers, 3rd edn. Elsevier, Amsterdam
- van Krevelen DW (1976) Properties of polymers—their estimation and correlation with chemical structure, 2nd edn. Elsevier, New York
- Askadskii AA (1996) *Polym Sci USSR* 9:471–487
- Askadskii AA, Slonimskii GL (1971) *Polym Sci USSR* 13: 2158–2160
- Gao H, Harmon JP (1997) *J Appl Polym Sci* 64:507–517
- Wiff DR, Altieri MS, Goldfarb IJ (1985) *J Polym Sci Part B: Polym Phys* 23:1165–1176
- Hopfinger AJ, Koehler MG (1988) *J Polym Sci Part B: Polym Phys* 26:2007–2028
- Boudouris D, Constantinou L, Panayiotou C (2000) *Fluid Phase Equilibria* 167:1–19
- Camelio P, Cypcar CC, Lazzeri V, Waegell B (1997) *J Polym Sci Part A: Polym Chem* 35:2579–2590
- Cypcar CC, Camelio P, Lazzeri V, Mathias LJ, Waegell B (1996) *Macromolecules* 29:8954–8959
- Katritzky AR, Sild S, Lobanov V, Karelson M (1998) *J Chem Inf Comput Sci* 38:300–304
- Katritzky AR, Sild S, Lobanov V, Karelson M (1998) *J Chem Inf Comput Sci* 38:1171–1176
- Cao CZ, Lin YB (2003) *J Chem Inf Comput* 43:643–650
- Knoll W (1998) *Annu Rev Phys Chem* 49:569–638
- Herve P, Vandamme LK (1994) *J Infrared Phys Technol* 35: 609–615
- Mekenyany O, Dimitrov S, Bonchev D (1983) *Eur Polym J* 1: 1185–1193
- Yang CJ, Jenekhe SA (1995) *Chem Mater* 7:1276–1285
- Brandrup J (1999) Polymer Handbook, 4th edn. Wiley, New York
- Polymer Database: <http://polymer.nims.go.jp/PoLyInfo/>
- Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, MonTgomery JA, Jr, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson BG, Chen W, Wong MW, Andres JL, Head-Gordon M, Replogle ES, Pople JA (2003) Gaussian 03, Revision B.05. Gaussian Inc, Pittsburgh PA
- Tang Q, Feng M (2002) Practical statistics and DPS Data Processing System. Beijing Science