Liquid Crystals
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713926090

Prediction of nematic transition temperatures in thermotropic liquid crystals by a heuristic method

Yueying Ren[a]; Huanxiang Liu[a]; Xiaojun Yao[a]; Mancang Liu[a]; Botao Fan[b]
[a] Department of Chemistry, Lanzhou University, Lanzhou 730000, China [b] Universite' Paris 7-Denis Diderot, ITODYS 1, rue Guy de la Brosse, 75005 Paris, France

To cite this Article Ren, Yueying , Liu, Huanxiang , Yao, Xiaojun , Liu, Mancang and Fan, Botao(2007) 'Prediction of nematic transition temperatures in thermotropic liquid crystals by a heuristic method', Liquid Crystals, 34: 11, 1291 — 1297

To link to this Article: DOI: 10.1080/02678290701663712

URL: http://dx.doi.org/10.1080/02678290701663712

Liquid Crystals, Vol. 34, No. 11, November 2007, 1291–1297

Prediction of nematic transition temperatures in thermotropic liquid crystals by a heuristic method

YUEYING REN†, HUANXIANG LIU†, XIAOJUN YAO*†, MANCANG LIU† and BOTAO FAN‡
†Department of Chemistry, Lanzhou University, Lanzhou 730000, China
‡Universite' Paris 7-Denis Diderot, ITODYS 1, rue Guy de la Brosse, 75005 Paris, France

(*Received 15 June 2006; accepted in revised form 15 July 2007*)

A quantitative structure–property relationship (QSPR) study was performed on 42 homogeneous thermotropic liquid crystals (TLCs) to predict their nematic transition temperatures (T_N) using molecular descriptors calculated by CODESSA software. A heuristic method (HM) was applied to select descriptors and generate models. The selected five descriptors reflect the constitutional, steric and electronic characteristics that affect the nematic transition phenomena. The high correlation coefficient, R^2, of 0.9881 for the training set indicates that a linear relationship exists between the structural information and T_N in TLCs. Satisfactory results were obtained. For the test set, this model gave an R^2 of 0.96, a root mean squared error of 6.3654 and an absolute average relative deviation of 9.2017%. The results provide a simple, practical and effective method for the prediction of T_N of TLCs and the approach can be extended to other QSPR investigations.

1. Introduction

In thermotropic liquid crystals (TLCs), the nematic phase is composed of rod-like molecules, the long axes of which are preferably oriented along a given direction; the phase is characterized by full translational disorder. The nematic LC phases are technologically the most important among all possible LC phases and, mainly because of their unique physicochemical properties, used in all commercially available LC displays [1].

As one of the earliest findings that indicated the existence of an LC phase, the nematic–isotropic (N–I) transition has been a topic of active and extensive theoretical and experimental studies. Generally speaking, the N–I transition is weak first order; however, it was found that if the molecules prefer to line up transverse to an applied field (an electric rather than a magnetic field) there is a critical field above which the phase transition is more closely a second-order one [2]. Marčelja [3] proposed a theory that accounts specifically for the effects of the end-chains on the anisotropic interactions between the molecules. This theory provides good agreement with experimental data on I–N transition temperatures (T_N) and entropies, and explains for the first time the odd–even effect of the end-chains on these quantities along a homologous series [3]. By extending this theory to the isotropic, density-dependent component of the molecular interaction, the predicted magnitudes of both $(T_N - T^*)/T_N$ and its density variation across the N–I transition, as well as the variation of T_N with pressure, agree with experimental values, where T^* indicates the supercooling temperature [4].

The N–I transition temperature (T_N) is a significant property in revealing both intramolecular and intermolecular aspects of TLCs and is very important in their processing and application.

The quantitative structure–property relationship (QSPR) approach has been widely used in the prediction of physical and chemical properties of organic compounds [5–7]. QSPR is based on the assumption that the variation of the behaviour of the compounds, as expressed by any measured physical or chemical properties, can be correlated with changes in molecular features of the compounds, termed descriptors. Accordingly, the QSPR approach attempts to establish simple mathematical relationships to describe the correlation of a given property to molecular structures for a set of compounds. The advantage of this approach over other methods lies in the fact that it requires only the knowledge of chemical structure and is not dependent on the experimental properties.

The main steps involved in QSPR include data collection, molecular geometry optimization, molecular descriptor generation, descriptor selection, model development and, finally, model performance evaluation. A

---

*Corresponding author. Email: xjyao@lzu.edu.cn

Liquid Crystals
ISSN 0267-8292 print/ISSN 1366-5855 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/02678290701663712

QSPR study cannot only develop a method for the prediction of the property of new compounds that have not yet been synthesized but also can identify and describe important structural features of molecules that are responsible for variations in molecular properties.

Recently, García *et al.* [8] used multiple linear regression (MLR) analysis to develop the first QSPR model for prediction of T_N for a homologous series of 42 TLCs. The nine-parameter model obtained is notable for its excellent fit to experimental data, showing a correlation coefficient R^2 of 0.954 and a standard deviation of 6.425% for the 29 compounds in the training set that were modelled. However, this model has some drawbacks. Firstly, inclusion of variables into a QSAR/QSPR should meet stringent criteria, from both a statistical and mechanistic (i.e. that the parameter relates in some manner to activity) viewpoint [9]. The ratio of observations to variables should be as high as possible, and at least 5:1 [10], whereas the ratio in this model is less than 5:1 (29:9≈3:1). From the statistics, if the number of variables is comparable to the number of training patterns, the parameters of the model may become unstable and unlikely to replicate if the study were to be repeated. In this sense, this model is too complex and will meet these problems. Secondly, some of descriptors are highly inter-correlated (>0.95) with one or two remainders (G and H, V and S, O and S) and, therefore, leads to redundancy of information. Depending on the nature of the regression technique, the presence of the redundant features can cause the system to focus attention on the idiosyncrasies of the individual samples and lose sight of the broad picture that is essential for generalization beyond the training set. Based on the above observations, it is highly necessary to investigate this data set again and develop a new, more acceptable model to further investigate the relationship between the property and structure.

One of the important problems for QSPR applications is the numerical representation (often called molecular descriptor) of the chemical structure. The built model performance and the accuracy of the results are strongly dependent on the way the structural representation is performed. Various numerical representations of the compounds have been proposed in QSPR studies, i.e. constitutional and topological descriptors, numerical code, quantum chemistry descriptors, etc. The software CODESSA, developed by Katritzky *et al.* [11], enables the calculation of a large number of quantitative descriptors based solely on the geometrical and electronic structural information and codes this chemical information into mathematical form. CODESSA combines diverse methods for quantifying the structural information about the molecule with advanced statistical analysis to establish molecular structure–property/activity relationships. CODESSA has been applied successfully in a variety of QSPR analyses [5–7].

In the present work, the CODESSA program was used for the calculation of the descriptors and for the statistical analysis to obtain multi-parameter QSPR equations describing the T_N of 42 TLCs. A heuristic method was utilized to establish a quantitative linear relationship between T_N and the molecular structure. The aim of the present study was to establish a QSPR model that could be used for the prediction of T_N for TLCs from their molecular structures alone, and at the same time, to seek for the important structural features related to the N–I transition of these compounds.

2. Materials and methods

The data set for this investigation consists of 42 TLCs, which have been used by García *et al.* [8] to construct an MLR model. All the molecules contain two aromatic rings linked by an ester group, COO–, with different terminal chains. The molecular structures as well as their corresponding T_N values are shown in table 1.

2.1. *The calculation of structural descriptors*

All molecules were drawn and pre-optimized using molecular mechanics force fields in the HyperChem 6.0 program [12]. A more precise optimization was performed with a semi-empirical AM1 method. All calculations were carried out at the restricted Hartree–Fock level with no configuration interaction using the Polar–Ribiere algorithm until rms gradient $0.1\,\text{kcal}\,\text{Å}\,\text{mol}^{-1}$. The resulting geometries were transferred into the MOPAC 6.0 program package to calculate optimized structural coordinates and net atomic charges [13]. Then the output files from MOPAC were transferred into the CODESSA software package to calculate constitutional, topological, geometrical, electrostatic and quantum chemical descriptors. These descriptors contain information about the connections between atoms, shape, branching, symmetry, distribution of charge and quantum-chemical properties of the molecule. Using the above process, the major problem in QSPR studies, i.e. the translation of molecular structure to computer-readable form while retaining as much as structural information as possible, is solved.

2.2. *Heuristic method (HM)*

The heuristic multilinear regression procedures available in the CODESSA program were used to perform a complete search for the best multilinear correlations with a multitude of descriptors. The heuristic method

Table 1. List of studied chemicals and experimental, calculated and predicted T_N by HM and MLR by García *et al.* [8].

No.	L	R	T_N (Exp.)/°C	T_N (HM)/°C	%ARD	T_N (MLR)/°C	%ARD
1	C_5H_{11}–	O–CH_3	42.2	43.0	−2.0	39.9	5.5
2*	C_5H_{11}–	O–C_2H_5	63.4	65.6	−3.5	47.4	25.2
3	C_5H_{11}–	O–C_3H_7	44	45.9	−4.4	50.2	−14.1
4	C_5H_{11}–	O–C_4H_9	57.7	49.2	14.7	52	9.9
5	C_5H_{11}–	O–C_5H_{11}	51.8	54.6	−5.3	50.8	1.9
6	C_5H_{11}–	O–C_6H_{13}	59.3	58.1	2.1	53.2	10.3
7	C_5H_{11}–	O–C_7H_{15}	57.4	56.9	0.9	55.6	3.1
8*	C_5H_{11}–	O–C_8H_{17}	60.6	53.2	12.2	56	7.6
9*	C_5H_{11}–	O–C_9H_{19}	58.4	57.7	1.2	54.9	6.0
10*	C_5H_{11}–	O–$C_{10}H_{21}$	60.3	59.8	0.8	54.1	10.3
11	C_5H_{11}–	O–$C_{12}H_{25}$	60.6	59.9	1.1	52	14.2
12	C_5H_{11}–	O–$C_{14}H_{29}$	60.9	58.6	3.7	52.9	13.1
13*	C6H13–	O–CH_3	38	30.9	18.8	43	−13.2
14	C_6H_{13}–	O–C_2H_5	51.8	51.0	1.5	50.9	1.7
15	C_6H_{13}–	O–C_3H_7	36	43.9	−21.9	53.2	−47.8
16*	C_6H_{13}–	O–C_4H_9	49.4	48.0	2.8	51.6	−4.5
17*	C_6H_{13}–	O–C_5H_{11}	45	53.3	−18.5	60.2	−33.8
18	C_6H_{13}–	O–C_6H_{13}	53.2	53.5	−0.5	52.2	1.9
19*	C_6H_{13}–	C_3H_7–	40	23.5	41.2	37.3	6.8
20	C_6H_{13}–	C_5H_{11}–	43.5	41.5	4.7	46.2	−6.2
21	C_6H_{13}–O–	C_5H_{11}–	63	64.8	−2.8	61.4	2.5
22	C_4H_9–O–	O–C8H17	89	86.7	2.6	85.3	4.2
23	C_4H_9–O–	O–C_9H_{19}	86	86.6	−0.7	88.1	−2.4
24	C_4H_9–O–	O–$C_{10}H_{21}$	87	85.8	1.4	85.7	1.5
25	C_4H_9–O–	O–$C_{12}H_{25}$	84.5	86.6	−2.5	88.1	−4.3
26*	C_5H_{11}–O–	O–CH_3	72	69.8	3.1	68.5	4.9
27	C_5H_{11}–O	O–C_2H_5	90.8	92.6	−2.0	79.8	12.1
28	C_5H_{11}–O	O–C_3H_7	78.5	76.6	2.4	84.1	−7.1
29	C_5H_{11}–O	O–C_4H_9	82	83.8	−2.2	81.4	0.7
30	C_5H_{11}–O	O–C_5H_{11}	81	83.0	−2.5	83.9	−3.6
31	C_5H_{11}–O	O–C_6H_{13}	84.5	84.8	−0.4	85.5	−1.2
32*	C_5H_{11}–O	O–C_7H_{15}	82	85.7	−4.5	87.1	−6.2
33	C_5H_{11}–O	O–C_8H_{17}	85	85.6	−0.7	88.5	−4.1
34	C_5H_{11}–O	O–C_9H_{19}	88	86.5	1.7	87.7	0.3
35	C_5H_{11}–O	O–$C_{10}H_{21}$	82	85.4	−4.2	87.1	−6.2
36*	C_5H_{11}–O	O–$C_{12}H_{25}$	80	86.4	−8.0	88.8	−11.0
37*	C_5H_{11}–O	O–$C_{14}H_{29}$	78.2	81.8	−4.6	88.3	−12.9
38*	C_5H_{11}–O	O–$C_{16}H_{33}$	76.5	76.8	−0.4	82.4	−7.7
39	C_5H_{11}–O	O–$C_{18}H_{37}$	74.7	75.6	−1.2	78.6	−5.2
40	C_6H_{13}–O–	O–CH_3	78.5	76.5	2.6	77.1	1.8
41	C_6H_{13}–O–	O–C_2H_5	95.9	94.8	1.1	87.4	8.9
42	$C_{10}H_{21}$–O–	O–C_6H_{13}	88.9	86.0	3.2	86.9	2.2

*Test set compounds

(HM) has the advantages of high speed and no restrictions on the size of the data set. It can either quickly give a good estimation about what quality of correlation to expect from the data, or derive several best regression models. Besides, it can demonstrate which descriptors have bad or missing values, which descriptors are insignificant and which descriptors are highly inter-correlated. This information is helpful in reducing the number of descriptors in QSAR/QSPR studies.

The HM for descriptor selection proceeds with a pre-selection of descriptors by sequentially eliminating descriptors that do not match any of the following criteria: (i) the F-test greater than one unit; (ii) R^2 value less than a value defined at the start (default 0.01); (iii) the student's t-test less than that defined (default 0.1); and (iv) duplicate descriptors having a higher squared inter-correlation coefficient than a predetermined level (usually 0.8). Thereafter all possible one-parameter regression models were tested and insignificant descriptors were removed. The remaining descriptors are listed in decreasing order of correlation coefficient when used in one-parameter correlations. After the pre-selection of descriptors, MLR models are developed in a stepwise procedure. Starting with the top descriptor from the list, two-parameter correlations are calculated. In the following steps new descriptors are added one-by-one until the pre-selected number of descriptors in the model is achieved. The final result is a list of the 10 best models according to the values of the *F*-test (*F*) and correlation coefficient (R^2). The goodness of the correlation is tested by *F*, R^2 and the standard deviation (*s*). The stability of the correlations was tested against the cross-validated coefficient (R^2_{cv}) by focusing on the sensitivity of the model to the elimination of any single data point. Briefly, for each data point, the regression is recalculated with the same descriptors but for the data set without this point. The obtained regression is used to predict the value of this point, and the set of estimated values calculated in this way is correlated with the experimental values.

The HM usually produces correlations 2–5 times faster than other methods, with comparable quality [14]. The rapidity of calculations from the HM renders it the first choice in practical research. Thus, in the present investigation, we used this method to select descriptors and build the linear model. A detailed discussion about the HM can be found in the literature [15].

3. Results and discussion

To enable comparison with results in the literature, the separation of the LCs into training and test sets is identical with that used by García *et al.* [8]. The training set of 29 compounds was used to adjust the parameters of the model, and the test set of 13 compounds was used to evaluate model prediction ability.

In order to identify possible outliers and clusters, principal component analysis (PCA) was performed by using the whole data set. Result of PCA gives two significant PCs (eigenvalues>1), which explains 78.64% of the total variance in the descriptors (48.64% and 30.00%, respectively). Figure 1 shows a score plot of the compounds in the training and test set. As can be seen from figure 1, there are no obvious clusters in the training and test set. Seven compounds, 14, 27, 40 and 41 in the training set and 2, 19 and 26 in the test set, were found to be farther away from the majority of compounds. There seems to be no evidence to suggest that these compounds are outliers. Therefore, they are retained in the data sets.

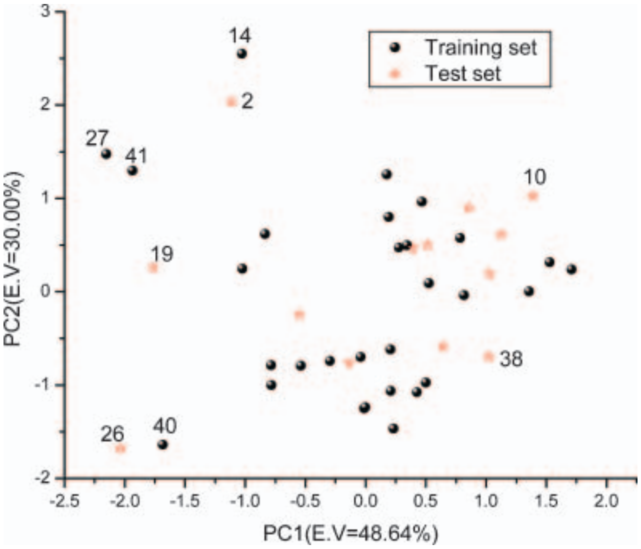

Figure 1. Principal component analysis of the structural descriptors for the data set.

When adding another descriptor did not significantly improve the statistics of a model, it was determined that the optimum subset size had been achieved. To avoid over-parameterization of the models, such as those which contain an excess of descriptors and are difficult to interpret in terms of physical interactions, an increase of the R^2 value of less than 0.02 was chosen as the breakpoint criterion. At the same time, the model with more than five parameters was not considered because the number of compounds should be five times at least the number of the parameters, generally [10].

After HM, five descriptors were selected to develop a linear model (shown in figures 2 and 3). The goodness-of-fit parameters for the obtained model are R^2=0.9754, F=182.44, s=3.0686, and R^2_{cv}=0.9615, respectively, which are better than corresponding values of the MLR model derived by García *et al.* [8], indicating that the training set is described relatively well by these five descriptors and that this model can be expected to be a better predictor of T_N. Using this model, the T_N values for the training set compounds were calculated. The estimated root mean square error (RMSE) and the

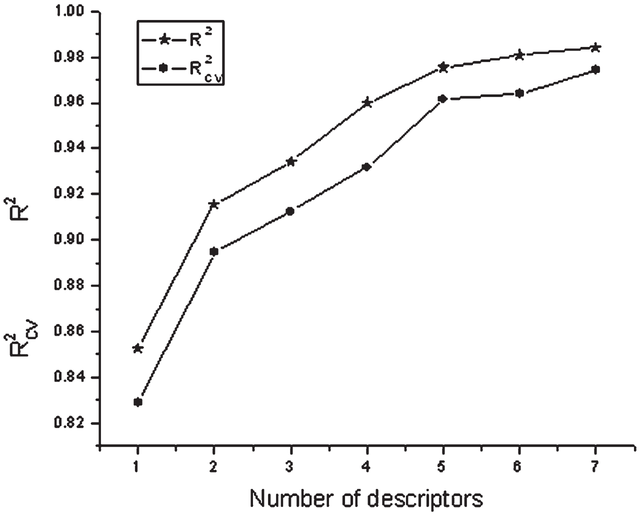

Figure 2. Influence of number of descriptors on R^2 and R^2_{cv} of the regression models.

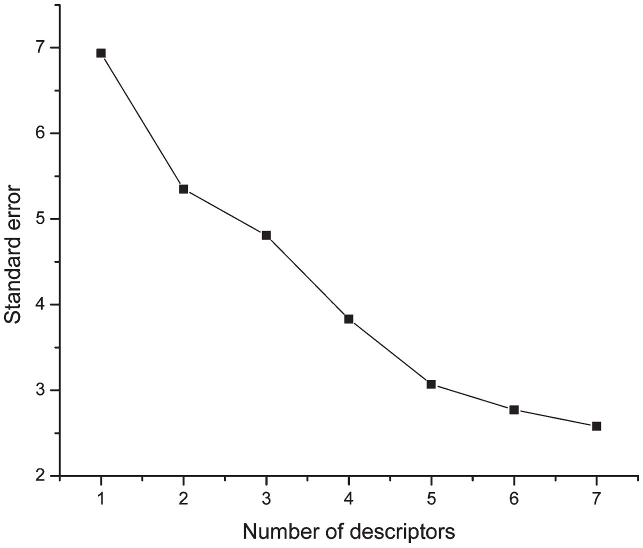

Figure 3. Influence of number of descriptors on the standard error (s) of the regression models.

absolute average relative deviation (AARD) were 2.7324 and 3.3508%, respectively, which are smaller than those of 5.5391 and 6.8322% in the work by García *et al.* [8] (our calculation).

In the present model, each compound is represented by a five-dimensional vector, components of which are the selected five descriptors. The names, chemical meaning and statistical characteristics of these descriptors were summarized in table 2, and the correlation matrix is shown in table 3. From table 3, it can be seen that all of the linear correlation coefficients of any two descriptors are less than 0.8, which means that the descriptors in the HM model are independent. The high correlation coefficient R^2 of 0.9754 indicates that a linear relationship exists between the structural information and the T_N of TLCs.

In the model obtained, RNO, the relative number of oxygen atoms in a molecule, reflects the constitutional factor influencing the generation of nematic phases. In the work of García *et al.* [8], the dummy descriptors L and R were employed to reflect the number of oxygen atoms present in the terminal chain. The principal moment of inertia of the molecules around the z-axis, I_C, is a geometrical descriptor. In rigid rotator approximation, I_C can be calculated as follows

Table 3. Correlation matrix of the selected five descriptors by HM.

	RNO	I_C	$E_{Max,\ S(C)}$	$E_{tot,(CO)}$	$E_{Min,C(CH)}$
RNO	1				
I_C	0.625	1			
$E_{Max,\ S(C)}$	0.440	0.240	1		
$E_{tot,(CO)}$	−0.442	−0.164	−0.755	1	
$E_{Min,C(CH)}$	0.454	−0.065	0.062	−0.079	1

$$I_C = \sum_i m_i r_{iz}^2, \tag{1}$$

where m_i are the atomic masses and r_{iz} denote the distance of ith atomic nucleus from the main rotational axis z of the molecule [16]. This descriptor characterizes the mass distribution in the molecule and represents the steric characteristic influencing the nematic phase phenomena. $E_{Max,\ S(C)}$ is the maximum atomic state energy for a C atom, $E_{Min,C(CH)}$ is the minimum coulombic interaction for a C–H bond and $E_{tot,(CO)}$ is the maximum total interaction for a C–O bond. The above three descriptors belong to quantum mechanical energy-related descriptors, which indicate the importance of the intramolecular electronic effects on the intermolecular electrostatic interactions of a molecule in determining the T_N of TLCs. According to the values of t-test, which reflects the significance of the parameter within a particular model (table 3), I_C is the most important parameter in this model and this indicates that steric characteristics play a significant role in the generation of nematic phases of TLCs. This agrees well with a general point of view that the observed phase is mainly determined by the geometrical shape of the molecule.

The model was then applied to the test set to evaluate its prediction capability. This model gave a predictive correlation coefficient, R^2, of 0.9216 for the test set. The predicted RMSE and AARD were 6.3654, and 9.2017%, respectively, which are better that those of 8.0791 and 11.5346% of MLR model by García *et al.* [8] (our

Table 2. Descriptors, coefficients, standard error, and t-test for the linear model by HM.

Descriptor	Chemical meaning	X	DX	t-test
	Intercept	−6.59E+04	9.50E+03	−6.9352
I_C	Moment of inertia C	−3.52E+04	1.76E+03	−19.9549
RNO	Relative number of O atoms	1.13E+03	9.37E+01	12.0163
$E_{Max,\ S(C)}$	Max. atomic state energy for a C atom	5.66E+02	8.15E+01	6.9513
$E_{Min,C(CH)}$	Min. coulombic interaction for a C–H bond	6.13E+02	1.10E+02	5.5534
$E_{tot,(CO)}$	Max. total interaction for a C–O bond	2.05E+02	5.18E+01	3.9642

calculation), confirming the predictive capability of the HM model we proposed.

The predicted results for the test set as well as calculated values for training set are listed in table 1 and shown in figure 4. From figure 4, it can be seen that several points show relatively large scatter around the straight line. For the training set, 93.10% of the compounds show absolute relative errors less than 5%, but two compounds, 4 and 15, show large relative errors of −14.68% and 21.93%, respectively. Within the test set, 61.53% of the compounds have absolute relative errors of less than 5%, and 69.23% show absolute relative errors less than 9%, with compounds 8, 13, 17 and 19 showing large relative errors of −12.23%, −18.76%, 18.44% and 41.16%, respectively. For MLR model derived by García *et al.* [8], 55.17% of the compounds in the training set show absolute relative errors less than 5% and 79.31% have absolute relative errors of less than 10%, with molecule 15 showing a large relative error of 47.78%. In the case of the test set in the work of García *et al.* [8], only 15.38% compounds show absolute relative errors less than 5% and 53.85% show absolute relative errors less than 10%, respectively, with molecule 17 having a large relative error of 33.78%.

Based on the above observations, it can be easily seen that the predicted results obtained using the HM model are more accurate than those of MLR model derived by García *et al.* [8]. In addition, the HM model is based on five descriptors that are independent of each other (as shown in the correlation matrix in table 3). As for the MLR model derived by García *et al.* [8], nine descriptors were used, some of which are high inter-correlated (>0.95) with one or two remainders (*G* and *H*, *V* and *S*, O and *S*), leading to redundancy of information. In this sense, the HM model we propose is easier to interpret, more reliable and more accurate than the MLR model.

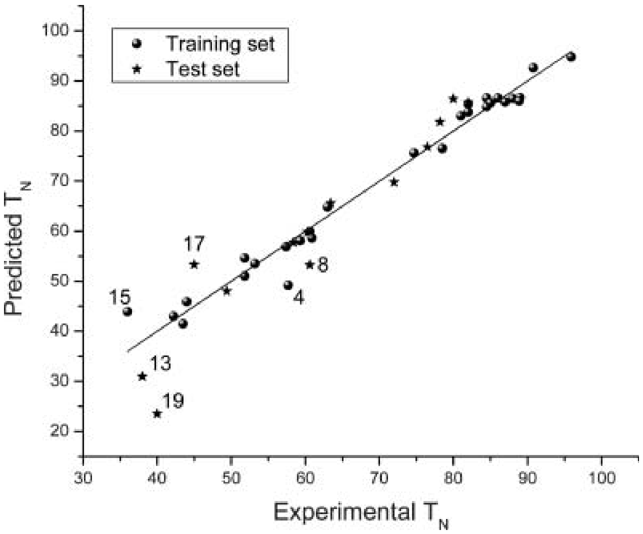

Figure 4. Plot of predicted vs. experimental transition temperature for the training and test set by heuristic method.

Further investigation on the results indicates that compounds 17 and 19 likely to exert a very large leverage on the statistical parameters for the test set. When compound 19 was discarded from the test set, an improvement in the predictive parameters can be observed, with R^2=0.9304, RMSE=4.4340 and AARD=6.0350%. In the case of discarding compound 17, the predictive parameters change to R^2=0.9657, RMSE=5.9312 and AARD=7.7776%. Initially, we suppose this may be due to the structural difference of these two compounds from the compounds in the training set. From the above discussion, it can be concluded I_C is the most important structural factor in the generation of nematic phases of TLCs. Therefore, we performed a careful study on the distribution of I_C for each molecule and the result is shown in figure 5. From figure 5, it can be seen that the value for molecule 19 is further away from those of the rest of the compounds. It is likely that this difference leads to the poor prediction for this compound. However, no obvious difference can be observed after carefully reviewing both the I_C distribution and the first two principal components for the compound 17. So, the reason for the poor prediction for compound 17 still remains unknown.

4. Conclusion

With molecular descriptors generated using the CODESSA program, a quantitative structure–property

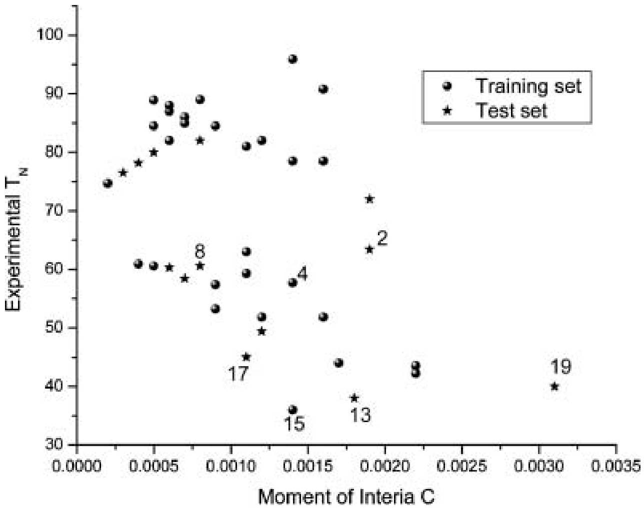

Figure 5. Distribution of I_C for different compounds.

relationship model was developed successfully to predict the nematic transition temperature of 42 homologous series of thermotropic liquid crystal molecules. The five descriptors selected by the heuristic method available in CODESSA reflect information affecting the generation of nematic phases, i.e. the constitutional, steric and electronic characteristics. The model should be able to predict the nematic transition temperature of unknown or unavailable compounds of this class, as indicated by the high correlation coefficient, small root mean squared error and absolute average relative deviation of the test set. Moreover, the results provide a simple, practical and effective method prediction of the nematic transition temperature of thermotropic liquid crystals.

Acknowledgement

We are grateful to Dr. Lin-Cheng Zhou for valuable discussion.

References

[1] P.J. Collings, M. Hird. *Introduction to Liquid Crystals – Chemistry and Physics*. Taylor and Francis, (1997).

[2] C.P. Fan, M.J. Stephen. *Phys. Rev. Lett.*, **25**, 501 (1970).

[3] S. Marčelja. *J. chem. Phys.*, **60**, 3599 (1974).

[4] R.B. Tao, P. Sheng, Z.F. Lin. *Phys. Rev. Lett.*, **70**, 1271 (1993).

[5] A.R. Katrizky, R. Jain, A. Lomaka, R. Petrukhin, M. Karelson, A.E. Visser, R.D. Rogers. *J. chem. Inf. Comput. Sci.*, **42**, 225 (2002).

[6] X.J. Yao, X.Y. Zhang, R.S. Zhang, M.C. Liu, Z.D. Hu, B.T. Fan. *Talanta*, **57**, 297 (2002).

[7] H.X. Liu, X.J. Yao, R.S. Zhang, M.C. Liu, Z.D. Hu, B.T. Fan. *J. phys. Chem. B*, **109**, 20565 (2005).

[8] M.V. García, R.N.G. Parra, A.M. Richa, J. Robles. *J. Mol. Struct. (Theochem)*, **727**, 63 (2005).

[9] M.T.D. Cronin, T.W. Schultz. *J. Mol. Struct. (Theochem)*, **622**, 39 (2003).

[10] J.G. Topliss, R.J. Costello. *J. med. Chem.*, **15**, 1066 (1972).

[11] A.R. Katritzky, V.S. Lobanov, M. Karelson. *CODESSA Version 2.0 Reference Manual* (1995–1997).

[12] *HyperChem, Release 6.0 for Windows*. Hypercube, Inc (2000).

[13] J.P.P. Stewart. *MOPAC 6.0*. Quantum Chemistry Program Exchange (QCPE), Program 455, Indiana University, Bloomington, IN (1989).

[14] A.R. Katritzky, R. Petrukhin, R. Jain, M. Karelson. *J. chem. Inf. Comput. Sci.*, **41**, 1521 (2001).

[15] A.R. Katritzky, S. Perumal, R. Petrukhin, E. Kleinpeter. *J. chem. Inf. Comput. Sci.*, **41**, 569 (2001).

[16] *Handbook of Chemistry and Physics*. CRC Press, Cleveland, OH (1974).