ISSN 1070-3632, Russian Journal of General Chemistry, 2006, Vol. 76, No. 4, pp. 615–620. © Pleiades Publishing, Inc., 2006.
Original Russian Text © R.G. Shestakova, T.R. Prosochkina, E.F. Tokunova, L.A. Tyurina, E.A. Kantor, 2006, published in Zhurnal Obshchei Khimii, 2006, Vol. 76, No. 4, pp. 648–653.

Structure–Mesomorphic Activity Relationship in the Series of Nitrogen-containing Heterocyclic Compounds

R. G. Shestakova, T. R. Prosochkina, E. F. Tokunova, L. A. Tyurina, and E. A. Kantor

Ufa State Technical University of Oil, Ufa, ul. Kosmonavtov 1, Bashkortostan, 540062 Russia
e-mail: physics@oil.rb.ru

Received January 28, 2005

Abstract — The structure–mesomorphic activity relationship was studied using the SARD (Structure Activity Relationship and Design) system for 300 nitrogen-containing heterocyclic compounds of which 178 compounds form a liquid crystal mesophase. Structural fragments exerting positive and negative effects on the manifestation of mesomorphic properties were revealed.
DOI: 10.1134/S1070363206040190

Revealing relationships between the structure and various properties of organic substances (including mesomorphic properties) is one of the basic problems of the modern organic chemistry. On the other hand, the search for compounds with mesomorphic activity involves screening of a large number of potentially active analogs. Today systematization of the available experimental data allows prediction of the mesomorphic activity. The effect of side, terminal, bridging, and cyclic substituents on the mesomorphic properties of various classes of substances is considered; however, for heterocyclic compounds it is difficult to correctly distinguish various effects of substitution even on a qualitative level [1]. Wide prospects for prediction of the mesomorphism are opened by qualitative mathematical methods SAR (Structure–Activity Relationship) and SPR (Structure–Property Relationship) and the corresponding quantitative methods QSAR and QSPR [2]. These methods are based on the pattern recognition, game, graph, and neural network theories, on elements of correlation and regression analyses, etc. [3, 4].

As a continuation of our previous studies [5], we made here an attempt to reveal the most efficient combinations of structural fragments responsible for the mesomorphic activity in molecules of nitrogen-containing heterocyclic compounds: pyrimidine, pyrazine, pyridazine, tetrazine derivatives, etc.

Analysis of the structure–mesomorphic activity relationship was made using the SARD system, which is a particular case of the SAR method and has been developed for estimating the effect of various structural parameters on the properties of chemical compounds and for constructing mathematical recognition models [6, 7]. The SARD program package includes the moduli of molecular design and prediction of heteroorganic compounds with preset properties. The SARD system was successfully used previously for predicting ~40 different types of activity: herbicidal, broncholytic, antiherpetic, toxicological, etc. A total of ~100 different classes of compounds were examined with this system to assess the probability of manifestation of the given type of activity [8]. By the mesomorphic activity is meant the probability of formation of a liquid crystal mesophase by a substance in the given temperature range.

The nitrogen-containing heterocyclic compounds under consideration can exhibit simultaneously both smectic and nematic properties. Therefore, to examine the structure–mesomorphic activity relationship, we formed two training sets, each containing two groups of substances with alternative properties. The first training set consists of 98 substances exhibiting smectic activity in a fairly wide temperature interval from 13 to 415°C and 122 substances showing no mesomorphic activity; the second training set consists of 80 substances exhibiting nematic activity in the temperature interval from 33 to 415°C and 122 substances exhibiting no mesomorphism [9–12].

The structure–mesomorphic activity model is constructed as follows. A sample of compounds with known parameters of the mesomorphic activity is subdivided into training and check samples. For each structure of the training sample, the SARD program algorithm automatically disintegrates the input structural formulas into fragments. In so doing, the space of the initial descriptors (separate atoms, atomic groups, functional groups, cyclic systems) is formed.

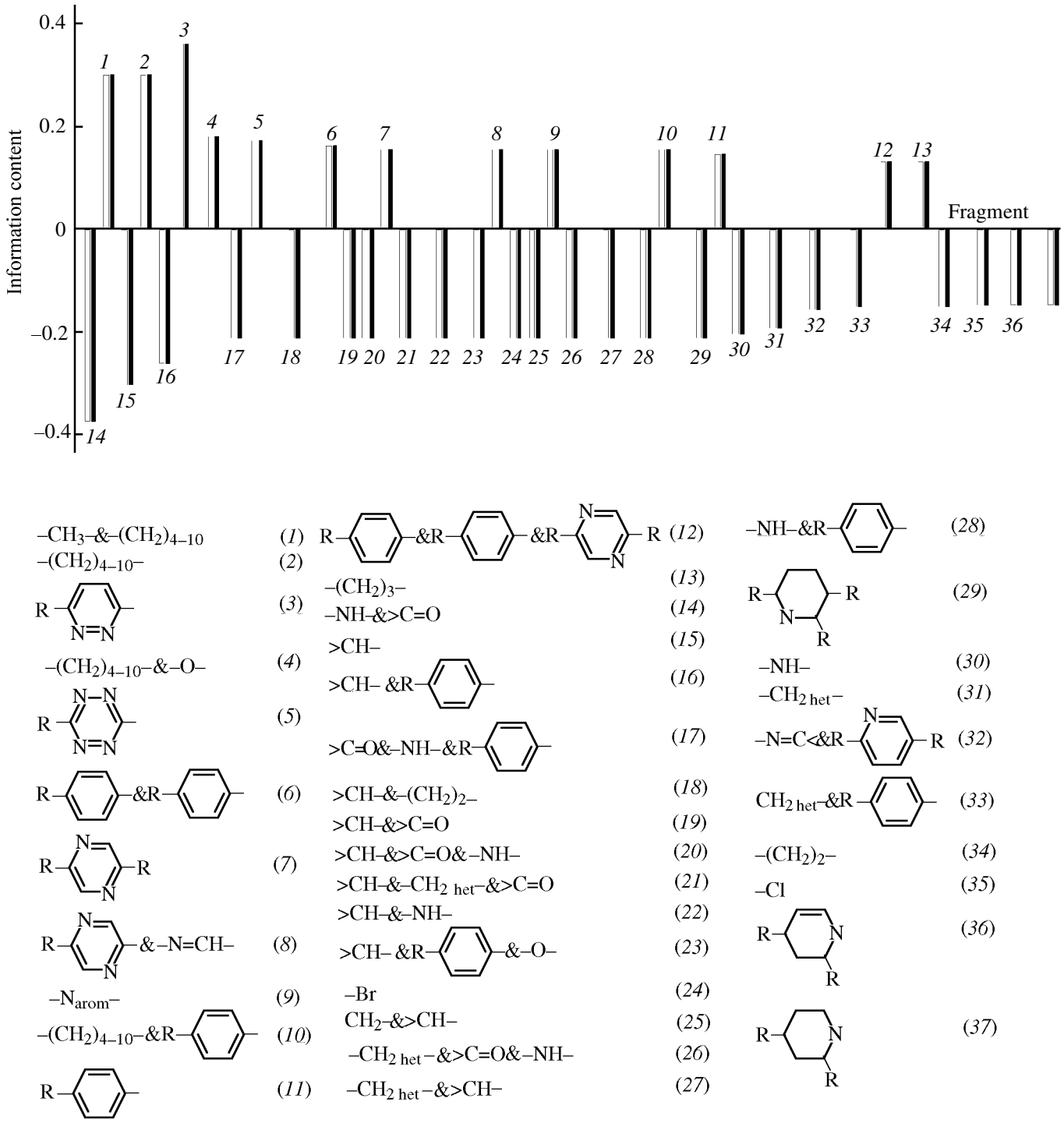


Fig. 1. Characteristics of the decisive set for the mesomorphic activity of nitrogen-containing heterocyclic compounds; & denotes conjunctive combination of fragments.

In the second step, the permanent, weakly variable, and strongly correlated descriptors are excluded. With the remaining descriptors, a series of regression models are constructed, the best model is chosen on the basis of statistical criteria, and its predicting power is checked using a check sample consisting of compounds that were not used in constructing the model.

As a result of a computer analysis, we assessed the effects of various structural elements on the meso-

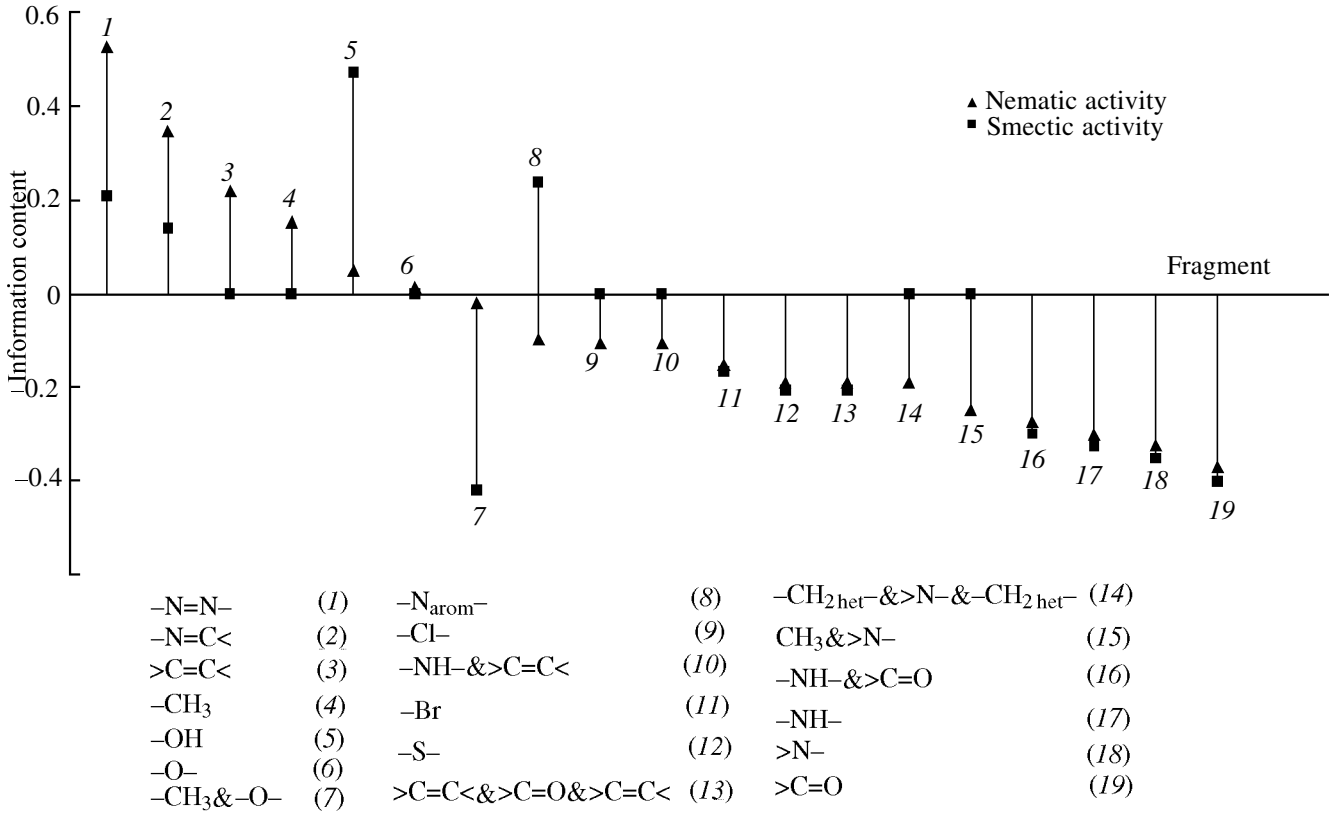


Fig. 2. Effect of fragments on the nematic and smectic activitites.

morphic behavior and constructed a mathematical model for predicting mesomorphic properties, $A = F(S)$, where A is the kind of mesomorphic activity and S is a complex of various combinations of structural fragments (decisive set of characteristics).

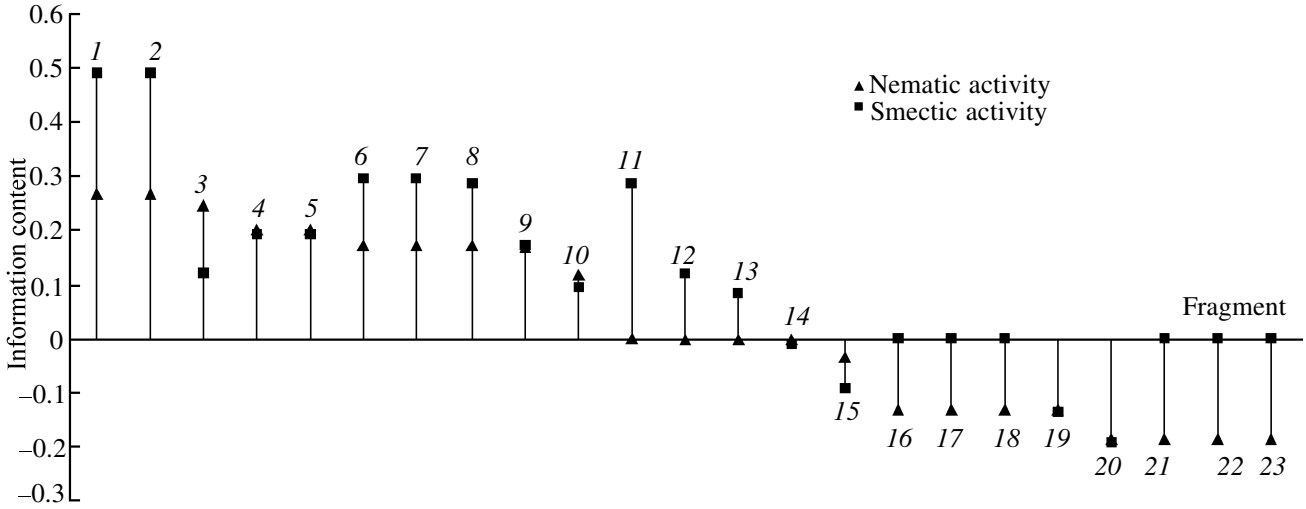
The effect of fragments on the mesomorphic activity was evaluated by the qualitative correlation coefficient (estimation of information content). The information content was estimated by the Juhl correlation coefficient of qualitative characteristics, varying from –1 to +1. The higher the absolute value of the information content, the higher the probability of the effect of this characteristic on the manifestation of the property under consideration. The sign of the information content corresponds to the positive or negative effect of a given characteristic on a given kind of mesomorphic activity.

Our study shows that the most informative from the viewpoint of the mesomorphic activity are alkyl and alkoxyl fragments, aromatic rings (including those containing N_{arom} atoms), and their combinations (Fig. 1). The negative information content is characteristic of the >CH– group, simultaneously present –NH– and >C=O groups, and a combination of the >CH– group with a substituted aromatic ring (Fig. 1).

Estimation of the information content of particular structural fragments from the viewpoint of the mesomorphic activity shows that the strongest effect on the manifestation of smectic properties is exerted by the hydroxy group –OH, nitrogen in the aromatic ring (N_{arom}), and azo group –N=N– (Fig. 2). The negative effect on the manifestation of the smectic activity is exerted by a combination of the terminal methyl group $–CH_3$ with the bridging oxygen –O– and by the fragments >C=O, –N=C<, >N–, and –NH– (Fig. 2).

The highest nematic activity is exhibited by compounds containing azo group –N=N–, terminal methyl groups $–CH_3$, hydroxy groups –OH, and bridging oxygen atom –O– (Fig. 2). The strongest negative effect on the manifestation of the nematic activity is exerted by the carbonyl group >C=O, tertiary nitrogen atom >N–, –NH– group, and combination of –NH– and >C=O groups (Fig. 2).

The effect of individual characteristics on the mesomorphic activity depends on surrounding fragments; therefore, when using data on individual char-



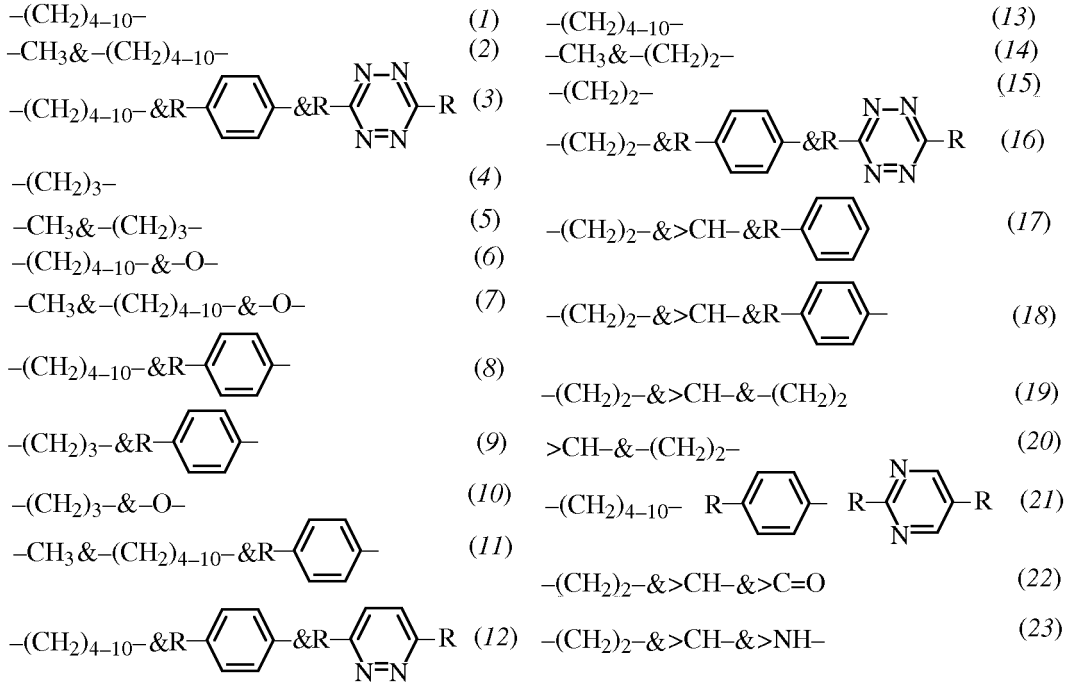


Fig. 3. Effect of fragments containing the CH_2 group on the nematic and smectic activities.

acteristics, it is necessary to consider the surrounding; it is appropriate to consider larger fragments.

Comparison of the estimated information content of the same characteristics with respect to two types of mesomorphic activity, smectic and nematic, shows that a mainly unambiguous effect (positive or negative) is exerted by the azo (–N=N–), azomethine (–CH=N–), and hydroxy groups (–OH), positively affecting the manifestation of the mesomorphic activity, and by the carbonyl group (>C=O), tertiary nitrogen atom (>N–), and amino group (–NH–), including their combinations, whose effect is negative (Fig. 2). Estimation of the information content of alkyl substituents and their conjunctions with cyclic fragments showed that long alkyl and alkoxyl substituents $–(CH_2)_{4–10}–$ and $–(CH_2)_{4–10}$&(–O–) in combination with the terminal $–CH_3$ group or benzene ring make the largest contribution to manifestation of the smectic activity (Fig. 3). The negative contribution is rather untypical of alkyl groups. In most cases, the smectic activity is not manifested with short alkyl fragments in combination with the >CH– group, namely: >CH–&$–(CH_2)_2–$, $–CH_2–$&>CH–&$–(CH_2)_2–$ (Fig. 3).

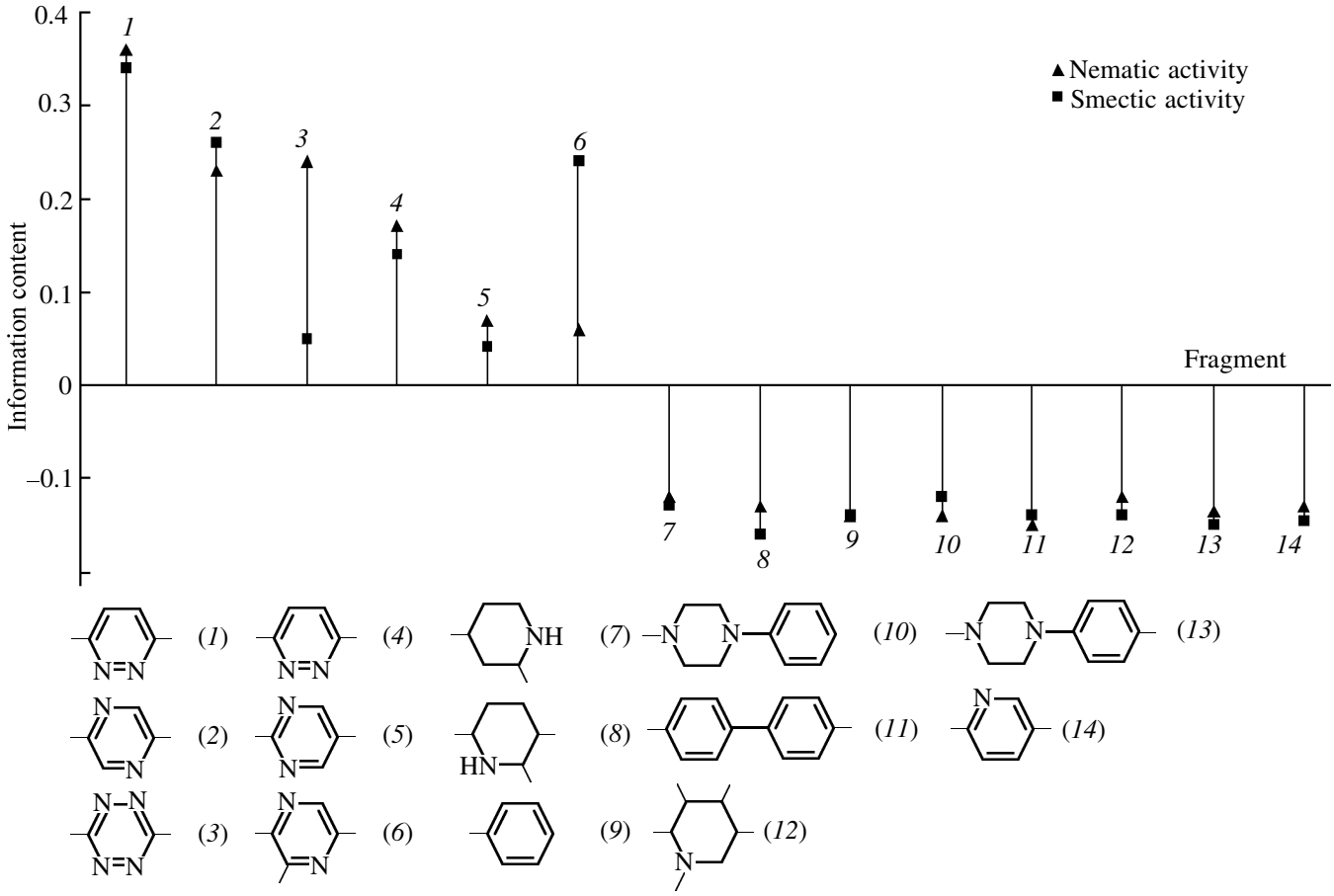


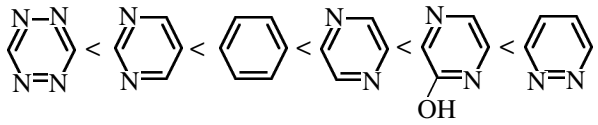
Fig. 4. Effect of cyclic fragments on the nematic and smectic activities.

For the nematic type of activity, alkyl substituents also exert a positive effect, but their information content is considerably lower (Fig. 3).

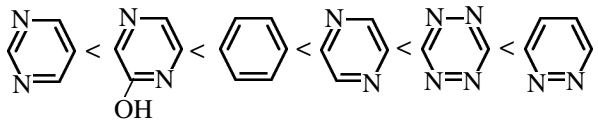
Thus, terminal alkyl and alkoxy substituents favor manifestation of smectic properties to a greater extent compared to nematic properties. This is confirmed by the experimental data [13].

Cyclic fragments, as a rule, exert a positive effect on the mesomorphic activity (Fig. 4). The cyclic fragments can be ranked in the following order with respect to the estimated information content:

Smectics



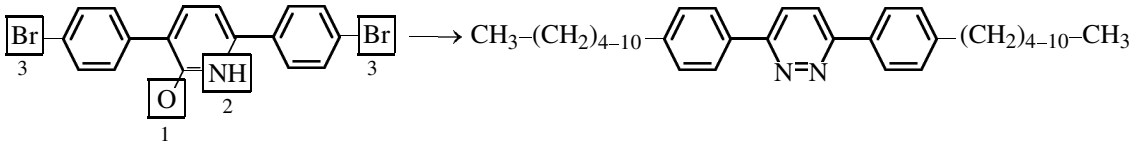
Nematics



All the compounds under consideration were classed using the constructed predictive model by two algorithms of the pattern recognition theory: "geometric approach" based on estimation of the distance (in terms of Euclidean geometry) between a given substance and activity standards and "voting" based on counting the numbers of positive and negative characteristics in the structure of a compound. The model is considered as adequate and the prediction results, as reliable, if the recognition level exceeds 70% [2]. The recognition of the training structures in the first set is 90 and 82% by the geometric approach (active and inactive compounds, respectively) and 93 and 77% by voting; for the second set, the respective quantities are 81 and 88% by the geometric approach and 71 and 76% by voting. Thus, the recognition level of the training sample is satisfactory, and the mathematical model can be used for predicting the mesomorphic activity of nitrogen-containing heterocyclic compounds. In so doing, we must pass the following steps: choice of a base structure for modification, determination of fragments that should be replaced in this structure, and replacement of these fragments. Usually the possible alternatives are examined using

the available theoretical concepts, experimental data, and synthetic accessibility.

An example of the modification of the base structure is given below:



The order of replacement of fragments 1, 2, and 3 was set on the basis of estimated contributions of these fragments to the mesomorphic activity. The new structure was generated by replacement of the >C=O and –NH– groups by the N_{arom} atoms and of the Br atoms by the $–(CH_2)_{4–10}–CH_3$ groups. The structure obtained, according to the experimental data [12], exhibits pronounced mesomorphic activity, which confirms the correctness and reliability of the recognition model constructed.

REFERENCES

1. Aver'yanov, E.M., *Stericheskie effekty zamestitelei i mezomorfizm* (Steric Effects of Substituents and Mesomorphism), Novosibirsk: Nauka, 2004.
2. Kadyrov, Ch.Sh., Tyurina, L.A., Simonov, V.D., and Semenov, V.A., *Mashinnyi poisk khimicheskikh preparatov s zadannymi svoistvami* (Computer Search for Chemical Substances with Preset Properties), Tashkent: Fan, 1989.
3. Gal'bershtam, N.M., Baskin, I.I., Palyulin, V.A., and Zefirov, N.S., *Usp. Khim.,* 2003, vol. 72, no. 7, p. 706.
4. Palyulin, V.A., Baskin, I.I., Radchenko, E.V., and Zefirov, N.S., *Bashk. Khim. Zh.,* 1997, vol. 4, no. 4, p. 46.
5. Prosochkina, T.R., Tyurina, L.A., Kantor, E.A., Murza, M.M., and Shestakova, R.G., *Russ. J. Gen. Chem.,* 2002, vol. 72, no. 9, p. 1444.
6. Tyurina, L.A., Kadyrov, Ch.Sh., and Simonov, V.D., *Itogi Nauki Tekh., Ser.: Org. Khim.,* 1989.
7. Kirlan, V.V., *Cand. Sci. (Chem.) Dissertation,* Ufa, 2003.
8. Karimova, F.S., Tyurina, L.A., Kolbin, A.M., Solominova, T.S., Gil'mkhanova, V.T., and Shagaleeva, Z.F., in *Trudy I Mezhdunarodnoi konferentsii "Khimiya i biologicheskaya aktivnost' azotistykh geterotsiklov i alkaloidov"* (Proc. I Int. Conf. "Chemistry and Biological Activity of Nitrogen-Containing Heterocycles and Alkaloids"), Moscow, 2001, p. 132.
9. Schubert, H., *Wiss. Z. Univ. Halle, Math. Naturwiss. R.,* 1970, vol. 19, no. 1, p. 1.
10. Ahmed, N.A. and Kitaigorodsky, A.I., *Acta Crystal logr., Sect. B,* 1972, vol. 28, no. 3, p. 739.
11. Laing, M. and Sommerville, P., *Acta Crystallogr., Sect. B,* 1976, vol. 32, no. 10, p. 2764.
12. Demus, D., Demus, H., and Zaschke, H., *Flüssige Kristalle in Tabellen,* Leipzig: Grundstoffindustrie, 1976.
13. Amerik, Yu.B. and Krentsel', B.A., *Khimiya zhidkikh kristallov i mezomorfnykh polimernykh sistem* (Chemistry of Liquid Crystals and Mesomorphic Polymer Systems), Moscow: Nauka, 1981.