crystals

MDPI

Article

Neuro-Evolutive Modeling of Transition Temperatures for Five-Ring Bent-Core Molecules Derived from Resorcinol

Elena Niculina Drăgoi [1], Irina Cârlescu [1], Răzvan Puf [2], Tudor Vasiliu [2] and Elena-Luiza Epure [1,*]

[1] "Cristofor Simionescu" Faculty of Chemical Engineering and Environmental Protection, "Gheorghe Asachi" Technical University, Bd. Prof. Dimitrie Mangeron, No. 73, 700050 Iaşi, Romania
[2] Centre of Advanced Research in Bionanoconjugates and Biopolymers, Romanian Academy Petru Poni (PP) Institute of Macromolecular Chemistry, 700487 Iasi, Romania
* Correspondence: lepure@tuiasi.ro

Abstract: Determining the phase transition temperature of different types of liquid crystals based on their structural parameters is a complex problem. The experimental work might be eliminated or reduced if prediction strategies could effectively anticipate the behavior of liquid crystalline systems. Neuro-evolutive modeling based on artificial neural networks (ANN) and a differential evolution (DE) algorithm was applied to predict the phase transition temperatures of bent-core molecules based on their resorcinol core. By these means, structural parameters such as the nature of the linking groups, the position, size and number of lateral substituents on the central core or calamitic wings and the length of the terminal chains were taken into account as factors that influence the liquid crystalline properties. A number of 172 bent-core compounds with symmetrical calamitic wings were selected from the literature. All corresponding structures were fully optimized using the DFT, and the molecular descriptors were calculated afterward. In the first step, the ANN-DE approach predicted the mesophase presence for the analyzed compounds. Next, ANN models were determined to predict the transition temperatures and whether or not the bent-core compounds were mesogenic. Simple structural, thermophysical and electronic structure descriptors were considered as inputs in the dataset. As a result, the models determined for each individual temperature have an R^2 that varied from 0.89 to 0.98, indicating their capability to estimate the transition temperatures for the selected compounds. Moreover, the impact analysis of the inputs on the predicted temperatures showed that, in most cases, the presence or not of liquid crystalline properties represents the most influential feature.

Keywords: artificial neural networks; differential evolution; bent-core molecules; liquid crystals; melting; isotropization; crystallization; DFT

Citation: Drăgoi, E.N.; Cârlescu, I.; Puf, R.; Vasiliu, T.; Epure, E.-L. Neuro-Evolutive Modeling of Transition Temperatures for Five-Ring Bent-Core Molecules Derived from Resorcinol. *Crystals* **2023**, *13*, 583. https://doi.org/10.3390/cryst13040583

Academic Editors: Ana Almeida and Pedro De Almeida

Received: 25 February 2023
Revised: 21 March 2023
Accepted: 23 March 2023
Published: 29 March 2023

1. Introduction

Displaying mesophases with distinctive characteristics is a well-known property of bent-core mesogens. Bent-core liquid crystals represent an essential class of achiral bent-core (BC) compounds packed into bent directions, parallel aligned to each layer. The layers present spontaneous polarisation, and the molecules can switch, a property especially useful for electro-optical devices such as liquid crystal displays or tactile devices. After discovering these remarkable materials' electro-optical, ferroelectric and anti-ferroelectric responses [1], the technological interest in BC liquid crystals has dramatically increased. So far, many banana-shaped molecules with different structural combinations have been synthesized to understand the relationship between chemical structure and mesomorphic behavior [1–6]. It has been found that structural parameters influencing the liquid crystalline properties vary according to (1) the type of central core, (2) the overall length of the molecular structure, (3) the symmetrical or asymmetrical calamitic wings, (4) the nature, position and orientation of the linking groups, (5) the position, size, number and electronic (polarisable) properties of

lateral substituents on the central core or calamitic wings and (6) the length of the terminal chains. Furthermore, the presence of electrical switching properties in bent-core systems depends not only on the shape of the molecules but also on other factors, such as local polar moments and the extent of conjugation that affect the intermolecular interactions [7–10].

The following units have been reported as the central part of the bent-core molecules: 1,3-disubstituted benzene, 2,7-disubstituted naphthalene, 1,3-disubstituted biphenyl, or heterocyclic units [11]. Of the 1,3-disubstituted benzene derivatives, resorcinol is the most widely used central core unit [1,6,12–17].

To induce a certain degree of flexibility into BC structure, the linking groups are inserted within the calamitic wings and vary from ester, azomethine, azo, double or triple bonds, respectively [1,7,18]. At the same time, the linking units promote the liquid crystalline behavior and prevent crystallization [11].

The introduction of different lateral substituents such as bromo [1,13], chloro [10,19–21], fluoro [16,22–31], cyano [2,20,32,33], nitro [10,27] or methyl groups [10,20,21] to the central core or the aromatic calamitic wings greatly influences the thermal and mesomorphic behavior. Generally, the presence of lateral substituents on the central core or wings of the mesogens lower transition temperatures, compared to their non-substituted analogues.

The development of smart materials based on liquid crystals with desired properties for a wide range of applications should be conducted more efficiently and planned, knowing that a little change in the molecular structure could drastically influence their physicochemical properties. Neural networks can be used to achieve this goal. Hence, a neuro-evolutive modeling procedure combining artificial neural networks (ANNs) with a differential evolution (DE) algorithm was applied to model the temperatures for the considered class of liquid crystals. The concept is similar to our previous work [34], where acceptable performance was obtained for some homopolymers' glass transition temperature (T_g) with saturated carbon chain backbone modeling using a neuro-evolutive approach, combining ANNs with bacterial foraging optimization. The ANNs were chosen because of their ability to model highly non-linear multiple input–multiple output relations. The flexibility and capability of ANNs to act as universal approximators allowed the application of this model for various types of use cases.

Although there are studies that use ANNs and other artificial intelligence techniques that model isotropization temperatures or liquid crystalline behavior of some bent-core liquid crystals (e.g., decision trees and multivariate adaptive regression splines) [35,36], nematic transition temperatures [37] or physical properties [38], to the authors' knowledge, the application of the neuro-evolutive technique for the determination of all phase transition temperatures (melting, isotropization, crystallization) has not been reported before. From the multitude of bio-inspired optimizers, the DE algorithm is a well-established approach that has proven its efficiency for various problems from different areas.

The high number of bent-core liquid crystal compounds reported so far motivated us to perform the current research and to determine the predictive model for transition temperatures. This might substantially contribute to the synthesis of compounds with targeted properties by avoiding the synthesis of non-mesomorphic compounds. This work represents a proof of concept that aims to demonstrate the potential of artificial intelligence techniques (ANNs and DE) in solving this critical issue. We focused on bent-core compounds based on resorcinol with symmetric arms. The bent-core compounds were chosen due to their electro-optical characteristics. The novelty of this work is supported by the prediction of liquid crystalline properties and the development of a series of neural models for all phase transitions of bent-core compounds based on resorcinol, regardless of their mesogenic characteristics.

This work is organized as follows. Section 2.1 presents the data gathering and analysis strategy, and details the structural, physical and energetic parameters considered inputs for the ANN-based models. Section 2.2 details the ANN-DE modeling methodology. Section 3.1 describes the main characteristics of the gathered dataset, and Section 3.2 focuses on the modeling and prediction strategy, with a detailed analysis of the impact of the considered

features on the model output (Section 3.3). Section 4 concludes the manuscript. A list of acronyms is presented at the end of the article.

2. Materials and Methods

2.1. Dataset

The modeling strategy considered in this work is data-dependent. Thus, the first step consists in creating and analyzing the database containing the compounds with targeted characteristics. The properties of the bent-core compounds are influenced by their structure; moreover, if they are liquid crystals, they can be either monotropic or enantiotropic. Unfortunately, non-mesomorphic bent-core compounds are rarely mentioned in the scientific literature because they are deemed irrelevant.

The resorcinol-based bent-core compounds that have reported experimental data were chosen to create the database (works [2,10,19–21,32,33,39–44], with details in the Supplementary Information Table S1). Subsequently, only the compounds with identical arms connected to the central core and 5 benzene rings in their structure were considered. The investigation was limited to pure compounds and not to mixtures. Since one of the objectives of the current study was to predict the mesophase properties based on the molecular structure, compounds that fulfill the previously mentioned selection criteria but do not present mesomorphic properties were also included in the database (Table S1).

The general molecular formula of bent-core derivative compounds is presented in Figure 1.

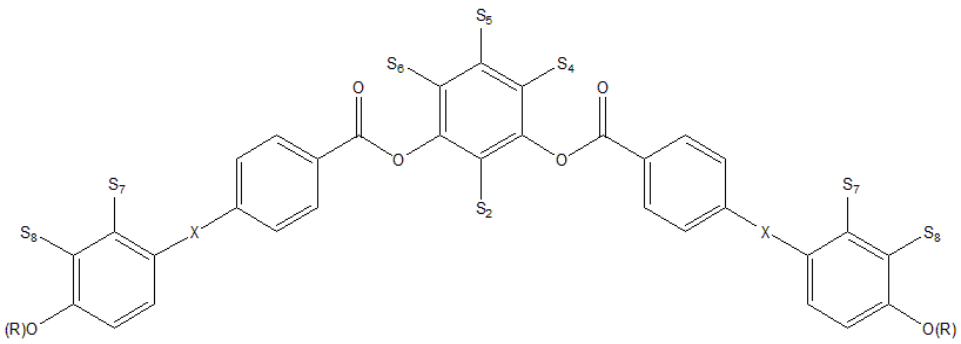


Figure 1. Schematic representation of the investigated bent-core compounds (X—linking group, R—terminal flexible chain). Substituents can be found in positions S2, S4, S5, S6, S7 and/or S8.

After data collection and analysis, the final dataset contains 172 bent-core compounds. The dataset consists of compounds with different linking group substituents on the central core and/or calamitic wings and with different types (alkyl or alkoxy group) and lengths of terminal chains. Linking group X is an azo (-N=N-), ester (-OC(=O)-) or imine (-N=CH-) group (Table S1). The substituents are, depending on the case, -F, -Cl, -Br, -CN, $-CH_3$, -OH or $-NO_2$, which can be found in the S2, S4, S5, S6, S7 and/or S8 positions.

The molecular structures corresponding to bent-core compounds in their most extensive form were optimized. The geometry optimization of the compounds was performed in the ground state at the density functional theory (DFT) level, using a B3LYP hybrid functional [45–47] and 6-31G basis set. All theoretical quantum–chemical computations were performed using the Gaussian 16 package. Finally, the descriptors were calculated using the Materials Studio package, version 4.0 [48].

2.2. Modeling

To construct models for all phase transitions of a BC compound based on resorcinol, whether or not it shows mesophase behavior, the first step of the developed strategy is determining the presence of liquid crystalline properties. In addition, if the mesophase property is identified, the neural model also determines the mono- or enantiotropic character. Finally, all these aspects are introduced into a new feature (named F21) that, along with

the other considered features (described in Section 2.1), represents the input for the models that predict the individual transition temperatures from one phase to another.

Since finding the optimal neural network is a problem-dependent complex task, applying neuro-evolution as a strategy to automatically determine the optimal parameters represents an excellent alternative solution to the classical manual topology identification. Neuro-evolution can evolve different ANN parameters (weights, topologies, activation functions or groups of learning rules) [49]. In this work, the DE algorithm is applied at the topology level to determine the number of hidden layers and neurons in each hidden layer for a fully connected Keras sequential model. DE is a metaheuristic based on the Darwinian principle of evolution [50]. It evolves a series of potential solutions (that are initially randomly generated) until a stop criterion is reached. The evolutionary process includes mutation (where a new population of mutants is created by applying the differential operator), crossover (where a new trial population is formed from individuals that combine characteristics from the corresponding parents belonging to the current and the mutated populations) and selection (where the best individuals are selected to create the new generation). Depending on the mutation and crossover operations and how the parameters that control them are set, DE has many variants. In this work, a DE/Best/2/Bin self-adaptive variant is used. This implies that the mutation strategy consists of 2 differential terms added to the so-far best individual in the population, the crossover version is binomial and the control parameters are set adaptively (by including them in the individuals). More details regarding DE variants can be found in [51].

Since DE only performs a topology determination in this work, the individuals that form the population contain the necessary encoded information (number of hidden layers and neurons in each hidden layer). The activation function for the neurons in the hidden layer is ReLU, while the output layer contains neurons with a linear activation function. The training is performed in batches using Adam optimizer [52] with a 0.05 learning rate. The loss function was selected as the mean squared error (MSE), and this metric is used to compute the fitness function that indicates the fit of a specific individual. About 70% of the data is randomly assigned to training and 30% to testing. From the training data, 20% is designated for model validation. Before training, the features were standardized by scaling them to the unit variance using the StandardScaler from the Scikit package. Thus, the entire implementation of the neuro-evolutive algorithm was performed in Python, supported by packages such as NumPy, Scikit, Keras and Pandas.

After identifying the best model for each considered case, the impact of the considered features on the model output was analyzed using the SHapley Additive exPlanations (SHAP) approach [53]. SHAP assigns to each feature of a black box model an importance value corresponding to a particular prediction. The resulting dataset (with the exact dimensions as the training subset) contains SHAP values. The sum of these SHAP values can be used to reproduce the model's predictions. Moreover, the SHAP values can reveal how the input variables influence the predictions at individual instances and across the entire subset.

3. Results and Discussions

3.1. Dataset Analysis

The features used in this work are simple ones. Even though many properties could significantly influence the phase transitions, the current work focuses on easy-to-identify parameters that do not need complex simulations and experiments.

With the help of the Material Studio software, a variety of structural (molecular weight, lengths, bending angle), thermophysical (van der Waals volume, enthalpy of formation) and electronic descriptors (HOMO-LUMO energies, dipole moment) of the minimum energy structures were obtained (Table 1). In addition, the type of linking group or topological information, such as substitution on the central core and/or wings, were also considered.

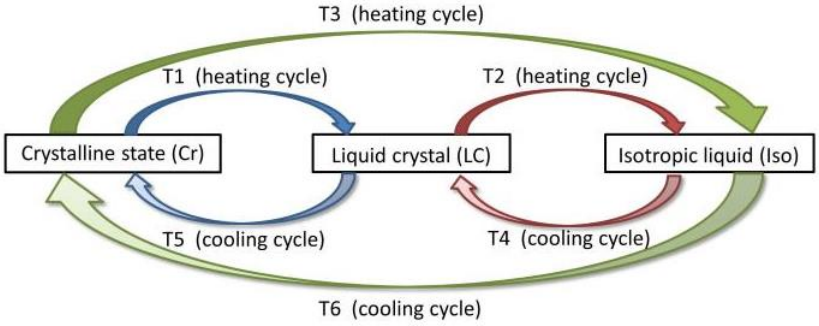
Table 1. List of the features included in ANN with the type of descriptor and a short description.

ANN Feature	Descriptor	Definition, Scope or the Value of the Descriptor
F0	substitution at S6	Depending on the type of atom or functional group present, the value of this feature is as follows: 1 (-H), 2 (-CN), 3 (-Br), 4 (-Cl), 5 (-F), 6 (-CH_3), 7 (-OH), 8 (-NO_2).
F1	substitution at S5	idem.
F2	substitution at S4	idem.
F3	substitution at S2	idem.
F4	substitution at S7	idem.
F5	substitution at S8	idem.
F6	types of linking group, X	The value of F6 is as follows: 1 (- N=N-), 2 (-OCO-), 3 (-N=CH-).
F7	arm's length (Å)	Length of the mesogen's wings in the most extended conformation with all-trans terminal alkyl chains.
F8	molecule length (Å)	Length of the long axe of mesogen, meaning the distance between the terminal C of each wing.
F9	bending angle (°)	The angle between the two wings of the molecule.
F10	van der Waals volume (A^3)	The sum of van der Waals atomic volumes, the real space occupied by the atoms.
F11	van der Waals surface area(A^2)	The contour obtained by the representation of the van der Waals radii of all covalently bound atoms and the overlap of these where it has happened.
F12	energy HOMO (eV)	The energy of the highest occupied molecular orbital.
F13	energy LUMO (eV)	The energy of the lowest unoccupied molecular orbital.
F14	dipole moment (Debye)	Total dipole moment.
F15	enthalpy of formation (kcal/mol)	The enthalpy change for the formation of 1 mol of a compound from its component elements.
F16	molecular weight (a.m.u.)	Size of the molecule.
F17	heat capacity (cal/K/mol) at 298 K	The amount of heat to be supplied to an object to produce a unit change in its temperature.
F18	entropy (cal/K/mol) at 298 K	A parameter that depends on the intrinsic energy, involving only contributions of the individual atoms in the molecule, ignoring interactions with atoms in other molecules [54].
F19	stiff segment length (Å)	Length of the mesogen's wing from the central core to the flexible terminal chain.
F20	flexible segment length (Å)	The length of the flexible terminal chain in the most extended conformation.

Both types of terminal chains (alkyl or alkoxy group) did not have specific descriptors as there are for the types and positions of the substituents. It was considered that terminal chains, through their type and size, will influence descriptors such as the length of the terminal chains, the length of the arm, the length of the flexible arm, the length of the molecule and its mass.

The competition between steric effects, excluded volume and interaction in or between layers leads to a particular ordering of the mesogens. A few reasons that influenced the choice of the descriptions used as an entering date in the ANN are as follows: kinetic stability is determined by the HOMO-LUMO energy gap; the enthalpy of formation indicates the stability of the isomer; the dipole moment is a crucial parameter for the stability of the mesophase; the length of the flexible terminal chains affects the organization and the fluidity of bent-core molecules in the mesophase or the spacing of consecutive smectic layers; and the molecular shape and, in some cases, the substituents influence the packing of the molecules, minimizing the free volume. As the length of the stiff segment to that of the flexible segment is a crucial parameter for rod-like mesogens, it was intended to know how these descriptors could affect the mesomorphism of bent-core molecules.

Scheme 1 presents the considered temperatures that were modeled using the ANN-DE strategy.



Scheme 1. The outcome temperatures (T1–T6) of the ANN. T1 or T3 are known as the melting point, T2 as the isotropization/clearing point and T5 or T6 as the crystallization point.

3.2. Modeling Results

After the dataset was gathered and analyzed, the ANN-DE methodology was applied to model and predict the phase transition temperatures. The data used in each modeling case varied depending on the available information. Thus, to select the exemplars used in the training and testing datasets, the compounds for which the modeled output was not present (N/A) were temporarily eliminated. For instance, there have been cases when the T1 temperatures for enantiomers were reported in the scientific literature but not the T2 temperatures. Therefore, exemplars were retained for the T1 model evaluation but removed for the T2 model evaluation. After that, the data was split into subsets and normalized, as described in Section 2.2.

Given that the objective was to predict phase transition temperatures (Scheme 1) without specifying whether or not mesophase properties are present, the first step of the modeling procedure involved identifying these properties. Thus, the first determined ANN was a multi-input–multi-output model that simultaneously predicted all T1-T6 temperatures (Scheme 1) based on which, through logical rules, the feature F21 was computed. This feature is 0 if the compound does not have liquid crystalline properties, 1 if it is a monotropic liquid crystal and 2 if it is an enantiotropic liquid crystal. Next, the neuronal models with a single output were built for the temperatures present in Scheme 1. In this case, the inputs for the ANN were those from Table 1, including the F21 one. Due to the stochastic nature of DE, it was executed 10 times for each case considered, and the solution with the highest fitness was chosen. The self-adaptive parameters are F (which controls the mutation phase) and Cr (which controls the crossover step), and both of them have values in the interval [0, 1]. The other parameters of the DE algorithm, population dimension and the number of generations, were set manually to 20 and 50, respectively.

A preliminary study indicated that F21 highly influences the outputs of the ANN models, and its elimination from the feature list raised the errors significantly. However, directly predicting F21 using the remaining features proved difficult, as the preliminary tests focusing on the direct application of different machine learning techniques (including ANNs, support vector machines and decision trees) provided unsatisfactory results. As such, the modeling and prediction of phase transition temperatures were performed in two stages using a series of ANN models linked, as presented in Figure 2. The first phase aims to predict F21 based on a general ANN model that focuses on simultaneously determining all T1-T6 temperatures. Based on these general predictions and using logical rules, F21 is predicted. In the second phase, the initial 21 features (F0–20) along with F21 are used to determine individual models for each temperature.

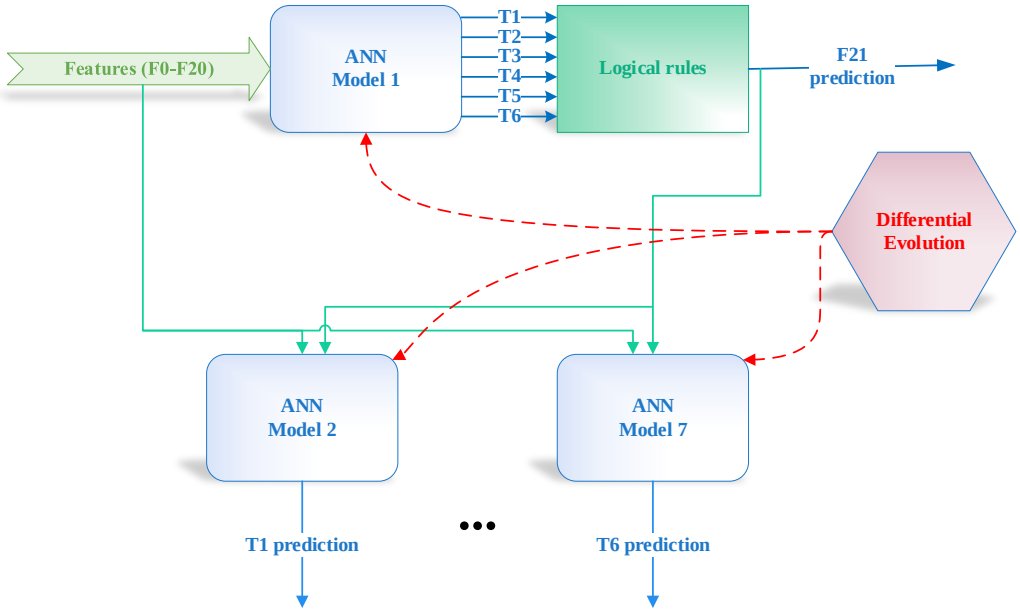


Figure 2. Modeling strategy for phase transition temperatures.

As previously mentioned, the modeling strategy's first step is identifying the lack or the presence of liquid crystalline properties corresponding to each compound, characterized by features (Features 0–20 from Table 1). To this mean, a multi-input–multi-output ANN model that estimates all six temperatures (Figure 2) is determined with the neuro-evolutive approach. Next, F21 is calculated based on the predictions generated and a set of logical rules. The best model selected in this case had two hidden layers with 9 and 20 neurons, respectively. The mean absolute error (MAE) varied between 9.7 and 13.3 with correlations from 0.81 to 0.96 in the training phase and between 22.02 and 32.24 with correlations between 0.15 and 0.83 in the testing phase. The lowest performance was obtained for T6, indicating that simultaneously considering all temperatures as outputs does not provide acceptable results in some cases. Next, the temperature predictions performed in this case were used in a logical rule to estimate F21 as follows: if T3 is 0 (tested as a lower value than 20 to include a high margin for the prediction errors), then the corresponding value of F21 is 2 (identifying the enantiotropic compounds with temperature of both cooling and heating); if T5 is 0 (tested similarly as in the case of T3) then the class is 0 (corresponding to the compounds without liquid crystalline properties); otherwise, F21 is 1 (corresponding to monotropic compounds with only cooling temperature). This strategy was applied instead of direct prediction due to the high errors introduced by the machine learning classifiers tested. The logical rule approach led to a misclassification rate of 4.3% in the training phase and 17.2% in the testing phase. In total, 8 exemplars from the 98 data points available (with complete values for all 6 temperatures analyzed) were wrongly identified. From those 8 exemplars, 6 correspond to class 2 that were improperly assigned to class 1.

In the next phase, the identified feature F21 is considered in determining individual ANN models for each temperature. Table 2 presents the main characteristics of the best models determined for each temperature, where the topology is represented using a notation in the form inp:H1:..Hn:out with inp indicating the number of inputs (corresponding to the model's features), Hi is the number of neurons in the hidden layer and out is the number of outputs. In Table 2, RMSE represents the root mean squared error, and MAPE represents the mean absolute percentage error. For the outputs where temperatures are not applicable (data set to 0), the MAPE has very high values. The reported MAPE is computed only for temperatures at which experimental values are stated (see previous explanations regarding the exclusion of N/A data). Lower values show better models for all three indicators, RMSE, MAE and MAPE (Equations (S1)–(S3) in Supplementary Information).

Table 2. The best ANN obtained for a single output modeling.

Output	Topology	RMSE Training	MAE Training	MAPE Training	RMSE Testing	MAE Testing	MAPE Testing
T1	22:15:5:8:1	6.3506	3.41059	0.047475	10.4202	6.38020	0.09367
T2	22:15:01	6.0332	2.76686	0.029141	13.9295	7.12744	0.08375
T3	22:17:8:1	5.4194	2.84321	0.053734	12.9596	5.48218	0.13090
T1 T4	22:15:01	6.5939	4.46069	0.041046	13.4473	13.02247	0.13327
T5	22:19:5:9:1	5.6258	3.84472	0.056967	15.2732	11.31098	0.11245
T6	22:04:01	11.6825	2.86585	0.070765	18.6110	8.071391	0.20796

Figure 3 shows the experimentally measured data versus the ANN predictions for the data in the testing phase. The statistical indicators show that the determined models perform well for all solutions. In terms of correlations, the predictions for T1 and T2 have the highest R^2 (>0.98), while the lowest was for T6 (R^2 = 0.89). Even for the single output models, T6 remains the most challenging temperature to model. This is because, for many compounds, T6 = 0 (is not applicable or corresponds to classes 1 and 2) (from 160 compounds analyzed, only 15 have values greater than 0). For the testing data, as Figure 3f shows, from 42 compounds with T6 = 0, for 6 cases, the predictions were greater than 20, indicating a high error. For the other compounds, no outliers were identified.

The explanation for this somewhat reduced ANNs performance for T3 and T6 is related to the fact that the compounds that exhibit these temperatures lack the characteristics of mesophase. The dataset used in this work is gathered from the available literature, where authors often do not mention the compounds that do not exhibit mesophases since they are considered irrelevant. On the other hand, the compounds with T3, T4 and T5 show monotropic behavior. This induced some difficulties for the neural network, given that some compounds present higher errors in the prediction phase.

3.3. Feature Analysis

After that, the SHAP method was applied for each identified model to explain how the predictions were made. Thus, the beeswarm plots were drawn and analyzed (Figures S1–S6, Supplementary Information). In a beeswarm plot, for each feature, every instance of the dataset appears as a colored point, red indicating relatively high values and blue relatively low values. The values falling inside the positive range indicate an upward tendency compared to the mean. Thus, the points on the scale with negative values indicate that the feature will provide a forecast that is less than the mean. Finally, the features are ranked from top to bottom based on the dataset's mean absolute SHAP values.

The SHAP data for T1 (Figure S1, Supplementary Information) showed that the input with the most significant influence on the model is Feature 21. This surprisingly high impact of Feature 21 indicates that the model efficiently correlated the rise of T1 for enantiotropic compounds compared with other types. Features 9 and 19 have opposite effects, so high values for Feature 9 (bending angle) will predict lower values for T1, while high values for Feature 19 (stiff segment length) will predict higher values for T1 than the absolute mean value. Scrolling through the SHAP representation from top to bottom, it can be observed that the importance of the descriptors in determining T1 decreases in the order Feature 9, Feature 19, Feature 4, Feature 12, Feature 13, Feature 0, Feature 2 and so on. As a general trend, T1 is influenced by bending angle, substitution in positions S7 and S4, and to a lesser extent by that in S8. In addition, high importance of the energy factors in establishing the mathematical model of T1 is observed. Focusing on Feature 4 (substitution in S7), it can be seen that large values, i.e., substitution with OH, will result in large values of T1. In contrast, small values, i.e., lack of a substituent, will decrease transition temperature T1.

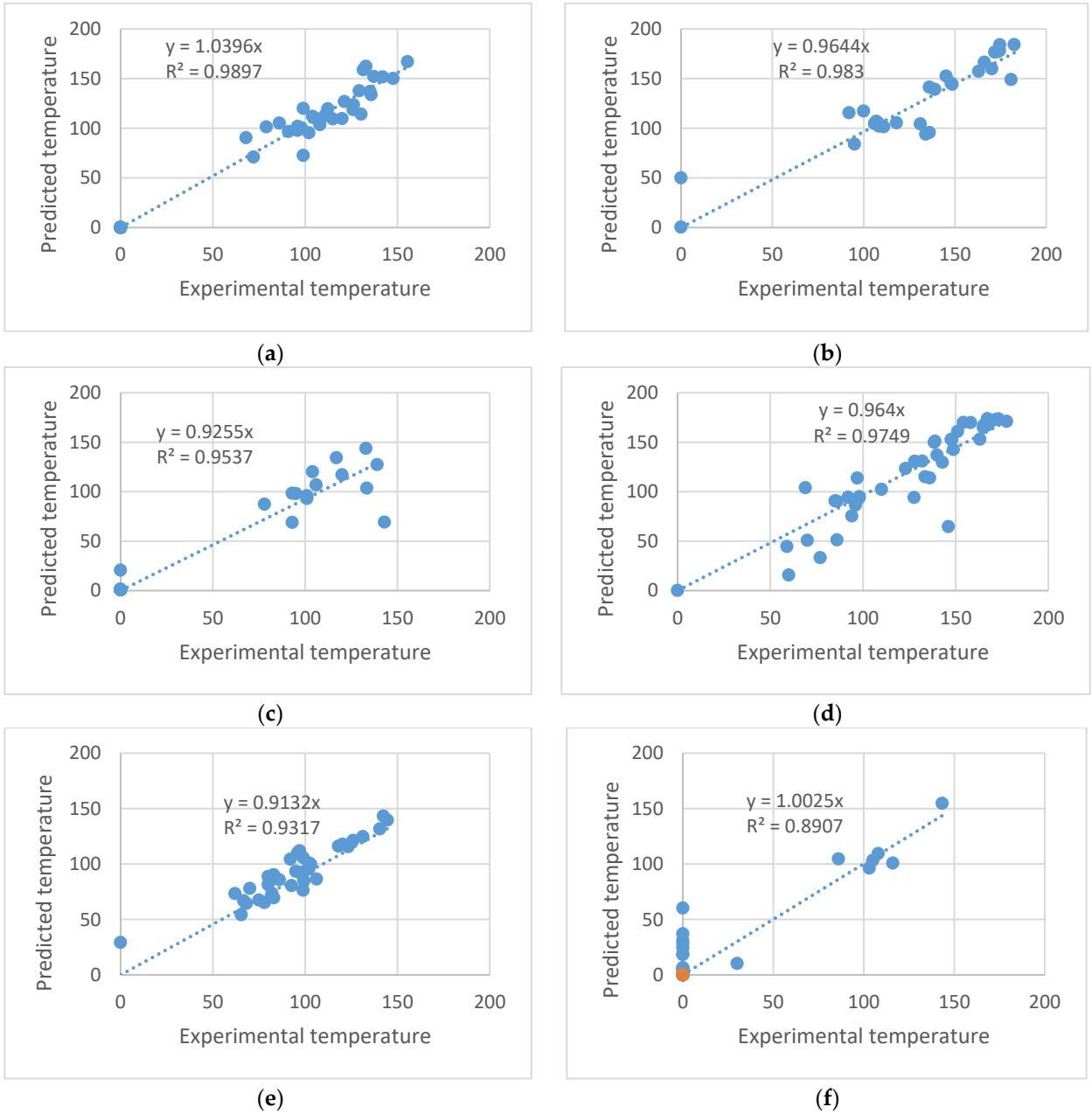


Figure 3. Experimental versus predicted temperatures (°C) for phase transition (**a**) T1; (**b**) T2; (**c**) T3; (**d**) T4; (**e**) T5; and (**f**) T6.

The SHAP analysis for the T2 model (Figure S2, Supplementary Information) reveals a different order of the descriptors than that of T1, except Feature 21, which again has the most significant influence. In this case, a higher relevance of the HOMO energy (Feature 12), the type of bonding (Feature 6) and the molecular mass (Feature 16) can be seen. The observations for Feature 21 are the same as for T1; its value 2 will lead to higher values of T2; otherwise, the value of T2 decreases. If the type of linking group is analyzed (Feature 6), the connection by imine (-N=CH-) group will increase the value of T2. Higher molecular mass values (Feature 16) predict lower values of T2.

As for the T1 and T2 models, the SHAP data show that F21 is the most crucial feature for modeling T3 transition temperature (Figure S3, Supplementary Information), maintaining the behavior observed in T1 and T2. Then in descending order of descriptor relevance in model prediction, that the following is found: HOMO energy (Feature 12), connection type (Feature 6), substitution in position S2 (Feature 3), LUMO energy (Feature 13), substitution

in S4 and S7 (Features 2 and 4) and entropy (Feature 18). High values of the HOMO energy (Feature 12), as in T2, positively impact T3. On the other hand, if the connection (Feature 6) is made through the imine (-N=CH-) group, then the predicted value of T3 will be lower.

Feature 21 significantly decreases its relevance in predicting T4 (Figure S4, Supplementary Information), ranking fifth compared to the previous cases. At first glance, the ranking of the stiff segment length descriptor (Feature 19) on position 1 seems surprising. In fact, the stiff segment length changes with the type of connection (Feature 6), a descriptor that also ranks high in the SHAP analysis. Further, also crucial in predicting the T4 transition temperature are the LUMO and HOMO energies (Features 13 and 12) and substitution in S8 and S4.

For the T5 model (Figure S5, Supplementary Information), Feature 21 is no longer in the first place. Instead, the most essential parameter this time is the LUMO energy (F13), followed by Feature 21, rigid segment length (Feature 19), bending angle (Feature 9), enthalpy (Feature 15), substitutions in positions 4 and 5 (Features 2 and 1) and dipole moment (Feature 14).

For the T6 model (Figure S6, Supplementary Information), Feature 21 is the most influential input. In some descriptors, chaotic behavior is observed, such as in the bond type (Feature 6) and the substitution at position S4 (Feature 2), where their large values influence the temperature prediction positively and negatively, respectively. We attribute this to implementation restrictions. It is noted that improving the quality of this model can be achieved by introducing more information from the experimental level.

It was found that the influence of the features on the T1–T6 model output is different from case to case. Although some features seem to have a higher importance in relation to others, their order varies from model to model, showing that the considered inputs distinctly influence the phase transition temperatures. These features are F21, substitution at positions 4 and 7, HOMO energy, LUMO energy, dipole moment and type of linking group, which can be correlated with the length of the rigid segment. It can be stated that the phase transition temperatures are influenced by lateral substituents on the central core in position 4 and by lateral substituents on the calamitic wings (in position 7), by the dipole moment of the mesogens and by the type of linking group. According to the results of the SHAP analysis, the flexible segment length, arm length, molecule length, van der Waals area or volume and molecular weight had the most negligible impact on the phase transition prediction. These descriptions are probably considerably more pertinent for a certain mesogen's self-assembly that generates a particular kind of mesophase.

The resulting ANN for all types of modeling performed in this work and the script that allows generating predictions for different values for the considered features (corresponding to other compounds than the ones used in this work) can be found at: https://elenadragoi.ro/CV/Documents/ModelsAndScripts.7z (accessed on 1 February 2023). Through their use, different predictions can be performed for additional compounds from the considered class, thus directing the focus toward the ones with desired properties.

4. Conclusions

The phase transition temperatures of five-ring bent-core molecules derived from resorcinol were predicted using neural models developed by a neuro-evolutive approach based on a differential evolution algorithm. The uniqueness of this work's findings arises from the fact that the identified models allow one to determine whether or not liquid crystalline characteristics are present and predict all the possible transition temperatures. This problem was challenging, bearing in mind the proposed strategy has never been tested until now. The structural, topological, thermophysical and electronic descriptors were obtained after the minimization of the structures at the DFT level. To reach the objectives set, two ANN modeling steps were required. In the first phase, the type and behavior of BC compounds were determined (property of liquid crystal, monotropic or enantiotropic behavior). All these aspects were included in the F21 feature that was further used in combination with the F0-F20 features in the second phase to determine the individual temperature models.

About 70% of the data were randomly assigned to training and 30% to testing. From the training data, 20% were designated for model validation. Regarding the analysis of features, only their importance was analyzed, and no feature selection strategy was applied to reduce the number of features. The idea was to use the same inputs for all temperatures and to be able to make a comparison in the same conditions. The best performance (indicated by R^2) was obtained for T1 and the lowest for T6. Finally, the analysis of features' influence on the model predictions was performed using the SHAP method. The results showed that the highest impact on all models except those for T4 (liquid–liquid crystal transition) and T5 (crystallization temperature) was for F21. Although some features seem more significant than others, and their order varies from model to model, showing that the considered inputs distinctly influence the temperature models. In addition, it was found that phase transition temperatures are influenced by the lateral substituents attached to the outer ring and on the central core, the dipole moment and the linking group type. According to the results of the SHAP analysis, the flexible segment length, arm length, molecule length, van der Waals area or volume and molecular weight had the least impact on the phase transition prediction, being more critical in the packing interactions.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cryst13040583/s1, Table S1: The molecular formula of the bent-core compounds included in the database; Equations (S1)–(S3)—equations of the accuracy indicators; Figure S1: the SHAP values for T1 model; Figure S2: the SHAP values for T2 model; Figure S3: the SHAP values for T3 model; Figure S4: the SHAP values for T4 model; Figure S5: the SHAP values for T5 model; Figure S6: the SHAP values for T6 model.

Author Contributions: Conceptualization, E.-L.E.; methodology, E.-L.E. and E.N.D.; software, E.N.D., R.P., T.V. and E.-L.E.; validation, E.N.D. and E.-L.E.; resources, E.N.D., R.P., T.V. and E.-L.E.; data curation, E.N.D., I.C. and E.-L.E.; writing—original draft preparation, E.N.D., I.C. and E.-L.E.; writing—review and editing, E.N.D., I.C. and E.-L.E.; visualization, E.-L.E.; supervision, E.-L.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data generated or analyzed during this study are included in this published article and its Supplementary Materials.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ANN—artificial neural networks; BC—bent-core; DE—differential evolution; DFT—density functional theory; HOMO—highest occupied molecular orbital; LUMO—lowest unoccupied molecular orbital; MSE—mean squared error; SHAP—SHapley Additive exPlanations; MAE—mean absolute error; MAPE—mean absolute percentage error; and RMSE—root mean squared error.

References

1. Alaasar, M. Azobenzene-containing bent-core liquid crystals: An overview. *Liq. Cryst.* **2016**, *43*, 2208–2243. [CrossRef]
2. Alaasar, M.; Prehm, M.; Poppe, S.; Tschierske, C. Development of Polar Order by Liquid-Crystal Self-Assembly of Weakly Bent Molecules. *Chem. A Eur. J.* **2017**, *23*, 5541–5556. [CrossRef] [PubMed]
3. Etxebarria, J.; Ros, M.B. Bent-core liquid crystals in the route to functional materials. *J. Mater. Chem.* **2008**, *18*, 2919–2926. [CrossRef]
4. Khan, R.K.; Turlapati, S.; Begum, N.; Mohiuddin, G.; Rao, N.V.S.; Ghosh, S. Impact of terminal polar substitution on elastic, electro-optic and dielectric properties of four-ring bent-core nematic liquid crystals. *RSC Adv.* **2018**, *8*, 11509–11516. [CrossRef] [PubMed]
5. Kohout, M.; Alaasar, M.; Poryvai, A.; Novotná, V.; Poppe, S.; Tschierske, C.; Svoboda, J. Photosensitive bent-core liquid crystals based on methyl substituted 3-hydroxybenzoic acid. *RSC Adv.* **2017**, *7*, 35805–35813. [CrossRef]
6. Pelzl, G.; Diele, S.; Weissflog, W. Banana-Shaped Compounds—A New Field of Liquid Crystals. *Adv. Mater.* **1999**, *11*, 707–724. [CrossRef]

7. Reddy, R.A.; Tschierske, C. Bent-core liquid crystals: Polar order, superstructural chirality and spontaneous desymmetrisation in soft matter systems. *J. Mater. Chem.* **2006**, *16*, 907–961. [CrossRef]
8. Takezoe, H.; Takanishi, Y. Bent-Core Liquid Crystals: Their Mysterious and Attractive World. *Jpn. J. Appl. Phys.* **2006**, *45*, 597–625. [CrossRef]
9. Weissflog, W.; Naumann, G.; Kosata, B.; Schröder, M.W.; Eremin, A.; Diele, S.; Vakhovskaya, Z.; Kresse, H.; Friedemann, R.; Krishnan, S.A.R.; et al. Ten isomeric five-ring bent-core mesogens: The influence of the direction of the carboxyl connecting groups on the mesophase behaviour. *J. Mater. Chem.* **2005**, *15*, 4328–4337. [CrossRef]
10. Yelamaggad, C.V.; Mathews, M.; Nagamani, S.A.; Rao, D.S.S.; Prasad, S.K.; Findeisen, S.; Weissflog, W. A novel family of salicylaldimine-based five-ring symmetric and non-symmetric banana-shaped mesogens derived from laterally substituted resorcinol: Synthesis and characterization. *J. Mater. Chem.* **2007**, *17*, 284–298. [CrossRef]
11. Ting, T.X.; Sarjadi, M.S.; Rahman, M.L. Influences of Central Units and Terminal Chains on the Banana-Shaped Liquid Crystals. *Crystals* **2020**, *10*, 857. [CrossRef]
12. Ciobanu, C.I.; Berladean, I.; Epure, E.-l.L.; Simion, A.; Lisa, G.; Boussoualem, Y.; Carlescu, I. Mesomorphic and Thermal Behavior of Symmetric Bent-Core Liquid Crystal Compounds Derived from Resorcinol and Isophthalic Acid. *Crystals* **2021**, *11*, 1215. [CrossRef]
13. Dunemann, U.; Schröder, M.W.; Reddy, R.A.; Pelzl, G.; Diele, S.; Weissflog, W. The influence of lateral substituents on the mesophase behaviour of banana-shaped mesogens. Part II. *J. Mater. Chem.* **2005**, *15*, 4051–4061. [CrossRef]
14. Ionica Ciobanu, C.; Drochioiu, G.; Carlescu, I.; Lisa, G.; Antoci, V.; Vasilache, V.; Scutaru, D. Thermal Behavior of Some Bent-Core Resorcinol Derivatives with Azo-Type Spacers and Variable Flexible Chain. *Lett. Org. Chem.* **2016**, *13*, 156–161. [CrossRef]
15. Šmahel, M.; Poryvai, A.; Xiang, Y.; Pociecha, D.; Troha, T.; Novotná, V.; Svoboda, J.; Kohout, M. Photosensitive bent-core nematic liquid crystals with various linking units in the side arms: Structure-properties relationships. *J. Mol. Liq.* **2020**, *306*, 112743. [CrossRef]
16. Weissflog, W.; NÃ¡dasi, H.; Dunemann, U.; Pelzl, G.; Diele, S.; Eremin, A.; Kresse, H. Influence of lateral substituents on the mesophase behaviour of banana-shaped mesogens. *J. Mater. Chem.* **2001**, *11*, 2748–2758. [CrossRef]
17. Weissflog, W.; Shreenivasa Murthy, H.N.; Diele, S.; Pelzl, G. Relationships between molecular structure and physical properties in bent-core mesogens. *Philos. Trans. A Math. Phys. Eng. Sci.* **2006**, *364*, 2657–2679. [CrossRef]
18. Shen, D.; Tschierske, C.; Diele, S.; Wirt, I. A novel class of non-chiral banana-shaped liquid crystals with ferroelectric properties. *Chem. Commun.* **1998**, *23*, 2573–2574. [CrossRef]
19. Alaasar, M.; Prehm, M.; Tschierske, C. Mirror symmetry breaking in fluorinated bent-core mesogens. *RSC Adv.* **2016**, *6*, 82890–82899. [CrossRef]
20. Keith, C.; Lehmann, A.; Baumeister, U.; Prehm, M.; Tschierske, C. Nematic phases of bent-core mesogens. *Soft Matter* **2010**, *6*, 1704. [CrossRef]
21. Kohout, M.; Kozmík, V.; Slabochová, M.; Tůma, J.; Svoboda, J.; Novotná, V.; Pociecha, D. Bent-shaped liquid crystals based on 4-substituted 3-hydroxybenzoic acid central core. *Liq. Cryst.* **2015**, *42*, 87–103. [CrossRef]
22. Achten, R.; Smits, E.A.W.; Reddy, R.A.; Giesbers, M.; Marcelis, A.T.M.; Sudhölter, E.J.R. Monofluorinated unsymmetrical bent-core mesogens. *Liq. Cryst.* **2006**, *33*, 57–65. [CrossRef]
23. Amaranatha Reddy, R.; Sadashiva, B.K. Influence of fluorine substituent on the mesomorphic properties of five-ring ester banana-shaped molecules. *Liq. Cryst.* **2003**, *30*, 1031–1050. [CrossRef]
24. Dantlgraber, G.; Shen, D.; Diele, S.; Tschierske, C. Antiferroelectric switchable mesophases of nonchiral bent-core liquid crystals containing fluorinated central cores. *Chem. Mater.* **2002**, *14*, 1149–1158. [CrossRef]
25. Heppke, G.; Parghi, D.D.; Sawade, H. A laterally fluoro-substituted "banana-shaped" liquid crystal showing antiferroelectricity. *Ferroelectrics* **2000**, *243*, 269–276. [CrossRef]
26. Korblova, E.; Walba, D.M.; Gong, T.; Reddy, A.; Zhu, C.; Shao, R.; Maclennan, J.E.; Glaser, M.A.; Clark, N.A. Design and synthesis of an achiral ferroelectric smectic liquid crystal. *Liq. Cryst. XV. SPIE* **2011**, *8114*, 160–168.
27. Mieczkowski, J.; Gomola, K.; Koseska, J.; Pociecha, D.; Szydlowska, J.; Gorecka, E. Liquid crystal phases formed by asymmetric bent-shaped molecules. *J. Mater. Chem.* **2003**, *13*, 2132–2137. [CrossRef]
28. Murthy, H.N.S.; Sadashiva, B.K. Influence of a fluorine substituent on the mesomorphic properties of unsymmetrical five-ring bent-core compounds. *J. Mater. Chem.* **2004**, *14*, 2813–2821. [CrossRef]
29. Nádasi, H.; Weissflog, W.; Eremin, A.; Pelzl, G.; Diele, S.; Das, B.; Grande, S. Ferroelectric and antiferroelectric "banana phases" of new fluorinated five-ring bent-core mesogens. *J. Mater. Chem.* **2002**, *12*, 1316–1324. [CrossRef]
30. Murthy, H.N.S.; Sadashiva, B.K. Ferroelectric and antiferroelectric switching behaviour in new unsymmetrical bent-core compounds derived from 3-hydroxybenzoic acid. *J. Mater. Chem.* **2005**, *15*, 2056–2064. [CrossRef]
31. Yelamaggad, C.V.; Shashikala, I.S.; Hiremath, U.S.; Liao, G.; Jakli, A.; Rao, D.S.S.; Prasad, S.K.; Li, Q. Fluorine containing nonsymmetrical five-ring achiral banana-shaped compounds with columnar and synclinic antiferroelectric layered phases. *Soft Matter* **2006**, *2*, 785–792. [CrossRef] [PubMed]
32. Alaasar, M.; Prehm, M.; Brautzsch, M.; Tschierske, C. 4-Methylresorcinol based bent-core liquid crystals with azobenzene wings—A new class of compounds with dark conglomerate phases. *J. Mater. Chem. C* **2014**, *2*, 5487–5501. [CrossRef]
33. Kovalenko, L.; Schröder, M.W.; Amaranatha Reddy, R.; Diele, S.; Pelzl, G.; Weissflog, W. Unusual mesomorphic behaviour of new bent-core mesogens derived from 4-cyanoresorcinol. *Liq. Cryst.* **2005**, *32*, 857–865. [CrossRef]

34. Epure, E.-L.; Oniciuc, S.D.; Hurduc, N.; Drăgoi, E.N. Artificial Neural Network Modeling of Glass Transition Temperatures for Some Homopolymers with Saturated Carbon Chain Backbone. *Polymers* **2021**, *13*, 4151. [CrossRef]
35. Antanasijević, J.; Antanasijević, D.; Pocajt, V.; Trišović, N.; Fodor-Csorba, K. A QSPR study on the liquid crystallinity of five-ring bent-core molecules using decision trees, MARS and artificial neural networks. *RSC Adv.* **2016**, *6*, 18452–18464. [CrossRef]
36. Antanasijević, J.; Pocajt, V.; Antanasijević, D.; Trišović, N.; Fodor-Csorba, K. Prediction of clearing temperatures of bent-core liquid crystals using decision trees and multivariate adaptive regression splines. *Liq. Cryst.* **2016**, *43*, 1028–1037. [CrossRef]
37. Xu, J.; Wang, L.; Zhang, H.; Yi, C.; Xu, W. Accurate quantitative structure–property relationship analysis for prediction of nematic transition temperatures in thermotropic liquid crystals. *Mol. Simul.* **2010**, *36*, 26–34. [CrossRef]
38. Cao, Y.; Yu, J.; Song, H.; Wang, X.; Yao, S. Prediction of electric conductivity for ionic liquids by two chemometrics methods. *J. Serb. Chem. Soc.* **2013**, *78*, 653–667. [CrossRef]
39. Scutaru, D.; Carlescu, I.; Bulai, E.-R.; Ciobanu, C.I.; Lisa, G.; Hurduc, N. *Bent-Core Liquid Crystals: Structures and Mesomorphic Properties. Liquid Crystals-Self-Organized Soft Functional Materials For Advanced Applications*; BoD–Books on Demand: Norderstedt, Germany, 2018.
40. Alaasar, M.; Prehm, M.; May, K.; Eremin, A.; Tschierske, C. 4-cyanoresorcinol-based bent-core mesogens with azobenzene wings: Emergence of sterically stabilized polar order in liquid crystalline phases. *Adv. Funct. Mater.* **2014**, *24*, 1703–1717. [CrossRef]
41. Yelamaggad, C.V.; Hiremath, U.S.; Nagamani, S.A.; Rao, D.S.S.; Prasad, S.K. A switchable salicylaldimine-based achiral bent-shaped mesogen: Synthesis and characterization. *J. Mater. Chem.* **2001**, *11*, 1818–1822. [CrossRef]
42. Nagaveni, N.; Roy, A.; Prasad, V. Achiral bent-core azo compounds: Effect of different types of linkage groups and their direction of linking on liquid crystalline properties. *J. Mater. Chem.* **2012**, *22*, 8948–8959. [CrossRef]
43. Alaasar, M.; Prehm, M.; Tschierske, C. Influence of halogen substituent on the mesomorphic properties of five-ring banana-shaped molecules with azobenzene wings. *Liq. Cryst.* **2013**, *40*, 656–668. [CrossRef]
44. Alaasar, M.; Prehm, M.; Brautzsch, M.; Tschierske, C. Dark conglomerate phases of azobenzene derived bent-core mesogens—Relationships between the molecular structure and mirror symmetry breaking in soft matter. *Soft Matter* **2014**, *10*, 7285–7296. [CrossRef]
45. Becke, A.D. Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals. *J. Chem. Phys.* **1997**, *107*, 8554–8560. [CrossRef]
46. Becke, A.D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652. [CrossRef]
47. Lee, C.; Yang, W.; Parr, R.G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B Condens Matter.* **1988**, *37*, 785–789. [CrossRef] [PubMed]
48. BIOVIA Materials Studio. Materials Studio. Available online: https://www.3ds.com/products-services/biovia/products/molecular-modeling-simulation/biovia-materials-studio/ (accessed on 24 February 2023).
49. Chen, C.Y. Evolving binary-weights neural network using hybrid optimization algorithm for color space conversion. *Sci. Iran.* **2015**, *22*, 1625–1634.
50. Storn, R.; Price, K. *Differential Evolution: A Simple and Efficient Adaptive Scheme for Global Optimization Over Continuous Spaces*; Berkley: Greenwich, CT, USA, 1995.
51. Dragoi, E.-N.; Curteanu, S.; Galaction, A.-I.; Cascaval, D. Optimization methodology based on neural networks and self-adaptive differential evolution algorithm applied to an aerobic fermentation process. *Appl. Soft Comput.* **2013**, *13*, 222–238. [CrossRef]
52. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
53. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 4768–4777.
54. Faure, G.; Delgado-Buscalioni, R.; Español, P. The entropy of a complex molecule. *J. Chem. Phys.* **2017**, *146*, 224106. [CrossRef] [PubMed]